Lecture Notes in Networks and Systems 759

Cengiz Kahraman · Irem Ucal Sari · Basar Oztaysi · Selcuk Cebi · Sezi Cevik Onar · A. Çağrı Tolga *Editors*

Intelligent and Fuzzy Systems

Intelligence and Sustainable Future Proceedings of the INFUS 2023 Conference, Volume 2

Springer

Lecture Notes in Networks and Systems 759

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Cengiz Kahraman · Irem Ucal Sari ·
Basar Oztaysi · Selcuk Cebi · Sezi Cevik Onar ·
A. Çağrı Tolga
Editors

Intelligent and Fuzzy Systems

Intelligence and Sustainable Future
Proceedings of the INFUS 2023 Conference, Volume 2

Springer

Editors
Cengiz Kahraman
Department of Industrial Engineering
Istanbul Technical University
Istanbul, Türkiye

Irem Ucal Sari
Department of Industrial Engineering
Istanbul Technical University
Istanbul, Türkiye

Basar Oztaysi
Department of Industrial Engineering
Istanbul Technical University
İstanbul, Türkiye

Selcuk Cebi
Department of Industrial Engineering
Yildiz Technical University
Istanbul, Türkiye

Sezi Cevik Onar
Department of Industrial Engineering
Istanbul Technical University
Istanbul, Türkiye

A. Çağrı Tolga
Department of Industrial Engineering
Galatasaray University
Istanbul, Türkiye

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-031-39776-9 ISBN 978-3-031-39777-6 (eBook)
https://doi.org/10.1007/978-3-031-39777-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

INFUS is an acronym for intelligent and fuzzy systems. INFUS 2019 was an on-site conference organized in Istanbul, Türkiye. INFUS 2020 and INFUS 2021 conferences were organized as online conferences because of pandemic conditions. INFUS 2022 conference was organized as both online and on-site conference in Izmir with the cooperation of Yasar University and Izmir Bakircay University. INFUS 2023 is the fifth conference of this series organized by Istanbul Technical University.

The theme of INFUS 2023 conference this year is *Intelligent and Sustainable Future*. Intelligence can be used in a wide variety of ways to manage environmental impacts and climate change such as clean sustainable supply chains, environmental monitoring and enforcement, and advanced weather and disaster forecasting. The emergence of artificial intelligence (AI) and its growing impact on many industries require research into how it can be used to achieve the Sustainable Development Goals. Applications of AI can create a sustainable and eco-friendly future. AI is a promising tool for the production of new materials that help in building a sustainable environment. The sustainability of biological diversity is another very important problem since many animal species are extinct or endangered. Therefore, intelligence can be used to study animal behavior patterns. Soil pollution is another important problem today as population growth, intensive farming, and other activities increase day by day. Since food production is the key to sustain human life, we can maintain environmental sustainability by monitoring crops and soils and maximize the crop yields, while having less impact on the environment through AI-augmented agriculture. The excessive consumption of natural resources by humans has a detrimental effect on water resources. The level of garbage accumulating in the oceans is higher than ever before. Artificial intelligence tools should be used to ensure environmental sustainability. Artificial intelligence can be used in automated garbage collection vehicles; it can help solve problems such as illegal fishing and discharge of industrial wastewater into water bodies and illegal dumping of solid wastes into the seas. The use of intelligence for a livable future has become a necessity. A program focusing on intelligence and sustainability future, which is the theme of this year's INFUS 2023 conference, is foreseen. INFUS 2023 aims to bring together the latest theoretical and practical intelligent and fuzzy studies on sustainable future in order to create a discussion environment.

Researchers from more than 30 countries such as Türkiye, Russia, China, Iran, Poland, India, Azerbaijan, Bulgaria, Spain, Ukraine, Pakistan, South Korea, UK, Indonesia, USA, Vietnam, Finland, Romania, France, Uzbekistan, Italy, and Austria contributed to INFUS 2023. Our invited speakers this year are Prof. Krassimir Atanassov, Prof. Vicenc Torra, Prof. Janusz Kacprzyk, Prof. Ahmet Fahri Özok, and Prof. Ajith Abraham, and Prof. Irina Perfilieva. It is an honor to include their invaluable speeches in our conference program. We appreciate their voluntary contributions to INFUS 2023, and we hope to see them at INFUS conferences for many years. This year, the number of submitted papers became 291. After the review process, about 40% of these papers have been rejected. More than 50% of the accepted papers are from other countries outside Türkiye.

We again thank all the representatives of their countries for selecting INFUS 2023 as an international scientific arena to present their valuable research results. We are honored and aware of our responsibility that our participants have chosen us in a highly competitive environment with hundreds of conferences in the same field and organized in close dates to each other. INFUS conference manages high-cost international conference participation processes for the benefit of the participants, with lower registration fees but more well-known expert invitations and rich social activities.

We also thank the anonymous reviewers for their hard works in selecting high-quality papers of INFUS 2023. Each of the organizing committee members provided invaluable contributions to INFUS 2023. INFUS conferences would be impossible without their efforts. We hope meeting all of our participants next year in Türkiye one more time with a new research theme at a new city and new social activities.

We would like to thank our publisher Springer Publishing Company, Series Editor Prof. Janusz Kacprzyk, Interdisciplinary and Applied Sciences and Engineering, and Editorial Director Thomas Ditzinger, last but not least, Project Coordinator Nareshkumar Mani for their supportive, patient, and helpful roles during the preparation of this book.

Cengiz Kahraman
Irem Ucal Sari
Basar Oztaysi
Selcuk Cebi
Sezi Cevik Onar
A. Çağrı Tolga

Organization

Program Committee Chairs

Cebi, Selcuk	Yildiz Technical University, Industrial Engineering, Istanbul, Türkiye
Kahraman, Cengiz	ITU, Industrial Engineering Department, Istanbul, Türkiye
Cevik Onar, Sezi	Istanbul Technical University, Istanbul, Türkiye
Oztaysi, Basar	Istanbul Technical University, Istanbul, Türkiye
Tolga, Çağrı	Galatasaray University, Department of Industrial Engineering, Istanbul, Türkiye
Ucal Sari, Irem	Istanbul Technical University, Industrial Engineering, Istanbul, Türkiye

Program Committee Members

Alkan, Nurşah	Istanbul Technical University, Department of Industrial Engineering, Maçka, Beşiktaş, Türkiye
Aydin, Serhat	National Defence University, Industrial Engineering Department, Istanbul, Türkiye
Boltürk, Eda	Istanbul Technical University, Istanbul, Türkiye
Cebi, Selcuk	Yildiz Technical University, Industrial Engineering, Istanbul, Türkiye
Dogan, Onur	Izmir Bakircay University, Department of Industrial Engineering, İzmir, Türkiye
Haktanır Aktaş, Elif	Bahcesehir University, Industrial Engineering, Istanbul, Türkiye
Kahraman, Cengiz	ITU, Industrial Engineering Department, Istanbul, Türkiye
Kutlu Gündoğdu, Fatma	National Defence University, Industrial Engineering Department, Istanbul, Türkiye
Otay, Irem	Istanbul Bilgi University, Department of Industrial Engineering, Istanbul, Türkiye
Cevik Onar, Sezi	Istanbul Technical University, Istanbul, Türkiye

Oztaysi, Basar	İstanbul Technical University, Istanbul, Türkiye
Tolga, Çağrı	Galatasaray University, Department of Industrial Engineering, Istanbul, Türkiye
Ucal Sari, Irem	Istanbul Technical University, Industrial Engineering, Istanbul, Türkiye

Contents

Artificial Intelligence

Monitoring Sleep-Related Breathing Disorders with Optical Flow

Olena Mikhnova(✉)

State Biotechnological University, Kharkiv, Ukraine
mikhnova@yahoo.co.uk

Abstract. Sleep-related disorders often appear in neurological practice. To detect the reasons for these states, overnight EEG monitoring is widely implemented, which along with its expense is quite hard to perform for children and elderly people. In cases where it may be replaced by video monitoring only, it should be done so to eliminate extra stress on patients. The author has analyzed current developments in this medical field, including devices, software and methods used for visual data processing. In the current paper a novel application to the technique based on optical flow video analysis has been proposed to find segments in video frames with abnormal flow changes for further experts engaging. By comparing the results obtained from intelligent video processing with the results from EEG monitoring, great correspondence of data has been found in terms of transitions that contain important episodes. Thus, the proposed method can significantly reduce working hours of high qualified doctors and eliminate unnecessary analyses or reveal previously unknown details about the existing disease.

Keywords: EEG monitoring · intelligent video analysis · medical data processing · optical flow · sleep-related disorders

1 Introduction

An area of neurobiology and medicine dedicated to sleep research, sleep disorders and their treatment, their effect on human health (somnology) is developing extensively during the last three decades [1, 2]. In part, this is because patients with such a severe neurological disorder as epilepsy often have to deal with sleep disorders as well. Sleep disorders related with movement (cramp, bruxism, leg muscle contraction, sleep startles, smacking, gulping, etc.), vocalization (e.g., groaning, snoring), breath (obstructive and central sleep apnea) and behaviour are very peculiar to epileptic seizures [3].

Diagnosis is made based on patient's medical history and polysomnography. Existing methods for apnea diagnostics are viewed in [4]. Video EEG monitoring implies installation of two video cameras in a patient's room: a camera for general view and a camera for the foreground, which are further used for detailed analysis of various movements and vocalizations in synchronization with EEG and data required from polysomnography sensors. Doctor has to analyze EEG pattern, polysomnography results, and to monitor constantly patient's activity during the night. While a human biorhythm is accustomed

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 3–10, 2023.
https://doi.org/10.1007/978-3-031-39777-6_1

to have a rest, high concentration of attention is required. This is one more reason of high cost for such medical examinations. The diagnosis and the quality of the whole further patient's life greatly depend on the doctor's qualification, who performs such an examination.

But, in case of breathing disorders monitoring, without any relation to electroencephalography, such conditions are easily detected with pulse oxymeter. Periods of blood saturation decrease with oxygen (desaturation) may be detected to make judgements on frequency and severity of breathing interruption (apnea). But, short-term breathing interruptions cannot be detected during the sleep phase as the oxygen remains in tissues. Thus, such method is not always applicable and requires sensors installation as in case of encephalography [5].

The goal of the current paper is to provide a novel approach for constant video control and automatic analysis of human body state in terms of breathing and brain saturation with oxygen with the help of optical flow video processing technology. The reasons for using specifically this artificial intelligence method and the existing methods are reviewed in the next section. The proposed solution to the current problem is described in Sect. 3. Practical results are given in Sect. 4 and future enhancement to the method is given in the last section of the paper.

2 Existing Methods and the Proposed Solution

Image and video processing is a rapidly developing field of research [6]. Though, many of recently appeared techniques rely on concepts that were investigated during the previous century, namely clustering and various classification schemes, advanced filters [7–9]. It is evident that the best artificial intelligence results obtained during the last decades in image processing find their application in medicine and biology [10].

In particular, much work has already been done in the field of video analysis techniques to monitor sleep disorders. 10 existing methods of machine learning for diagnostic and monitoring of apnea are observed in [4]. 5 other methods are mentioned in [11], among which are the following: variance analysis, gradient sum, spectral analysis, wavelet transform. Precisely in the latter paper, the authors have used only face region tracking for detection. They have compared their method with all the aforementioned, which turned to be better, though the rest body parts can bear crucial information as well. Their method has not yet been validated in clinical practice.

Assumption concerning night apnea syndrome can be made based on human photo analysis. Australian specialists have developed diagnostic model of high precision that uses 3D-photography to predict whether a person has an apnea. They assumed particular facial and neck characteristics: neck width, width and length of the lower jaw bone, face width and distance between eyes, presence of abnormalities. Upon data analysis, the researchers have found such facial characters that are peculiar to people with apnea, they have united these critical features into a single predicting algorithm. Their research reveals that the night apnea syndrome can be diagnosticated with 98% of accuracy [12].

Similar to the current research, the one has been made in Spain [13]. But here, the authors use numerical thresholds estimated from the results of polysomnography as the total number of events and sleep time. The calculation is very simple. They assume that the amount of airflow is defined as the relative maximum minus the relative minimum of an airflow, divided by the duration of fall multiplied by the local average of the air flow. Thus, they compare this characteristic with a certain threshold all the time to detect a respiratory event. We propose a bit more sophisticated approach based on pixel brightness derivative changes. We observe these minor changes constantly, but when they stop happening, it symbolizes breathing abnormality. In video processing field, this is called an optical flow analysis.

By now there are many algorithms for optical flow computation. First, this is the algorithm of Lucas and Kanade (1981), algorithm of Horn and Schunck (1981), algorithm of Black and Anandan (1996), and there are many others. The full list can be found in Middleburry estimation data base where the quality rank for each of currently existing methods is given (in accordance to different types of video data) along with their performance estimation that is computed without any consideration to normalization of computer processor characteristics and other hardware accelerators [14]. This data base is continuously updated. By December 2009 it consisted of 24 algorithms, by December 2012 it already had 77 algorithms available, and by the end of 2013 90 algorithms were available in its arsenal. By the time this article has been preparing for printing (September 2021), 191 algorithms were presented there. One peculiarity inherent to this data base is that all the algorithms present there are tested on one and the same video samples, which gives a possibility of their correct comparison with each other by different parameters.

3 Optical Flow Implementation in Overnight Monitoring

The idea that lies underneath an optical flow paradigm grounds upon relatively permanent intensity in a local region of an image when the latter is changing in time. Thus, if a great change of intensity occurs in some local region, this will mean movement of some particular image segment. In other words, image segment correlation is considered in terms of their intensity change from frame to frame. In a general form, optical flow equation can be written as follows (1). The next formula describes optical flow of motion between pixels of consequent video frames that are shot at time t and $t + \Delta t$.

$$\begin{gathered} I(x + \Delta x, y + \Delta y, t + \Delta t) \\ \approx I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t, \end{gathered} \tag{1}$$

Δ means time increment and movement of corresponding coordinates between frames; $I(x, y, t)$ is the intensity of a pixel with coordinates (x, y).

When the velocity of intensity change in a number of images is homogeneous, motion is hardly estimated by optical flow. This concept allows evaluation of significant changes only. Moreover, camera flashes may have negative impact as the method significantly relies on intensity. But when we consider medical imagery with artificial light and quite constant characteristics of background brightness and camera flashes, this is a real plus for this method to be taken as the basis.

A number of approaches to optical flow computation are described in [15] along with their classification. Despite their similarity, only a few are applicable to the current medical problem and here are the reasons. First, presence of multiple constraints to achieve better results enhances the number of equations, complicating the computations. But on the other hand, image background is unchangeable in medical imagery that gives a way to use less sophisticated methods to reduce time and processing resources needed.

Such intensity-based differential methods as global and local methods can be successfully applied to video EEG monitoring as they derive moving pixel velocity from spatiotemporal intensity derivatives. Horn and Schunck algorithm belongs to this group. It assumes that the intensity of pixels from the same region of interest remains almost unchanged during the objects movement. But, motion may result in image feature change in a local region during an affine transformation. Therefore, there should be some degree of variation to the pixel intensities inside the same region:

$$\int_D \left((\nabla I v + I_t)^2 + \lambda^2 tr\left((\nabla v)^T \nabla v \right) \right) dx \tag{2}$$

with the domain of interest D. $\nabla I = \left(I_x, I_y \right)$ and I_t are the first order partial derivatives of corresponding intensities, and $v = (u, v)$ is an image velocity.

Existence of multiple motions within a local region adds a number of equations to be solved. For instance, aperture problem can be solved by introducing the second order derivatives to pixel intensities (3) [16–18]:

$$(\nabla \nabla I) v^T = -\nabla I_t \tag{3}$$

When we obtain a matrix of image intensities with their second order derivatives that change over time along with additional constraints, the algorithm starts requiring more computational resources. To overcome this problem, Gauss divergence theorem is used. Though, there are some restrictions for its usage also. In the current case with medical imagery, we may argue that the benefit we have is the stability of background conditions and scene. To estimate optical flow in this situation, to perfect the results and ease calculations, intensity derivatives may be computed upon object contour extraction. Such an image presegmentation is done in [19, 20].

Object border detection can be done with the help of such filtering techniques as Gabor filter, supervised and unsupervised pixel classification filters, Markov random field-based filters, watershed-based segmentation filters. A bit more sophisticated techniques for object segmentation include: Fourier descriptors, Zernike moments, Freeman chain code, wavelet transformation, Roberts cross operator, Sobel operator, Kirsch operator, Prewitt operator, Laplace operator, Canny edge detector. As an optical flow itself, many of the aforementioned techniques are based on intensity gradient analysis.

Analysis of existing optical flow detection based on segmented regions shows great scattering of flow along the object edges. In other words, different velocities tend to appear at border pixels of the same object. That is why smoothing velocity is applied at object borders to obtain some unique vector that presents full object motion [15]. The below formula (4) demonstrates commonly used approach to minimization of velocity differences along the edges of detected regions. Formula (5) computes predicted normal velocity which is subtracted from the product of estimated velocity and the unit vector $v \cdot \hat{n}$.

$$\int \left(\frac{\partial v}{\partial S} \right)^2 + \beta \left(v \cdot \hat{n} - \| v_\perp \|_2 \right)^2 dS, \tag{4}$$

$$v_{\perp} = \frac{S_t \nabla S}{\|\nabla S\|_2^2} \tag{5}$$

where β is a weight applied to the squared difference of velocities computed for the region edge S.

Such a concept of unique velocity for each image region is beneficial from the point of medical imagery where objects are clearly determined from the background. It also gives an opportunity of error elimination from missegmented regions.

4 Practical Results

In this section we show the results that can be obtained from medical video data processing upon optical flow implementation. First, we took a video of a sleeping child with homogeneous content, where the only movement that exists is caused by the child's breathing (Fig. 1). In Fig. 1, optical flow changes are demonstrated for the consecutive frames taken from video of a sleeping child. The slight strokes in Fig. 1 show whether the changes took place or not.

Fig. 1. Slight video frame changes derived from optical flow analysis

For any system of optical flow analysis, the longer the strokes are in each picture of a moving object, the greater changes occur to it. Figure 2 is an example of significant movement derived due to the same optical flow change detector applied to another video of a sleeping child, which contain more diverse and a bit more heterogeneous content.

An alternative way to key frame extraction has been demonstrated by authors of this paper in their previous research work [21] concerning significant frame detection from video using Voronoi tessellations for image partitioning with further segment comparison in terms of texture, color and shape. Currently applied optical flow approach to medical imagery is more sophisticated and does not necessarily need video frame presegmentation step. Though, many optical flow computation techniques imply presegmentation too. To be true, this step is crucial for the majority of video analysis techniques that imply artificial reasoning.

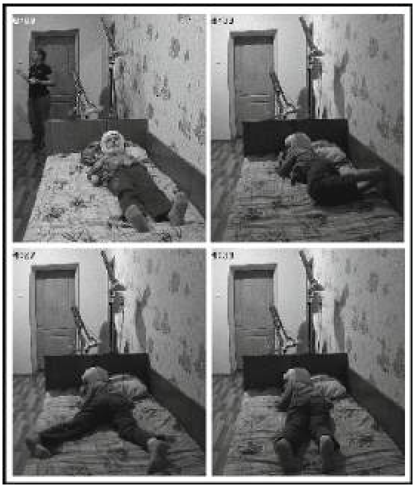


Fig. 2. Key frames extracted with significant changes derived upon optical flow analysis

The application of optical flow to medical video processing, and key frame extraction in particular, has not been tested in terms of quality in relation to other key frame extraction and video processing methods. First, it is connected with subjectivity of such tests that rely on human perception. The only thing we know for sure is that the proposed technique provides adequate results, but at the same time requires considerable computational resources. It utilized almost the same color (illumination) characteristics, but lacks texture and shape feature consideration. Though, the latter are indirectly taken into account due to extended monitoring of illumination flow change at the object edges with their transformation in time, which in some particular cases provide much more significant information than any other video processing technique that relies on a more diverse range of image features.

5 Conclusions and Future Work

The proposed solution to sleep-related breathing disorders enables patients not to be hooked up to equipment to perform their state monitoring. The technique can also be implemented as part of video EEG-monitoring in epilepsy treatment to aid medical personnel and reduce their close overnight attention to wave control. Though, comparison of intensity derivatives between pixels of consecutive frames is a procedure vulnerable to computational resources, which can be partially resolved by appropriate hardware usage. That is why despite clear benefits for implementing the proposed solution, it is hardly to be used in a near to real time mode.

As areas for the future research, the following basic directions should be named: graphical analysis procedure enhancement with future expert system development that should propose some medical treatment and provide diagnosis with a definite level of accuracy; numerical comparison of computational time for the proposed solution and other existing optical flow methods on the same hardware; verifying the proposed solution on non-medical test collections to check whether it is applicable for other domains.

Aside from medical imagery application, as an artificial intelligence technique, optical flow can also be used for image segmentation, time to collision prediction, velocity computation (time of expansion or time of reduction), motion compensated coding, any kind of flow measurements. The issues that should be primarily in focus are the illumination uniformity of the regions of interest (which movement is under consideration) and brightness constancy of the background scene to ensure pure motion estimation. All these set some restrictions to the domain of application, but it totally matches the needs for overnight monitoring.

References

1. Sateia, M.J.: International classification of sleep disorders-third edition. Chest **146**(5), 1387–1394 (2014). https://doi.org/10.1378/chest.14-0970
2. Kotterba, S.: Schlafstörungen bei neurologischen Erkrankungen. Der Nervenarzt **86**(6), 759–771 (2015). https://doi.org/10.1007/s00115-014-4204-6
3. Krenzer, M., Oertel, W., Trenkwalder, C.: Praktische Diagnostik und Therapie des Restless-legs-Syndroms. Der Nervenarzt **85**(1), 9–18 (2014). https://doi.org/10.1007/s00115-013-3888-3
4. Tkachenko, D.A., Krush, I.V., Mykhalko, V.G., Petrenko, A.I.: Machine learning for diagnosis and monitoring of sleep apnea. Syst. Res. Inf. Technol. **4**, 43–58 (2020). https://doi.org/10.20535/SRIT.2308-8893.2020.4.04
5. Duchna, H.-W.: Schlafbezogene Atmungsstörungen. Pneumologie **60**(9), 568–575 (2006). https://doi.org/10.1055/s-2006-944248
6. Rashid, F., Singh, G., Tulu, A., Abate, J.: Analysis of different image processing strategies. Turk. J. Comput. Math. Educ. **13**(01), 180–189 (2022)
7. Bodyanskiy, Ye.V., Shafronenko, A.Y., Klymova, I.N.: Online fuzzy clustering of incomplete data using credibilistic approach and similarity measure of special type. Radio Electron. Comput. Sci. Control (1), 97–104 (2021). https://doi.org/10.15588/1607-3274-2021-1-10
8. Tulyakova, N., Trofymchuk, O.: Adaptive myriad filter with time-varying noise - and signal-dependent parameters. Radioelectron. Comput. Syst. **2**, 217–238 (2022). https://doi.org/10.32620/reks.2022.2
9. Moskalenko, V., Kharchenko, V., Moskalenko, A., Petrov, S.: Model and training method of the resilient image classifier considering faults, concept drift, and adversarial attacks. Algorithms **15**(384), 1–24 (2022). https://doi.org/10.3390/a15100384
10. Dastres, R., Soori, M.: Advanced image processing systems. Int. J. Imaging Robot. **21**(1), 27–44 (2021)
11. Kopaczka, M., Özkan, Ö., Merhof, D.: Face tracking and respiratory signal analysis for the detection of sleep apnea in thermal infrared videos with head movement. In: Battiato, S., Farinella, G.M., Leo, M., Gallo, G. (eds.) ICIAP 2017. LNCS, vol. 10590, pp. 163–170. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70742-6_15
12. Eastwood, P.R., et al.: Predicting sleep apnea from 3-dimensional face photography. J. Clin. Sleep Med. **16**(4), 493–502 (2020). https://doi.org/10.5664/jcsm.8246
13. Abad, J., et al.: Automatic video analysis for obstructive sleep apnea diagnosis. SLEEP **39**(8), 1507–1515 (2016). https://doi.org/10.5665/sleep.6008
14. Baker, S., Scharstein, D., Lewis, J.P., Roth, S.: A database and evaluation methodology for optical flow. Int. J. Comput. Vision **92**(1), 1–31 (2011). https://doi.org/10.1007/s11263-010-0390-2

15. Beauchemin, S.S., Barron, J.L.: The computation of optical flow. ACM Comput. Surv. **27**(3), 433–467 (1995). https://doi.org/10.1145/212094.212141
16. Uras, S., Girosi, F., Verri, A., Torre, V.: A computational approach to motion perception. Biol. Cybern. **60**, 79–87 (1988). https://doi.org/10.1007/BF00202895
17. Tretiak, O., Pastor, L.: Velocity estimation from image sequences with second order differential operators. IEEE Proc. ICPR 20–22 (1984)
18. Haralick, R.M., Lee, J.S.: The facet approach to optic flow. In: DARPA Image Understanding Workshop, pp. 84–93 (1982)
19. Duncan, J.H., Chou, T.: On the detection of motion and computation of optical flow. IEEE Trans. Pattern Anal. Mach. Intell. **14**(3), 346–352 (1992). https://doi.org/10.1109/34.120329
20. Perrone, J.A.: Simple technique for optical flow estimation. J. Opt. Soc. Am. **7**(2), 264–278 (1990). https://doi.org/10.1364/JOSAA.7.000264
21. Hassanien, A.E., Fouad, M.M., Manaf, A.A., Zamani, M., Ahmad, R., Kacprzyk, J. (eds.): Multimedia Forensics and Security. ISRL, vol. 115. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-44270-9

The Present and Future Role of Artificial Intelligence in the Promotion of Sustainable Practices in the Construction Sector

Aditya Singh(✉)

Lovely Professional University, Phagwara 144411, PB, India
aditya.11602217@lpu.in

Abstract. In this paper, the focus was given on the general concept of Artificial Intelligence and to understand the way Artificial Intelligence is changing the world, especially the construction sector. Sustainability is essential in the construction industry and the paper highlighted how AI can help it to achieve it in its own ways. A number of research papers which were published in the recent years were studied in order to find out the research gaps. Data from different sources were collected in order to perform graphical analysis, to further understand the advantages and challenges at present involved while implementing AI to accomplish sustainable goals. Through this paper, it will help in making people understand the future trend as well regards to this area.

Keywords: Smart Construction · Construction Sector · Artificial Intelligence · Sustainable Practices · Construction Projects · Civil Engineering

1 Introduction

Artificial Intelligence influences not only machines but also computers to impersonate the mind of humans because of the decision making as well as problem solving abilities of humans. According to John McCarthy it can be said that, AI allows the creation of intelligent machines, particularly intelligent computer programs, in a way it is associated to comparable job of utilizing computers to comprehend the intelligence of humans, and nonetheless it doesn't need to restrict itself to the ways which are physically noticeable by humans. In other words, AI is a field which has the ability to merge robust datasets in addition to computer science in order to allow the ability to solve problems, then it also includes deep learning in addition to machine learning, which are its sub-sub field and subfield respectively. Then Artificial Intelligence can be classified as two types, which are [20]:

- Strong Artificial Intelligence
- Weak Artificial Intelligence

Strong Artificial Intelligence: It is created with the help of Artificial Super Intelligence as well as Artificial General Intelligence. As the name suggests, Artificial Super

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 11–18, 2023.
https://doi.org/10.1007/978-3-031-39777-6_2

Intelligence in principle is expected to exceed the abilities in addition to intelligence of human beings and research is still going on in this direction. On the other hand, Artificial General Intelligence in theory has intelligence be the same at the human beings, future planning, learning and problem solving abilities due to self-aware consciousness will be like that of a human [20].

Weak Artificial Intelligence: It is also known as Artificial Narrow Intelligence, where it is trained in a way to be able to perform particular tasks. Such type of Artificial Intelligence is present in everywhere in today's world. Some examples are autonomous automobiles, Apple's Siri, IBM Watson as well as Amazon's Alexa [20].

1.1 Artificial Intelligence Required for Smart Construction

The construction sector is experiencing a digital transformation with the help of AI. Machine Learning as well as AI are technologies which are aiding all the stages of the construction sector including preconstruction stage, construction stage and post construction stage, to reach new heights in comparison to the past. Wherever AI is applied in the construction sector, noticeable difference is being observed as the cost in addition to time is being significantly reduced to complete the tasks. In the construction sector, designing in addition to planning stage are expected gain the most with the help of AI. In the case of construction industry on an international level, the market in Europe is expected to have the highest growth rate. The technological shift is set to affect in a positive way throughout the course of the project to all the stakeholders comprising the owners, contractors as well as service providers. Other sectors like manufacturing as well as transportation have already started implementing AI in their workplace to improve themselves further, similarly the construction sector also needs to work on applying AI in their system in order to avoid lagging behind with the other related sectors with time. Since, the technological sift is at the emerging level in the construction sector, so it will be highly beneficial for those companies which are upgrading their technologies with the changing time [17].

1.2 Is There Any Need to Use Advanced Technologies Like AI in the Construction Sector in Order to Promote Sustainable Practices?

In the recent years, it has been observed that the construction industry is being moving towards construction 4.0 because of the latest technological advancements. One of the other factors are that the other sectors are developing at a rapid rate like computer science, electronics, robotics, etc. Of course, civil engineering domain is also slowly and steadily being affected by technologies like Artificial Intelligence, Machine Learning, and Robotics and so on, this can be observed in traffic and transportation engineering which is a branch of civil engineering, where Intelligent Transport System has come into the picture thanks to the advanced technologies like fuzzy logic, AI, sensors, etc. [21, 23] Similarly, digitization, digitalization and digital transformation are slowly taking place in the construction sector in the recent years and of course, Artificial Intelligence will play an important role not only in the current development but also in the future development of Construction 4.0, where the efficiency, quality of work, complexity in design, and time saving will be drastically improved [22]. Plus, the wastage of materials in construction projects, can be reduced greatly, which will also decrease the carbon footprint to an extent. This will definitely help in promoting sustainable practices upto a certain degree in the construction sector and projects.

1.3 Objectives

In this paper, following objectives are considered:

- To understand the concept of Artificial Intelligence and its need to promote sustainable construction practices.
- To understand the gaps in the current research and application of AI in the various stages of a construction project.
- To understand the current and future market trend of AI in the construction sector.

In this paper, Sect. 1 is Introduction, where general concept of AI and its types are explained as well as how AI is contributing towards smart construction and towards sustainable construction practices. Sections 2 and 3 are about Literature Review and Research Gaps respectively. This is followed by Sect. 4 as Methodology, Sect. 5 as Results and Discussion, Sect. 6 as Advantages and Challenges of Applying AI, and last section is Sect. 7 which is Conclusion, in this paper.

2 Literature Review

Oluleye et al. (2023) worked on a review paper, where they adopted AI in order to improve the execution of systemic circularity in the construction sector [1]. Liu (2023) worked on applying AI in order to incorporate a new uncertainty construction model for green supply chain management to aid not only medium sized but also small sized businesses [2]. Baduge et al. (2022) worked on a review paper, where they studied building as well as construction 4.0 and they worked on smart vision in addition to AI for them, where they applications as well as methods of deep learning and machine learning in them [3]. Pan and Zhang (2021) studied construction engineering as well as management, and they worked finding out on AI's roles in them [4]. Egwim et al. (2021) used AI in order to forecast the delays in construction projects [5]. Sacks et al. (2020) worked on AI and BIM in the area of construction technology [6]. Gracia, et al. (2023) studied smart as well as green buildings, and then worked on applying the techniques of AI in them [7]. Xiang et al. (2022) studied green building engineering and they conducted a research to evaluate them in regards to sustainability on the grounds of energy consumption as well as AI [8]. Nguyen et al. (2023) studied the recycled aggregate concrete and they worked to forecast in addition to perform sensitivity analysis of its mechanical properties using the AI algorithms [9]. Zhang et al. (2022) studied the industry of architecture engineering construction management or facility management, and they worked on incorporating the applications of AI and BIM in them [10].

3 Research Gaps

Based on the recent published scientific and research or review papers in the recent years, it was found that more research has been going on in applying of AI in other sectors like medical and healthcare, etc. compared to the construction sector. Plus, insufficient research has been found in India in the recent years regarding the use of AI in the construction sector to promote sustainable construction practices.

4 Methodology

In this section how Artificial Intelligence is applied in various phases of a construction project, like Preconstruction Stage, Construction Stage and Post Construction Stage are explained through Fig. 1.

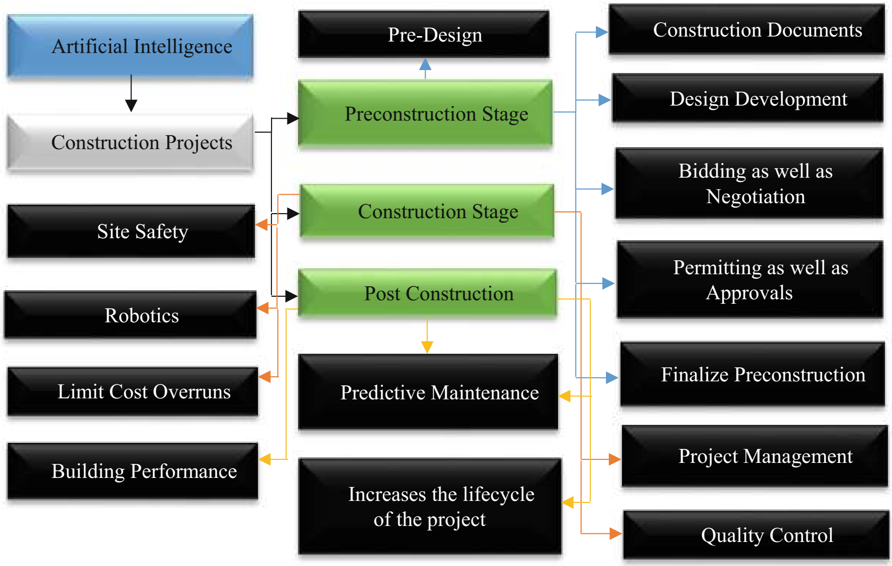


Fig. 1. Showing the application of AI in the various phases of a construction project [18, 19].

5 Result and Discussion

In this section, data from several sources were collected from World Economic Forum, Markets and Markets, etc., to perform graphical analysis, to support the study (Tables 1 and 2).

Table 1. Expected global employment by 2025 due to AI [24]

Global Employment Due To AI	Expected Market Size in $
Un-employment by 2025	85
Employment by 2025	97

Table 2. Comparison of expected market size of AI in global construction market by different sources [11–16]

Value of AI In Global Construction Market	In USD
Market value evaluated in 2017 by MM	329,300,000
Market value evaluated in 2018 by GVR	398,600,000
Market value evaluated in 2020 by ATR	400,000,000
Market value evaluated in 2021 by AMR	496,400,000
Market value evaluated in 2021 by VMR	610,640,000
Market value evaluated in 2022 by NMW	709,090,000
Expected Market value in 2023 by MM	1,831,000,000
Expected Market value in 2025 by GVR	2,307,630,000
Expected Market value in 2027 by ATR	3,500,000,000
Expected Market value in 2029 by NMW	4,882,790,000
Expected Market value in 2030 by VMR	5,045,680,000
Expected Market value in 2031 by AMR	8,600,000,000

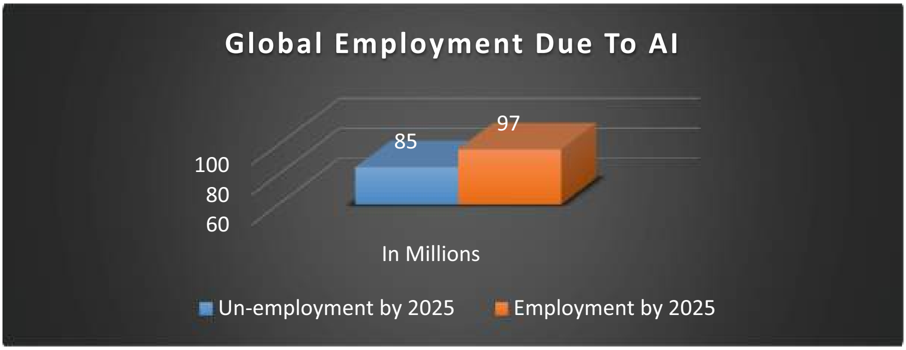


Fig. 2. Expected global employment by 2025 due to AI [24]

6 Advantages and Challenges of Applying AI

In this section major advantages and challenges of implementing AI in the construction sector, which are as follows:

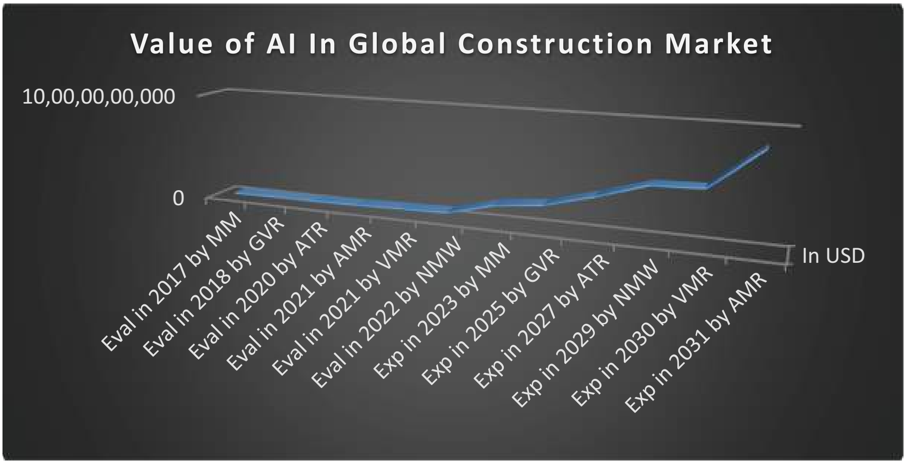


Fig. 3. Comparison of expected market size of AI in global construction market by different sources [11–16]

- Based on Fig. 2, the employment is expected to increase by 12 million by the year 2025, this can also be seen as a positive sign of implementing AI in all the sectors.
- According to Fig. 3, the value of AI in the worldwide construction market is expected to increase sharply in the coming years, which is a positive trend and it will be beneficial for construction companies to invest in the AI more and more with time.
- According to Fig. 2, it is shown that 85 million people will be unemployed by the year 2025 because of AI in all the sectors, similar trend can be noticed in the construction sector, where unskilled and semiskilled labourers and employees will be out of jobs, which will be a huge problem.
- Based on Fig. 3, it can be said that the construction companies and engineers who are used to traditional ways of construction will have trouble in growing in the future.
- According to Fig. 2 and Fig. 3, it will be a smart move by civil engineers to train themselves to work with AI in projects and same goes for the construction companies, this can also help in promoting sustainable construction practices.

7 Conclusion

The general concept of AI and major types of AI were explained in this paper. Then, numerous scientific and research as well as review papers published in the recent years were considered to find out the gaps in the research in the paper. Use of AI in various stages of a construction project was shown and some graphical analyses were performed by the author by collecting data from different sources in order to support the study. It is clear that AI has a lot of scope in the future and will create new types of jobs. Of course, semiskilled and unskilled employees and labourers will have unemployment issues. AI will help in promoting sustainable construction practices upto an extent. In the future, more research needs to be done to help in the development of AI technologies, which will further help the construction sector to enable smart and sustainable construction practices.

References

1. Oluleye, B.I., Chan, D.W.M., Antwi-Afari, P.: Adopting artificial Intelligence for enhancing the implementation of systemic circularity in the construction industry: a critical review. Sustain. Prod. Consumption **35**, 509–524 (2023)
2. Liu, B.: Integration of novel uncertainty model construction of green supply chain management for small and medium-sized enterprises using artificial intelligence. Optik **273**, 170411 (2023)
3. Baduge, S.K., et al.: Artificial intelligence and smart vision for building and construction 4.0: machine and deep learning methods and applications. Autom. Constr. **141**, 104440 (2022)
4. Pan, Z.: Roles of artificial intelligence in construction engineering and management: a critical review and future trends. Autom. Constr. **122**, 103517 (2021)
5. Egwim, C.N., Alaka, H., Toriola-Coker, L.O., Balogun, H., Sunmola, F.: Applied artificial intelligence for predicting construction projects delay. Mach. Learn. Appl. **6**, 100166 (2021)
6. Sacks, R., Girolami, M., Brilankis, I.: Building information modelling, artificial intelligence and construction tech. Dev. Built Environ. **4**, 100011 (2020)
7. Gracia, D., Uriarte, M., Yepez, E., Fernandez, J., Iribarne, L., Ayala, R.: Review of artificial intelligence techniques in green/smart buildings. Sustain. Comput. Inform. Syst. **38**, 100861 (2023)
8. Xiang, Y., Chen, Y., Xu, J., Chen, Z.: Research on sustainability evaluation of green building engineering based on artificial intelligence and energy consumption. Energy Rep. **8**, 11378–11391 (2022)
9. Nguyen, T.-D., Cherif, R., Mahieux, P.-Y., Lux, J., Mokhtar, A.A., Arteaga, E.: Artificial intelligence algorithms for prediction and sensitivity analysis of mechanical properties of recycled aggregate concrete: a review. J. Build. Eng. **66**, 105929 (2023)
10. Zhang, F., Chan, A.P.C., Darko, A., Chen, Z., Li, D.: Integrated applications of building information modeling and artificial intelligence techniques in the AEC/FM industry. Autom. Constr. **139**, 104289 (2022)
11. Allied Market Research. https://www.alliedmarketresearch.com/artificial-intelligence-in-construction-market-A12996. Accessed 25 Apr 2023
12. Globe Newswire. https://www.globenewswire.com/news-release/2022/12/07/2569499/0/en/Artificial-Intelligence-AI-In-Construction-Market-is-expected-to-generate-a-revenue-of-USD-5045-68-Million-by-2030-Globally-at-35-70-CAGR-Verified-Market-Research.html#:~:text=As%20per%20the%20deep%20research,35.70%25%20from%202022%20to%202030. Accessed 25 Apr 2023
13. Grand View Research, https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-in-construction-market, last accessed 2023/04/25
14. All The Research. https://www.alltheresearch.com/report/93/artificial-intelligence-in-construction-market. Accessed 25 Apr 2023
15. Yahoo! Finance. https://finance.yahoo.com/news/artificial-intelligence-construction-market-report-153422883.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvLmluLw&guce_referrer_sig=AQAAANunW8UQ-M7K4k36joRhxtH188ULDnYt4K9OAlWlJLXs94f7hW9WlNMGncjOVE2NOJpJ3fEeu_peAOepQ5XuZg4X0fK2UkvRvN-wcyaNVhlRbu_n8zRqTU2cksTWO_0lNY1b8Ch-Uvb2KJKv8sX0aBPMPOUiKkhArWCZEEqmQ5lR. Accessed 25 Apr 2023
16. Markets and Markets. https://www.marketsandmarkets.com/Market-Reports/ai-in-construction-market-235710349.html. Accessed 25 Apr 2023
17. Appinventiv. https://appinventiv.com/blog/ai-in-construction/. Accessed 25 Apr 2023
18. Constructconnect. https://www.constructconnect.com/blog/ai-in-construction-has-landed#:~:text=Building%20performance%3A%20AI%20could%20be,for%20better%20design%20and%20construction., last accessed 2023/04/25

19. ESDS. https://www.esds.co.in/blog/how-ai-with-machine-learning-is-changing-the-future-of-construction/#:~:text=AI%20in%20construction%20has%20the,%3B%20and%2C%20business%20model%20transformation. Accessed 25 Apr 2023
20. IBM. https://www.ibm.com/topics/artificial-intelligence. Accessed 25 Apr 2023
21. Singh, A.: Importance of fuzzy logic in traffic and transportation engineering. In: Vasant, P., Zelinka, I., Weber, G.W. (eds.) Intelligent Computing & Optimization, ICO 2021. LNNS, vol. 371, pp. 87–96. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93247-3_10
22. Singh, A.: The significance of digitalization of the construction sector. In: Vasant, P., Weber, GW., Marmolejo-Saucedo, J.A., Munapo, E., Thomas, J.J. (eds.) Intelligent Computing & Optimization, ICO 2022. LNNS, vol. 569, pp. 1069–1077. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19958-5_100
23. Bhowmik, S., Singh, A., Misengo, C.: A case study on intelligent transport system using traffic lights. Our Heritage **67**(7), 96–110 (2019)
24. Data Prot. https://dataprot.net/statistics/ai-statistics/#:~:text=Nine%20out%20of%20ten%20leading,billion%20will%20be%20by%202023. Accessed 25 Apr 2023

Honeywell Experion Hive as Breakthrough Approach in Control Systems Evolution

Nodirdek Yusupbekov[1], Farukh Adilov[2], Maksim Astafurov[2](✉), and Arsen Ivanyan[2]

[1] Tashkent State Technical University, Uzbekistan, Uzbekistanskaya-2 Prospect, 100095 Tashkent, Republic of Uzbekistan

[2] "XIMAVTOMATIKA" Ltd, Uzbekistan, Afrosiyob Street, 1000311 Tashkent, Republic of Uzbekistan

{Farukh.Adilov,Maksim.Astafurov,Arsen.Ivanyan}@himavtomatika.uz

Abstract. Considered evolution of the control systems and input-output processing concepts in order to compare with breakthrough technologies from recognized world-class leader in industrial automation – Honeywell LEAP™ (Lean Execution of Automation Projects) and Experion HIVE (Highly Integrated Virtual Environment).

Keywords: Automated control · Remote access · Remote administration · SCADA · Remote control · Process industry · Cyber security · Industrial internet of things · Controllers · IO modules

1 Introduction

Since the appearance of the first microprocessor and computer controls in production, their evolution has not stopped for a second - from bulky and low-power, occupying entire cabinets and devouring mega-watts of electricity, to miniature and super-efficient capable of processing enormous amounts of data in real time while being no bigger than an average book.

Over time, not only the size or performance of controls changed, but also their architecture - becoming more and more decentralized, distributed, it got rid of the so-called "single point of failure" - dependence on the CPU (PLC, ALU, etc. - names many, but the meaning is the same) - part of the primary data processing functions was transferred to the so-called input-output cards - a separate electronic unit that collects process parameters from sensors or transmits a control action to a field device. The introduction of I/O cards has significantly affected many key system parameters:

- fault tolerance of the system (because now in case of an error (dividing by zero, extracting the square root of a negative value, etc.), the failure occurred only in a separate module and did not shut down the entire system).
- performance (a considerable number of resources previously spent by the CPU itself were released).
- scalability (the number of processed field devices has multiplied, since now the data is transmitted in digital format and via a separate data bus).

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 19–25, 2023.
https://doi.org/10.1007/978-3-031-39777-6_3

Some time later, so-called "smart" sensors became widespread, capable of not only processing the measured value, storing many settings and calculating corrections, but also analyzing their own state, thereby predicting possible malfunctions and the need for maintenance in advance. Having access to the Internet, such sensors easily inform the manufacturer about the lack of consumables or the replacement of failed components [1].

Another of the qualitative steps in the development of control systems was the use of virtualization tools - one super-powerful and capacious computer is able to maintain dozens of virtual systems of various configurations in operation, significantly reducing maintenance costs, losses from equipment downtime during backup/recovery, deployment of new systems, etc. Placing the server and all client stations on the same physical layer provides significant savings on all budget lines - from the relatively low cost of "thin clients" compared to conventional workstations, energy costs and, importantly, reducing the carbon footprint of CO_2 emissions.

More recently, cloud services have successfully complemented virtualization technologies by adding flexibility in terms of the required performance of the control system and operating costs, as well as increasing the already high reliability due to multiple redundancy of equipment on the side of the cloud service provider.

It would seem that there is much more to improve technical equipment, but progress does not stand still for a second. Honeywell launches the new Experion HIVE software and hardware complex.

2 Experion HIVE

Experion HIVE – (Highly Integrated Virtual Environment) – represents synthesis of three main components IT HIVE, IO HIVE and Control HIVE integrating with Customer's facilities (Fig. 1).

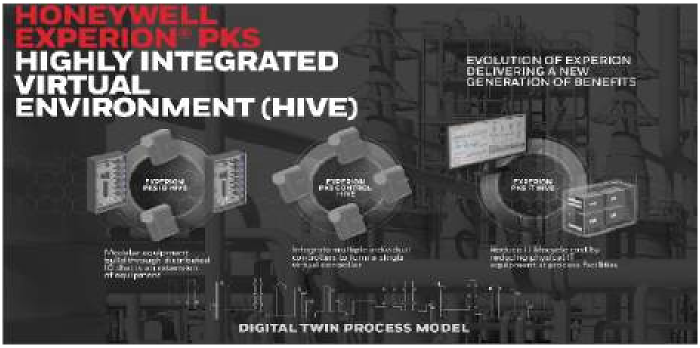


Fig. 1. Experion HIVE – the core components (IT, IO and Control).

IT HIVE – technology which allows wisely choose IT workload distribution at the site by selecting tasks must be done locally and those that could be transferred to datacenter. Giving following benefits [2, 7, 8]:

- Less infrastructure (significantly reduce networking and compute resources kept on site).
- Simplified management and enhanced security (centralized IT workloads reducing number of personnel with system administrative rights).
- Reliability (based on time-tested Honeywell (FTE) over existing WANs instead of using dedicated and expensive WAN infrastructure) (Figs. 2 and 3).

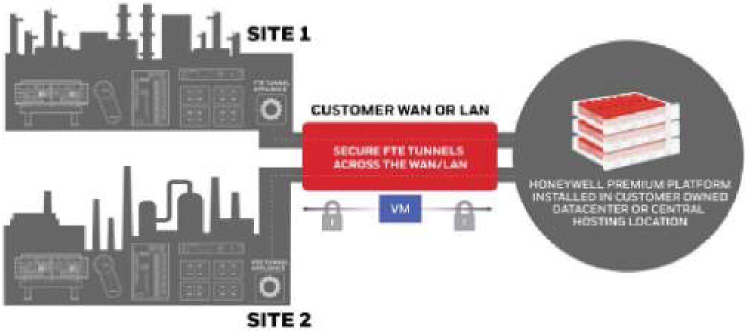


Fig. 2. IT HIVE – FTE tunnels over existing WANs.

- Increased availability (in case of any system failure replication technology enables the rapid recovery of a centrally hosted workload at the remote site).

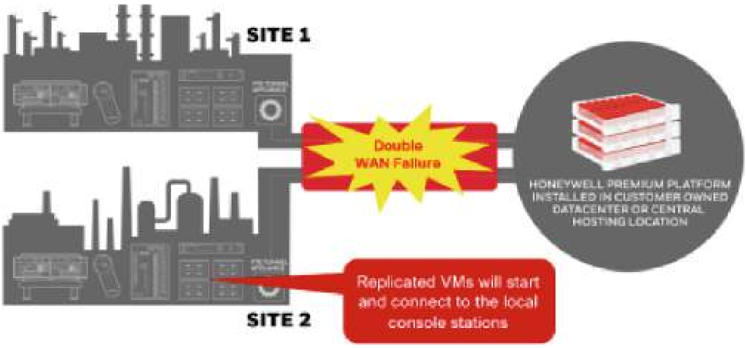


Fig. 3. IT HIVE – Replication in case of double WAN failure.

- Disaster recovery (in case of server or workstation hardware failure it could be quickly restored on another physical host within a short period of time in compare with procurement and installation of new hardware on site).

IO HIVE – next generation of control technology based on Honeywell's LEAP™ (Lean Execution of Automation Projects). This approach is delivering a broad range of features to facilitate modular and parallel project execution, major of them are following [3, 8, 9]:

- Flexibility - compare with traditional architecture where controller and IO have direct physical connection and one controller unable to use IO which belongs to another, new approach makes assignment much easier as control strategy will automatically find its relevant IO without limitations. Each IO can be shared with 20 controllers.
- CapEx & OpEx reduction – without strict connection between IO module and Controller, all the IO channels physically connected to each cabinet can be efficiently used with any controller. In addition, exact IO+ spares can be delivered without wastage due to unused spare channels. Depending on spares percentage and number of IO modules per cabinet this reduces 10% to 17% of IO channels that are unused with the traditional IO link design (Fig. 4).

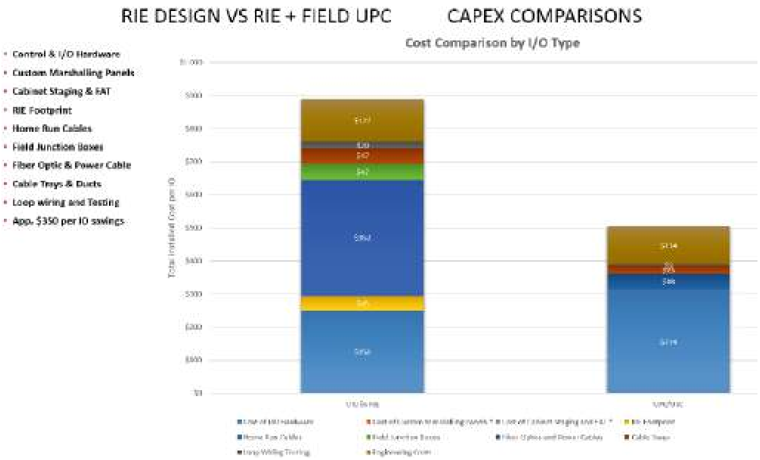


Fig. 4. IO HIVE – Capital Expenditures comparisons.

- Design standardization - since IO and Controller are not strictly linked to each other, the IO cabinets can be designed and built independent of the Controller cabinets. This is true for both indoor or outdoor (remote) locations. This eliminates exhaustive engineering hours spent on prototype designs reducing man-hours by 10% to 15%
- Commissioning without control system - Experion controllers CEE (or Control Execution Environment) can be simulated on engineering a laptop, plugged directly into the remote cabinet in field and fully complete certain commissioning activities as if the rest of the control system is available. This feature makes possible commissioning costs reduction by 20% to 30%.

Control HIVE – outstanding technology allowing multiple individual controllers act as a cluster of shared compute resources with universal access to I/Os. It allows significant flexibility in all project execution phases: design, commissioning & start-up or expansion of existing site facilities [4, 5].

One of the key features of Experion Control HIVE is smart assignment and load balancing - control strategies that in conventional way always assigned to one CEE only may result CEE usage unbalanced – some of them due to complex logic inside will be overloaded, while others are underutilized. In the Control HIVE - they devoid of this disadvantage strategies are automatically allocated to available controller CEE within the HIVE in an optimal manner (Fig. 5).

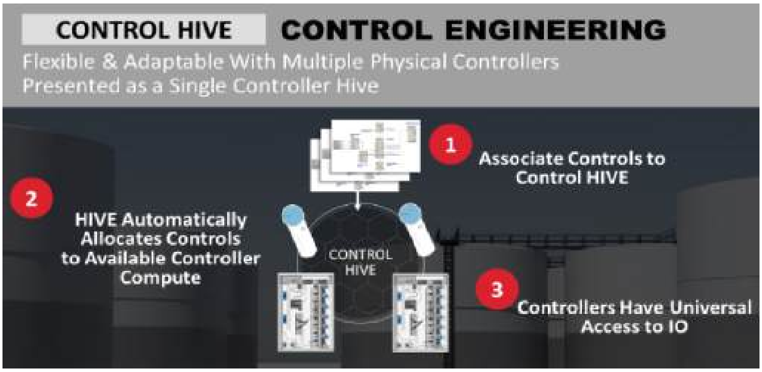


Fig. 5. Experion PKS Control HIVE – overview.

Automated smart assignment and load balancing functionality are also flexible and could be based on system and/or user-defined rules.

Another enhancement in system that was unavailable in previous versions is infinite availability. Traditional primary & backup redundant pair is limited to one way of control transfer – from primary to backup in case if primary got failed by any reason. In Control HIVE control transfers automatically to any other controller with available compute resources, so this process can be repeated continuously as long, as resources remain available in the HIVE (Fig. 6).

Fig. 6. Experion PKS Control HIVE – system view.

Control HIVE provides universal software-based control where the control application is independent from the used physical platform. This means control strategies can be run on any hardware, from dedicated embedded controllers such as the C300 or on traditional PCs and servers (Fig. 7).

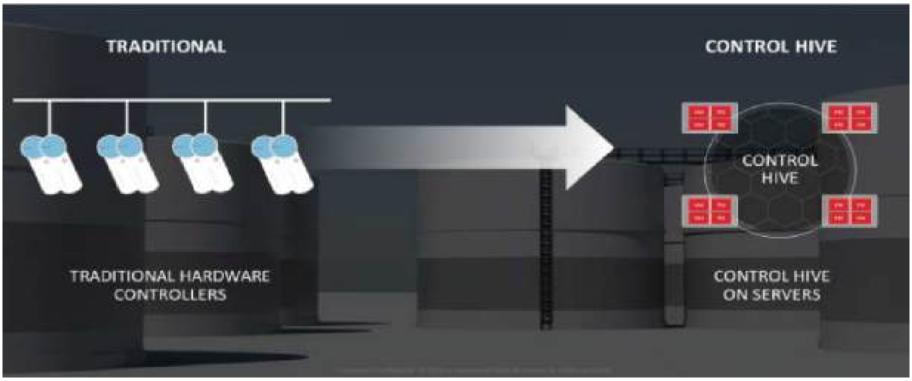


Fig. 7. Experion PKS Virtual Control HIVE overview.

This allows to combine Control HIVE along with IT HIVE allow control to be on any hardware platform and in any location from – adjacent to the process unit, to the control room, or to remote data center.

3 Conclusion

"Industrial facilities lose about $650 billion per year in unplanned downtime and shutdowns; only 25% of maintenance downtime is planned; and about 40% of today's global installed base of automation systems are at least 20 years old," said Michael Khilla, offering manager for Experion controls and engineering tools at Honeywell. The technology considered above has all possibilities to reduce operating expenses (OpEx) and capital expenditures (CapEx) by decreasing the risk of unexpected downtime as well as outsourcing IT workloads in big automation projects [7].

Honeywell Experion HIVE is suitable for any sites, small and medium enterprises will even get more benefits from this system. After all, users do not always manage to use the full potential of the controllers, so Experion PKS HIVE will help better plan the available resources at site without rebuilding the entire process [6].

References

1. Yusupbekov, N., Abdurasulov, F., Adilov, F., Ivanyan, A.: Improving the efficiency of industrial enterprise management based on the forge software-analytical platform. In: Arai, K. (eds.) Intelligent Computing. LNNS, vol. 283, pp. 1107–1113 (2022). https://doi.org/10.1007/978-3-030-80119-9_74
2. Experion® PKS IT HIVE. Product Information Note, Honeywell International Inc., November 2020
3. Experion® PKS IO HIVE. Product Information Note. Honeywell International Inc., November 2020
4. Experion® PKS Control HIVE. Product Information Note. Honeywell International Inc., November 2020

5. Control Network Module. Product Information Note. Honeywell International Inc., April 2022
6. Honeywell Experts - About new technologies for digital enterprise transformations. Control Eng. Russia # **6**(84) (2019)
7. Experion HIVE takes sting out of digitalization. Jim Montague, 24 May 2022. https://www.controlglobal.com/events/honeywell-users-group-2022/news/11288020/honeywell-process-solutions-experion-hive-takes-sting-out-of-digitalization
8. Leveraging New Automation Approaches Across the Plant Lifecycle. Larry O'Brien, 22 June 2017. https://www.arcweb.com/blog/leveraging-new-automation-approaches-across-plant-lifecycle
9. Honeywell Experion PKS HIVE Swarms Process Control. Mark Sen Gupta, 12 July 2019. https://www.arcweb.com/blog/honeywell-experion-pks-hive-swarms-process-control

Pathfinding with Dynamic Obstacles

Anshul Patil[1], Shreshtha Kamboj[1], Anu Bajaj[1](✉), and Ajith Abraham[2]

[1] Thapar Institute of Engineering and Technology, Patiala, India
er.anubajaj@gmail.com
[2] Faculty of Computing and Data Sciences, Flame University, Pune, India

Abstract. Pathfinding algorithms are often at the center of any AI movement system. It is required to provide agents mobility in developing a realistic artificial intelligence (AI) for real-world scenarios. Its main goal is to find the optimized path from source to destination by avoiding obstacles. Finding a path through a maze is a major issue as it can be complex due to maze size, number of obstacles and have multiple routes that may lead to dead end. This paper proposed a search algorithm that finds the shortest path with dynamically created obstacles between the source and destination. It uses the A star algorithm and Manhattan distance as heuristic measure. The experiment was performed on grids of 10 × 10 to 100 × 100 size with step size of 10 × 10 on different obstacles ranges. The algorithm's overall performance is respectable, and it can determine the quickest route between any two points. The proposed algorithm will always return a path for less than 30% obstacles and in cases where feasible paths can be generated.

Keywords: Pathfinding · A* algorithm · maze solving · strategy games · dynamic obstacles

1 Introduction

Pathfinding is one of the popular Artificial Intelligence (AI) problem [11–15]. It is the study of finding a route from a source to a destination. It is very essential in real world scenario to solve any agent's mobility related problem. The shortest path study helped AI acquire the ability to think causally, learn, and think like a human. In other words, Pathfinding algorithm enables to answer the question "How do I get from source to destination?". A number of potential solutions are involved in finding a path from source (S) to destination (D). The main computational task is the application of a search algorithm to discover a path. However, wherever possible the solution need to achieve the following goals [1]:

1. The route from point S to point D.
2. The most effective way to get past obstacles.
3. The method for locating a potential optimized path.
4. A quick way to locate the path.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 26–33, 2023.
https://doi.org/10.1007/978-3-031-39777-6_4

Some pathfinding algorithms handle only one of these problems, others address all of them. However, in certain situations, none of these problems could be solved by any method. The fastest route may not always be the one with the least distance. Finding the quickest route between two points while avoiding obstacles is the goal of the pathfinding algorithm [19]. The challenge is to find the shortest total weight path between any two nodes in maze. The hardest part of pathfinding in the real world is knowing how to avoid barriers and search for the best route through various locations. In this paper, a variant of A star algorithm is proposed that considers the difficulty levels of the obstacles. The obstacles can be dynamically changed. In other words, one can manually alter the obstacles to upscale the difficulty levels.

The organization of the paper is as follows: Sect. 2 presents the related work and the proposed algorithm is presented in Sect. 3. Section 4 discusses and analyzes the results in detail with the conclusion in Sect. 5.

2 Related Works

There are a number of classical works proposing computation of the shortest path between two points. These works also adopt various algorithms other than A Star. Finding the shortest path is a challenge that has been addressed by a number of search algorithms, such as Barnouti et al. [1]. Proposed A* Search Algorithm and discussed Breadth-First Search, Depth-First Search, Depth-Limited Search, Best-First Search, and Hill-Climbing Search algorithms. There are several heuristics to find the optimal cost. Manhattan distance is the quickest heuristic out of the three popular heuristics: Manhattan, Euclidean, and Diagonal Shortcut. It adds the distance travelled in both directions to reach the destination. Manhattan Block get its name because one might move that way on city blocks [2].

It's possible that an entirely new sort of heuristic will produce even quicker timings than before. The majority of heuristics primarily consider the distance between the agent's current location and the goal. This novel form of heuristic concentrates on direction, directing the agent to further explore [3]. A* algorithm has been widely applied in many applications like it was used in the physical robot to carry out the search and rescue task in a real maze [8, 18]. It also performed well in a case study of bicycle pathfinding [16]. An improved version of A* with complete path coverage was proposed for unmanned ships [17].

In comparison to other naive search methods like the wall follower approach, graph search algorithms are significantly better. In a large search space, the uninformed search algorithms outperform the informed search algorithms [4]. Even though A* is currently the best search algorithm available, it must be utilized carefully to avoid resource waste. The A* algorithm needs a lot of memory to keep track of each search's progress, especially when searching across expansive and complex surroundings. A* search can perform better by optimizing the underlying search space, using less memory, enhancing heuristic functions, and adding new data structures [5].

HPA* algorithm was proposed for pathfinding to overcome the limitations of A* algorithm which produced the compromised at the expense of reducing the runtime [3]. Another study improved the algorithm that selectively expands certain nodes in grid map which are jump points. The path connecting two jump points never expands the intermediate nodes. It was found that A* gets speed up by using these jump points for searching. Also this jump point pruning technique in many cases may prove faster than HPA* algorithm [6, 10]. Uninformed algorithms like BIDDFS do not have any knowledge about the goal but in cases where there are large number of obstacles, they have exceptional execution and in games where one is unaware of the Finish point they may prove to be useful [9].

3 Proposed Algorithm

A-star (A*) is a widely used algorithm for pathfinding in graphs or grids. It combines the greatest features of heuristic search and Dijkstra's method [7]. In the proposed work, the source, destination and the percentage of obstacles are taken as inputs from the user. On the other hand, the obstacles can be generated dynamically at run time. A heuristic function is used to direct the search in the direction of the target node, and an evaluation function selects the best path out of all the open nodes. Here, the Manhattan distance is used for calculating the distance between two points as the sum of the absolute differences of their coordinates.

$$f(x) = g(x) + h(x) \tag{1}$$

$$h(x) = abs(x1 - x2) + abs(y1 - y2) \tag{2}$$

The algorithm initially adds the starting node to an open set, and assigns it a new score f(x), which is a sum of two values: the cost to reach that node from the starting node g(x), and the estimated cost of reaching the goal node from that node, h(x). The estimate h(x) is based on a heuristic function, which provides a lower bound of the actual cost. A common heuristic function used in grid-based maps is the Manhattan distance.

Then, it examines each of the neighboring nodes of the current node, calculates a score for them, and adds them to the open set. The neighboring nodes with the lowest score are selected as the next current node, and the steps are repeated until the goal node is reached. Finally, it returns the optimal path by reporting the computational time and the coordinates of the obstacles used. The working of the algorithm is present in Algorithm 1.

Algorithm 1 Proposed Algorithm for dynamic obstacles

Input: Start position S, Finish position F, List of Obstacles (O)
Output: Co-ordinates of optimal path, Runtime of Algorithm

1. Create a new priority queue PQ and assign PQ(0)=S.
2. Create the dictionaries parent P and cost C to record the path's origin and cost. Set C[S] to 0 and P[S] to None.
3. while PQ is full
4. Get the item with the lowest priority (i.e., lowest cost) and assign it to Current.
5. If Current is F return P and C.
6. Else, use the Manhattan distance to determine the cost from start and priority of each neighbor of Current. Update P and C and add the neighbor to the PQ with the computed priority if C does not include the neighbor or if the new cost is less than the previous cost.
7. If F is not in P, there is no path to it. Otherwise, use P to retrace the trip.
8. Plot the optimal path and the obstacles.

4 Data Analysis and Results

The analysis is done using a laptop with Intel(R) Core (TM) i7-11370H CPU 3.30 GHz with 16.0 GB RAM. The Python shell version 3.10.11 is utilized to build the application program. The path between source and destination will be discovered by the updated A* search algorithm. Locations for the source and destination must be chosen. The programme will identify the quickest path in light of the available data. The proposed algorithm is applied on different grid sizes ranging from 10 * 10 to 100 * 100. The difficulty level is verified by updating the percentage of number of obstacles. A sample of runs on 10 * 10 grid with 10, 20 and 30 obstacles are shown in the Figs. 1, 2 and 3, where left hand side shows the plot of optimal path in a grid of randomly generated obstacles between the start position and finish position. Black dots represent obstacles, Blue dot represents the starting node, Green dot represents finishing node, Red line joining the start and finish represents optimal path. Furthermore, right hand side shows the step nodes or co-ordinates that are in the optimal path. It also gives the coordinates of obstacles in case of randomly generating obstacles and in case they are added dynamically. Figure 4 shows a different scenario of dynamically changing the source and destination.

It is observed that that the algorithm does not always provide a path if the obstacles are increased more than 30%. This may be due to the use of Manhattan distance as heuristic function. In other words, Manhattan distance heuristic has a tendency to overestimate the actual distance to the target especially when there are impediments in the search space. The performance of the proposed algorithm can be improved by employing a different, more precise heuristic function that accounts for the presence of barriers in the search space. The Euclidean distance or the Chebyshev distance, for instance, can be used as a heuristic function and can be more accurate for specific types of barriers. On the other hand, if number of obstacles are increased beyond a higher percentage of grid size, then there are lesser chances of finding a path and eventually the obstacles might block the path entirely.

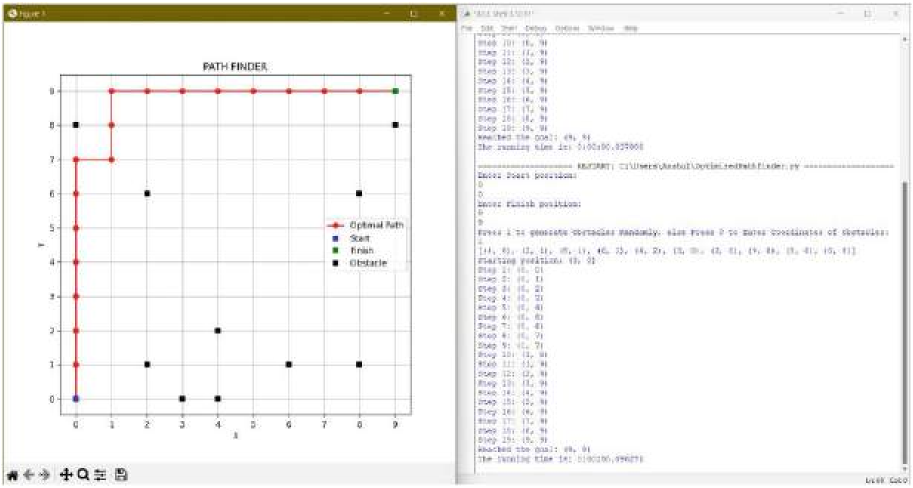


Fig. 1. Optimal path produced by the proposed algorithm for 10 obstacles

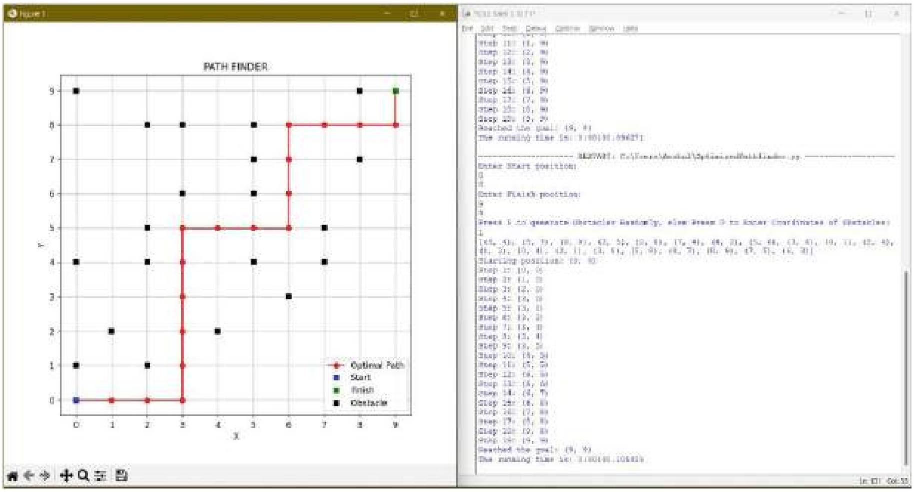


Fig. 2. Optimal path produced by the proposed algorithm for 20 obstacles

Table 1 shows the computational time for each combination after taking 10 runs. Observations showed that if the search space is not too big, the increase in execution time is not much noticeable as shown in Table 1. On the other hand, the runtime of A* algorithm increases with the number of obstacles as they can block certain paths and force the algorithm to take longer routes, leading to more node explorations and a longer runtime. Though exceptions were found in the case for 20 × 20 and 100 × 100 grids where runtime of 30% obstacles was found to be lesser than that of 20%. Future work will address this instability and try to improve the efficiency in future by either introducing better heuristics or implementing it with a different algorithm.

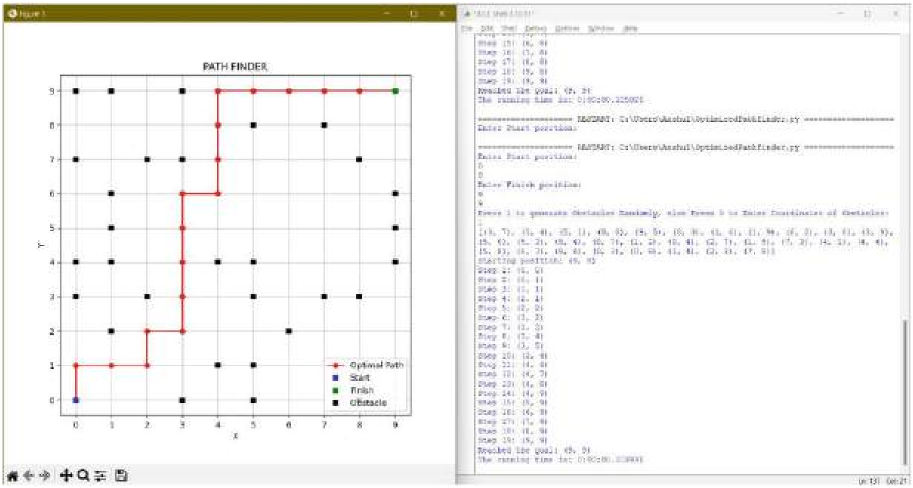


Fig. 3. Optimal path produced by the proposed algorithm for 30 obstacles

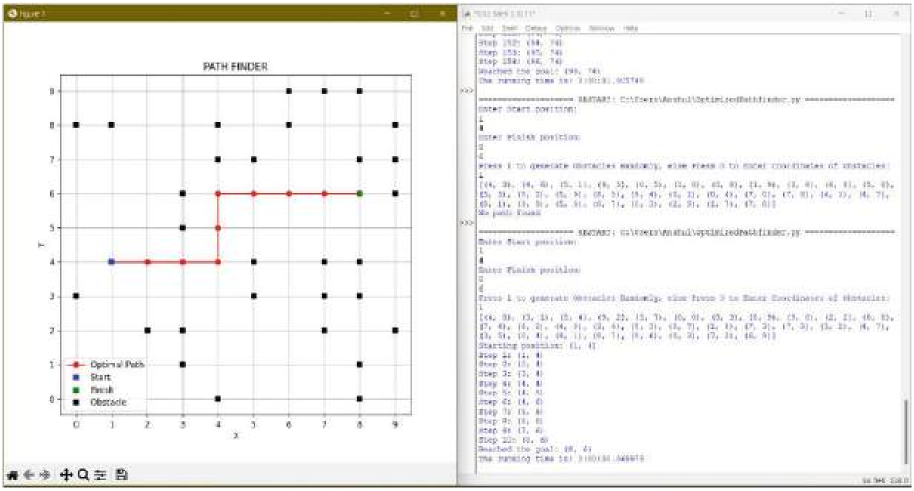


Fig. 4. Optimal path produced by the proposed algorithm with different start and finish point

Table 1. Computational Time of the algorithm for different grids and obstacles (in seconds)

No. of Obstacles	10%	20%	30%
10 * 10	0.0731822	0.0711304	0.0789092
20 * 20	0.1518094	0.1573552	0.1554294
30 * 30	0.237834	0.2612726	0.2806406
50 * 50	0.4338996	0.4424164	0.499105
100 * 100	1.1681428	1.2419254	1.0690206

5 Conclusions

A variant of A star algorithm has been proposed to find the path for dynamic obstacles. The performance of the algorithm is verified on a range of grids 10 * 10 to 100 * 100 with an increment of 10 * 10. The difficulty level was increased by getting a dynamic obstacles input from the users. The algorithm performed well up to 30% obstacles. In future, the robustness of the algorithm will further be improved by applying different heuristic function or hybridizing the A star algorithm with genetic algorithm for better search in case of large number of obstacles.

References

1. Barnouti, N.H., Al-Dabbagh, S.S.M., Naser, M.A.S.: Pathfinding in strategy games and maze solving using A* search algorithm. J. Comput. Commun. **4**(11), 15 (2016)
2. Leigh, R., Louis, S.J., Miles, C.: Using a genetic algorithm to explore A*-like pathfinding algorithms. In: 2007 IEEE Symposium on Computational Intelligence and Games, pp. 72–79. IEEE, April 2007
3. Foead, D., Ghifari, A., Kusuma, M.B., Hanafiah, N., Gunawan, E.: A systematic literature review of A* pathfinding. Procedia Comput. Sci. **179**, 507–514 (2021)
4. Kaur, N.K.S.: A review of various maze solving algorithms based on graph theory. IJSRD **6**(12), 431–434 (2019)
5. Cui, X., Shi, H.: A*-based pathfinding in modern computer games. Int. J. Comput. Sci. Netw. Secur. **11**(1), 125–130 (2011)
6. Mathew, G.E.: Direction based heuristic for pathfinding in video games. Procedia Comput. Sci. **47**, 262–271 (2015)
7. Sun, Y.-J., Ding, X.-Q., Jiang, L.N.: Heuristic pathfinding algorithm based on Dijkstra. In: 3rd Annual International Conference on Electronics, Electrical Engineering and Information Science (EEEIS 2017), pp. 421–425. Atlantis Press, September 2017
8. Liu, X., Gong, D.: A comparative study of A-star algorithms for search and rescue in perfect maze. In: 2011 International Conference on Electric Information and Control Engineering, pp. 24–27. IEEE, April 2011
9. Noori, A., Moradi, F.: Simulation and comparison of efficiency in pathfinding algorithms in games. Ciência e Natura **37**, 230–238 (2015)
10. Harabor, D., Grastien, A.: Online graph pruning for pathfinding on grid maps. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 25, no. 1, pp. 1114–1119, August 2011

11. Ali, M., Pant, M., Abraham, A.: Simplex differential evolution. ACTA Polytechnica Hungarica **6**(5), 95–115 (2009)
12. Izakian, H., Tork Ladani, B., Zamanifar, K., Abraham, A.: A novel particle swarm optimization approach for grid job scheduling. In: Prasad, S.K., Routray, S., Khurana, R., Sahni, S. (eds.) ICISTM 2009. CCIS, vol. 31, pp. 100–109. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00405-6_14
13. Dasgupta, S., Das, S., Biswas, A., Abraham, A.: On stability and convergence of the population-dynamics in differential evolution. AI Commun. **22**(1), 1–20 (2009)
14. Xhafa, F., Abraham, A.: Metaheuristics for Scheduling in Industrial and Manufacturing Applications, Studies in Computational Intelligence, vol. 128. Springer, Cham (2008). https://doi.org/10.1007/978-3-540-78985-7
15. Abraham, A.: Intelligent systems: architectures and perspectives. In: Abraham, A., Jain, L.C., Kacprzyk, J. (eds.) Recent Advances in Intelligent Paradigms and Applications, pp. 1–35. Springer, Cham (2003). https://doi.org/10.1007/978-3-7908-1770-6_1
16. Ariyanto, R., Rohadi, E., Kirana, A.P.: Implementing a star for bicycle route finding system using OSM and GraphHopper: case study: Batu, Indonesia. In: 2022 International Conference on Electrical and Information Technology (IEIT), pp. 307–312. IEEE (2022)
17. Guo, B., Kuang, Z., Guan, J., Hu, M., Rao, L., Sun, X.: An improved a-star algorithm for complete coverage path planning of unmanned ships. Int. J. Pattern Recogn. Artif. Intell. **36**(03), 2259009 (2022)
18. Fujita, H.: Effectiveness of robot motion block on A-Star algorithm for robotic path planning (2022)
19. Andreychuk, A., Yakovlev, K., Surynek, P., Atzmon, D., Stern, R.: Multi-agent pathfinding with continuous time. Artif. Intell. **305**, 103662 (2022)

Industry 4.0 Applications

Application of Digital Twin Intelligent Technology for Live Process Control System of Industrial Facility

Nodirbek Yusupbekov[1], Farukh Adilov[2], and Arsen Ivanyan[2](✉)

[1] Tashkent State Technical University, Tashkent, Uzbekistan
[2] LLC "XIMAVTOMATIKA", Tashkent, Republic of Uzbekistan
{Farukh.Adilov,Arsen.Ivanyan}@himavtomatika.uz

Abstract. Our research is dedicated to the development and application of workable concept and mechanism of total digitization of real live process control system operated in industrial facility using modern virtualization and IT-technologies. Principally it is about creation of clone or digital twin of process control system (PCS) or intelligent information control system (IICS) in the cloud which is constantly connected and synchronized with real industrial asset. This firstly gives total management and permanent improvement of lifecycle of real asset and secondly ensures timely troubleshooting and safe operation of the system reducing or absolutely avoiding any unexpected shutdown of industrial facility. Research and technology used is built on cyber secure remote infrastructure and this important aspect specially raised in actual paper. Concept of digital twin for PCS properly connected to the concept of all process assets digitization and our research in the program of digital twins' development for industrial assets with final simulation of holistic twin for whole industrial facility.

Keywords: Digital Prime · Digital Twin · process control system (PCS) · intelligent information control system (IICS) · Real-Time Database (RTDB) · synchronization

1 Introduction

Every intelligent information control system (IICS), process control system (PCS), integrated control and safety system (ICSS) or whatever another conceptual term used for upper-level automation system installed in industrial facility already contains key intelligent (or SMART) and IT-elements to be easily digitized and integrated to appropriate hardware-software platform where Digital twins are being created.

From this point of view conceptually and technologically there is no research challenges and new scientific-engineering discoveries and achievements required to digitize this asset in the cloud for 10–15 years already.

But to create fully functional and conceptually correct Digital Twin important element is to ensure real-time data flow process and real-time synchronization between physical asset and its' digital twin. As far as technical performance and availability of

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 37–44, 2023.
https://doi.org/10.1007/978-3-031-39777-6_5

every PCS is very high this real-time synchronization requirement for many years has been bottleneck for workable implementation of digital twin theory for control systems.

Meanwhile creation of digital twin for PCS from year to year became critical need because of number of issues and demands coming from real industrial plant operation (refer to Fig. 1).

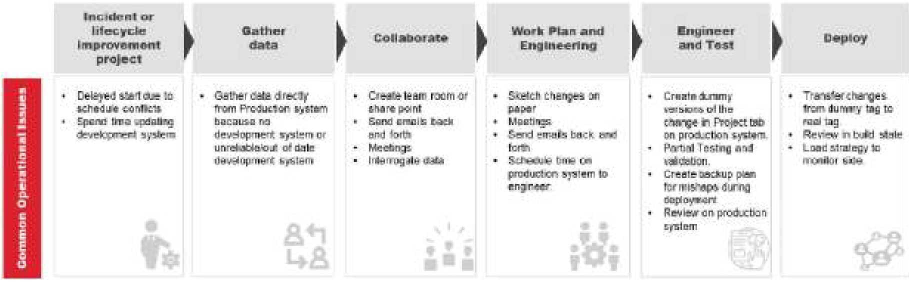


Fig. 1. Common operation issues in every industrial plant.

Lack of, or inadequate, development systems lead to uncontrolled changes on the production system. When production system becomes test system, it has the following affects:

– Increases safety risks
– Slower digital maturity progression
– Waiting for planned downtime to plan, test and implement when activities are already on the critical path
– Missed or delayed optimization opportunities
– When a plant faces incipient problems, the lack of a test bed equates to a longer optimization cycle. Compromising yield, productivity, and safety.
– Creates gaps in documentation.

Objective of this paper is to present and describe digital twin technology implemented for live PCS/ICSS and show the advantages of this solution in resolving of highlighted industrial challenges. Taking into account above mentioned difficulty in real-time synchronization for live control system resolving of this challenge becomes really unique scientific-engineering task and because of this our research on this field becomes more valuable and ensure interest of specialists and originality of the paper.

So, application of digital twin technology obviously seems to be the key to resolve these problems but how to implement it as real workable solution?

In next sections our paper will present exact Methods and result of implementation of digital twin technology to live ICSS with Conclusion what is achieved by this implementation.

2 Methods and Results

The base methodology for PCS Digital twin has been constructed on the fundamental principles of Industrial Internet of Things (IIOT) [1–3] and Virtualization [4–6] technologies.

The Industrial Internet of Things is the intelligent application of digital technology to solving the automation needs of an analog world. There are two fundamental pillars to IIoT: digital automation systems, and the Internet itself. The IIoT edge computing environment consists of a wide range of devices including sensors, actuators, controllers, and human machine interfaces. These devices are near the production process and may communicate directly with cloud-based services or via an edge gateway that acts as a data concentrator and/or filter and protocol converter.

The IIoT is of a new form of automation system architecture that balances the computational and life-cycle benefits of cloud computing with the requisite on-premises, appliance-hosted capabilities necessary to provide safe, secure, and long-lasting automation for complex manufacturing systems and processes [7].

Virtualization is the creation of a virtual version of something, such as an operating system, a server, a storage device, or network resource. There are many different types of virtualizations.

Virtualization, a technology which has been used in the data centers of the world, is delivering proven savings to the IT industry. Virtualization technology can now also be applied to open industrial control systems to reduce their lifecycle cost and improve manageability. Plant production teams globally are trying to reduce their operating costs in any way that they can, while maintaining plant safety and production levels. At the same time, companies are coping with less staff and having to bring new projects online faster and for less cost than ever before. For existing projects and systems, production teams need to:

- Reduce hardware and operating system (OS) changes
- Improve computer platform resource utilization
- Make the system easier to maintain

Virtualization has matured as a technology from a performance and reliability perspective and is ready to solve the above challenges facing the industrial control industry.

The term virtualization is the separation of hardware or software requests from the actual physical hardware or operating system resources that are responsible for their completion.

Virtualization places a buffer or Abstraction Layer between the component making the request (the requestor) and the one responsible for completing it. This prevents the requestor from having to know exactly how to complete the request on a given piece of hardware or operating system. Therefore, at a high level, virtualization provides:

- The ability for multiple requestors to simultaneously make “virtual” hardware and software requests in total isolation from one another
- The insulation of the requestor from changes to underlying hardware or operating system

Virtualization technology allowed to create real working mechanism to digitize control system asset in the cloud infrastructure and simulate it in autonomous mode. Such cloud simulated platforms, also called virtual engineering platforms, and also could be named as virtual laboratories [8, 9] 10 years ago sensitively helped designers and manufacturers of automation systems of industrial facility to resolve their problems in challenges which they faced mostly during design and construction of new industrial facilities (please refer to Fig. 2).

Virtual lab systems play a vital role in most industrial control settings. Systems replicating the operating environment without connecting to the live process provide a safe method to test changes or updates without risking disruption or safety.

They do, however, have their limitations and three of them are critical.

The first is the difficulty in keeping the lab system current. Over time changes to the system or process will see differences between the lab system and live environment emerge. Failure to adjust for modifications, updates, upgrades, and expansions will reduce the reliability and utility of the lab system: Its tests will only accurately illustrate the impact of changes on the live system if it accurately reflects that system. A digital twin of the operating environment needs to be an identical twin to be of use.

This gives rise to a further weakness, for it means that each new test requires a new lab system to be built, ensuring it accurately reflects the current operation. For smaller projects and testing of hotfixes or back-ups, the work and cost involved mean this is rarely practical. Such changes go untested.

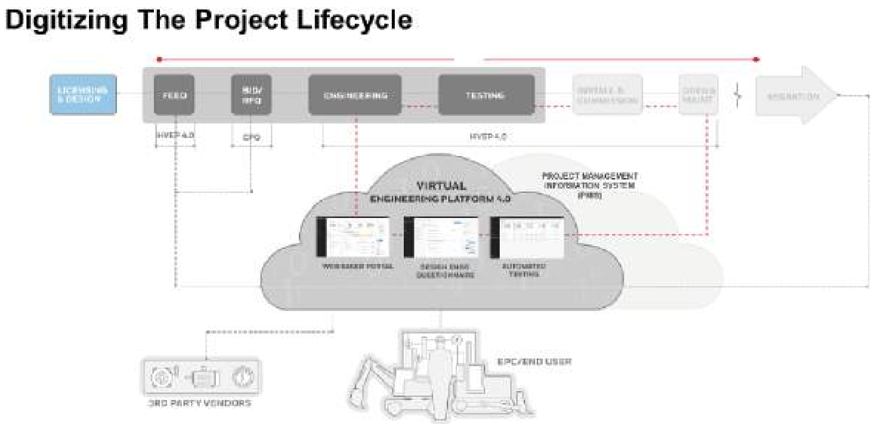


Fig. 2. Virtual Engineering Platform as forerunner of ICSS Digital Twin

There is a final, less discussed factor that decreases its utility, too: Lack of access. There is little use in a lab system if those that need it can't get to it. With dispersed operations, assets, and people, that again can make lab testing impractical even where it may be possible.

One way around this weakness is to network the lab system in some way, which is what many plants do to allow users remote access. That can solve the problem of access, but it opens a can of worms regarding security. Such non-standard networks present a key area of vulnerability for operators, especially when the labs might run on earlier operating system versions or require exceptions for patches [10].

It is critical to be clear: The danger is not the lab system itself. By definition, the lab system should have no connection to the live process; it simply replicates the operating environment, effectively a simulation of the control system. However, such ad-hoc connectivity can provide malicious actors with a route in. To take an early example, it was the vulnerability of the retailer's network providing access to the heating, ventilation, and air conditioning systems that ultimately gave hackers access to Target's payment systems in 2013 – still one of the biggest security breaches in history.

In an environment where almost four in ten industrial control systems were targeted by malicious activity last years, operators face a difficult choice: Either run a completely un-networked and air-gapped lab system, there further limiting its utility to the business; or run the risk of networking it and potentially leaving themselves vulnerable to attacks.

So full functional ICSS digital twin solution overcame these barriers. ICSS Digital Twin was built to address all the challenges and weaknesses in traditional lab systems – including security.

ICSS digital twin provides a "lab system as a service" that is always available and consistently accurate. A subscription service, it is continually updated to reflect changes to the operating environment, providing a consistent digital twin. Because it's continuously maintained and hosted by manufacturer of control system, it is always available for users to test even minor changes, patches, and upgrades.

It has a wide range of use cases:

– Functional reviews and impact analysis
– Remote factory acceptance tests (FAT)
– Training
– Documenting digital changes.

Its features include anomaly detection, automated change management, configuration compliance checks. For engineering and design, it helps eliminate the need for on-site reworks, ensure execution quality, simplify system documentation updates, and provide a historical overview of the control changes. For project execution and testing, it centralizes planning and resource allocation and provides project visibility, improving collaboration, helping cut travel costs and time, reducing risks, and ensuring compliance with detailed design specifications.

Apart from all this, it addresses the two fundamental weaknesses of traditional on-site lab systems: Their tendency to deviate from the current system setup and their impracticability for smaller systems. But it also resolves the central conflict such systems present regarding access and security. Figure 3 illustrates simplified conceptual vision of ICSS Digital Twin.

The benefits of a hosted, software as a service (SaaS) solution for a lab system are two-fold. The first – accessibility – is inherent. A hosted solution provides access to authorized users across the enterprise where and when it's needed. Users have access to a consistent (and consistently updated) digital twin for their own use and to collaborate with others, regardless of their location.

The second is achieved through industry-leading levels of cyber security to give users complete confidence and control.

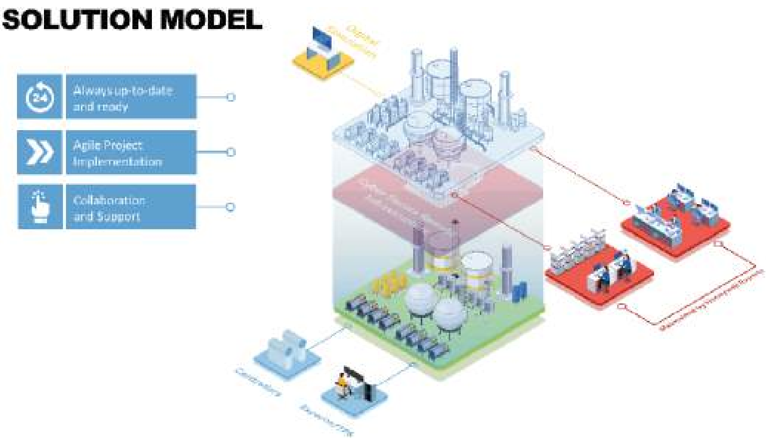


Fig. 3. Simplified conceptual vision of ICSS Digital Twin

As with a traditional lab system, the digital twin is entirely separate and independent of the actual control system. There is no communication between the two. But in addition to this, the solution also provides the highest standards of security.

First, virtual platform, which provides the service's infrastructure, is continually updated. There's no need for users to maintain the system. Effectively centralizing the lab system solution, there's no need to update the virtual machines or apply patches. Users don't even need anti-malware protection; it is built into the virtual platform infrastructure itself rather than dependent on the virtual machine. The centralized approach doesn't just simplify the management of the solution for users; it also ensures this level of security is consistently maintained.

Moreover, each virtual machine is entirely logically separated. There are no network paths between different subscribers nor between the virtual machine and the customer's operating system. Access can be managed using multi-factor authentication – giving users complete oversight and control of who can connect and view the solution.

3 Conclusion

Referring to picture illustrated in Fig. 1, ICSS Digital Twin replies on initially set challenges and targets which can be summarized as findings and added values of Digital Twin technology implementation for live industrial PCS as illustrated in Fig. 4.

Combining the best of a traditional lab system with the latest SaaS (software-as-a-service) solutions for connectivity, ICSS Digital Twin offers a solution that achieves the best of both worlds: readily available access and tight security.

It also provides a cost-effective solution to address the traditional limitations of lab systems, allowing users to test more frequently and have greater confidence in the results. It provides a digital twin that is always current, always available, and always secure. Once the industry adopts such an approach, it will never go back.

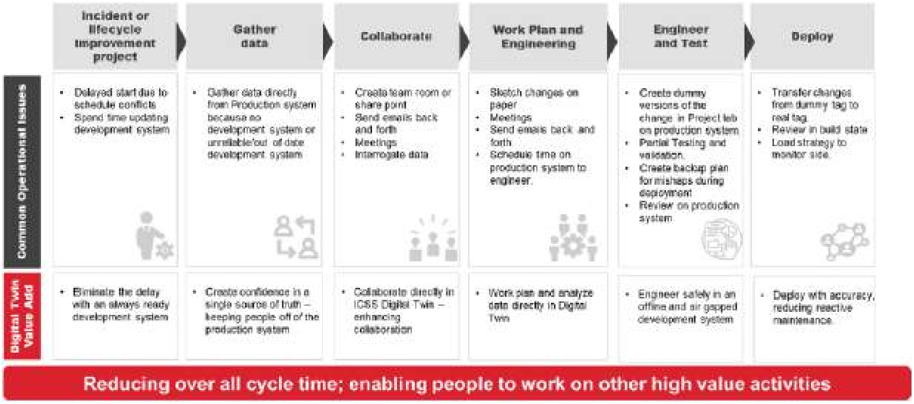


Fig. 4. ICSS Digital Twin as solution to resolve issues in industrial operation.

Future research of digital twin technology in this area should be definitely concentrated on establishing unified mechanisms of digital twin technology applications for live industrial PCS/ICSS allowing to collect statistics of such references and further improvements of digital twin technology in industry.

References

1. Yusupbekov, N., Abdurasulov, F., Adilov, F., Ivanyan, A.: Application of industrial Internet of Things technologies at modern manufacturing enterprises. Soft Meas. Calculations **12**(12), 26–30 (2019)
2. Yusupbekov, N., et al.: Using of sentience platform for integration of intelligent systems and devices into cloud. Chem. Technol. Control Manage. **4–5**(12), 67–71 (2018)
3. Adilov, F., Ivanyan, A.: Creation of a digital infrastructure of an industrial enterprise based on advanced information and communication technologies on the example of the Ustyurt gas and chemical complex. In: International Scientific and Practical Conference on the Topic "Using Modern Information and Communication Technologies in the Implementation of Reforms in New Uzbekistan" 2021, pp. 337–342, Andijon (2021)
4. Aliev, I., Gazul, S., Bobova, A.: Virtualization technologies and platforms: comparative overview and updated performance tests. In: VII International Conference "Safety Problems of Civil Engineering Critical Infrastructures" (SPCECI 2021), AIP Conference Proceedings, vol. 2700(1), p. 040048 (2021). https://doi.org/10.1063/5.0125249
5. Ltd. Huawei Technologies Co.: Virtualization Technology. Cloud Computing Technology. Posts&Telecom Press, Beijing (2023)
6. Weibin, S., et al.: Cloud-edge computing-based ICICOS framework for industrial automation and artificial intelligence: a survey. J. Circ. Syst. Comput. **32** (2023). https://doi.org/10.1142/S0218126623501682
7. Yusupbekov, N., Sattarov, S., Adilov, F., Ivanyan, A.: Industrial Internet of things – new level in evolution of automation systems. Chem. Technol. Control Manage./J. Korea Multimedia Soc. Special Issue **5**, 21–25 (2016)
8. Falah, M., Sukaridhoto, S, Rasyid, M., Wicaksono, H.: Design of virtual engineering and digital twin platform as implementation of cyber-physical systems. In: Procedia Manufacturing 52, 5th International Conference on System-Integrated Intelligence, pp. 331–336. Published by Elsevier B.V. (2020). https://doi.org/10.1016/j.promfg.2020.11.055

9. Hertwig, M., Lentes, J., Barwasser, A., Schuseil, F.: Crowd engineering platform—functions supporting co-creation in product development. In: Intelligent and Transformative Production in Pandemic Times (2023). https://doi.org/10.1007/978-3-031-18641-7_66
10. Yusupbekov, N., Adilov, F., Ivanyan, A., Astafurov M.: Remote access and management of plants experience during pandemics time across the world. In: Kahraman, C., et al. (eds.) INFUS 2022. LNNS, vol. 505, pp. 308–316. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09176-6_36

Level of Automation Assessment for Controlled-Environment Hydroponic Agriculture via Fuzzy MCDM Method

Sarper Alem[1], Biset Toprak[2], Buke Tolga[3], and A. Çağrı Tolga[4](✉)

[1] Graduate School of Science and Engineering, Galatasaray University, Ortakoy, 34349 Istanbul, Turkey

[2] Department of Industrial Engineering, Istanbul Sabahattin Zaim University, Halkalı, 34303 Istanbul, Turkey
biset.toprak@izu.edu.tr

[3] Department of Accounting and Finance, Galatasaray University, Ortakoy, 34349 Istanbul, Turkey
btolga@gsu.edu.tr

[4] Department of Industrial Engineering, Galatasaray University, Ortakoy, 34349 Istanbul, Turkey
ctolga@gsu.edu.tr

Abstract. The increasing world population and urbanization rates, and inevitable climate change tendencies have threatened food security. Integrating emerging technologies such as robotics, artificial intelligence (AI), and the Internet of Things (IoT) into agriculture through automation is crucial to sustainably meet global food demand by efficiently utilizing available arable land and natural resources. Although Controlled Environment Agriculture (CEA), where every ecological component may be monitored, controlled, and processes are executed with higher-level automation, is an emerging solution for urban areas; establishing automation levels for energy management has been identified as one of the major challenges in developing an efficient and viable CEA system. In this study, three levels of automation, manual, semi-automation, and full automation, were evaluated for controlled-environment hydroponic agriculture (CEHA). In this regard, 15 criteria were determined, and TODIM (an acronym in Portuguese-TOmada de Decisão Iterativa Multicritério) method was applied to incorporate risk factors in the decision-making process. In order to cope with the vagueness in data collection, fuzzy sets were integrated into the method. As a result of the application, the semi-automation alternative showed the best performance with respect to the evaluation criteria.

Keywords: Controlled Environment Agriculture (CEA) · Vertical Farming · Hydroponics · Vertical Farming · Fuzzy TODIM

1 Introduction

It is estimated that the world population, which is expected to reach 9.7 billion in 2050, will increase the global food demand by 70% [12]. Although the growing world population requires the maximization of agricultural output per unit area, agricultural production may decrease by 10% due to climate change [3]. Therefore, smart agriculture

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 45–52, 2023.
https://doi.org/10.1007/978-3-031-39777-6_6

systems that combine major emerging technologies with agriculture are vital to feeding the global population by efficiently utilizing available land and natural resources [22]. Controlled environment agriculture (CEA) is an indoor farming practice that uses advanced technologies to control growth parameters and automate the system with the aim of offering improved product quality, optimized plant yields, and human labour [6, 13, 17]. Vertical farming (VF) is a soilless method of cultivation practised in vertically stacked layers in a controlled environment. All vertical farming systems use one of the three soil-free growing options: hydroponics, aeroponics, and aquaponics. Hydroponic is the most common vertical farming technology. In this method, nutrient and vitamin solutions are frequently monitored and controlled to maintain accurate chemical composition [5, 14, 16, 17]. It has the advantage of reducing water usage by up to 60–70% and nutrient usage by 50% compared to traditional agriculture [15]. In vertical farming systems, higher security, self-adaptation, and scalability can be achieved with IoT systems, while energy optimization can be achieved with automation [18]. Since advanced technical systems and skilled human labour are vital for building an efficient and flexible manufacturing system, decisions regarding automation level require a deep understanding of automation from industry and research perspectives [8]. There are various classifications for mechanization level and automation level in the literature. The level of automation (LoA) refers to the sharing between humans and machines with varying degrees of human involvement. According to previous taxonomies, the LoA falls into three categories: manual, semi-automation, and full automation [8].

In the literature, there are various studies on the automation of CEA and evaluating different CEA alternatives. De Silva *et al.* [7] introduced an automated framework that collects data from the soilless farming system, analyses the data, and executes water quality adjustments. Adhau *et al.* [2] built a self-controlled automated hydroponics system using an AVR microcontroller board to monitor, automate, and transfer data for IoT applications. Shristava *et al.* [16] developed an IoT platform through a mobile app to design and implement an automated VF system. Büyüközkan *et al.* [4] evaluated three vertical farming technologies with an integrated multi-criteria decision-making (MCDM) approach using Pythagorean fuzzy sets. Tolga and Başar [20] evaluated basic, IoT, and automated vertical farms using WEDBA and MACBETH methods and observed that the automation alternative outperformed the others. In literature, establishing automation levels for energy management is indicated as one of the most significant challenges in developing efficient and viable CEA systems [13].

To the best of the authors' knowledge, there is no study in the literature evaluating automation levels for CEA. This study evaluates the manual, semi-automation, and full automation levels for the controlled environment hydroponic agriculture (CEHA) system on the fuzzy multi-criteria decision-making (MCDM) framework. For this purpose, evaluation criteria were determined by a literature review and then finalized by inquiries from two experts from two different CEA companies in Istanbul. Then, the manual, semi-automation, and full automation alternatives were evaluated by trapezoidal fuzzy extended TODIM (an acronym in Portuguese-TOmada de Decisão Iterative Multicritério) method.

This study consists of 5 sections, including introduction. The following section introduces the evaluation criteria for the LoA assessment of CEHA. Section 3 presents the application procedure of the proposed method. In Sect. 4, a real case application for LoA assessment for CEHA in Istanbul is presented. The last section concludes the study and presents suggestions for future studies.

2 Evaluation Criteria

The evaluation criteria were determined by literature review and finalized after the expert opinions from two companies, dealing with controlled-environment hydroponic farming. The criteria set consists of qualitative and quantitative criteria. The criteria between C1-C4 were categorized as capital expenses (CapEx), whereas the criteria between C5-C8 were categorized as operational expenses (OpEx). In addition, qualitative criteria (C9 to C15) were also considered to evaluate the alternatives.

C1) Main setup: The equipment cost for farming activities such as racks, pumps, pipes, ladders, etc. **C2) Sensor cost:** The implementation cost of the sensors for the data collection from the system parameters such as crop situation, temperature, and humidity, pH, etc. **C3) Robotic automation cost:** The implementation cost of the robotics for the automatized tasks that will be executed by robots. **C4) Application control system cost:** The system and software costs required to control robots, sensors, and growth parameters. **C5) Labor cost:** The labor requirement is dependent on the automation level of the system. **C6) Renting cost:** Different area sizes are considered for different automation levels. This criterion defines the unit rental cost for each LoA. **C7) Energy consumption cost:** The water and electricity consumption level of the system. **C8) Maintenance cost:** The maintenance cost of the sensors, machines, robots, and tools in the system. **C9) Venture capital attractiveness:** The level of automation is an important criterion for venture capital firms' decisions on their investment. **C10) R&D capabilities:** The capability of sustaining R&D activities for increasing product quality and shortening crop cycle by controllable growth parameters and developing new growing methods. **C11) Production scale:** The total output amount of the system per unit time. **C12) Production hygiene:** The condition of the harvested products, which do not contain chemical and soil residues and are suitable for eating without the need for washing. **C13) Product variety:** The variety of crops that can be grown at the different automation levels. **C14) Product quality:** The compliance of quality standards of agricultural products and quality metrics such as color, size and shape, and flavor. **C15) Situation awareness:** The capability of providing situational awareness of different levels of automation.

3 Fuzzy Extended TODIM

Decisions about automation levels that have been considered a structural choice category within the manufacturing strategy [11] require the consideration of many factors. Since CEA is a fresh concept, the inconvenience of obtaining accurate data may be addressed by using fuzzy sets, which was initially introduced by Zadeh [21]. This study will extend the TODIM method with trapezoidal fuzzy numbers due to its practicality. A trapezoidal fuzzy number is defined as $\tilde{A} = (a_1, a_2, a_3, a_4)$, $a_1 \leq a_2 \leq a_3 \leq a_4$. with a membership function of $\mu_{\tilde{A}}(x) : R \to [0, 1]$, which equals to:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & , x\in [a_1, a_2] \\ 1 & , x\in [a_2, a_3] \\ \frac{a_4-x}{a_4-a_3} & , x\in [a_3, a_4] \\ 0 & , others \end{cases} \quad (1)$$

The distance between any two fuzzy numbers, $\tilde{A}$ and $\tilde{B}$, is calculated by the following equation:

$$d(\tilde{A}, \tilde{B}) = \sqrt{\frac{[(a_1 - b_1)^2 + 2(a_2 - b_2)^2 + 2(a_3 - b_3)^2 + (a_4 - b_4)^2]}{6}} \quad (2)$$

For the defuzzified value (κ) of any fuzzy number and its maximum distance to zero (ψ) is calculated as follows: [9]:

$$\kappa(\tilde{A}) = \frac{1}{3}\left(a_1 + a_2 + a_3 + a_4 - \frac{a_3a_4 - a_1a_2}{(a_3 + a_4) - (a_1 + a_2)}\right) \quad (3)$$

$$\psi(\tilde{A}) = \begin{cases} \tilde{A} & \text{if } \kappa(\tilde{A}) > 0 \\ \tilde{0} & \text{if } \kappa(\tilde{A}) \leq 0 \end{cases} \quad (4)$$

where $\tilde{0} = (0, 0, 0, 0)$.

TODIM, developed by Gomes and Lima [10], is a prospect theory-based method that incorporates risk factors in the decision-making process. Since choosing the level of automation is a strategic decision, we use this method to take into account risk factors in our evaluation process.

In our decision-making process, there are k decision-makers ($D = \{D_1, D_2, ..., D_k\}$), n alternatives ($A = \{A_1, A_2, ..., A_n\}$), and m criteria ($C = \{C_1, C_2, ..., C_m\}$). A_i represents the ith alternative, where C_j represents the jth criterion. Each criterion has different weights assigned by decision-makers.

The steps of fuzzy TODIM are as follows [19]:

Step 1: Establish the decision matrix $\mathbf{DM}_z$.

$$DM_z = \left[\tilde{U}^t_{ij}\right]_{n\times m} = \begin{bmatrix} \tilde{U}^t_{11} & \tilde{U}^t_{12} & \dots & \tilde{U}^t_{1m} \\ \tilde{U}^t_{21} & \tilde{U}^t_{22} & \dots & \tilde{U}^t_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{U}^t_{n1} & \tilde{U}^t_{n2} & \dots & \tilde{U}^t_{nm} \end{bmatrix} \quad (5)$$

where $\tilde{U}^t_{ij}$ denotes the performance value of alternative a_i on criterion c_j assigned by the zth decision maker, where $1 \leq i \leq n, 1 \leq j \leq m$, and $1 \leq t \leq k$. Surveys were conducted with the decision makers considering the linguistic scale in Table 1 for the evaluation of qualitative criteria. In contrast, data from iFarm [1] was used as a reference for the quantitative criteria, and calculations were made considering the alternatives.

Table 1. Preference ratings in case of linguistic terms for alternatives.

Linguistic Terms	Fuzzy Number
Very Poor (VP)	(0, 0, 0, 1)
Poor (P)	(0, 1, 2, 3)
Medium Poor (MP)	(2, 3, 4, 5)
Medium/Fair (M)	(4, 5, 5, 6)
Medium Good (MG)	(5, 6, 7, 8)
Good (G)	(7, 8, 9, 10)
Very Good (VG)	(9, 10, 10, 10)

Table 2. Importance terms in case of linguistic terms for criteria.

Linguistic Terms	Fuzzy Number
Lowest (LT)	(0.0, 0.0, 0.0, 0.1)
Very low (VL)	(0.0, 0.1, 0.2, 0.3)
Low (L)	(0.2, 0.3, 0.4, 0.5)
Medium (M)	(0.4, 0.5, 0.5, 0.6)
High (H)	(0.5, 0.6, 0.7, 0.8)
Very High (VH)	(0.7, 0.8, 0.9, 1.0)
Highest (HT)	(0.9, 1.0, 1.0, 1.0)

Step 2: Establish the average decision matrix $\overline{G}$.

$$\tilde{U}_{ij} = \left(\left(\tilde{U}_{ij}^{1} \oplus \tilde{U}_{ij}^{2} \oplus \ldots \oplus \tilde{U}_{ij}^{k}\right)/k\right) \tag{6}$$

$$\overline{G} = \left[\tilde{U}_{ij}\right]_{n \times m} \tag{7}$$

where $\tilde{U}_{ij}$ represents the average performance value of alternative a_i on criterion c_j, $1 \leq i \leq n, 1 \leq j \leq m.$

Step 3: Establish the weighting matrix W_t for the criteria for tth decision maker considering the linguistic terms in Table 2.

$$W_t = \left[\tilde{w}_j^t\right]_{m \times 1} = \begin{bmatrix} \tilde{w}_1^t \\ \tilde{w}_2^t \\ \vdots \\ \tilde{w}_m^t \end{bmatrix} \tag{8}$$

where $\tilde{w}_j^t$ represents the weight of jth criterion c_j assigned by the tth decision maker, $1 \leq j \leq m, 1 \leq z \leq k.$

Step 4: Compute the normalized $\overline{\mathbf{N}}^{\nu}$ matrix for criteria c_j, where $1 \leq j \leq m$. The normalization process for benefit-type and cost-type criteria is applied as follows:

$$\overline{\mathbf{N}}^{\nu} = \left[\tilde{U}_{ij}^{\nu}\right]_{n \times m} = \begin{cases} \left[\tilde{U}_{ij} \oslash \tilde{U}_{ijmax}^{c_j}\right]_{n \times m} & \text{for benefit-type criteria} \\ \left[\tilde{U}_{ijmin}^{c_j} \oslash \tilde{U}_{ij}\right]_{n \times m} & \text{for cost-type criteria} \end{cases} \tag{9}$$

Step 5: Calculate the relative criteria weights $\tilde{w}_{rc_j}$ for each criterion by taking the greatest criterion weight as a reference criterion weight, $\tilde{w}_r$.

$$\left[\tilde{w}_{rc_j}\right]_{1 \times m} = \left[\tilde{w}_{c_j} \oslash \tilde{w}_r\right]_{1 \times m} \tag{10}$$

where $\tilde{w}_{c_j}$ is criterion weights, and $1 \leq j \leq m$.

Step 6: Calculate the dominance degrees of any alternative A_{a_i} over each alternative A_i according to each criterion c_j; $\tilde{\phi}_{c_j}(A_{a_i}, A_i)$.

$$\tilde{\phi}_{c_j}(A_{a_i}, A_i) = \begin{cases} \sqrt{\frac{\tilde{w}_{rc_j}}{\sum_{j=1}^{m} \tilde{w}_{rc_j}} d(\tilde{U}^v_{a_i j}, \tilde{U}^v_{ij})}, & \text{if } \kappa(\tilde{U}^v_{a_i j}) - \kappa(\tilde{U}^v_{ij}) > 0 \\ 0, & \text{if } \kappa(\tilde{U}^v_{a_i j}) - \kappa(\tilde{U}^v_{ij}) = 0 \\ -\frac{1}{\theta}\sqrt{\frac{\sum_{j=1}^{m} \tilde{w}_{rc_j}}{\tilde{w}_{rc_j}} d(\tilde{U}^v_{a_i j}, \tilde{U}^v_{ij})}, & \text{if } \kappa(\tilde{U}^v_{a_i j}) - \kappa(\tilde{U}^v_{ij}) < 0 \end{cases} \tag{11}$$

where θ is attenuation factor of risk, (A_{a_i}, A_i) is the the referred alternative over other i alternatives, $d(\tilde{U}^v_{a_{ij}}, \tilde{U}^v_{ij})$ is the distance between performance values of each criterion j to each alternative i. $\kappa(\tilde{U}^v_{a_i j})$, and $\kappa(\tilde{U}^v_{ij})$, obtained using Eq. 3, denote the crisp values of related performance values, respectively.

Step 7: Compute the dominance of each alternative over other alternatives by using the Eq. below:

$$\tilde{\eta}(A_{a_i}, A_i) = \sum_{j=1}^{m} \tilde{\phi}_{c_j}(A_{a_i}, A_i)\ \forall (a_i, c_j) \tag{12}$$

where $1 \leq i \leq n$, and $1 \leq j \leq m$.

Step 8: Transpose the dominance values and constitute the $\left[\tilde{\eta}(A_{a_i}, A_{a_i})\right]_{n\times n}$ matrix for ease of computation.

Step 9: Calculate the overall dominance degree of each alternative i from the other alternatives according to all criteria.

$$\tilde{\xi}_i = \frac{\sum \tilde{\eta}(A_{a_i}, A_i) - min \sum \tilde{\eta}(A_{a_i}, A_i)}{max \sum \tilde{\eta}(A_{a_i}, A_i) - min \sum \tilde{\eta}(A_{a_i}, A_i)} \tag{13}$$

where $1 \leq i \leq n$.

Step 10: Rank the alternatives in descending order of overall value, attaining the crisp values of each alternative i, $\kappa(\tilde{\xi}_i)$ using Eq. 3.

In the next section, a real case application with the fuzzy TODIM method is presented and the results are delivered.

4 Application

In this study, three automation levels, manual, semi-automation, and full automation, for CEHA were evaluated via the fuzzy TODIM method. In the manual CEHA system, every process is performed by human labor. In the semi-automated CEHA system, seeding and harvesting processes are performed by human labor while planting and packaging are performed by robots and sensors. In the fully automated CEHA system, each process is performed by robots and sensors. Linguistic evaluation of the alternatives and criteria was performed with two CEHA experts through surveys. Moreover, the cost items were calculated by using the iFarm startup cost calculator for leafy greens [1]. For the manual, semi-automation, and full automation alternatives, calculations were made based on the assumption of 500 sqm, 2500 sqm, and 5000 sqm area, respectively. Because the size of the area directly affects the calculation, total costs were transformed to a unit cost of 1 sqm for each alternative. Using data from iFarm, the main setup, labor cost, energy cost, maintenance cost, and human labor/hour were calculated based on the system requirements of each alternative. The labor cost was calculated using the human labor/hour data considering seeding, planting, harvesting, and packaging processes. Since planting and packaging in semi-automation and seeding, planting, harvesting, and packaging in full automation are done by technology; labor costs are ignored in the calculations of these processes. At first, we made an economic analysis based on Equal Uniform Annual Cost (EUAC) method. As searched from the economic markets, the interest rate for dollar is found as 9% per year, continuously compounded. As a result of the calculations; manual, semi-automation and full automation alternatives were ranked according to their increasing annual costs, respectively. In contrast, as a result of the fuzzy TODIM method according to the evaluation criteria, the semi-automation alternative performed best, followed by full automation and manual. The result is as expected since the both decision makers give high importance to venture capital attractiveness, R&D capabilities, and production hygiene.

5 Conclusion

The development of CEA practices is essential to meeting the projected growth in global food demand sustainably and efficiently. In this regard, evaluating LoA for CEA, an emerging method in urban areas, is critical for entrepreneurs and agribusinesses. In this study, we presented a fuzzy MCDM framework to evaluate LoA for CEHA enterprises. After the data collection process was completed, the EUAC analysis was performed, and the ranking of the alternatives according to their increasing annual costs was obtained as manual, semi-automated, and full automation, respectively. On the other hand, when the alternatives were evaluated via the fuzzy TODIM method according to the determined criteria, the semi-automation alternative showed the best performance, followed by full automation and manual. In future investigations, it might be possible to increase the number of decision-makers. In addition, the criteria set might be extended to include environmental, economic, and social sustainability.

References

1. Vertical farming technology, ifarm. https://ifarm.fi/
2. Adhau, S., Surwase, R., Kowdiki, K.H.: Design of fully automated low cost hydroponic system using labview and avr microcontroller. In: 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), pp. 1–4 (2017)

3. Beacham, A.M., Vickers, L.H., Monaghan, J.M.: Vertical farming: a summary of approaches to growing skywards. J. Horticult. Sci. Biotechnol. **94**(3), 3277–283 (2019)
4. Büyüközkan, G., Göçer, F., Uztürk, D.: A novel pythagorean fuzzy set integrated choquet integral approach for vertical farming technology assessment. Comput. Ind. Eng. **158**, 107384 (2021)
5. Chaudhry, A.R., Mishra, V.P.: A comparative analysis of vertical agriculture systems in residential apartments. In: 2019 Advances in Science and Engineering Technology International Conferences (ASET), pp. 1–5 (2019)
6. Chaux, J.D., Sanchez-Londono, D., Barbieri, G.: A digital twin architecture to optimize productivity within controlled environment agriculture. Appl. Sci. **11**(19), 8875 (2021)
7. De Silva, P.P., De Silva, P.A.: Ipanera: an industry 4.0 based architecture for distributed soilless food production systems. In: 2016 Manufacturing & Industrial Engineering Symposium (MIES) (2016)
8. Frohm, J., Lindström, V., Stahre, J., Winroth, M.: Levels of automation in manufacturing. Ergonomia Int. J. Ergon. Human Fact. **30**(3), 1–37 (2008)
9. Ghorabaee, M.K., Zavadskas, E.K., Amiri, M., Turskis, Z.: Extended edas method for fuzzy multi-criteria decision-making: an application to supplier selection. Int. J. Comput. Commun. Control **11**(3), 358–371 (2016)
10. Gomes, L., Lima, M.: Todim: basics and application to multicriteria ranking of projects with environmental impacts. Found. Comput. Decis. Sci. **16**(4), 113–127 (1992)
11. Lindström, V., Winroth, M.: Aligning manufacturing strategy and levels of automation: a case study. J. Eng. Technol. Manag. **27**(3), 148–159 (2010)
12. Ng, A.K., Mahkeswaran, R.: Emerging and disruptive technologies for urban farming: a review and assessment. J. Phys. Conf. Ser. **2003**(1), 012008 (2021)
13. Shamshiri, R., et al.: Advances in greenhouse automation and controlled environment agriculture: a transition to plant factories and urban agriculture. Int. J. Agric. Biol. Eng. **11**(1), 1–22 (2018)
14. Rajan, P., Lada, R.R., MacDonald, M.T.: Advancement in indoor vertical farming for microgreen production. Am. J. Plant Sci. **10**(08), 1397–1408 (2019)
15. Saad, M.H.M., Hamdan, N.M., Sarker, M.R.: State of the art of urban smart vertical farming automation system: advanced topologies, issues and recommendations. Electronics **10**(12), 1422 (2021)
16. Shrivastava, A., Nayak, C.K., Dilip, R., Samal, S.R., Rout, S., Ashfaque, S.M.: Automatic robotic system design and development for vertical hydroponic farming using IoT and big data analysis. Mater. Today: Proc. **80**, 3546–3553 (2021)
17. Srivani, P., Manjula, S., et al.: A controlled environment agriculture with hydroponics: variants, parameters, methodologies and challenges for smart farming. In: 2019 Fifteenth International Conference on Information Processing (ICINPRO), pp. 1–8. IEEE (2019)
18. Teo, Y.L., Go, Y.I.: Techno-economic-environmental analysis of solar/hybrid/storage for vertical farming system: a case study, Malaysia. Renew. Energy Focus **37**, 50–67 (2021)
19. Tolga, A.C., Basar, M.: Hydroponic system evaluation in urban farming via fuzzy edas and todim methods. J. Intell. Fuzzy Syst. **39**(5), 6325–6337 (2020)
20. Tolga, A.C., Basar, M.: The assessment of a smart system in hydroponic vertical farming via fuzzy MCDM methods. J. Intell. Fuzzy Syst. **42**(1), 1–12 (2021)
21. Zadeh, L.A.: Fuzzy sets. Information and control **8**(3), 338–353 (1965)
22. Zhang, Z., Rod, M., Hosseinian, F.: A comprehensive review on sustainable industrial vertical farming using film farming technology. Sustain. Agric. Res. **10**(1), 46 (2020)

Navigating the Ethics of the Metaverse: A Fuzzy Logic Approach to Decision-Making

Rahib Imamguluyev[1](✉), Niyar Umarova[2], and Rena Mikayilova[3]

[1] IT and Engineering Department, Odlar Yurdu University, Baku AZ1072, Azerbaijan
rahib.aydinoglu@gmail.com

[2] "General Disciplines" Department, Azerbaijan Academy of Labor and Social Relations, Baku, Azerbaijan
niyar61@list.ru

[3] Department of Digital Technologies and Applied Informatics, Azerbaijan State University of Economics, Baku, Azerbaijan
rana.mikayilova@unec.edu.az

Abstract. The Metaverse is rapidly becoming a part of our daily lives, and with it come a host of ethical concerns. It is essential to address these issues to ensure a sustainable virtual future. In this article, we propose a fuzzy logic approach to decision-making in the Metaverse. This approach allows for more nuanced and adaptable responses to uncertain and complex situations. We examine how fuzzy logic can be applied to various ethical challenges, including privacy and security, social and cultural implications, economic and labor issues, and legal and regulatory challenges. We also discuss the benefits and limitations of using fuzzy logic in the Metaverse, as well as ways that companies and regulators can integrate this approach into their decision-making processes. Ultimately, we argue that prioritizing ethics and adopting a fuzzy logic approach to decision-making in the Metaverse are crucial steps towards ensuring a responsible and sustainable virtual future. By doing so, we can create a virtual world that is inclusive, equitable, and just for all.

Keywords: Fuzzy logic · Metaverse · Virtual world · Decision-making · Societal transformation · Ethical concerns · Privacy and security

1 Introduction

The Metaverse is a concept that has captured the imagination of many in recent years. It describes a virtual world that merges the physical and digital spaces, creating a new reality that could transform how we live, work, and socialize. The Metaverse would allow users to interact with each other, and with digital objects and spaces, in a seamless and immersive way [1].

While the Metaverse presents many exciting possibilities, it also raises a number of ethical concerns. As more people begin to participate in this virtual world, there is a need to ensure that it is a safe and equitable space for everyone. This article will explore

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 53–60, 2023.
https://doi.org/10.1007/978-3-031-39777-6_7

the ethical challenges that may arise in the Metaverse, with a focus on four key areas: privacy and security, social and cultural implications, economic and labor issues, and legal and regulatory challenges. By examining these challenges, we can begin to develop a deeper understanding of the issues at stake and work towards building a Metaverse that is both innovative and ethical [2].

1.1 Importance of Addressing Ethical Challenges in the Metaverse

As the Metaverse grows in popularity and becomes more integrated into our lives, it is important to address the ethical challenges that arise in this virtual world. The Metaverse is not just a game or a social platform; it has the potential to transform the way we interact with each other and with the world around us. It offers new opportunities for innovation, creativity, and collaboration, but it also presents new risks and challenges. One of the most significant ethical challenges in the Metaverse is privacy and security. As users interact with each other and with virtual environments, they are generating a vast amount of personal information that can be vulnerable to exploitation or hacking. Additionally, the Metaverse blurs the lines between public and private spaces, which can raise questions about the appropriate level of surveillance and control.

Social and cultural implications are also a concern in the Metaverse. The virtual world can be a powerful tool for exploring different identities and perspectives, but it can also be used to perpetuate harmful stereotypes and biases. Furthermore, the Metaverse can be a site of cultural appropriation and exploitation, as users can appropriate cultural symbols and practices without understanding their meaning or context.

Economic and labor issues are also a concern in the Metaverse. Virtual economies and labor markets can create new opportunities for entrepreneurship and employment, but they can also perpetuate exploitation and wage gaps. Moreover, there is a risk that virtual work could be used to circumvent labor laws and regulations, leading to further exploitation and inequality [3].

Finally, legal and regulatory challenges arise as the Metaverse continues to develop. There is a need for clear and consistent regulations to ensure that the Metaverse is a safe and equitable space for all. However, the complex and constantly evolving nature of the Metaverse makes it difficult to develop regulatory frameworks that are effective and responsive.

Overall, addressing the ethical challenges of the Metaverse is critical to ensuring that this virtual world is a safe and equitable space for all. A thoughtful and proactive approach is needed to navigate the complex and evolving ethical landscape of the Metaverse.

2 Overview of the Fuzzy Logic Approach to Decision-Making

Fuzzy logic is a method of reasoning that deals with imprecise or uncertain information. It allows for decision-making in situations where traditional logic models may not be effective. Fuzzy logic is based on the idea that variables can have degrees of membership to a set, rather than simply being a part of the set or not. This approach enables a more nuanced understanding of complex systems that are difficult to define precisely [14–16].

Fuzzy logic systems typically involve defining variables and their membership functions, rules that describe the relationship between variables, and a method of combining these rules to produce a final output. These systems can be used in a wide range of applications, including control systems, data analysis, and decision-making [16].

In the context of the Metaverse, a fuzzy logic approach to decision-making can help address the unique challenges of a constantly evolving virtual world with uncertain and imprecise information. By taking into account factors such as cultural context, user feedback, and the potential for harm or offense, a fuzzy logic model could enable a more flexible and adaptable approach to ethical decision-making in the Metaverse.

3 Ethical Challenges in the Metaverse

There are several ethical challenges in the Metaverse that need to be addressed. Some of these challenges include:

Privacy and Security: The Metaverse presents risks to personal information, such as identity theft, data breaches, and unauthorized access. Additionally, the use of virtual reality and other technologies in the Metaverse raises concerns about surveillance and the potential for abuse of personal information [4].

Social and Cultural Implications: The Metaverse raises questions about cultural appropriation, appropriate behavior, and the spread of misinformation. As virtual worlds can be accessed by people from all over the world, it is important to ensure that users respect each other's cultural norms and do not engage in behavior that may be offensive or harmful [5].

Economic and Labor Issues: The Metaverse presents new economic and labor challenges, such as the risk of exploitation and wage gaps in virtual economies and labor markets. It is important to ensure that the benefits of the Metaverse are distributed fairly and that users are not exploited for their labor or contributions to the virtual world [6].

Legal and Regulatory Challenges: The Metaverse raises questions about the legal and regulatory frameworks that should govern the virtual world. There is a need for clear guidelines on issues such as intellectual property, user-generated content, and the protection of personal information [7].

Overall, it is important to address these ethical challenges in the Metaverse to ensure that the virtual world is a safe and equitable space for all users. A fuzzy logic approach to decision-making could help address these challenges in a more adaptable and responsive way.

3.1 Privacy and Security

Privacy and security are significant ethical challenges in the Metaverse. As users engage with the virtual world, they create and share personal information, which could be compromised if not adequately protected. Additionally, virtual worlds are prone to hacking and other security breaches, which can lead to data theft, identity theft, and other harmful consequences [8].

Fuzzy logic could be useful in addressing these challenges by allowing for a more nuanced approach to security. For example, a fuzzy logic model could be used to determine the appropriate level of security measures needed for different types of personal information. The model could take into account factors such as the sensitivity of the information, the potential harm if the information is compromised, and the likelihood of a security breach occurring. This would allow for a more adaptable and responsive approach to security that recognizes that not all information requires the same level of protection [9].

3.2 Social and Cultural Implications

The Metaverse also raises important questions about social and cultural implications. As virtual worlds become more advanced and realistic, users may become more immersed in the environment and form strong emotional connections with their digital avatars and communities. This can lead to situations where individuals may act in ways that are different from their real-world personalities, or where cultural norms may be misunderstood or disregarded [10].

One ethical challenge in this area is the potential for cultural appropriation. In the Metaverse, users may have access to a range of cultural artifacts, from clothing and jewelry to music and dance. However, the use of these artifacts may be seen as inappropriate or even offensive if they are not used in a respectful way. For example, a non-Indigenous user wearing traditional Indigenous clothing or using traditional Indigenous music may be seen as appropriating Indigenous culture.

Another challenge is the need to reinforce societal norms in the Metaverse. As with any online community, there may be individuals who engage in behavior that is considered harmful or inappropriate. However, in the Metaverse, this behavior can be amplified due to the immersive nature of the environment. As a result, it is important to establish clear guidelines and rules of conduct to ensure that users feel safe and respected [11].

Finally, the Metaverse presents a unique opportunity to explore and challenge social norms in a safe and controlled environment. For example, users may be able to experiment with different gender identities or cultural experiences that are not available to them in the real world. However, it is important to consider the potential for harm or offense and to ensure that users are fully informed about the potential consequences of their actions.

3.3 Legal and Regulatory Challenges

As the Metaverse continues to develop and expand, legal and regulatory challenges are becoming increasingly important to address. One of the primary issues is the lack of clear legal frameworks governing virtual worlds, which can lead to uncertainty and ambiguity in terms of user rights and responsibilities.

For example, it may be unclear who is responsible for enforcing laws and regulations in the Metaverse, and what legal recourse users have in the case of disputes or illegal activities. Additionally, there may be challenges in enforcing intellectual property rights and preventing copyright infringement in the virtual world [12].

Regulators and policymakers are also grappling with questions around data protection and privacy in the Metaverse. As virtual environments become more sophisticated, they will likely collect and process vast amounts of user data, which raises concerns around data ownership, access, and use.

A fuzzy logic approach to decision-making could be particularly useful in navigating these legal and regulatory challenges. The flexibility and adaptability of fuzzy logic models could help regulators and policymakers respond to the constantly evolving nature of the Metaverse, while still upholding important legal and ethical principles [13].

4 Fuzzy Logic and Its Relevance to Decision-Making in the Metaverse

Fuzzy logic is a mathematical framework that allows for decision-making in situations with uncertain or imprecise information. Unlike traditional Boolean logic, which is binary and relies on strict true/false distinctions, fuzzy logic operates on a spectrum of values between 0 and 1. This allows for more nuanced and flexible decision-making that can take into account multiple factors and degrees of certainty [14].

In the context of the Metaverse, where regulations are still being developed and the virtual world is constantly evolving, fuzzy logic can be particularly useful in addressing ethical challenges. For example, in the area of privacy and security, a fuzzy logic model could take into account factors such as the sensitivity of personal information, the potential harm if the information is compromised, and the likelihood of a security breach occurring. This would allow for a more nuanced approach to security that recognizes that not all information requires the same level of protection [15].

Similarly, in the area of social and cultural implications, a fuzzy logic model could take into account factors such as cultural context, user feedback, and the potential for harm or offense in determining the appropriateness of certain types of behavior or content in the Metaverse. This would allow for a more flexible approach that recognizes the complex and constantly evolving nature of social and cultural norms [16].

Overall, fuzzy logic offers a way to navigate the complex and uncertain terrain of the Metaverse in a way that is adaptable and responsive to the unique challenges of a virtual world.

4.1 Examples of How Fuzzy Logic Can Be Applied to Ethical Challenges in the Metaverse

Privacy and Security: A fuzzy logic model could be used to determine the appropriate level of security measures needed for different types of personal information in the Metaverse. The model could take into account factors such as the sensitivity of the information, the potential harm if the information is compromised, and the likelihood of a security breach occurring. This would allow for a more nuanced approach to security that recognizes that not all information requires the same level of protection.

Social and Cultural Implications: A fuzzy logic model could be used to determine the appropriateness of certain types of behavior or content in the Metaverse. The model could take into account factors such as cultural context, user feedback, and the potential for harm or offense. This would allow for a more flexible approach that recognizes the complex and constantly evolving nature of social and cultural norms.

Economic and Labor Issues: A fuzzy logic model could be used to determine the appropriate level of compensation for virtual labor in the Metaverse. The model could take into account factors such as the complexity of the task, the rarity of the skill, and the potential for exploitation. This would allow for a more equitable approach to compensation that recognizes the unique challenges of virtual economies and labor markets.

Legal and Regulatory Challenges: A fuzzy logic model could be used to determine the appropriate level of regulation needed for different types of behavior or content in the Metaverse. The model could take into account factors such as the potential harm to users, the impact on society, and the potential for abuse. This would allow for a more nuanced approach to regulation that recognizes the unique challenges of a virtual world.

5 Advantages of Fuzzy Logic for Decision-Making in the Metaverse

Handling imprecise and uncertain information: Fuzzy logic allows decision-making in situations where information is uncertain or imprecise, which is particularly relevant in the constantly evolving and complex context of the Metaverse.

Flexibility: Fuzzy logic is flexible and allows for more nuanced decision-making that can adapt to the unique challenges of the virtual world.

Adaptable: Fuzzy logic can be updated and refined as the Metaverse evolves and new ethical challenges arise.

Customizable: Fuzzy logic models can be customized to specific situations or challenges, allowing for more targeted and effective decision-making.

Reflects human thinking: Fuzzy logic is based on the way humans think and make decisions, making it intuitive and accessible for decision-makers.

Can handle multiple variables: Fuzzy logic can handle multiple variables and criteria simultaneously, allowing for a more comprehensive and accurate assessment of a situation.

Overall, fuzzy logic provides a more nuanced and adaptable approach to decision-making in the Metaverse that can better address the complex and constantly evolving ethical challenges of a virtual world.

5.1 Implications for Companies and Regulators

Existing laws and regulations may not directly apply to the Metaverse due to its A fuzzy logic approach to decision-making in the Metaverse has implications for companies and regulators. Companies that operate in the Metaverse have a responsibility to address the ethical challenges that arise in this virtual world. They can use fuzzy logic to develop more nuanced and responsive approaches to addressing these challenges. For example, companies can use fuzzy logic to determine the appropriate level of security measures needed for different types of personal information or to assess the appropriateness of certain types of behavior or content based on cultural context and user feedback. Regulators also have an important role to play in governing the Metaverse. They can use fuzzy logic to develop regulatory frameworks that are more adaptable and responsive to the unique challenges of a virtual world. For example, regulators can use fuzzy logic to determine the appropriate level of regulation needed for different types of activities in the Metaverse or to assess the potential impact of new technologies or business models on ethical considerations.

Overall, companies and regulators should work together to address the ethical challenges of the Metaverse in a way that promotes safety, fairness, and equity for all users. A fuzzy logic approach to decision-making can be a useful tool in achieving this goal.

6 Conclusion

In conclusion, the Metaverse presents exciting opportunities for innovation and societal transformation, but also poses significant ethical challenges. To overcome these challenges, a fuzzy logic approach to decision-making can be useful because it allows decision-making in situations with uncertain or imprecise information, which is especially relevant in the ever-evolving virtual world. Fuzzy logic can be applied to areas such as privacy and security, social and cultural impacts, economic and labor issues, and legal and regulatory issues. Companies and regulators should prioritize ethics and apply a fuzzy logic approach to decision-making in the Metaverse to create a fairer and more ethical virtual world for all. Future studies will develop a fuzzy logic model for this issue.

References

1. Shi, F., et al.: A new technology perspective of the Metaverse: its essence, framework and challenges. Digital Commun. Netw. (2023). ISSN 2352–8648. https://doi.org/10.1016/j.dcan.2023.02.017
2. Yoo, K., Welden, R., Hewett, K., Haenlein, M.: The merchants of meta: a research agenda to understand the future of retailing in the metaverse. J. Retail. (2023). ISSN 0022-4359. https://doi.org/10.1016/j.jretai.2023.02.002
3. Weking, J., Desouza, K.C., Fielt, E., Kowalkiewicz, M.: Metaverse-enabled entrepreneurship. J. Bus. Venturing Insights **19**, e00375 (2023). ISSN 2352-6734. https://doi.org/10.1016/j.jbvi.2023.e00375
4. Vidal-Tomás, D.: The illusion of the metaverse and meta-economy. Int. Rev. Finan. Anal. **86**, 102560 (2023). ISSN 1057-5219. https://doi.org/10.1016/j.irfa.2023.102560
5. Zallio, M., Clarkson, P.J.: Designing the metaverse: a study on inclusion, diversity, equity, accessibility and safety for digital immersive environments. Telemat. Inf. **75**, 101909 (2022). ISSN 0736-5853. https://doi.org/10.1016/j.tele.2022.101909
6. Huynh-The, T., Pham, Q.V., Pham, X.Q., Nguyen, T.T., Han, Z., Kim, D.S.: Artificial intelligence for the metaverse: a survey. Eng. Appl. Artif. Intell. **117**(Part A), 105581 (2023). ISSN 0952-1976. https://doi.org/10.1016/j.engappai.2022.105581
7. De Felice, F., De Luca, C., Di Chiara, S., Petrillo, A.: Physical and digital worlds: implications and opportunities of the metaverse. Procedia Comput. Sci. **217**, 1744–1754 (2023). ISSN 1877-0509. https://doi.org/10.1016/j.procs.2022.12.374
8. Oh, H.J., Kim, J., Chang, J.J., Park, N., Lee, S.: Social benefits of living in the metaverse: the relationships among social presence, supportive interaction, social self-efficacy, and feelings of loneliness. Comput. Human Behav. **139**, 107498 (2023). ISSN 0747-5632. https://doi.org/10.1016/j.chb.2022.107498

9. Imamguluyev, R., Aliyeva, A.: Analysis of intelligent interfaces based on fuzzy logic in human-computer interaction. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M.B., Sadikoglu, F. (eds.) 15th International Conference on Applications of Fuzzy Systems, Soft Computing and Artificial Intelligence Tools – ICAFS-2022. ICAFS 2022. Lecture Notes in Networks and Systems, vol. 610. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-25252-5_94
10. Golf-Papez, M., et al.: Embracing falsity through the metaverse: the case of synthetic customer experiences. Bus. Horizons **65**(6), 739–749 (2022). ISSN 0007-6813. https://doi.org/10.1016/j.bushor.2022.07.007
11. Qamar, S., Anwar, Z., Afzal, M.: A systematic threat analysis and defense strategies for the metaverse and extended reality systems. Comput. Secur. **128**, 103127 (2023). ISSN 0167-4048. https://doi.org/10.1016/j.cose.2023.103127
12. Huynh-The, T., et al.: Blockchain for the metaverse: a review. Future Gener. Comput. Syst. 143, 401–419 (2023). ISSN 0167-739X. https://doi.org/10.1016/j.future.2023.02.008
13. Pamucar, D., Deveci, M., Gokasar, I., Delen, D., Köppen, M., Pedrycz, W.: Evaluation of metaverse integration alternatives of sharing economy in transportation using fuzzy Schweizer-Sklar based ordinal priority approach. Decis. Supp. Syst., 113944 (2023). ISSN 0167-9236.https://doi.org/10.1016/j.dss.2023.113944
14. Akram, M., Zahid, K., Kahraman, C.: Integrated outranking techniques based on spherical fuzzy information for the digitalization of transportation system. Appl. Soft Comput. **134**, 109992 (2023). ISSN 1568-4946. https://doi.org/10.1016/j.asoc.2023.109992
15. Abdullayev, T., Imamguluyev, R., Umarova, N.: Application of fuzzy logic model for optimal solution of light reflection value in lighting calculations. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F.M. (eds.) ICSCCW 2021. LNNS, vol. 362, pp. 384–391. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-92127-9_53
16. Zadeh, L.A., Aliev, R.: Fuzzy logic theory and applications: part I and part II, p. 61 (2018). https://doi.org/10.1142/10936

Intelligent Health Applications

Investigation the Activity of Medical Institutions for hospital care in Bulgaria with the InterCriteria Analysis Method

Boyko Mirazchiyski[1,2], Evdokia Sotirova[3,4](✉), and Veselina Bureva[4]

[1] Medical Faculty, Prof. Asen Zlatarov University, 8010 Burgas, Bulgaria
dirmbal@abv.bg

[2] University Hospital for Active Treatment, Address 73 Stefan Stambolov, 8000 Burgas, Bulgaria

[3] Faculty of Public Health and Health Care, Prof. Assen Zlatarov University, 8010 Burgas, Bulgaria
esotorova@btu.bg

[4] Laboratory of Intelligent Systems, Prof. Asen Zlatarov University, 8010 Burgas, Bulgaria
vbureva@btu.bg

Abstract. The activity of the hospitals of the medical establishments is analyzed annually. In the paper an information about the activity of the inpatient facilities of the medical institutions - for hospital care and other medical institutions (mental health centers, centers for skin-venereal diseases and complex oncology centers) was investigated.

To analyze the data in the study, an intelligent multi-criteria decision making tool named Intercriteria Analysis (ICA) method was applied.

The paper presents three applications of the ICA method: for the analysis of the bed stock in the medical institutions in Bulgaria; for the analysis of the dependencies of the medical institutions according to the bed stock and patient status in the Burgas region; for the analysis of data related to the passed patients in the wards and clinics of the University Hospital for Active Treatment in Burgas - 32 in number, of which 15 therapeutic, 10 surgical and 7 wards with Covid beds.

Keywords: InterCriteria Analysis · Intuitionistic Fuzzy Logic · Medical institutions · Patients

1 Introduction

Information on the health network by types of medical and health facilities and their territorial distribution in Bulgaria, as well as on the medical staff by specialties and categories, is analyzed annually. The statistical activity in the field of public health and health care is carried out by National Statistical Institute [12, 13] and the "National Health Care Data and e-Health" Directorate of National Center for Public Health and Analyses [11].

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 63–70, 2023.
https://doi.org/10.1007/978-3-031-39777-6_8

Inpatient facilities include hospitals, centers for skin and venereal diseases, mental health centers, and comprehensive cancer centers. Hospitals are multi-profile and specialized. In 2021, 1915843 patients (27.9 per 100 people of the population) were admitted to the inpatient facilities of medical institutions in Bulgaria. The frequency of hospitalizations is greatest in multi-profile hospitals for active treatment (13.5 per 100 people). This is followed by private hospitals (10.0 per 100 people), specialized hospitals (3.1 per 100 people, of which 1.8 per 100 people in specialized hospitals for active treatment, 1.1 per 100 people in rehabilitation hospitals) [14].

The subject of the present study is to analyze statistical data on the medical facilities in Bulgaria, the medical facilities on the territory of the Municipality of Burgas and data related to the largest hospital in the city of Burgas - UMBAL Burgas AD.

The aim of the study is to find dependencies between the multiple criteria by which the medical institutions in Bulgaria are evaluated annually, which will support the decision-making process.

As a tool to achieve the set objective InterCriteria Analysis (ICA) method is chosen [3]. In [16] the ICA approach was applied for analysis of a health data connected with in-patient and out-patient health care establishments in Bulgaria for 2010–2018. The ICA method has been successfully used to analyze oncological data [17] and establish dependencies between parameters related to malignant melanoma [18] and malignant neoplasms [15, 19]. The next applications can be for the other types of cancers [9, 20], ovarian cancer [7], cervical cancer [8], etc.

The paper is organized as follows. Section 1 is introduction to the topic. Section 2 is a short presentation of the ICA method. Section 3 presents the application of the ICA method. Three applications of the ICA method are demonstrated. In Sect. 3.1 the ICA method was applied to data for the medical facilities according bed fund in Bulgaria; in Sect. 3.2 the ICA method was used for analysis of the medical facilities dependences according bed fund and patient status in Burgas region; and in Sect. 3.3 the ICA was used for analysis of passed patients in University General Hospital for Active Treatment in Burgas AD. Section 4 presents the conclusions. Section 5 and Sect. 6 are acknowledgments and references.

2 Presentation of the ICA Approach

The ICA approach is an intelligent multi-criteria decision making tool, based on the theories of indexed matrices [1] and intuitionistic fuzzy sets [2]. It is introduced in [3].

The approach is applied on an indexed matrix, the columns of which contain the evaluated objects (O_i) and the rows the evaluation criteria (C_j), i, j ($1 \approx i \approx m$, $1 \approx j \approx n$). The elements of the matrix $a_{Ci,Oj}$ are real numbers or other objects that are commensurable with respect to a relation R with the other a-objects, so that for each m, k, l $R(a_{Cm,Ok}, a_{Cm,Ol})$ is defined.

Through ICA approach, criteria are compared with each other by comparing the values obtained by measuring or evaluating multiple objects according to these criteria. As a result, the dependencies between each pair of criteria are obtained in the form of intuitionistic fuzzy pairs (IFPs), [4].

For establishing the degree of dependence between IFPs of criteria was used the scale from [5]. It sets 11 degrees of dependence, which are shown in Fig. 1.

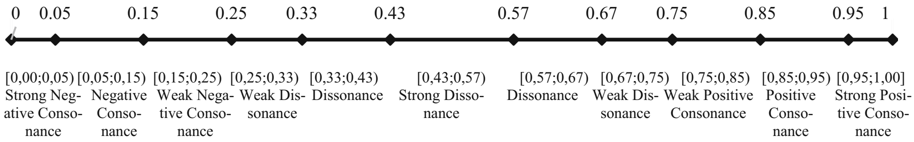


Fig. 1. Scale for determination of the degree of the dependences between criteria

3 Application of the ICA and Discussion

For the application of the ICA method the software developed under a free license was used [10]. Three applications of the method are presented below.

3.1 Application of the ICA Method for Analysis of the Medical Facilities According Bed Fund in Bulgaria

Information for a twelve-year period about the healthcare network in Bulgaria by types of medical and healthcare facilities was analyzed. The object of study in the research is the dynamics of the bed fund in all medical and health care facilities in the country: hospital medical facilities (multi-profile hospitals (C1), specialized hospitals (C2), centers for skin and venereal diseases (C3), complex oncology centers (C4), mental health centers (C5)), outpatient medical facilities (diagnostic-consultative centers (C6), medical centers (C7), dental centers (C8), medical and dental centers (C9), independent medical-diagnostic and medical-technical laboratories (C10)) and other medical and health care facilities (C11) for the period 2010–2021.

The ICA method was applied over index matrix with 11 rows (for the type of medical facility) and 12 columns (for bed fund). As a result, an index matrix with 55 IFPs of values were obtained (Table 1). The IFPs give the correlations of each pair of criteria (types of the medical facilities in Bulgaria) according bed fund.

Three pairs in positive consonance: “multi-profile hospitals”-“medical centers” with evaluation $\langle 0.92; 0.076\rangle$ “centers for skin and venereal diseases”-“other medical and health care facilities” with evaluation $\langle 0.86; 0.00\rangle$ and “multi-profile hospitals”-“medical and dental centers” with evaluation $\langle 0.86; 0.12\rangle$ This means, that these healthcare facilities have a similar development of the bed fund over the years.

Four pairs in weak positive consonance were obtained, the other IFPs are in weak dissonance, dissonance or strong dissonance.

3.2 Application of the ICA Method for Analysis of the Medical Facilities Dependences According Bed Fund and Patient Status in Burgas Region

Information for a four-year period (2018–2021) was analyzed for the activity of medical facilities in Burgas region - number of beds as of December 31, average annual number of beds, patients admitted, discharged sick, died, passed sick and days spent in bed. The data on the medical facilities cover: public medical facilities (PMFs), including multi-profile hospitals (MPs), specialized hospitals (SHs) and dispensaries (Ds) and private medical facilities (PrMFs) in Burgas region - number of beds as of December 31, average annual

Table 1. IFPs of the correlations between the types of health establishments in Bulgaria according bed fund for 2010 − 2021.

⟨ , ν⟩	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1	⟨1.00; 0.00⟩	⟨0.61; 0.39⟩	⟨0.03; 0.83⟩	⟨0.61; 0.39⟩	⟨0.24; 0,67⟩	⟨0.80; 0.20⟩	⟨0.92; 0.08⟩	⟨0.52; 0.39⟩	⟨0.86; 0.12⟩	⟨0.00; 0.00⟩	⟨0.09; 0.91⟩
C2	⟨0.61; 0.39⟩	⟨1.00; 0.00⟩	⟨0.30; 0.56⟩	⟨0.64; 0.36⟩	⟨0.53; 0.38⟩	⟨0.41; 0.59⟩	⟨0.56; 0.44⟩	⟨0.47; 0.44⟩	⟨0.53; 0.45⟩	⟨0.00; 0.00⟩	⟨0.42; 0.58⟩
C3	⟨0.03; 0.83⟩	⟨0.30; 0.56⟩	⟨1.00; 0.00⟩	⟨0.42; 0.44⟩	⟨0.68; 0.15⟩	⟨0.18; 0.68⟩	⟨0.08; 0.79⟩	⟨0.24; 0 8.53⟩	⟨0.06; 0.79⟩	⟨0.14; 0.00⟩	⟨0.86; 0.00⟩
C4	⟨0.61; 0.39⟩	⟨0.64; 0.36⟩	⟨0.42; 0.44⟩	⟨1.00; 0.00⟩	⟨0.50; 0.41⟩	⟨0.59; 0.41⟩	⟨0.59; 0.41⟩	⟨0.26; 0.65⟩	⟨0.50; 0.48⟩	⟨0.00; 0.00⟩	⟨0.48; 0.52⟩
C5	⟨0.24; 0,67⟩	⟨0.53; 0.38	⟨0.68; 0.15⟩	⟨0.50; 0.41⟩	⟨1.00; 0.00⟩	⟨0.14; 0.77⟩	⟨0.21; 0.70⟩	⟨0.35; 0.47⟩	⟨0.18; 0.71⟩	⟨0.09; 0.00⟩	⟨0.71; 0.20⟩
C6	⟨0.80; 0.20⟩	⟨0.41; 0.59⟩	⟨0.18; 0.68⟩	⟨0.59; 0.41⟩	⟨0.14; 0.77⟩	⟨1.00; 0.00⟩	⟨0.82; 0.18⟩	⟨0.44; 0.47⟩	⟨0.82; 0.17⟩	⟨0.00; 0.00⟩	⟨0.20; 0.80⟩
C7	⟨0.92; 0.08⟩	⟨0.56; 0.44⟩	⟨0.08; 0.79⟩	⟨0.59; 0.41⟩	⟨0.21; 0.70⟩	⟨0.82; 0.18⟩	⟨1.00; 0.00⟩	⟨0.53; 0.38⟩	⟨0.82; 0.17⟩	⟨0.00; 0.00⟩	⟨0.14; 0.86⟩
C8	⟨0.52; 0.39⟩	⟨0.47; 0.44⟩	⟨0.24; 0.53⟩	⟨0.26; 0.65⟩	⟨0.35; 0.47⟩	⟨0.44; 0.47⟩	⟨0.53; 0.38⟩	⟨1.00; 0.00⟩	0.53; 0.39⟩	⟨0.09; 0.00⟩	⟨0.30; 0.61⟩
C9	⟨0.86; 0.12⟩	⟨0.53; 0.45⟩	⟨0.06; 0.79⟩	⟨0.50; 0.48⟩	⟨0.18; 0.71⟩	⟨0.82; 0.17⟩	⟨0.82; 0.17⟩	⟨0.53; 0.39⟩	⟨1.00; 0.00⟩	⟨0.02; 0.00⟩	⟨0.09; 0.89⟩
10	⟨0.00; 0.00⟩	⟨0.00; 0.00⟩	⟨0.14; 0.00⟩	⟨0.00; 0.00⟩	0.09; 0.00⟩	⟨0.00; 0.00⟩	⟨0.00; 0.00⟩	⟨0.09; 0.00⟩	⟨0.02; 0.00⟩	⟨1.00; 0.00⟩	⟨0.00; 0.00⟩
11	⟨0.09; 0.91⟩	⟨0.42; 0.58⟩	⟨0.86; 0.00⟩	⟨0.48; 0.52⟩	⟨0.71; 0.20⟩	⟨0.20; 0.80⟩	⟨0.14; 0.86⟩	⟨0.30; 0.61⟩	⟨0.09; 0.89⟩	⟨0.00; 0.00⟩	⟨1.00; 0.00⟩

number of beds, patients admitted, discharged sick, died, passed sick and days spent in bed. The data was extracted from the site of the Burgas Regional Health Inspection [6].

The ICA method was applied over index matrix with 5 rows (for the type of medical facility) and 28 columns (for beds and patient movement). As a result, index matrices with membership part and non-membership part of the Intuitionistic fuzzy pairs were obtained. They give the correlations of each pair of criteria (types of the medical facilities) in the form of intuitionistic fuzzy pairs of values (Table 2).

From the obtained 10 IFPs one is in strong positive consonance ("specialized hospitals"-"private medical facilities, total" with evaluation ∫0.966; 0.013∇. That means that the beds and patient movement in the specialized hospitals and in the private medical facilities have almost identical behavior over the time. Three IFPs are in positive consonance ("public medical facilities, total"-"private medical facilities, total", "public medical facilities, total"-"specialized hospitals", "public medical facilities, total"-"multi-profile hospitals". That means that they have a very similar behavior over time. The other six IFPs are in weak positive consonance, i.e. there is a similar trend over the years for the bed fund and patient movement.

Table 2. IFPs of the correlations between the types of health establishments in Burgas region for 2018 − 2021.

⟨ , ν⟩	MPs	SHs	Ds	PMFs, total	PrMFs, total
MPs	⟨1.00;0.00⟩	⟨0.83;0.16⟩	⟨0.76;0.22⟩	⟨0.90;0.10⟩	⟨0.84;0.14⟩
SHs	⟨0.83;0.16⟩	⟨1.00;0.00⟩	⟨0.85;0.13⟩	⟨0.93;0.07⟩	⟨0.97;0.01⟩
Ds	⟨0.76;0.22⟩	⟨0.85;0.13⟩	⟨1.00;0.00⟩	⟨0.83;0.15⟩	⟨0.84;0.14⟩
PMFs, total	⟨0.90;0.10⟩	⟨0.93;0.07⟩	⟨0.83;0.15⟩	⟨1.00;0.00⟩	⟨0.93;0.05⟩
PrMFs, total	⟨0.84;0.14⟩	⟨0.97;0.01⟩	⟨0.84;0.14⟩	⟨0.93;0.05⟩	⟨1.00;0.00⟩

3.3 Application of the ICA Method for Analysis of Passed Patients in UMBAL Burgas

Here, the ICA method is applied to detect dependencies among data on the number of patients who passed in the departments and clinics of University General Hospital for Active Treatment in Burgas AD (UMBAL Burgas) for the period 2017–2022. The UMBAL Burgas is the largest hospital in South-Eastern Bulgaria. It has 680 beds for active treatment of internal, neurological, pediatric, obstetric and gynecological diseases and a maternity ward [21]. Data for 32 clinics and departments were analyzed:

- 15 therapeutic department: internal department of endocrinology, department of gastroenterology, department of cardiology, department of internal diseases, department for long-term treatment, internal department of nephrology, department of rheumatology, first department of pediatrics, second department of pediatrics, department of nervous diseases, department of infectious diseases, department of physical and rehabilitation medicine, department of anesthesiology and intensive care, department of skin and venereal diseases, department of neonatology.
- 10 surgical department: maternity department, gynecology department, first general surgery department, second general surgery department, vascular surgery clinic, neurosurgery department, orthopedics and traumatology department, urology clinic, otolaryngology clinic throat, department of eye diseases.
- 7 departments with COVID beds: long-term treatment unit with Covid beds, gastroenterology with Covid beds, rheumatology with Covid beds, internal medicine with Covid beds, endocrinology with Covid beds, nephrology with Covid beds, cardiology with Covid beds.

A total of 168,839 patients underwent surgery during the period 2017–2022. The average number of passed patients was 28148.83 ± 3450.86.

Application of the ICA Method by Years. The ICA method was applied over an index matrix with 6 rows (for years) and 32 columns (for clinics and departments). As a result, index matrices with membership part and non-membership part of the IFPs that give the correlations of each pair of years were obtained (Table 3).

Table 3. IFPs of the correlations between years in Burgas for 2017 − 2022.

⟨ , ν⟩	2017	2018	2019	2020	2021	2022
2017	⟨1.00;0.00⟩	⟨0.97;0.03⟩	⟨0.95;0.04⟩	⟨0.86;0.09⟩	⟨0.76;0.18⟩	⟨0.84;0.11⟩
2018	⟨0.97;0.03⟩	⟨1.00;0.00⟩	⟨0.95;0.03⟩	⟨0.87;0.08⟩	⟨0.780.17⟩	⟨0.850.10⟩
2019	⟨0.95;0.04⟩	⟨0.95;0.03⟩	⟨1.00;0.00⟩	⟨0.890.07⟩	⟨0.78;0.18⟩	⟨0.850.11⟩
2020	⟨0.860.09⟩	⟨0.87;0.08⟩	⟨0.890.07⟩	⟨1.00;0.00⟩	⟨0.85;0.15⟩	⟨0.900.08⟩
2021	⟨0.76;0.18⟩	⟨0.78;0.17⟩	⟨0.78;0.18⟩	⟨0.85;0.15⟩	⟨1.00;0.00⟩	⟨0.86;0.13⟩
2022	⟨0.84;0.11⟩	⟨0.85;0.10⟩	⟨0.85;0.11⟩	⟨0.90;0.08⟩	⟨0.86;0.13⟩	⟨1.00;0.00⟩

From the obtained 15 IFPs two are in strong positive consonance (“2017”–“2018” with evaluation $\langle 0.97;\ 0.03\rangle$ and (“2018”–“2019” with evaluation $\langle 0.95;\ 0.03\rangle$. The number of patients who passed in the departments and clinics of UMBAL Burgas is similar and in 2019 the dependence begins to weaken slightly. Nine IFPs are in positive consonance, four are in weak positive consonance. The other pairs have a very similar behavior over time.

Application of the ICA Method by Departments. The ICA method was applied over 3 index matrices: (1) one with 15 rows as criteria (for therapeutic departments) and 6 columns (for years), (2) one with 10 rows as criteria (for surgical departments) and 6 columns (for years), and one with 7 rows as criteria (for departments with Covid beds) and 6 columns (for years). As a result, 3 index matrices with 105 IFPs, 45 IFPs and 21 IFPs respectively were obtained. For the therapeutic departments two IFPs are in strong positive consonance: “Department of Cardiology”-“Department of Rheumatology” and “First Department of Pediatrics” -“Department of Physical and Rehabilitation Medicine”. For the surgical department one IFP is in strong positive consonance: “Maternity department”-“Clinic of Urology”. For the departments with Covid beds there is absolutely the same tendency for the passed patients in the Deptment for long-term treatment with Covid beds, Internal Medicine with Covid beds and Endocrinology with Covid beds on the one hand, and for Gastroenterology with Covid beds, Rheumatology with Covid beds and Cardiology with Covid beds on the other hand.

4 Conclusion

The study contains data related to the health system in the Republic of Bulgaria, Burgas region and the city of Burgas. Using the ICA method, hidden trends between types of medical institutions and patient movements were identified. Determined dependencies between the criteria by which evaluated the studied facilities are in the form of IFPs.

In the decision-making process, some of the criteria between which have strong dependencies with other criteria can be ignored, which leads to speeding up or cheapening the overall decision-making process, which is always beneficial, and in certain cases can be vital.

In a future study, the authors plan to propose an option in which the part of the data that does not carry additional information can be reduced, i.e., redundant information is reduced.

Acknowledgment. This research was funded in part by the European Regional Development Fund through the Operational 268 Programme "Science and Education for Smart Growth" under contract UNITe. BG05M2OP001–1.001–0004 269 (2018–2023)."

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Atanassov, K. (ed.) Index Matrices: Towards an Augmented Matrix Calculus. SCI, vol. 573. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10945-9
2. Atanassov, K. (ed.): On Intuitionistic Fuzzy Sets Theory. Springer, Berlin (2012). https://doi.org/10.1007/978-3-642-29127-2
3. Atanassov, K., Mavrov, D., Atanassova, V.: InterCriteria decision making: a new approach for multicriteria decision making, based on index matrices and intuitionistic fuzzy sets. Issu. Intuitionistic Fuzzy Sets Gener. Nets **11**, 1–8 (2014)
4. Atanassov, K., Szmidt, E., Kacprzyk, J.: On intuitionistic fuzzy pairs. Notes Intuitionistic Fuzzy Sets **19**(3), 1–13 (2013)
5. Atanassov, K., Atanassova, V., Gluhchev, G.: InterCriteria analysis: ideas and problems. Notes Intuitionistic Fuzzy Sets **21**(1), 81–88 (2015)
6. Burgas Regional Health Inspection, Activity of medical facilities. https://rzi-burgas.bg/statisticheski-danni/. Accessed 14 Mar 2023
7. Dobrev, P., Yordanov, A., Strashilov, S.: Synchronous primary cervical carcinoma and ovarian fibroma: challenge in surgery. Gazzetta Medica Italiana-Archivio per le Scienze Mediche **179**(5), 375–377 (2020)
8. Dobrev, P., Strashilov, S., Yordanov, A.: Metastasis of malignant melanoma in ovarium simulating primary ovarian cancer: a case report. Gazetta Medica Italiana – Archivio per le Scienze Medicine **180**, 867–869 (2021)
9. Ferlay, J., et al.: Cancer statistics for the year 2020: an overview. Int. J. Cancer **149**(4), 778–789 (2021)
10. InterCriteria Analysis Portal, Software. https://intercriteria.net/software/
11. National Center of Public Health and Analyses, Ministry of Health (2020). https://ncpha.government.bg/index/3035-deinost-na-lechebnite-zavedenia.html. Accessed 31 Jan 2023
12. National Statistical Institute, Statistics on public health and health and safety at work, https://www.nsi.bg/sites/default/files/files/metadata/Zdr_2_Methodology_EN.pdf
13. National Statistical Institute, Healthcare 2022, Republic of Bulgaria. https://www.nsi.bg/sites/default/files/files/publications/Zdraveopazvane_2022.pdf
14. National Statistical Institute, Health 2022. Republic of Bulgaria, Sofia (2023)
15. Sotirov, S., Bozova, G., Vasilev, V., Krawczak, M.: Clustering of InterCriteria analysis data using a malignant neoplasms of the digestive organs data. In: Atanassov, K.T., Atanassova, Vassia, Kacprzyk, Janusz, Kałuszko, Andrzej, Krawczak, Maciej, Owsiński, Jan W., Sotirov, Sotir S., Sotirova, Evdokia, Szmidt, Eulalia, Zadrożny, S.ławomir (eds.) IWIFSGN 2019 2019. AISC, vol. 1308, pp. 193–201. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77716-6_18

16. Sotirova, E., Vasilev, V., Sotirov, S., Bozov, H.: InterCriteria analysis of public health data in bulgaria. In: Kahraman, C., Onar, S.C., Oztaysi, B., Sari, I.U., Selcuk Cebi, A., Tolga, C. (eds.) INFUS 2020. AISC, vol. 1197, pp. 910–915. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-51156-2_105
17. Sotirova, E., Petrova, Y., Bozov, H.: InterCriteria analysis of oncological data of the patients for the city of burgas. Notes Intuitionistic Fuzzy Sets **25**(2), 96–103 (2019)
18. Sotirova, E., Bozova, G., Bozov, H., Sotirov, S., Vasilev, V.: Application of the InterCriteria analysis method to a data of malignant melanoma disease for the burgas region for 2014–2018. In: Atanassov, K.T., et al. (eds.) IWIFSGN 2019 2019. AISC, vol. 1308, pp. 166–174. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77716-6_15
19. Sotirova, E., Bozova, G., Bozov, H., Sotirov, S., Vasilev, V.: Application of the InterCriteria analysis method to a data of malignant neoplasms of the digestive organs for the burgas region for 2014–2018. In: 2019 Big Data, Knowledge and Control Systems Engineering (BdKCSE), Sofia, Bulgaria, pp. 1–6 (2019)
20. Stoyanov, V., Petkov, D., Bozdukova, P.: Pott's puffy tumor: a case report. Trakia J. Sci. **18**(1), 93–96 (2020). ISSN: 1313-3551
21. UMBAL Burgas AD. https://www.mbalburgas.com/bg/umbal-burgas-dnes. Accessed 13 Mar 2023

On an Intelligent Hybrid Intuitionistic Fuzzy Method for Breast Cancer Classification

Evdokia Sotirova[1,2], Hristo Bozov[3,4], Radovesta Stewart[1], Sotir Sotirov[1,5](✉), and Pavel Dobrev[3,4]

[1] Laboratory of Intelligent Systems, Prof. Asen Zlatarov University, 8010 Burgas, Bulgaria
{esotorova,ssotirov}@btu.bg
[2] University Hospital for Active Treatment, 73 Stefan Stambolov, 8000 Burgas, Bulgaria
[3] Prof. Asen Zlatarov University, 8010 Burgas, Bulgaria
pavel_dobrev_@abv.bg
[4] Oncology Complex Center - Burgas, 86 Demokratsiya Blvd, Burgas 8000, Bulgaria
[5] Southeast Digital Innovation Hub, 2 Serdika Str. TRIA City Center, Floor 2, Burgas, Bulgaria

Abstract. Breast cancer is the most common type of cancer in women. It is a disease in which abnormal cells begin to grow and multiply uncontrollably. They form a tumor mass that can be either benign (non-cancerous) or malignant.

In the present work information for patients with breast cancer was analyzed. The data contains measurements (parameters), calculated from a digitized fine-needle aspirate image of a breast mass: type of formation, radius, perimeter, area, texture, smoothness, symmetry, concavity, etc.

For data analysis an InterCriteria Analysis method is used. The method uses indexed matrices and intuitionistic fuzzy estimations. By the method the correlations between each pairs of parameters, explaining the formation in the breast mass were obtained. The obtained correlations are in a form of intuitionistic fuzzy pairs with values in the [0, 1] interval.

The aim of the study is to propose a method for reducing the input data about breast cancer at the inputs of a Deep learning neural network. This can be easily done, but the goal here is to achieve a reduction in the number of neural network inputs without affecting the classification accuracy of the data. For this purpose, the obtained intuitionistic fuzzy pairs were used, showing the degree of connection between the measured parameters.

Keywords: InterCriteria Analysis · Intuitionistic Fuzzy Logic · Deep learning neural network · Breast cancer

1 Introduction

Breast cancer is currently the most common type of cancer in women worldwide [1]. Breast cancer occurs mainly in middle-aged and older women [12]. Almost a third of all newly diagnosed types of cancers in women per year are breast cancer [12]. According the American Cancer Society's for 2023 in the United States there will be diagnosed about 297790 new cases of invasive breast cancer in women from witch 55720 new cases will be with non-invasive or pre-invasive breast cancer [12]. This shows the extreme importance of early diagnosis of breast cancer [9].

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 71–80, 2023.
https://doi.org/10.1007/978-3-031-39777-6_9

The breast cancer is a disease in which abnormal cells begin to grow and multiply uncontrollably [11]. They form a tumor (mass of cells) that can be either benign or malignant. A benign tumor is not malignant [14]. Malignant tumors are cancerous. They are a mass of cells that grow uncontrollably and spread locally or form metastases [10, 14].

In the current investigation information for patients with breast cancer was analyzed. The data was extracted from public domain with free license [6]. The data contains measurements (parameters), calculated from a digitized fine-needle aspirate image of a breast mass: type of formation, radius, perimeter, area, texture, smoothness, symmetry, concavity, etc. For data analysis an InterCriteria Analysis (ICA) approach [4] and Deep learning neural network [16] were used.

The ICA method uses indexed matrices [2] and intuitionistic fuzzy estimations [3]. By means of the developed software for ICA, freely available at http://intercriteria.net/software, data sets obtained by the evaluation of many objects according to many criteria can be analyzed. By the method the correlations between each pairs of parameters, explaining the formation in the breast mass were obtained. The obtained correlations are in a form of intuitionistic fuzzy pairs (IFPs, [5]) with values in the [0, 1] interval. Multi-criteria decision-making with medical and health care data has been implemented in the [7, 8, 13, 15, 17–20, 24].

Artificial neural networks (ANNs) are computer systems that are inspired by biological neural networks in the human brain. They are designed to model and process information in a similar way to the human brain, using a large number of interconnected artificial neurons.

They can be used to solve various computer science problems, such as:

- Data classification - can classify information such as object recognition, voice recognition, face recognition, and more.
- Forecasting and optimization - can make predictions about future events and optimize processes, such as in finance, logistics and marketing.
- Content generation - can generate text, images and other types of content such as automatically generating labels, product descriptions and more.
- Emotion recognition - can recognize emotions in text, voice recordings and images, such as in social networks and customer service systems.

The aim of the study is to propose a method for reducing the input data about breast cancer at the inputs of a deep learning neural network [16]. This can be easily done, but the goal here is to achieve a reduction in the number of neural network inputs without affecting the classification accuracy of the data. For this purpose, the obtained intuitionistic fuzzy pairs from ICA method were used.

The paper is organized as follows. Section 2 presents the application of the InterCriteria Analysis method. Section 3 presents the method for reducing the input data at the inputs of an ANN. Section 4 presents the conclusions. Section 5 and Sect. 6 are acknowledgments and references.

2 Application of the ICA Method

The dataset from [6] was organized in index matrix with 10 rows (for the mean square error of the parameters for evaluation the breast cancer) and 569 columns (for each patient). The evaluating parameters (criteria) are as follows: C1: radius of the tumor (mean of distances from center to points on the perimeter); C2: texture of the tumor (standard deviation of gray-scale values); C3: perimeter of the tumor (mean size of the core tumor); C4: area of the tumor; C5: smoothness of the tumor (mean of local variation in radius lengths); C6: compactness of the tumor (perimeter^2/area of the tumor - 1.0); C7: concavity of the tumor (mean of severity of concave portions of the contour); C8: concave points of the tumor (mean for number of concave portions of the contour); C9: symmetry points of the tumor; C10: mean of fractal dimension points of the tumor.

After applying the ICA approach, an index matrix with membership part and non-membership part of the 45 Intuitionistic fuzzy pairs were obtained. In Table 1 the membership part of the Intuitionistic fuzzy pairs is shown.

Table 1. Membership part of the IFPs of the correlations between parameters for evaluation the breast cancer.

ρ	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	1.000	0.614	0.981	0.992	0.549	0.669	0.732	0.783	0.540	0.376
C2	0.614	1.000	0.617	0.615	0.508	0.592	0.617	0.604	0.537	0.480
C3	0.981	0.617	1.000	0.978	0.561	0.687	0.748	0.798	0.550	0.392
C4	0.992	0.615	0.978	1.000	0.546	0.667	0.730	0.780	0.538	0.374
C5	0.549	0.508	0.561	0.546	1.000	0.745	0.681	0.698	0.690	0.708
C6	0.669	0.592	0.687	0.667	0.745	1.000	0.859	0.826	0.695	0.673
C7	0.732	0.617	0.748	0.730	0.681	0.859	1.000	0.888	0.655	0.586
C8	0.783	0.604	0.798	0.780	0.698	0.826	0.888	1.000	0.645	0.547
C9	0.540	0.537	0.550	0.538	0.690	0.695	0.655	0.645	1.000	0.648
C10	0.376	0.480	0.392	0.374	0.708	0.673	0.586	0.547	0.648	1.000

Three pairs are in strong positive consonance, that show a very strong dependence: “radius of the tumor”-“area of the tumor” with evaluation $\langle 0.992; 0.006\rangle$, “radius of the tumor”-“perimeter of the tumor” with evaluation $\langle 0.980; 0.018\rangle$ and “perimeter of the tumor”-“area of the tumor” with evaluation $\langle 0.978; 0.021\rangle$. Two pairs are in positive consonance, four are in weak positive consonance, that means, that they have a similar tendency. The other pairs are in weak dissonance, dissonance or strong dissonance that mean, that they are independent from each other.

3 Application of the Deep Learning Neural Network

The neural network structure is 9-7-10-10-1, it has more than three layers, which is a deep neural network (DNN). For training the DNN for classification, the breast carcinoma data taken from [6] is used. The inputs of the DNN are the ten criteria from Sect. 2 (C1, C2,…, C9, C10).

The structure of the neural network is shown in Fig. 1

Fig. 1. Neural network structure

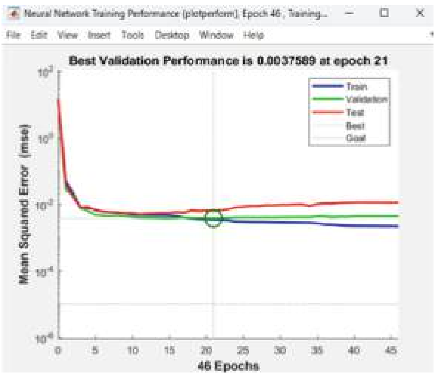


Fig. 2. Neural network training process

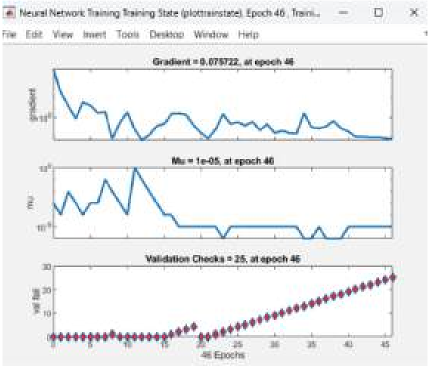


Fig. 3. Gradient and validation checks during training

The neural network in training with all data is trained for 46 epochs. The data is divided into three parts. A neural network training part that covers 70% of all data. The second part is 20% and is used to test the training. The third part is about validating the obtained data, and here there is one peculiarity.

The vectors that are used for verification are perceived as "external", i.e. the neural network was not trained with them. In our case, these three parts are randomly taken from the total number of vectors. In Fig. 2 the learning process of the neural network is shown.

As seen in Fig. 2 the best training result is at 21 epochs, where the root mean square error is 0.00375897.

During its training, several of its parameters are changed, the main ones being the gradient and validation checks - given in Fig. 3.

As proof of the good training, the R coefficients are given, which for the total training are 0.93546, for the whole process are 0.91817, for the testing and for the verification are 0.85607 and 0.87548 (Fig. 4).

The size of the neural network is not optimized. This is clearly seen in the ICA shown. There are consonance coefficients that are very high. This means a very large correlation between the given parameters and, accordingly, an excess of information. This can be removed by reducing some of the inputs where large positive resonance exists. In this case, one of the connections (the parameter and the input of the neural network, respectively) can be removed (see Table 2).

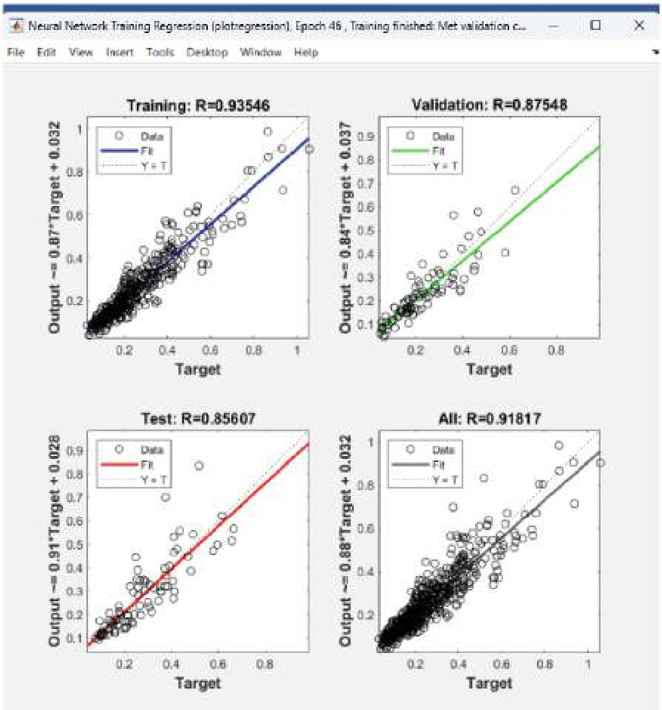


Fig. 4. R coefficients of ANN

The process by which we conduct this study is as follows: we start removing one by one the neural network inputs (with the corresponding parameter) that have high positive consonance and very low negative consonance. If there are two pairs with a same high or low consonance, each value of the pair is compared and if the next closest consonance is high, it is removed and the parameter with the next lower consonance remains [21–23].

When reducing the number of parameters, several things become apparent:

Not all parameters are necessary in a study, even though we can measure them;
The saturation of information about a process supports the process of classification and informed decision-making by doctors;
Part of the information can be removed if in the process of removal one (or more than one) of the parameters is reduced for parameters with a large positive consonance;
This does not seriously affect the accuracy of the neural network in making a decision, due to removing some of the error.

Table 2. Structure, IFPairs, Mean square error and Number of weight coefficients of the ANN.

N:	Structure of ANN	IFPairs	Mean square error	Number of weight coefficients
1	10-7-10-10-1		0.0033837	143
2	9-7-10-10-1 without "area of the tumor"	"radius of the tumor" - "area of the tumor": $\langle 0.992067;\ 0.00697418\rangle$	0.0033837	136
3	9-7-10-10-1 without "radius of the tumor"	"radius of the tumor" - "area of the tumor": $\langle 0.992067;\ 0.00697418\rangle$	0.0039585	136
4	8-7-10-10-1 without "area of the tumor" and "perimeter of the tumor"	"radius of the tumor" - "perimeter of the tumor": $\langle 0.980835;\ 0.0180512\rangle$	0.0040561	129
5	8-7-10-10-1 without "area of the tumor" and "perimeter of the tumor"	"perimeter of the tumor" - "area of the tumor": $\langle 0.978112;\ 0.0213867\rangle$	0.0410753	129
6	7-7-10-10-1 without "area of the tumor", "concavity of the tumor" and "perimeter of the tumor"	"concavity of the tumor" - "concave points of the tumor": $\langle 0.887528;\ 0.112243\rangle$	0.00430847	122
7	7-7-10-10-1 without "area of the tumor", "compactness of the tumor", and "perimeter of the tumor"	"compactness of the tumor" - "concavity of the tumor": $\langle 0.859037;\ 0.14014\rangle$	0.0049873	122

(*continued*)

Table 2. (*continued*)

N:	Structure of ANN	IFPairs	Mean square error	Number of weight coefficients
8	6-7-10-10-1 without "area of the tumor", "compactness of the tumor", "concave points of the tumor", and "perimeter of the tumor"	"compactness of the tumor" - "concave points of the tumor": ⟨0.825986; 0.173222⟩	0.00510438	115
9	5-7-10-10-1 without "area of the tumor", "compactness of the tumor", "concave points of the tumor", "radius of the tumor" and "perimeter of the tumor"	"perimeter of the tumor" - "concave points of the tumor": ⟨0.79778; 0.201329⟩	0.00435831	108

4 Conclusion

The research presents an approach to reduce the number of inputs of a neural network using the obtained strongly correcting IFPs by applying the method of intercriteria analysis. This leads to a reduction in the number of weight coefficients in the neural network and its faster learning. On the other hand, the size of the used memory is reduced without significantly affecting the accuracy of the classification of the analyzed data. The approach has been applied to data related to breast cancer, but can be used analogously for other types of data.

Acknowledgment. This research was funded in part by the European Regional Development Fund through the Operational 268 Programme "Science and Education for Smart Growth" under contract UNITe. BG05M2OP001–1.001–0004 269 (2018–2023)."

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Akram, M., Iqbal, M., Daniyal, M., Khan, A.U.: Awareness and current knowledge of breast cancer. Biol. Res. **50**, 1–23 (2017)
2. Atanassov, K. (ed.): Index Matrices Towards an Augmented Matrix Calculus. Studies in Computational Intelligence Series, vol. 573, Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10945-9

3. Atanassov, K.: On Intuitionistic Fuzzy Sets Theory. Springer, Berlin (2012). https://doi.org/10.1007/978-3-642-29127-2
4. Atanassov, K., Mavrov, D., Atanassova, V.: InterCriteria decision making: a new approach for multicriteria decision making, based on index matrices and intuitionistic fuzzy sets. Issues Intuitionistic Fuzzy Sets Gener. Nets **11**, 1–8 (2014)
5. Atanassov, K., Szmidt, E., Kacprzyk, J.: On intuitionistic fuzzy pairs. Notes Intuitionistic Fuzzy Sets **19**(3), 1–13 (2013)
6. Breast Cancer Dataset. https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset
7. Dobrev, P., Strashilov, S., Yordanov, A.: Metastasis of malignant melanoma in ovarium simulating primary ovarian cancer: a case report. Gazetta Medica Italiana – Archivio per le Scienze Med. **180**, 867–869 (2021)
8. Dobrev, P., Yordanov, A., Strashilov, S: Synchronous primary cervical carcinoma and ovarian fibroma: challenge in surgery. Gazzetta Medica Italiana-Archivio per le Scienze Mediche **179**(5), 375–377 (2020)
9. Elmore, J.G., Armstrong, K., Lehman, C.D., Fletcher, S.W.: Screening for breast cancer. Jama **293**(10), 1245–1256 (2005)
10. Hermanek, P., Sobin, L.H. (eds.): TNM Classification of Malignant Tumours. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-82982-6
11. Key, T.J., Verkasalo, P.K., Banks, E.: Epidemiology of breast cancer. Lancet Oncol. **2**(3), 133–140 (2001)
12. Key Statistics for Breast Cancer. https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html
13. Mirazchiyski, B., Bakalivanov, L., Petrova-Geretto, E.: Efficient hospital management based on risk assesment principles. Knowl. Int. J. **46**(4), 693–702 (2021)
14. Patel, A.: Benign vs malignant tumors. JAMA Oncol. **6**(9), 1488 (2020)
15. Petrova-Geretto, E., Tsankova, M., Petrova, Z., Mirazchiyski, B.: Opportunities for application of risk based approach and control in obstetrics and gynecology practice. Knowl. Int. J. **42**(4), 639–647 (2020)
16. Samek, W., Montavon, G., Lapuschkin, S., Anders, C.J., Müller, K.R.: Explaining deep neural networks and beyond: a review of methods and applications. Proc. IEEE **109**(3), 247–278 (2021)
17. Sotirova, E., Petrova, Y., Bozov, H.: InterCriteria Analysis of oncological data of the patients for the city of Burgas. Notes Intuitionistic Fuzzy Sets **25**(2), 96–103 (2019)
18. Sotirova, E., Bozova, G., Bozov, H., Sotirov, S., Vasilev, V.: Application of the InterCriteria analysis method to a data of malignant melanoma disease for the burgas region for 2014–2018. In: Atanassov, K.T., et al. (eds.) IWIFSGN 2019 2019. AISC, vol. 1308, pp. 166–174. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77716-6_15
19. Sotirova, E., Bozova, G., Bozov, H., Sotirov, S., Vasilev, V.: Application of the InterCriteria analysis method to a data of malignant neoplasms of the digestive organs for the Burgas region for 2014-2018. In: 2019 Big Data, Knowledge and Control Systems Engineering (BdKCSE), Sofia, Bulgaria, pp. 1–6 (2019)
20. Sotirov, S., Bozova, G., Vasilev, V., Krawczak, M.: Clustering of InterCriteria analysis data using a malignant neoplasms of the digestive organs data. In: Atanassov, K.T., et al. (eds.) IWIFSGN 2019 2019. AISC, vol. 1308, pp. 193–201. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77716-6_18
21. Sotirov, S., Petrova, Y., Bozov, H., Sotirova, E.: A hybrid algorithm for multilayer perceptron design with intuitionistic fuzzy logic using malignant melanoma disease data. In: Cengiz Kahraman, A., Tolga, C., Onar, S.C., Cebi, S., Oztaysi, B., Sari, I.U. (eds.) Intelligent and Fuzzy Systems: Digital Acceleration and The New Normal - Proceedings of the INFUS 2022 Conference, Volume 1, pp. 665–672. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-09173-5_77

22. Sotirov, S., Atanassova, V., Sotirova, E., Bureva, V., Mavrov, D.: Application of the intuitionistic fuzzy InterCriteria analysis method to a neural network preprocessing procedure. In: 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), Gijon, Spain, 30 June–03 July 2015, pp. 1559–1564 (2015)
23. Sotirov, S., et al.: Application of the intuitionistic fuzzy InterCriteria analysis method with triples to a neural network preprocessing procedure. Comput. Intell. Neurosci. **2017**, Article ID 2157852 (2017)
24. Stoyanov, V., Petkov, D., Bozdukova, P.: Pott's puffy tumor: a case report. Trakia J. Sci. **18**(1), 93–96 (2020). ISSN: 1313-3551

Assessment of Cancer Detection from CT Scan Images Using Hybrid Supervised Learning Methods

Panuganti Sai Sri Vaishnav and Bhupinder Singh(✉)

Lovely Professional University, Phagwara, Punjab 144411, India
bhupinder.28636@lpu.co.in, bhupinder0741@gmail.com

Abstract. Cancer is one of the deadliest and most ubiquitous diseases, and it claims a lot of lives every year. Radio diagnosis has a significant impact on medical image processing. Among all the main cancers, lung and liver cancers are the most prevalent and has the highest mortality rate these days. The current condition makes it difficult and time-consuming to manually identify the cancer tissue. Assessing the tumor load, planning therapies, making predictions, and tracking the clinical response can all be done using the segmentation of lesions in CT scans. In order to address the current problem of liver cancer, a Hybridized Fully Convolutional Neural Network, which is potent method for analyzing liver cancer in terms of semantic segmentation, has been proposed in this research for liver tumor segmentation. Dilation and erosion are two processes that are combined in this procedure. The goal of the proposed technique is to draw attention to the tumor region shown on the CT scan. And as part of the algorithm for the identification of lung cancer, the segmentation of the lung region of interest using mathematical morphological procedures is recommended. For picture pre-processing, this algorithm also employs techniques like median filtering. Support vector machines are used to classify CT scan pictures and compute geometric attributes from the recovered region of interest. Here, Computed Tomography (CT) pictures are used to identify malignant liver tumors. The image is improved using anisotropic diffusion filters and segmented using straightforward morphological techniques.

Keywords: SVM Classifier · Median Preprocessing · Median Filtering · Image Segmentation · Feature Extraction · Watershed Algorithm

1 Introduction

Worldwide, cancer is a major cause for concern, and patient outcomes are significantly improved by early identification. The accuracy of conventional diagnostic techniques can be constrained, but computed tomography (CT) scans are a frequently used diagnostic tool for finding cancer. In order to increase the precision of cancer identification from CT scan pictures, recent developments in machine learning have given rise to hybrid supervised learning methods, which mix different machine learning algorithms. In order to examine the present level of research on the evaluation of cancer detection from CT

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 81–89, 2023.
https://doi.org/10.1007/978-3-031-39777-6_10

scan pictures using hybrid supervised learning approaches, this review paper's goal is to provide an overview. Study will pay special attention to how these techniques can be used to find liver and lung cancer, two of the most prevalent cancers worldwide. The idea of supervised learning and how it can be used to identify cancer from CT scan images will be introduced initially. The limits of conventional diagnostic approaches and the possible advantages of applying machine learning techniques will next be covered. The various kinds of hybrid supervised learning techniques, such as support vector machines, artificial neural networks, convolutional neural networks, deep belief networks, and recurrent neural networks, will be thoroughly reviewed. Study will also go through recent research that have used similar techniques to the analysis of CT scan pictures for the identification of lung and liver cancer. With the help of hybrid supervised learning techniques, Research seek to present a thorough overview of the state of research on cancer diagnosis from CT scan pictures in this review. For researchers and clinicians engaged in the study of cancer diagnosis and therapy, study anticipate that this document will be a useful resource.

Objective and Originality

This study's primary goal is to identify the tumor site and evaluate it using a CT scan image and Hybrid Supervised Learning Algorithms. The uniqueness of this study comes in its emphasis on the evaluation of cancer utilizing a CT scan image to evaluate cancer in the lungs and liver using hybrid supervised learning methods. The goal of this study is to identify the tumor components by extracting new features. Additionally, in order to obtain accurate and precise results, the study seeks to identify various methodologies for tumor part detection.

Organization

The structure of the essay is as follows. First, Study presents a review of the existing literature on lung and liver tumours, stressing the drawbacks of earlier studies and the knowledge gaps that this work seeks to fill. Then, Study goes over the specifics of how Study chose our sample, collected our data, and conducted our statistical analysis. In the discussion section, Study interprets these results in the context of the body of knowledge and examine how they may affect public health and policy. Finally, Study summarise our key findings and provide recommendations for further study.

2 Literature Survey

Two of the most prevalent cancers in the world are liver and lung cancer, and effective treatment depends on early detection. Growing interest in the use of machine learning methods, particularly hybrid supervised learning approaches, to increase the precision of cancer detection has been observed in recent years. This literature review aims to present an overview of the state of knowledge on the effectiveness of hybrid supervised learning approaches for the detection of lung and liver cancer in their early stages.

Lung Cancer Detection

In a 2018 study, lung nodules in CT images were classified as benign or malignant using a hybrid supervised learning method that used an artificial neural network (ANN) and support vector machine (SVM). The results showed that the hybrid technique outperformed both the SVM and ANN alone, with an accuracy of 93.3% [15–18]. Convolutional neural networks (CNN) and support vector machines (SVM) were utilized in a hybrid technique in a 2019 study to categories lung nodules as benign or cancerous. The combination approach outperformed CNN and SVM alone with an accuracy of 91.2%. In a 2020 study, lung cancer was identified from CT images using a hybrid deep learning approach integrating CNN and recurrent neural network (RNN). The hybrid approach outperformed CNN and RNN individually with an accuracy of 92.87%.

Liver Cancer Detection

Deep belief networks (DBN) and support vector machines (SVM) were utilized in a study that was published in 2018 to categories liver tumors as benign or malignant. The accuracy of the hybrid technique, which outperformed both DBN and SVM alone, was 93.25% [19]. In a different 2019 study, liver cancers were identified from CT images using a hybrid technique that included CNN and extreme learning machine (ELM). The hybrid approach outperformed CNN and ELM separately with an accuracy of 97.42%. In a 2020 study, liver cancer was identified from CT images using a hybrid method of deep learning integrating CNN and RNN. The hybrid approach outperformed both CNN and RNN alone, with an accuracy of 96.7%. [2–4] Overall, these research' findings show that hybrid supervised learning techniques can increase the precision of lung and liver cancer diagnosis. For the purpose of early detection and treatment, the combination of various machine learning approaches enables a more reliable and precise classification of malignant cells or tissues. The efficacy of these hybrid approaches on bigger datasets and in actual clinical situations needs to be confirmed through additional study.

2.1 Problem Formulation

A binary classification challenge can be used to formulate the issue of identifying lung or liver cancer from CT scan images utilizing hybrid supervised learning techniques. The objective is to create a model that can correctly identify if a particular CT scan image contains cancer or not. A CT scan image, which is a three-dimensional volume of intensity values indicating the tissue densities of the lung or liver, serves as the model's input. The model's output is a binary classification label that designates whether the image contains cancer or not.

The model can be trained using a dataset of CT scan pictures, where each image is classified as either malignant or non-cancerous, to produce a hybrid supervised learning model for this purpose. Convolutional neural networks (CNNs) and decision trees can be used in combination by the model to learn the features that are most suggestive of malignancy in CT scan pictures. To increase its accuracy in identifying cancer, the model can potentially incorporate additional data, such as patient demographics, clinical history, or imaging modalities. By combining labelled and unlabeled data, the hybrid model can be trained using methods like semi-supervised learning or transfer learning. Metrics including accuracy, sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve can be used to assess the model's performance. The model can be further validated on a separate test set to ensure its generalizability to new data.

Overall, the problem of detecting lung or liver cancer from CT scan images using hybrid supervised learning methods is a challenging and important problem that can have a significant impact on cancer diagnosis and treatment.

2.2 Research Gap

Using hybrid supervised learning techniques, there is a research gap in the detection of lung or liver cancer utilizing CT scan pictures. Despite the fact that there has been a lot of study on utilizing machine learning algorithms to identify cancer from CT scan pictures, there are still significant research gaps when it comes to creating hybrid supervised learning techniques for identifying lung or liver cancer. [5–10] Here are some possible areas for further research in this field:

Lack of Annotated Datasets: The lack of annotated datasets is one of the main obstacles to creating supervised learning models for cancer detection. Despite the fact that there are some publicly accessible databases, they might not be sufficiently broad or diverse to capture all the subtleties of the condition. Consequently, creating a larger and more varied dataset for the diagnosis of liver or lung cancer can be a potential study area.

Using a Variety of Imaging Techniques: CT scans are just one imaging technique that can be used to find cancer. The accuracy of cancer detection can be increased by combining CT scans with other imaging modalities like MRI or PET scans, which can provide extra information. An intriguing research area is the creation of hybrid models that can efficiently combine data from several imaging modalities.

Examining Several Feature Extraction Methods: The development of machine learning models for cancer diagnosis begins with feature extraction. While conventional feature extraction methods like scale-invariant feature transform (SIFT) and histogram of oriented gradients (HOG) have been employed, there may be additional methods that can extract more subtle information from CT scans. A promising study area might be the investigation of various feature extraction methods. Table 1 shows the Literature summary of Latest Research Contributions.

Creating Models that are Easy to Understand: Deep learning models have excelled in many tasks, but they can be difficult to interpret. An intriguing study area is the development of hybrid models, which blend deep learning models with models that are easier to grasp, like decision trees or rule-based models. Figure 1 describes Feature Extraction. Overall, there are various research gaps that can be filled in the domain of building hybrid supervised learning techniques for identifying lung or liver cancer from CT scan pictures.

Table 1. Literature summary of Latest Research Contributions

Citation	Research Methodology	Conclusion
Cai et al. [1]	Simple linear regression, intra-class correlation coefficient and bland altman analysis	Model achieved 95% confidence interval in bland-altman analyses
Dasa et al. [4]	Watershed transform and Gaussian mixture model techniques	Model has performed with a classification accuracy of 99.38%
Frid-Adar et al. [2]	DL - GAN	Model performed with accuracy of 78.6%
Dong et al. [7]	Hybridized Fully Connected Neural Network and Semantic Segmentation	The algorithm provided measurements of the liver volume that were 97.22% very accurate
Das et al. [3]	Adaptive clustering and spatial fuzzy clustering approach	The method detected the lesion with accuracy of 89.15%
Vaidehi Nayantara et al. [11]	Conventional and deep learning-based methods	Use image fusion and radiological methods for better accuracy
Khan et al. [12]	Deionizing, deblurring and segmentation methods, SVM	Calculate ROI (region of interest) to get tumor part

3 Research Methodology

The following are potential study procedures for employing hybrid supervised learning techniques to identify lung or liver cancer from CT scan images:

Data Gathering: Compiling a sizable database of CT scan images for people with liver or lung cancer. With an equal number of positive (malignant) and negative (non-cancerous) cases, the dataset should be well-balanced. Image normalisation and noise or artefact removal are done during preprocessing on CT scan pictures. Image scaling, contrast normalisation, and image cropping are a few examples of this. Extraction of features from CT scan pictures that are useful for identifying cancer. Combining supervised and unsupervised learning methods, such as semi-supervised learning or transfer learning, can be used to create the hybrid model. [11, 13].

Here Study have worked on two datasets for lung cancer detection and liver cancer detection.

About Lung Cancer Detection Dataset: Yusuf Dede's Lung Cancer Dataset on Kaggle is a collection of CT scan pictures of lung nodules from lung cancer patients. A total of 704 CT scans—320 normal and 384 malignant—are included in the dataset. The scans are kept in the common DICOM format for medical imaging. The dataset's goal is to make it possible for researchers and machine learning experts to create algorithms for identifying and treating lung cancer from CT scan pictures. Because they can show the presence, size, and location of lung nodules, CT scans are a crucial diagnostic tool for identifying lung cancer [21].

About Liver Cancer Detection Dataset: A collection of liver CT scan pictures with tumour regions tagged may be found in Andrew Meijer van Dixhoorn's Liver Tumour Segmentation Dataset on Kaggle. The dataset consists of 131 CT scans in total, with annotations for tumour, liver, and background regions for each scan. The scans are kept in the common DICOM format for medical imaging. The dataset's goal is to make it possible for researchers and machine learning experts to create algorithms for automatically segmenting liver tumours from CT scan pictures. As they can show the presence, size, and location of liver tumours, CT scans are a crucial diagnostic tool for detecting liver cancer [22].

Model Training: Using the labelled, preprocessed CT scan pictures to train the hybrid model. Several optimization methods, including stochastic gradient descent (SGD) and adaptive moment estimation, can be used to train the model (Adam).

Model Evaluation: Analysing the hybrid model's performance on a different test set after training. The model can be assessed using performance metrics including accuracy, sensitivity, specificity, and area under the ROC curve.

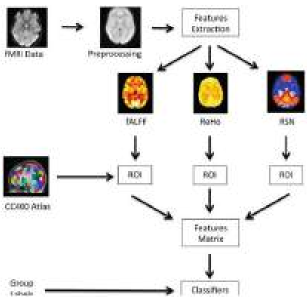


Fig. 1. Feature Extraction

Statistical Analysis: Analysing the results statistically to assess the importance of the conclusions. Using visualization tools like heatmaps, feature significance plots, and confusion matrices, one may interpret the results and present them in a form that is simple to follow and understand.

Comparing the performance of the created hybrid model with current state-of-the-art techniques for detecting lung or liver cancer [8]. Overall, the research methodology combines data collection, preprocessing, feature extraction, hybrid model development, model training, model evaluation, statistical analysis, interpretation, and visualization to identify lung or liver cancer from CT scan images using hybrid supervised learning methods [9].

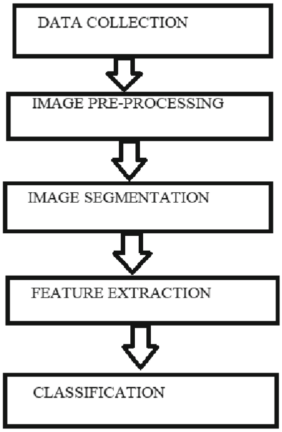


Fig. 2. Block Diagram represents Research Methodology

Using methods like CNNs or other deep learning models, which can learn features directly from the photos, this can be accomplished. Creating a hybrid supervised learning model by fusing various models and methods will increase the precision of cancer detection. Figure 2 represents Block Diagram represents Research Methodology [6].

4 Discussion

The suggested methodology offers a workable solution to the challenge of automatically segmenting medical images according to disorders. This method is tested on several CT scans of patients with liver cancer of the hepatocellular and metastatic carcinoma types [10]. Spatial fuzzy c-means clustering was used to automatically segment the lesion, defining the afflicted area. The study uses a set of extracted feature vectors to test the effectiveness of several classifiers, which were then categorized using MLP and C4.5 decision tree classifiers. Compared to the MLP classifier, the C4.5 classifier performs better with the LBP-HF texture descriptor [1, 12].

In order to lower the risk of liver failure after hepatic surgery, a precise preoperative volumetric calculation of the excisional and remaining liver is crucial (255). A specific quantity of remaining liver volume, suitable blood supply, and biliary drainage must be satisfied for the patient's safety. It can be straightforward to get this data using the planning platform developed in this work because surgeons generally have to make the taxing mental effort of mentally reconstructing 260 3D images [13]. A unique surgical planning may be established using the manual mapping and shifting of the liver resection plane margins offered by the MI-3DVS. The importance of virtual hepatectomy in liver surgery is increasing as procedures get more complex [5].

5 Conclusion

In conclusion, hybrid supervised learning methods have shown promising results in enhancing the accuracy of cancer diagnosis when used to analyses the identification of cancer from CT scan pictures. Convolutional neural networks (CNNs) and support vector machines (SVMs) are two examples of machine learning methods that can be combined to improve the performance of cancer detection models.

According to a review of numerous studies, hybrid models may perform cancer detection tasks with a high degree of accuracy, with some studies finding sensitivity and specificity values of up to 98%. These findings show how hybrid models may aid radiologists and other healthcare specialists in correctly diagnosing cancer. To validate these results and increase the generalizability of these models across various demographics and datasets, additional study is nonetheless required. The accuracy of cancer detection may also be increased by combining hybrid models with additional imaging modalities and clinical data [12].

Overall, the evaluation of cancer detection from CT scan images utilizing hybrid supervised learning techniques is an exciting field of research with the potential to greatly enhance cancer diagnosis and ultimately patient outcomes.

6 Future Scope

The potential for applying AI and ML to detect liver and lung cancer is enormous and encouraging. Here are some potential directions for future study:

Multimodal Imaging: Using a variety of imaging techniques, including MRI, PET, and CT scans, to increase the precision of cancer detection. Using explainable AI, creating models that can explain the characteristics and elements that go into the prediction of lung and liver cancer models that can explain the characteristics and elements that go into the prediction of lung and liver cancer using explainable AI. Precision medicine is the creation of AI and machine learning models that can forecast a patient's response to various cancer treatments based on their unique genetic and phenotypic characteristics. [14, 20].

Real-time Diagnosis: Using AI and ML algorithms, developing real-time models that can diagnose cancer that can diagnose cancer in real time. This can make it easier for people in isolated and underserved places to get a cancer diagnosis.

Collaboration Research: Sharing data, models, and algorithms with other researchers and organizations to enhance cancer diagnosis and treatment. Overall, the potential for employing AI and ML to detect lung and liver cancer in the future is exciting and shows significant promise for enhancing cancer detection and treatment results.

References

1. Cai, W., Fan, Y., Hu, H., Xiang, N., Fang, C., Jia, F.: Postoperative liver volume was accurately predicted by a medical image three dimensional visualization system in hepatectomy for liver cancer. Surg. Oncol. 26(2), 188–194 (2017). ISSN 0960-7404. https://doi.org/10.1016/j.suronc.2017.03.006
2. Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., Greenspan, H.: GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing **321**, 321–331 (2018). ISSN 0925-2312. https://doi.org/10.1016/j.neucom.2018.09.013

3. Das, A., Das, P., Panda, S.S., et al.: Detection of liver cancer using modified fuzzy clustering and decision tree classifier in CT images. Pattern Recognit. Image Anal. **29**, 201–211 (2019). https://doi.org/10.1134/S1054661819020056
4. Das, A., Acharya, U.R., Panda, S.S., Sabut, S.: Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques. Cogn. Syst. Res. **54**, 165–175 (2019). ISSN 1389-0417. https://doi.org/10.1016/j.cogsys.2018.12.009
5. Armato, S.G., Giger, M.L., MacMahon, H.: Automated detection of lung nodules in CT scans: preliminary results. Med. Phys. **28**, 1552–1561 (2001). https://www.ncbi.nlm.nih.gov/pubmed/11548926
6. Armatur, S.C., Piraino, D., Takefuji, Y.: Optimization neural networks for the segmentation of magnetic resonance images. IEEE Trans. Med. Image. **11**, 215–220 (1992). http://www.neuro.sfc.keio.ac.jp/publications/pdf/sundar.pdf
7. Dong, X., Zhou, Y., Wang, L., Peng, J., Lou, Y., Fan, Y.: Liver cancer detection using hybridized fully convolutional neural network based on deep learning framework. IEEE Access **8**, 129889–129898 (2020). https://doi.org/10.1109/ACCESS.2020.3006362
8. Cheran, S.C., Gargano, G.: Computer aided diagnosis for lung CT using artificial life models. In: El-Baz, A., Gimel'farb, G., Falk, R., Abo El-Ghar, M. (eds.) Proceeding of the 7th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, 25–29 September 2005, pp. 329–332. IEEE Computer Society, Romania (1991). ISBN: 0-7695-2453-2
9. Gurcan, M.N., Sahiner, B., Petrick, N., Chan, H., Kazerooni, E.A., et al.: Lung noduledetection on thoracic computed tomography images: preliminary evaluation of a computer aided diagnosis system. Med. Phys. **29**, 2552–2558 (2002)
10. Hara, T., Fujita, H., Lee, Y., Yoshimura, H., Kido, S.: Automated lesion detection methods for 2D and 3D Chest X-ray images. In: Proceedings of the 10th International Conference on Image Analysis and Processing, Venice, Italy, 27–29 September 1999, pp. 768–773 (1999). ISBN: 0-7695-0040-4
11. Vaidehi Nayantara, P., Kamath, S., Manjunath, K.N., Rajagopal, K.V.: Computer-aided diagnosis of liver lesions using CT images: a systematic review. Comput. Biol. Med. **127**, 104035 (2020). ISSN 0010-4825. https://doi.org/10.1016/j.compbiomed.2020.104035
12. Khan, R.A., Luo, Y., Wu, F.X.: Machine learning based liver disease diagnosis: a systematic review. Neurocomputing **468**, 492–509 (2022). ISSN 0925-2312. https://doi.org/10.1016/j.neucom.2021.08.138
13. ASCO. Liver cancer: statistics. cancer.net (2021). https://www.cancer.net/cancer-types/liver-cancer/statistics. Accessed 15 May 2021

Unlocking Insights into Mental Health and Productivity in the Tech Industry Through KDD and Data Visualization

Fariba Farid[1] and Yaser Donyatalab[2](✉)

[1] Department of Management, University of Nabi Akram, Tabriz, Iran
[2] Industrial Engineering Department, University of Moghadas Ardabili, Ardabil, Iran
yaser.d1991@gmail.com

Abstract. In the last few decades, mental illness has increasingly become one of the most crucial and prevalent public health problems in the competitive career world. The reason is that mental health disorders have been a key factor in the comprehensive health and culture of a company including employees' satisfaction, productivity, efficiency, and performance, to name but a few. It is an axiom that the workplace environment is a major contributor to workers' mental well-being. This article applies the Knowledge Discovery in Databases (KDD) approach to analyze a dataset on employees' mental health evaluation in the tech sector. The survey utilizes data from Open Sourcing Mental Illness (OSMI), a non-profit organization that conducted the largest survey of mental health within the tech sector. Over 1484 responses were collected via Google Forms during 2017, 2018, 2020, and 2021, contributing to the dataset scoring very highly in access and provision. The methodology, which starts with data collection, extraction, and processing, is the most suitable fit for this dataset. According to the survey results and visualization by Python programming, the largest group of respondents were men, and a majority of employees believed that mental health disorders have negative impacts on their productivity.

Keywords: Mental Health Disorders · Knowledge Discovery in Databases (KDD) · KDD Analysis of Mental Health Data in the Tech Sector

1 Introduction

In recent years, mental illness disorders have been a widely discussed issue that has become a significant global health concern. Many individuals are facing mental health disorders, which can lead to symptoms like sadness, anxiety, sleep disturbances, low energy, and exhaustion [1]. According to the World Health Organization (WHO), more than 350 million people are affected by depression. The financial impact of mental health issues was approximately US$2.5 trillion in 2010, expected to increase to US$6.0 trillion by 2030 [2]. Due to the progress in modern society, the number of people affected by mental health disorders, including anxiety, depression, and stress, has increased significantly.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 90–98, 2023.
https://doi.org/10.1007/978-3-031-39777-6_11

Furthermore, mental health problems have a significant impact on a company's overall health and culture, which includes employee satisfaction, productivity, efficiency, and performance [3] . In the past, the workplace has been identified as a major contributor to an individual's mental well-being [4] . Ensuring employees' well-being at work is considered a crucial factor in economic and commercial growth [5]. Research has shown that work-related stressors, such as high job demands, occupational uncertainty, and effort-reward imbalance, increase the risk of developing common mental health issues such as depression and anxiety. Workplace stress can also lead to reduced productivity, increased turnover rates, and decreased morale and burnout, as highlighted by several researchers [6]. A recent study by the World Health Organization (WHO) estimated that mental health disorders result in $1 trillion in lost productivity globally each year. Some studies as [7] have shown that work-related stress and its impact on employee mental health can vary depending on the nature of the job. Certain occupations, such as those in the IT industry that involve high demands, have been associated with mental health issues. When employees are unable to handle external demands and expectations that exceed their capacity, skills, and knowledge to perform their tasks, it can lead to increased mental stress [8] . Such stressful situations can have detrimental effects on an individual's health, including physical and mental health problems like depression, anxiety, emotional exhaustion, immune deficiency disorders, and cardiovascular diseases [9].

The field of mental health has recently benefited greatly from the advances in information technology, particularly in the areas of Artificial Intelligence (AI) and Machine Learning (ML). These tools have proven to be effective in anticipating mental health disorders and creating personalized healthcare plans [10]. By using these tools, organizations can gain a better understanding of their workers' mental health, implement proactive measures, and recognize the type of therapy a person may need at an early stage. Overall, AI and ML have the potential to revolutionize the field of mental health by providing more personalized and accessible care, early detection and prevention of mental health disorders, and evidence-based policies and programs. However, it is important to ensure that these technologies are developed and implemented ethically, with the protection of individual privacy and confidentiality as a top priority.

This paper uses data from Open Sourcing Mental Illness (OSMI), a non-profit organization that conducted a comprehensive survey of mental health in the tech industry. The aim of the survey is to increase awareness of mental illness and disorders in the workplace and eliminate the stigma associated with these conditions. OSMI also provides guidance to organizations on identifying appropriate resources to address these issues. The survey collected over 1484 responses via Google Forms during 2017, 2018, 2020, and 2021, with questions covering personal, workplace-related, and general information. To gain insights into mental illness, the article uses the Knowledge Discovery in Databases approach proposed by Fayyad et al. (1996) [11], which involves processing, experimenting, evaluating, and analyzing the data. The article notes that missing values exist in the survey data, but the variables considered in this analysis have mostly complete data. The KDD methodology aims to overcome the challenges posed by the type and process of the data being analyzed.

The paper begins by providing an introduction to mental health disorders related to work, along with a discussion of the various methods of utilizing artificial intelligence to address these disorders. The next section of the article delves into the methodology and applications that have been used to analyze datasets from Open Sourcing Mental Illness (OSMI). Finally, the conclusion of the paper has been presented by summarizing findings and suggesting areas for future research.

2 Literature Review

2.1 Workplace Mental Disorders (WMD)

The concept of workplace mental health goes beyond simply identifying the presence or absence of mental disorders. It also involves identifying factors that contribute to mental well-being and motivation [12]. The World Health Organization (WHO) defines mental health as a state of well-being where individuals can realize their full potential, handle normal life stresses, be productive in their work, and contribute positively to their community [13]. Various studies have explored the complex link between mental health and work productivity in developed nations [14]. Mental disorders have been shown to lead to reduced labor market participation, with individuals with mental illnesses being less likely to work, and those who do work earning less than their counterparts without mental illnesses [15].

Although some studies have attempted to review the literature on the relationship between interior office space and employee well-being, they have used broad terms such as 'office,' 'workplace design,' or 'architecture' [16]. The effect of the physical workplace on employees' mental health depends on their psychological and physiological reactions to its features [17]. Sander et al. [18] found that employees' responses to the physical work environment can be evaluated based on their cognitive, emotional, and relational reactions to the entire office environment or specific features. These features can be classified as resources (i.e., salutogenic) or demands (i.e., pathogenic) depending on their potential impact on employees' mental health [19]. Previous research has demonstrated that physical workplace characteristics, such as light, noise, and air quality, can affect employees' mental health, resulting in stress, fatigue, or mood changes [16]. Mental disorders are diagnosable conditions that can be characterized by changes in thought, mood, or behavior (or a combination of these) that can lead to anxiety or difficulty functioning [20]. Examples of mental disorders include anxiety disorders, mood disorders, psychotic disorders, and eating disorders, to mention but a few.

2.2 Mental Health Disorders and Artificial Intelligence Methods

Artificial intelligence is a philosophical approach that aims to enhance human intelligence by utilizing intelligent machines that can solve problems more effectively than humans alone. The use of AI is particularly valuable in medical diagnosis, where it can significantly reduce the time and effort required for accurate diagnosis. AI is widely used in healthcare to improve disease detection, monitor disease progression, optimize medication dosages, and identify new treatments. In the field of mental health, AI has the potential to revolutionize diagnosis and improve the accuracy of pre-diagnosis screening and risk modeling. Machine learning, a subset of AI, has emerged as a powerful tool in the mental health field, enabling the analysis of complex metadata and facilitating faster and more accurate predictions based on advanced statistical and probabilistic techniques. With ML, patterns in big datasets can be quickly accessed, leading to significant advances in fields such as bioinformatics. As per reference [21] , patient data gathered from everyday activities can be used to create predictive models that examine behavior, emotions, and other characteristics captured by sensors in mobile devices and wearables. This can provide valuable insights for mobile health interventions. Shatte et al. [22] conducted a literature review on mental health research and identified detection and diagnosis, treatment and support, and research and clinical administration as the most popular research domains.

In the fields of artificial intelligence and machine learning, data mining (also known as knowledge discovery in databases) refers to the process of extracting implicit, potentially valuable information from data that was previously unknown. This process involves using statistical techniques to identify trends and other relationships in large databases [23].

3 Methodology and Application

3.1 Methodology

The process of knowledge discovery in a dataset (KDD) involves extracting useful information from data. As stated by Fayyad et al. [11], KDD is a repetitive and interactive procedure that encompasses defining objectives and understanding the application, selecting relevant data, cleaning and preprocessing the data, eliminating irrelevant variables, applying suitable data mining algorithms, and ultimately uncovering understandable patterns or relationships within the dataset to aid in making informed decisions.

3.2 Application

This survey used 1484 data points sourced from Open Sourcing Mental Illness (OSMI) collected via Google Forms during 2017, 2018, 2020, and 2021. Analyzing the attributes of the questions answered in the data was crucial for the survey, and this was done by closely examining the content. The survey consisted of several types of questions, including general, workplace-related, and personal. Large data sets can be affected by the presence of special characters, missing values, and blank spaces, which can impact the performance of the model. Therefore, it is important to handle such characters carefully to preserve the accuracy of the information during the preparation process. Although the data source was generally consistent, there were some inconsistencies across different years, which were acknowledged by OSMI in an effort to ensure accurate visualizations. The majority of the variables in the OSMI dataset are represented as strings. To optimize performance and facilitate analysis, all datasets were preprocessed by converting string types to boolean data types and cleaning the data using the following steps: The first step is converting questions with two binary answers and no null responses, such as self-employment, to a scale of No = 0 and Yes = 1. The second step is transforming ordinal variables, such as employee numbers, to a scaled range between [−1, 1], with the lowest value of the range divided by 6, yielding a range of 7 with a "nan". Next step is coding questions with [Y, N, N/A], [Y, N, Maybe], and [Y, N, nan] as [1, 0, −1], and equating responses of [I don't know] and [Not applicable to me] to "nan" responses for questions such as employer-discussed mental health or revealing mental disorders to coworkers/employers. For the following step, physical health is coding as [−1] and mental health as [1], and age years classifying into three groups: Group1 (18–35 years), Group2 (35–50 years), and Group3 (50–70 years). Then, the bad data is grouped in gender, such as incomplete, inconsistent, and inaccurate information, into females and coded females as 1 and males as 0. Besides, living countries are classified into three groups: Group1 (Developed Countries), which is Group of Seven (G7), coded as [1]; Group2 (Developing Countries), which is Group of Twenty (G20), coded as [0]; and Group3 consisting of others, such as Bangladesh, Israel, and so on, coded as [−1], and racial options are coded such as White, American Indian or Alaska Native, Black or African American, I prefer not to answer, More than one of the above, and Asian, as a range between [−1, 1]. Furthermore, individuals who live and work in different countries are categorized as part of the living countries step. In addition, string data about current and past disorders, with anxiety disorders are coded as [1] and otherwise [−1]. Finally, some individuals did not accurately report their disorder experiences in the past and present. These incorrect answers were removed to improve data accuracy, resulting in a cleaned dataset consisting of 1458 users. Additionally, 61 people who claimed to no longer suffer from mental disorders were identified, indicating that they had been cured over time.

Data visualization is a method that presents data in a visually appealing and easy-to-understand manner using graphs, charts, and other visual aids. This approach can simplify complex data and make it more accessible. In this paper, Python libraries such as Matplotlib, Plotly, and Seaborn were utilized for data visualization purposes. Matplotlib is a widely used data visualization library in the Python programming language and is used for creating visual representations of data, such as graphs and charts. It is an essential part of the Python data science stack and is compatible with other popular libraries like NumPy and Pandas. Additionally, Seaborn is a visualization library that is built on top of matplotlib and provides users with easy-to-use tools for common data visualization tasks, such as mapping colors to specific variables or creating faceted plots. Plotly is another data visualization library that can be implemented to analyze and present data through the creation of various types of charts and graphs, including statistical charts, scientific charts, 3D charts, multiple axes, and dashboards. With Plotly, users can easily create and customize interactive visualizations to gain insights and communicate their findings effectively. Based on the analysis, the majority of the statistical population in this study are men, accounting for approximately 65% of the participants. Women make up around 31% of the survey participants, while approximately 2.5% of people preferred not to disclose their gender, as well as most employees reside in the United States of America (Fig. 1).

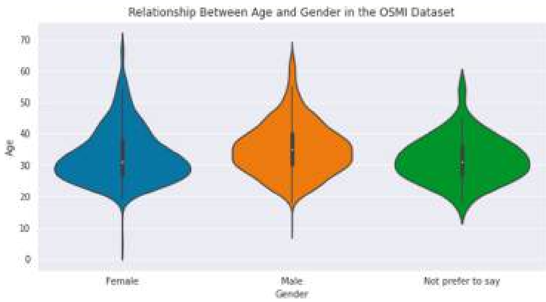


Fig. 1. Gender and Age relations

To investigate the relationship between gender and age in the OSMI dataset, a violin plot is created using the Seaborn library. The plot shows that the density of male participants was wider over the age of 30, indicating that more men are aged 30 or older. In contrast, the density of female participants is wider around the age of 30, suggesting that women who participated in the survey tended to be around 30 years old. This information can provide insights into the demographic composition of the OSMI dataset and may be useful for further analysis (Fig. 2).

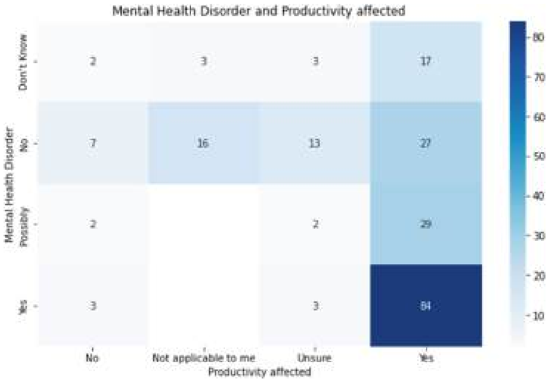


Fig. 2. Mental Health Disorders and Productivity Relationships

To conduct a preliminary analysis using preprocessed OSMI data, a heatmap was utilized as a graphical representation of a data matrix. The heatmap displays each value as a color, with the intensity and magnitude of the color indicating the distribution of the data. The heatmap revealed that individuals who have experienced mental health disorders and perceive negative effects on their work productivity had a much higher density and magnitude of data compared to others (Fig. 3).

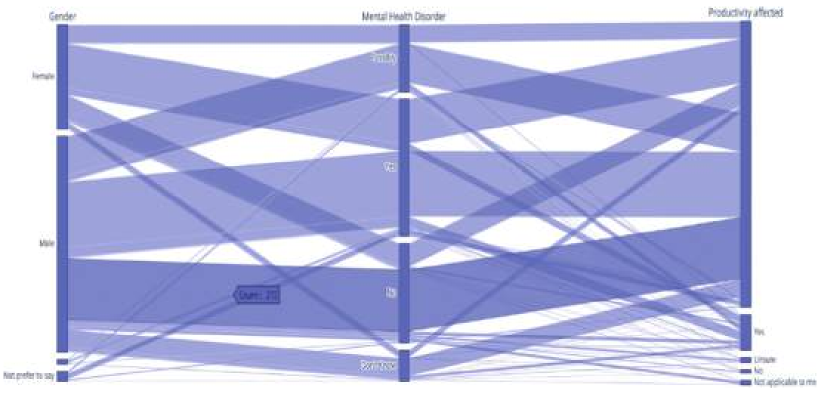


Fig. 3. "mental health disorders", "productivity affected", and "gender" relationships

Another type of visualization diagram that enables the exploration of relationships among several categorical variables in a single plot is known as a Basic Parallel Category Diagram. This diagram includes categorical variables such as "mental health disorders", "productivity affected", and "gender" displayed on the axis. The individual data points are connected with horizontal lines that represent each categorical variable. Additionally, the color of the lines can illustrate a third categorical variable, providing a visual representation of relationships among multiple variables. For instance, the diagram demonstrates that 272 men from the statistical population who experienced mental health disorders perceived a negative impact on their job productivity due to the disorder.

4 Conclusion

In recent years, data mining and data science techniques have been extensively utilized across various scientific disciplines, yielding significant outcomes. This research paper presents vital insights into mental health issues at workplaces, employing data science methods. These findings are crucial for human resource departments to select healthy candidates for hiring or to plan for educational interventions to enhance employee productivity and performance in today's highly competitive world. Future research should focus on exploring multiple practical and easily implementable machine learning algorithms to produce useful outcomes that can be utilized by companies' human resource departments to promote the health and well-being of their employees.

References

1. Rehm, J., Shield, K.D.: Global burden of disease and the impact of mental and addictive disorders. Curr. Psychiatry Rep. **21**(10), 1–7 (2019)
2. Bloom, D.E., et al.: The global economic burden of noncommunicable diseases (2012)
3. Appiah, S., Barnard, S., Deiven, J.: Density-based clustering of workplace effects on mental health (2017)
4. Leka, S., Jain, A., World Health Organization, et al.: Health impact of psychosocial hazards at work: an overview. World Health Organization **2** (2010)
5. Atilola, O.: Partaking in the global movement for occupational mental health: what challenges and ways forward for sub-Sahara Africa? Int. J. Ment. Heal. Syst. **1–10**, 1–10 (2012)
6. Brunner, B., Igic, I., Keller, A.C., Wieser, S.: Who gains the most from improving working conditions? Health-related absenteeism and presenteeism due to stress at work. Eur. J. Health Econ. **20**, 1165–1180 (2019)
7. Shockey, T.M., Zack, M., Sussell, A.: Health-related quality of life among US workers: variability across occupation groups. Am. Public Health Assoc. **107**(8), 1316–1323 (2017)
8. Brohan, E., Henderson, C., Little, K., Thornicroft, G.: Employees with mental health problems: survey of UK employers' knowledge, attitudes and workplace practices. Epidemiol. Psychiatr. Sci. **19**(4), 326–332 (2010)
9. Noblet, A., LaMontagne, A.D.: The role of workplace health promotion in addressing job stress. Health Promot. Int. **21**(4), 346–353 (2006)
10. Bzdok, D., Meyer-Lindenberg, A.: Machine learning for precision psychiatry: opportunities and challenges. Biolog. Psychiatry Cogn. Neurosci. Neuroimag. **3**(3), 223–230 (2018)

11. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. AI Mag. **17**(3), 37 (1996)
12. Forooraghi, M., Miedema, E., Ryd, N., Wallbaum, H.: Scoping Review of Health in Office Design Approaches. Emerald Publishing Limited (2020)
13. World Health Organization, et al.: Promoting mental health: concepts, emerging evidence, practice: summary report. World Health Organization (2004)
14. Chopra, P.: Mental health and the workplace: issues for developing countries. Int. J. Ment. Health Syst. **3**, 1–9 (2009)
15. Luciano, A., Meara, E.: Employment status of people with mental illness: national survey data from 2009 and 2010. Psychiatr. Serv. **65**(10), 1201–1209 (2014)
16. Bergefurt, L., Weijs-Perrée, M., Appel-Meulenbroek, R., Arentze, T.: The physical office workplace as a resource for mental health–a systematic scoping review. Build. Environ. **207**(2022)
17. Danielsson, C.B.: Office experiences, vol. 2, no. 5, pp. 605–628 (2008)
18. Sander, E.Li.J., Caza, A, Jordan, P.J.: Psychological perceptions matter: developing the reactions to the physical work environment scale. Build. Environ. **148**, 338–347 (2019)
19. Roskams, M., Haynes, B.: Salutogenic workplace design: A conceptual framework for supporting sense of coherence through environmental resources. J. Corp. Real Estate **22**(2), 139–153 (2020)
20. Murphey, D., Barry, M., Vaughn, B.: Mental Health Disorders. Adolescent Health Highlight. Publication# 2013-1.ERIC (2013)
21. Mohr, D.C., Zhang, M., Schueller, S.M.: Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Ann. Rev. Clin. Psychol. **13**, 23–47 (2017)
22. Shatte, A.B.R., Hutchinson, D.M., Teague, S.J.: Machine learning in mental health: a scoping review of methods and applications. Psychol. Med. **49**(9), 1426–1448 (2019)
23. Zhou, Z.-H.: Three perspectives of data mining. Psycholog. Med. (2003)

Prioritization of User Strategies for mHealthcare App Selection

Gülçin Büyüközkan[1](✉) and Fethullah Göçer[2]

[1] Industrial Engineering Department, Galatasaray University, Ortaköy, Istanbul, Turkey
gulcin.buyukozkan@gmail.com
[2] Industrial Engineering Department, Kahramanmaraş Sütçü İmam, Kahramanmaraş, Turkey

Abstract. The healthcare industry is gigantic, and as such, very competitive, and thus one of the more obvious categories of mobile applications (apps) is the healthcare apps and need little explanation. mHealthcare (mobile healthcare) is a general term for the use of mobile technologies and other sensor technologies in healthcare. User strategies to select a suitable app to handle the needs of healthcare for a patient is a complicated procedure that requires different aspects to think about. Hence, the problem of prioritization of user strategies for mHealthcare app selection is a multicriteria decision-making (MCDM) problem. MCDM methods are simple, effective and efficient methods to be used in solving many different types of problems related to the ranking and selection of alternatives. Analytic Hierarchy Process (AHP) is a widely used method among MCDM problems. This method is used by combining AHP with an intuitionistic fuzzy set to prioritize a suitable user strategy. Intuitionistic fuzzy sets can tackle the problem of describing vagueness and uncertainty of an event in the traditional model. Finally, a numerical example is conducted for mHealthcare app selection to illustrate the function of the intuitionistic fuzzy AHP method.

Keywords: Healthcare · Mobile Applications · MCDM · AHP · Intuitionistic Fuzzy

1 Introduction

Medical devices become health monitoring systems that can be called "smart devices," and these can communicate with remote physicians and connect to the internet. Industry 4.0 enables the development of patient-specific systems, provides essential advantages in the field of medical services. Providing patient-specific healthcare offers high quality because each service is unique. Although the costs of these services are currently a problem, they will be widely used in a few years with the novel Industry 4.0 solutions. The healthcare industry is gigantic, and as such, very competitive and thus one of the more obvious categories of mobile applications (apps) is the healthcare apps and need little explanation. mHealthcare (mobile healthcare) is a general term for the use of mobile technologies and other sensor technologies in healthcare. While there are some significant concerns about the unregulated status and correctness of mHealtcare apps for

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 99–107, 2023.
https://doi.org/10.1007/978-3-031-39777-6_12

patients, it is a rather promising and rapidly developing area. That's why it's crucial to prioritize various strategies to ensure users are getting apps that are best suited to their environment. Evaluating and prioritizing strategies can be complex, but if this vital step cannot be assessed carefully, the resulting outcome can negatively impact the health of consumers.

Decision-making activity is a daily life action that is mostly associated with choosing the best candidate among many. Multiple-criteria decision-making (MCDM) techniques can be used to solve the given problems. The exact depiction of decision makers' linguistic statements is an essential step in the MCDM literature. Zadeh [1] developed the fuzzy set concept, in which verbal judgments are asked rather than collecting crisp values. An extension of Fuzzy set is Intuitionistic Fuzzy Set (IFS), which is principally intended to address hesitation. Thus, it is extra advantageous than fuzzy sets under uncertain environments [2]. In this study, mHealthcare selection strategies and the criteria to evaluate these strategies are determined in line with the literature review and experts' opinions. The criteria weights are calculated using a fuzzy logic approach, and an MCDM approach is applied to prioritize the strategies. Furthermore, to reduce individual bias MCDM approaches are carried out in a group decision making (GDM) setting [3]. A four-step methodology is employed to solve the given problem; the list of strategies for mHealtcare app selection is determined, prioritization is done by evaluating the strategy using the defined criteria, the expert opinions and literature review define the criteria, the weights of each criterion is determined by the application of a fuzzy logic. The significant contribution and the originality of the proposed study is identifying evaluation strategies for mHealthcare app selection and developing criteria to evaluate these strategies. For as much as the authors know, no study in the literature uses the MCDM methods to find a solution to prioritization and evaluation of user strategies for mHealthcare app selection.

The paper also comprises of five subsequent parts. The following part of the introduction section is the review of the extant literature on the given subject and the method used. In the third part of the paper, after the presentation of the research purpose, information is presented about the methodology to be used in defining app selection strategies for mHealthcare. In the fourth part of the study, an application is made for an app strategy selection and the outcomes are interpreted. The last section of the paper embraces a brief assessment to conclude the study.

2 Literature Review

In the literature, there is an insignificant number of research that has studied the most suitable mHealthcare apps selection. Assessing and selecting mHealthcare apps is a type of MCDM research including multiple and conflicting criteria such as quality, cost, responsiveness, compatibility, and functionality, and so on. Many authors have used several types of integrated MCDM methodologies. Rajak and Shaw [4] developed an integrated methodology combining AHP and TOPSIS in the fuzzy environment in order to evaluate mHealth applications. Jiang and wang [5] presented a personal healthcare system to identify, diagnosis, monitor, and provide aids for patients with heart diseases. Chahal and Rudnick [6] proposed a decision-making algorithm that utilize decision criteria involving the strategic direction of the health institutes and the hospitals. Lee and Kim [7] presented a novel application selection methodology to consider the needs of target users altogether. This research demonstrated if selected health applications based on the needs of target users are able to influence health-related aspects. Due to space requirements, the extant literature is not revealed here.

3 Developed Methodology

The steps of the proposed methodology are presented in this subsection.

Step 1: Define hierarchy, criteria, and alternatives.

Let $A_i = \{A_1, A_2, \ldots, A_m\}, i = 1, 2, ..., m,$ and $C_j = \{C_1, C_2, \ldots, C_n\}$ be the sets represents the m alternative evaluated with respect to the n criteria. And let $w_j = \{w_1, w_2, \ldots, w_n\}$ indicate the weight of the vector for n criteria, $w_j \geq 0, j = 1, 2, ..., n$, ve $\sum_{\mathrm{j}=1}^{n} w_j = 1$.

Step 2: Calculate the weight of each decision-maker.

The weights (λ_k) of K decision-makers ($k = 1, 2, ..., K$) are calculated Eq. (1). The method found by Boran et al. [8] is adapted for the intuitionistic fuzzy evaluation of the decision-makers.

$$\lambda_k = \left[\mu_k + \pi_k\left[\frac{\mu_k}{1-\pi_k}\right]\right] / \sum_{k=1}^{K}\left[\mu_k + \pi_k\left[\frac{\mu_k}{1-\pi_k}\right]\right], \sum_{k=1}^{K}\lambda_k = 1 \quad (1)$$

Step 3: Get the judgments of decision-makers.

The linguistic term scale for Intuitionistic Fuzzy AHP is used here to prepare the pairwise comparison matrix. Linguistic terms in the intuitionistic fuzzy set form for pairwise comparison are provided in Abdullah and Najib [9]. After the hierarchy is created, verbal variables are transformed into an intuitionistic fuzzy set matrix.

Step 4: Creating a GDM Matrix.

The intuitionistic fuzzy weighted average (IFWA) operator [10] is used to combine distinct views of decision makers into a group view. The established R_{ij} matrix of criteria characterizes the weights for the corresponding intuitionistic fuzzy AHP values and the weight of the matrix alternatives obtained on all criteria with respect to each alternative.

$$R_{ij} = \left(1 - \prod_{k=1}^{K}\left(1-\mu_{ij}^{(k)}\right)^{\lambda_K}, \prod_{k=1}^{K}\left(\left(\upsilon_{ij}^{(k)}\right)^{\lambda_K}\right), \prod_{k=1}^{K}\left(1-\mu_{ij}^{(k)}\right)^{\lambda_K} - \prod_{k=1}^{K}\left(\left(\upsilon_{ij}^{(k)}\right)^{\lambda_K}\right)\right) \quad (2)$$

Step 5: Performing the Consistency Check (C.R.)

The checking for the consistency is applied to confirm the matrix in the form of pairwise comparison are neither random nor irrational. The consistency ratio (C.R.) of a pairwise comparison is established by the use of Eq. (3). For intuitionistic fuzzy AHP weights, C.R. is estimated using the random index (R.I.) [11] values.

$$\text{C.R.} = \frac{RI - \frac{\sum_{ij}^{\pi}(x)}{n}}{n-1} \quad (3)$$

n is the number of the matrix element and $\pi_{ij}(x)$ is the hesitation value. C.R is acceptable if less than or equal to 0.10. Otherwise, decisions are inconsistent and evaluation is collected once again from decision makers.

Step 6: Combining Multilevel Weights.

From the lowest level to the highest level in the hierarchy, it is multiplied by intuitionistic Fuzzy AHP criteria weights and intuitionistic Fuzzy AHP sub-criteria weights to find overall criteria weights. Using the general criteria weights ($\tilde{w}_j$) and the total weight of each alternative on the sub-criteria ($\tilde{w}_{ij}$), the total weight of each alternative on the objective ($\tilde{W}_i$) is calculated with Eq. (4).

$$\tilde{W}_i = \oplus_{j=1}^{n}\left(\tilde{w}_j \otimes \tilde{w}_{ij}\right) \tag{4}$$

Step 7: Ranking Criteria and Alternatives.

Overall weights are calculated with the Eq. (5), using the Entropy weights [9]. The smaller the value of $\overline{w}_i$, the greater the sequence.

$$\overline{w}_i = -\frac{1}{nln2}[\mu_j ln\mu_j + v_j lnv_j - \left(1-\pi_j\right)\ln\left(1-\pi_j\right) - \pi_j ln2] \tag{5}$$

4 Application

The mobile devices, which are among the easiest ways to reach information, are progressing with unbelievable speed, and different types of apps entering our lives, are increasing day by day. Therefore, mobile app developers are highly encouraged to work on new tools and techniques. The winner in this fast-growing market will be developers and entrepreneurs who choose and use the right tools and methods to respond to user needs in the quickest and easiest way. In this study, the problem of prioritization of user strategies for mHealthcare app selection is examined and the criteria that are significant for the most suitable app selection strategies are determined. In line with the detailed literature review [4, 12] and expert opinions, 5 different criteria are determined for the application and theoretically 4 different strategies are chosen. The criteria determined as Technical (C_1): the effect of application service quality (increase) (C_{11}), the effect of application responsiveness (increase) (C_{12}), the effect on technical performance (increase) (C_{13}); and Operational (C_2): the effect on the functionality of the application (increase) (C_{21}), and the effect on application costs (decrease) (C_{22}). However, it should not be forgotten that the application concept and requirements may differ. The tools and techniques to be used should be selected according to the application requirements, following the criteria listed above. The Intuitionistic Fuzzy AHP method is used because of the contained uncertainty due to the nature of the problem and the better handling of the uncertain linguistic expressions of the experts. In line with the detailed literature review and expert opinions, four possible user strategies have been identified based on the applicability in the short term for the selection of mHealthcare apps: Strategy 1: Augmenting the learnability and usability among users (thus learning and using the selected app is carried out with the minimum number of efforts), Strategy 2: Augmenting the availability of development and maintenance for the app (the aim is to achieve a more sustainable enhancement of the selected app), Strategy 3: Augmenting the use of able security systems (enabling physical and encrypted security management), Strategy 4: Augmenting the stability to continue to function over a long period of time (thus bringing resistance to the amplification of changes in selected app).

4.1 Numerical Results

Based on the given structure created for the problem, the solution is made using the AHP method. The matrices obtained from the solution are prepared according to the steps of the methodology. In the GDM process, the ideas of all decision-makers should be merged as a group notion without a loss of data in order to obtain the aggregated decision matrix. Hence, the IFWA aggregation operator developed by Xu [10] is used. The importance of decision-makers is thought as verbal statements. Linguistic terms to establish the priorities of experts are stated in intuitionistic fuzzy values. The priority weights of experts are presented in Table 1. In decision-making problems, the importance of each criterion are not equal to each other. The influences of the criteria is at different priorities for all decision-makers.

Table 1. Decision Makers' Individual Assessments on Criteria

	DM	C_1	C_2	C_3	C_4	C_5
	DM_1	EI	1/EMI	EMI	EMI	MI
C_1	DM_2	EI	1/ LI	EMI	LI	MRI
	DM_3	EI	1/EMI	EMI	EMI	MI
.	.	.	.	.	.	.
	DM_1	MI	AMI	MI	MRI	MI
C_5	DM_2	LI	1/EMI	LI	AMI	1/AMI
	DM_3	MI	I	1/I	1/I	1/I

Absolutely More Important (AMI); Extremely More Important (EMI); Very More Important (VMI); More Important (MRI); Equally Important (EI); Very Important (VI); Important (I); Less Important (LI); Medium Importance (MI).

Consequently, intuitionistic fuzzy values assigned to the criteria by decision-makers must be merged. After the criteria weights and the combined decision matrix are created, the aggregated weighted decision matrix is obtained. Pairwise comparison matrices of criteria are given in Table 2. Due to space requirements, the whole values is undisclosed. Besides, pairwise comparison matrices of strategies are calculated for each criterion. Consistency ratios are calculated to make pairwise comparison matrices acceptable. The matrixes are acceptable because the consistency rate of all matrices is less than 0.10. Intuitionistic Fuzzy Criteria Weights and Crisp Criteria Weights is given in Table 3. In the last stage of the AHP method, the priorities of each strategy are given in Table 4 by multiplying the priorities of the criteria and the strategic priorities found with respect to the criteria. In the AHP method, the priority weights range from 0 to 1. Among these values, the closest strategies to 1 is determined as the appropriate solution. In this study, it has been determined that the first priority for a suitable strategy is presented to be Strategy 4: Augmenting the stability to continue to function over a long period (thus bringing resistance to the amplification of changes in the selected app). Also, the Strategy 2: Augmenting the availability of development and maintenance for the app (the aim is to achieve a more sustainable enhancement of the selected app) has the second priority of given strategies to choose the best suitable user strategies for mHealthcare app selection.

Table 2. Individual Priorities of Each Decision Maker

DM	Preferences			$[\boldsymbol{\mu}(\boldsymbol{x}),$	$\boldsymbol{v}(\boldsymbol{x})]$	λ_k
DM_1	–	VI	I	0.654	0.335	0.398
DM_2	E	–	VI	0.605	0.349	0.255
DM_3	I	I	–	0.500	0.450	0.347

Very Unimportant (VU); Unimportant(U); Equal (E); Important (I); Very Important (VI).

Table 3. Intuitionistic Fuzzy Criteria Weights and Crisp Criteria Weights

	$\tilde{w}_{jj}$		$\tilde{w}_j \otimes \tilde{w}_{jj}$			
	$[\boldsymbol{\mu},$	$v]$	$[\boldsymbol{\mu},$	$v]$	w_j	Rank
C_1	0.157	0.175	0.018	0.349	0.204	3
C_2	0.238	0.128	0.028	0.313	0.221	2
C_3	0.210	0.097	0.024	0.288	0.225	1
C_4	0.134	0.242	0.014	0.450	0.173	5
C_5	0.167	0.246	0.017	0.453	0.176	4

Table 4. Prioritized Strategies

		$\boldsymbol{S_1}$	$\boldsymbol{S_2}$	$\boldsymbol{S_3}$	$\boldsymbol{S_4}$
$\tilde{W}_i$	$[\boldsymbol{\mu}.$	0.090	0.174	0.130	0.146
	$v]$	0.461	0.350	0.315	0.294
	$\overline{w}_i$	0.268	0.319	0.314	0.321
	Rank	4	2	3	1

4.2 Discussions and Comparisons

The decision making process constitute a significant part of the developing and advancing world of today. People, institutions or businesses are constantly making decisions under rapidly changing and increasingly difficult working conditions. That is the case, decision making problems have also become complicated. The fact that the number of alternatives and criteria is getting higher, and the criteria contradict each other makes it very difficult for decision makers to choose between them. Therefore, MCDM methods have been used to evaluate multiple criteria to reach a decision among multible alternatives. MCDM methods are not just executed as tools, techniques or algorithms, but also include human-specific judgments situations. If there are criteria having different measurement units, elimination of these differences are essential. For example, although the measurement unit of qualitative and quantitative criteria is different, this difference in the AHP method is eliminated with the respective scale of 1–9. When literature review about MCDM methods is done, it is seen that there are many different methods used for the solution of decision problems. A comparison analysis is also conducted to examine the impact of the changes on the results with respect to different MCDM methods. In order to observe how possible changes in relative methods will effect the final ranking should be investigated. For this purpose, TOPSIS, VIKOR, and COPRAS MCDM methods are examined under Intuitionistic Fuzzy Environment. In the analysis, the comparison is made to help providing information in order to see the variations of the alternative rankings. When the results are examined, it is observed that the mentioned changes in the methods slightly changed the ranking of some alternatives. However, the results for the first and last alternatives have remained the same except for the COPRAS technique. A summary of these candidate performance values (rankings) for different MCDM situations is shown in Fig. 1. The results provides a robust validation of the proposed method.

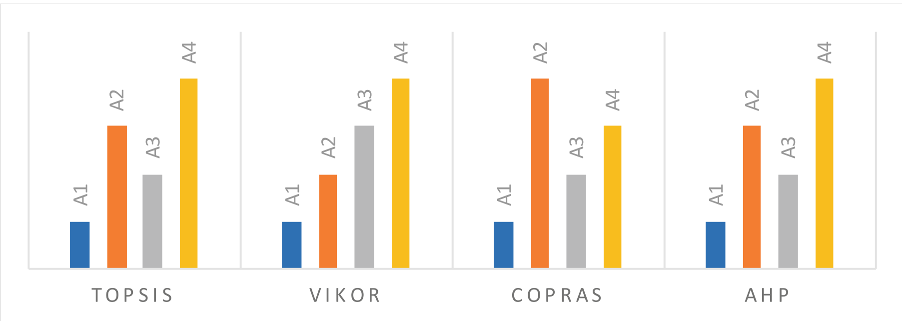


Fig. 1. Comparisons of Different MCDM Methods

5 Conclusion

Technology dominated developments are taking place in health services. Which is one of the largest and fastest-growing markets in the world. Digital technologies in healthcare services have application areas such as education. Research. Care. Treatment and early diagnosis. It is envisaged that health services will be reconstructed depending on the developing technology. The number of mobile applications developed based on the use of smartphones on healthcare services is increasing day by day. Accordingly. These applications are used extensively by users worldwide. MCDM approaches is generally applied in circumstances where there are not one but more contradicting criteria and aims to select the best one among multible alternatives. Accordingly. The results of the MCDM methods are frequently based on the preferences of decision-makers. To determine the importance of criteria in multi criteria situations, AHP has been proposed. According to the information obtained from the decision-makers. The selection of the most suitable alternative was made with the intuitionistic fuzzy AHP method considering the five evaluation criteria. In the intuitionistic fuzzy AHP method. The significance of decision-makers is identified in linguistic terms. These terms are stated in the form of IFS values. Later. The criteria weights and alternative valuations of regarding criteria are delivered by the decision-makers in verbal terms stated in IFS values. The IFWA aggregation operator is then used to aggregate IFS values.

Some potential directions for forthcoming studies to follow-up are exist from the developed research. In the study. The IFS theory has been applied to GDM problems. Meanwhile the recommended technique is an operative process to deal with vagueness instigated by human thoughts, it is foreseen to be applied to other selection problems, such as investment, portfolio, and supplier selection problems, etc. in the future. Additionally, the established technique validates how the given approach is to be applied to real-world decision-making problems. Thus, the presented assessment outline might be able to handle other selection problems and also other MCDM problems.

Acknowledgments. The authors would like to thank the experts whose insightful contribution have improved the paper greatly. This research has financial support from the Galatasaray University Research Fund (FOA-2021-1059 and FOA-2023-1181).

References

1. Zadeh, L.A.: Fuzzy sets. Inf. Control. **8**, 338–353 (1965). https://doi.org/10.1016/S0019-9958(65)90241-X
2. Büyüközkan, G., Göçer, F., Feyzioğlu, O.: Healthcare website evaluation using intuitionistic fuzzy choquet approach. J. Mult. Log. Soft Comput. **30**, 215–237 (2018)
3. Büyüközkan, G., Göçer, F., Karabulut, Y.: A new group decision making approach with IF AHP and IF VIKOR for selecting hazardous waste carriers. Measurement **134**, 66–82 (2019). https://doi.org/10.1016/j.measurement.2018.10.041
4. Rajak, M., Shaw, K.: Evaluation and selection of mobile health (mHealth) applications using AHP and fuzzy TOPSIS. Technol. Soc. **59**, 101186 (2019). https://doi.org/10.1016/j.techsoc.2019.101186

5. Jiang, S.-Y., Wang, L.-H.: Enhanced machine learning feature selection algorithm for cardiac arrhythmia in a personal healthcare application. In: 2018 IEEE Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics (PrimeAsia), pp. 39–42. IEEE (2018)
6. Chahal, A., Rudnick, A.: Selecting digital health technologies for validation and piloting by healthcare providers: a decision-making perspective from ontario. Int. J. Technol. Assess. Health Care. **35**, 1–4 (2019). https://doi.org/10.1017/S0266462318003720
7. Lee, J., Kim, J.: Can menstrual health apps selected based on users' needs change health-related factors? A double-blind randomized controlled trial. J. Am. Med. Informatics Assoc. **26**, 655–666 (2019). https://doi.org/10.1093/jamia/ocz019
8. Boran, F.E., Genç, S., Kurt, M., Akay, D.: A multi-criteria intuitionistic fuzzy group decision making for supplier selection with TOPSIS method. Expert Syst. Appl. **36**, 11363–11368 (2009). https://doi.org/10.1016/j.eswa.2009.03.039
9. Abdullah, L., Najib, L.: Sustainable energy planning decision using the intuitionistic fuzzy analytic hierarchy process: choosing energy technology in Malaysia. Int. J. Sustain. Energy. **35**, 360–377 (2016). https://doi.org/10.1080/14786451.2014.907292
10. Xu, Z.: Intuitionistic preference modeling and interactive decision making. Springer (2011). https://doi.org/10.1007/978-3-642-28403-8
11. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**, 234–281 (1977)
12. Silvera-Tawil, D., Hussain, M.S., Li, J.: Emerging technologies for precision health: an insight into sensing technologies for health and wellbeing. Smart Heal. **15**, 100100 (2020)

Fuzzy Numbers and Analysis of Radiological Images

Irma Ibrišimović[1,2], Bratislav Iričanin[3], Nataša Milosavljević[4], Ljubo Nedović[2], and Nebojša Ralević[2](✉)

[1] Faculty of Science, Depertment of Mathematics, University of Tuzla, Tuzla, Bosnia and Herzegovina

[2] Faculty of Technical Sciences, Department of Fundamental Sciences, University of Novi Sad, Novi Sad, Serbia
{nljubo,nralevic}@uns.ac.rs

[3] Faculty of Electrical Engineering, Department of Applied Mathematics, University of Belgrade, Belgrade, Serbia
iricanin@etf.rs

[4] Faculty of Agriculture, Department of Mathematics and Physics, University of Belgrade, Belgrade, Serbia
natasam@agrif.bg.ac.rs

http://mat.pmf.untz.ba/,
http://ftn.uns.ac.rs/n1386094394/faculty-of-technical-sciences,
https://www.etf.bg.ac.rs/en/departments/the-department-of-applied-mathematics, https://www.agrif.bg.ac.rs/

Abstract. Fuzzy sets are a mathematical notion that allows us to represent uncertainty by assigning a degree of membership to a value in a range. This degree of membership represents the degree to which a particular value belongs to the fuzzy set. Fuzzy numbers are fuzzy sets that satisfy specific conditions, and can be applied in the interpretation of the behavior of neural networks. In many real-world problems, the output of a neural network may not be a precise numerical value, but rather a range of values that are subject to uncertainty. In the context of neural networks, fuzzy numbers can be used to represent uncertainty in the output of the network. Fuzzy numbers are used in applications where the representation of ambiguity and uncertainty in numerical data is explicitly desired. This paper will consider discretizations and defuzzifications triangular fuzzy numbers with the use of MATLAB® software. MATLAB® functions were applied to the analysis of kidney images, where the fuzzy numbers represent the tendency value of the degree of redness, which serves as an indicator for classifying the image as normal or abnormal based on different characteristics.

Keywords: Fuzzy numbers · fuzzy number representation · defuzzification · discretizations

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 108–115, 2023.
https://doi.org/10.1007/978-3-031-39777-6_13

1 Introduction

This paper will describe the application of fuzzy number in the interpretation of the work of neural networks with the use of MATLAB® software. Medical imaging is an essential tool for diagnosing and treating various medical conditions. However, the interpretation of medical images can be a challenging task due to the complexity and the variability of data. Fuzzy logic provides a powerful framework for dealing with uncertainty and vagueness in data that is common in medical imaging. Fuzzy numbers are a key component of fuzzy logic and can be used to represent uncertain and vague data more precisely.

The first part of the paper refers to the definition of the fuzzy numbers and methods [11] that are essential for implementation in the MATLAB® software. The second part of the paper concerns the application of triangular fuzzy numbers to medical diseases, more precisely in this case to kidney disease. Medical image segmentation based on fuzzy numbers [25], fuzzy modeling for medical image processing [21], behavioral neural networks on fuzzy systems [10] and triangular fuzzy numbers for medical image segmentation [22] are key elements of this paper, because the simulation of their application was done in MATLAB® software. In the end, the analysis of the image was done and as an input it gives graphic fuzzy numbers that show the tendency of the value of the degree of redness, which represents an abnormality in the observed photographs (see Fig. 2).

2 The Definition of Fuzzy Number

A fuzzy number is a mathematical concept used in uncertainty theory. It is a number that has an indeterminate value that is between two limits. In other words, a fuzzy number is a number that has uncertainty in its value, which can be described using a membership function. A membership function is a mathematical tool used to describe uncertainty in values. It determines how much an element is connected to a group. For the case of fuzzy numbers, the membership function is used to describe how closely a value is related to a number. Fuzzy numbers often are used in various fields, such as finance, engineering, data science, etc. For example, in finance they are used for risk modelling and market price estimation. In engineering, they are used to model uncertainty in design and planning. In data science, they are used to analyze data that is not obtained completely precise. There are various operations that can be applied to fuzzy number, such as addition, subtraction, multiplication, and division, which are performed using membership functions (see [1–3,11]).

Definition 1. *If $X \neq \emptyset$ and $\mu_A : X \rightarrow [0,1]$, then the ordered pair (X, μ_A) is called the* ***fuzzy set*** *A. Mapping μ_A is* ***membership function*** *and is often identified with the fuzzy set.*

Definition 2. *Fuzzy set A is* ***normalized*** *if $\mu_A(x) = 1$, for someone $x \in X$.*

Definition 3. *Set* ${}^{\nu}A = [A]_{\nu} = \{x \in X \mid \mu_A(x) \geq \alpha\}$ *is called* α***-cut*** *fuzzy set* A*.*

Definition 4. *Fuzzy set* A *defined above* $\overline{\mathbb{R}}$ *is* ***convex*** *if every* α*-cut is convex.*

Definition 5. *Fuzzy set* A *defined above* $\overline{\mathbb{R}}$ *is* ***fuzzy number*** *if it is normalized and convex.*

Definition 6. *The set supp* $A = \{x \in X \mid \mu_A(x) > 0\}$ *is called* ***support*** *of the fuzzy set* A*.*

Definition 7. *The set* $\ker A = \{x \in X \mid \mu_A(x) = 1\}$ *is the* ***kernel*** *of the fuzzy set* A*.*

3 Methods

In this paper we focus on triangular fuzzy numbers which are commonly used in fuzzy logic. We provide MATLAB® implementations for the discretization and defuzzification of these fuzzy numbers. Discretization involves converting a fuzzy numbers into a set of discrete values, which can be useful for analysis and interpretation. Defuzzification involves calculating a crisp value that represents the central tendency of the fuzzy numbers, and can be useful for decision-making and classification (see [5,6,8]).

Definition 8. ***The triangular fuzzy number*** A *is the fuzzy number the membership function of which is:*

$$A(x) = \begin{cases} \frac{x-a}{b-a}, \, a < x \leq b, \\ \frac{c-x}{c-b}, \, b < x < c, \\ \quad 0, \, x \leq a \; or \, c \leq x \end{cases} .$$

The triangular fuzzy number A is often presented as an ordered triplet of numbers $A = (a, b, c)$ (see [11]). Triangular fuzzy number whose core is $\{b\}$ (a is not necessary indicate support) is usually marked with $\hat{b}$.

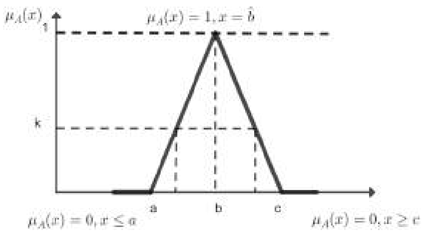


Fig. 1. The membership function of triangular fuzzy numbers

We apply these methods to medical images to demonstrate their usefulness in analyzing and interpreting medical data. For example, we can use triangular fuzzy numbers to represent the degree of redness (the observed area is marked by red circles) in a certain area of a kidney image. We can then discretize this fuzzy number and use the center of gravity method to calculate central tendency of the fuzzy number, which can help in the classification of the kidney image as normal or abnormal, based on the various features of the image.

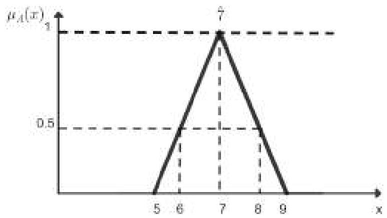
Suppose we have a set of kidney images and we want to determine if a particular image is normal or abnormal based on the degree of redness in a certain area of the image. We can use fuzzy numbers to represent the degree of redness in the area. Assume that the degree of redness can be represented by a triangular fuzzy number with the membership function shown below, where a, b, and c are the lower, middle, and upper bounds of the fuzzy number. Suppose that we have a kidney image with the following degree of redness in the area of interest. To apply the center of gravity method, we first need to discretize the fuzzy number. We can use the discrete set of values for the fuzzy number. Then we can calculate the degree of membership of each discrete value using the triangular membership function. The degree of membership of the value x in the support fuzzy number can be calculated as follows

$$trimf(x; a, b, c,) = max\left\{min\left\{\frac{x-a}{b-a}, \frac{c-x}{c-b}\right\}, 0\right\},$$

where $trimf$ represents the label for triangular membership function. For example, the degree of membership for the $x = 6$ and fuzzy number $\hat{7} = (5, 7, 9)$ is calculated

$$\mu_A(6) = max\left\{min\left\{\frac{6-5}{7-5}, \frac{9-6}{9-7}\right\}, 0\right\} = 0.5.$$

Next, we can calculate the center of gravity (Cog) of the fuzzy number using the following formula



$$Cog = sum(i * \mu(i))/sum(\mu(i)),$$

where i is the index of the discrete value, $\mu(i)$ is the degree of membership of the value i in the fuzzy number, and the summation is over all discrete values.

Finally, we can use this central tendency value to classify the kidney image as normal or abnormal based on some predetermined criteria. For example, if the central tendency value is less than a certain threshold, we can classify the image as normal, otherwise, we can classify it as abnormal (see [1,7–10]).

This is a simple example of how we can use fuzzy numbers and the center of gravity method to analyze and interpret medical images. Fuzzy logic and its related techniques can have a significant impact on medical images processing. Medical images often are complex, and there may be uncertainty and ambiguity in the interpretation of image features. Fuzzy logic allows us to represent this uncertainty and ambiguity through fuzzy numbers, which can then be discretized and defuzzified to provide more meaningful information. For example, for the case of kidney images, the degree of redness in a certain area may be an important indicator of kidney health. However, there may be variation in the degree of redness, and it may be difficult to determine a precise threshold for what constitutes abnormal redness. Fuzzy logic allows us to represent this variation and uncertainty using fuzzy numbers, and we can then use techniques such as the center of gravity method to extract meaningful information from the fuzzy numbers (see [12,15,21]). Overall, the application of fuzzy logic to medical images can improve the accuracy and the reliability of image analysis and interpretation, which can ultimately lead to better diagnosis and consequently better treatment of medical conditions.

To use the center of gravity method for a triangular fuzzy number, you can use the following code in MATLAB®:

- Define triangular fuzzy numbers and then define the range of values for discretization of triangular fuzzy number (a, b, c) with the feature $a = (min_{range}), c = (max_{range})$ and b is randomly created between a and c.
- Calculate membership function values, where $trimf(x, A)$ is a function to calculate the membership function values for each value in the range $\mu_A = trimf(x, A)$.
- Calculate the center of gravity $cog = sum(x. \cdot \mu_A)/sum(\mu_A)$.
- We then define a range of values for discretization and use the $trimf$ function, and finally, we use the center of gravity formula to calculate the central tendency of the fuzzy number (see [4,13,14,16–19]).

An expansive mass in the left kidney can be seen in Fig. 2, which is a CT (Computer Tomography) scan of the abdomen. Pathohistologically, it was confirmed that it was a renocellular carcinoma. For each description, it is recommended that the physician observe the following changes: location, shape, size, limitation, intensity, and heterogeneity/homogeneity.

Certainly, in order to better distinguish changes during CT and MRI (Magnetic Resonance Imaging) scans, it is recommended to use a contrast agent, and to scan the patient at different time intervals determined by the speed of the passage of the contrast agent through the vascular structures. The CT device is essentially based on the various abilities of tissues to absorb or reflect X-rays, which are then analyzed on detectors and presented to us in the form of images through computer processing. The fuzzy numbers represents the degree of redness in the circled region of the kidney image. The base of the fuzzy numbers represents the range of redness values that can be observed, with the peak of the fuzzy number representing the most likely or central value. The graph of the fuzzy numbers represents the degree of uncertainty in the redness measurement (see [15,20–25]).

The discretization and defuzzification of the fuzzy numbers can then be performed using methods such as the center of gravity method. This can provide a central tendency value for the degree of redness, which can be used in the

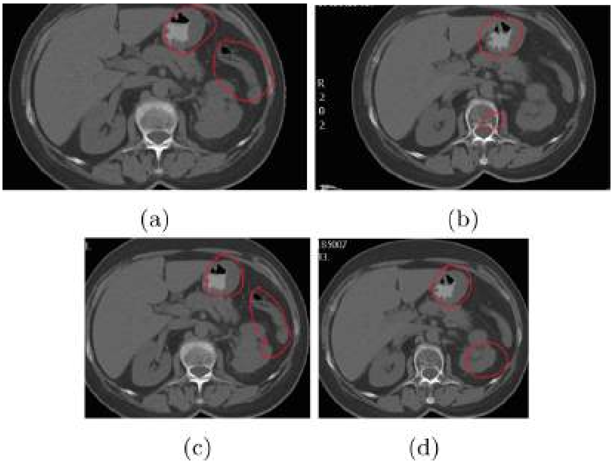


Fig. 2. Original pictures

classification of the kidney image as normal or abnormal based on the various features of the image (see Fig. 3).

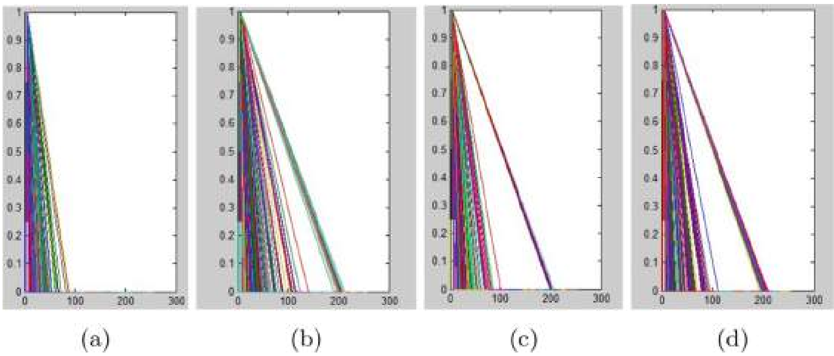


Fig. 3. Fuzzy number representing degree of redness on Fig. 2

4 Conclusion

The application of fuzzy number in the interpretation of the work of neural networks can be useful for assessing uncertainty in data and for understanding the role of individual variables in decision-making process. Fuzzy logic makes it possible to evaluate variables that are not completely defined.

In this paper it was shown that MATLAB® software can be a useful tool for implementing fuzzy logic and neural networks. MATLAB® offers many functions and tools that allow users to easily model and analyze complex systems.

This paper presents examples of the application of these methods to areas of medical imaging and improve the diagnosis and treatment of various medical conditions. Our results demonstrate that the use of fuzzy number and their discretizations and defuzzifications can provide a powerful tool for the analysis of medical images. By representing uncertain and vague data in a more precise manner, we can improve the accuracy and the reliability of medical imaging analysis. The MATLAB® implementations are used in this paper to analyze and interpret medical data with greater accuracy and reliability, demostrating that fuzzy logic and neural networks can be successfully applied in different situations. The library of original medical images recorded/obtained in the radiology office in Tuzla was analyzed. Based on the marked red part in Fig. 2, determined by the application of the mentioned procedure, the success of the answer to the question whether the resulting changes in parts of the kidney are abnormal was tested. The direction of further research will include the publication of adequate results and comparison with appropriate research.

Finally, this paper provides useful guidelines and examples for the application of fuzzy logic and neural networks in real-world situations. These tools are useful in many fields, including engineering, science, economics, and others, where we encounter uncertainties and complex systems.

Acknowledgements. The authors has been supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through the project no. 451-03-47/2023-01/200156: "Innovative scientific and artistic research from the FTS (activity) domain".

References

1. Ross, T.J.: Fuzzy logic with engineering applications. John Wiley and Sons (2010)
2. Jang, J.S.R., Sun, C.T., Mizutani, E.: Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Prentice Hall (1997)
3. Haykin, S.: Neural networks: a comprehensive foundation. Prentice Hall (1999)
4. MATLAB®. The MathWorks, Inc. (2023). https://www.mathworks.com/products/matlab.html
5. Negoita, M.G.: Fuzzy logic-based control systems: fuzzy logic controller design for uncertain nonlinear systems. Springer (2012)
6. Chen, H., Liu, Y.: A fuzzy neural network approach to real estate price forecasting. Appl. Soft Comput. **24**, 649–656 (2004)
7. Bezdek, J.C.: Pattern recognition with fuzzy objective function algorithms. Plenum Press (1981)
8. Dubois, D., Prade, H.: Fuzzy sets and systems: theory and applications. Academic Press (1980)
9. Mamdani, E.H., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man Mach. Stud. **7**(1), 1–13 (1975)
10. Kosko, B.: Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence. Prentice Hall (1992)
11. Klir, G.J., Yuan, B.: Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, New Jersey (1995)

12. Pal, N.R., Pal, S.K.: A review on image segmentation techniques. Pattern Recogn. **26**(9), 1277–1294 (1993)
13. Pedrycz, W.: Fuzzy sets in engineering design. John Wiley and Sons (1993)
14. Ripley, B. D.: Pattern recognition and neural networks. Cambridge University Press (1996)
15. Bezdek, J.C., Hathaway, R.J.: Exponential radial basis functions and fuzzy clustering. Fuzzy Sets Syst. **128**(1), 123–140 (2002)
16. Hagan, M.T., Demuth, H.B., Beale, M.H.: Neural network design. Pws Pub. (1996)
17. Lee, C.C.: Fuzzy logic in control systems: fuzzy logic controller. Part I. IEEE Trans. Syst. Man Cybern. **20**(2), 404–418 (1990)
18. Negoita, M. G., and Negoita, C. V.: Fuzzy systems engineering: theory and practice. Springer (2017). https://doi.org/10.1007/b102051
19. Tadeusiewicz, R.: Fuzzy control: fundamentals, stability and design of fuzzy controllers. Springer (1993). https://doi.org/10.1007/3-540-31766-X
20. Wang, L. X.: Adaptive fuzzy systems and control: design and stability analysis. Prentice Hall (1994)
21. Babuška, R., Kohout, L.: Fuzzy modeling for medical image processing. J. Intell. Rob. Syst. **37**(3), 261–275 (2003)
22. Chen, S., Yang, C.: Fuzzy clustering based on triangular fuzzy number for medical image segmentation. J. Med. Syst. **31**(1), 39–46 (2007)
23. Kumar, R.P., Singh, G.K.: A review on medical image segmentation using soft computing techniques. J. Med. Syst. **39**(12), 1–14 (2015)
24. Li, X., Zhang, L., Liu, Z., Liang, Z., Wei, D.: Fuzzy clustering based on triangular fuzzy number and adaptive threshold for brain image segmentation. J. Med. Syst. **39**(11), 1–9 (2015)
25. Wu, Y., Tian, J.: Fuzzy number based medical image segmentation. J. Med. Syst. **31**(6), 481–490 (2007)

Analyzing and Responding to Google Maps Reviews with a Chatbot in Healthcare

Ekin Akkol[1] and Onur Dogan[1,2](✉)

[1] Department of Management Information Systems, Izmir Bakircay University, 35665 Izmir, Turkey
{ekin.akkol,onur.dogan}@bakircay.edu.tr
[2] Department of Mathematcis, University of Padua, 35121 Padua, Italy
onur.dogan@unipd.it

Abstract. This paper aims to explore how Google Maps reviews for an education and research hospital can be analyzed and responded to using a chatbot. The study highlights the importance of customer feedback in improving hospital services and describes how classification algorithms can be used to collect and analyze reviews. It compares five algorithms to analyze reviews. The chatbot designed in this study responds to reviews and offers personalized suggestions to patients using the most accurate one among five algorithms. The findings suggest that automated chatbot responses can save time and resources while improving the hospital's online reputation. The study concludes that implementing a chatbot for Google Maps reviews can enhance patient satisfaction and lead to better overall service quality. Among the five classification algorithms used within the scope of the study, it was determined that Naive Bayes and Neural Networks algorithms gave the highest accuracy rate with 79% when categorizing the comments according to the subject and performing the sentiment analysis at the same time. However, other algorithms' success rates are similar, and the chatbot responds to people by using the results of the algorithm with the highest success rate for each newly entered sentence.

Keywords: Sentiment analysis · Chatbot · Text mining · Google maps reviews · Healthcare

1 Introduction

Customer feedback is a critical component of improving service quality in hospitals. In today's digital age, Google Maps reviews serve as an essential source of feedback that can help hospitals improve their services. Analyzing and responding to Google Maps reviews can enhance the hospital's online reputation, improve patient satisfaction, and increase trust among potential patients. Natural language processing (NLP) is a branch of artificial intelligence that focuses on enabling computers to understand and process natural language in a way that is similar to how humans understand and use language. NLP techniques are used to extract information from text, classify text, and perform various other

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 116–123, 2023.
https://doi.org/10.1007/978-3-031-39777-6_14

language-related tasks. Text mining, on the other hand, is a process of deriving useful information and insights from unstructured text data. Text mining techniques typically involve natural language processing, statistical analysis, and machine learning to extract insights from text data. These techniques can be used to collect and analyze Google Maps reviews to identify sentiments, themes, and keywords in reviews, which can provide valuable insights into the hospital's services [3].

Sentiment analysis is a specific application of NLP and text mining that involves identifying and extracting opinions, emotions, and attitudes from text data [18]. Sentiment analysis techniques are used to analyze social media data, online or offiine customer reviews, and other forms of text data to understand how people feel about a particular topic, product, or service. For instance, an NLP algorithms for sentiment analysis can detect whether patients are satisfied with the hospital's services, identify common complaints, and highlight areas where the hospital can improve its services. Analyzing Google Maps reviews can help the hospital understand patient needs and expectations, which can inform service improvement efforts.

Responding to Google Maps reviews is an effective way to engage with patients and demonstrate the hospital's commitment to improving services. However, responding to reviews can be time-consuming and resource-intensive. To address this challenge, this study proposes the use of a chatbot to automate responses to reviews. The chatbot can be designed to respond to reviews based on predefined templates, which can save time and resources while ensuring that patients receive personalized responses.

The chatbot can also offer personalized suggestions to patients based on their review content. For example, if a patient complains about long wait times, the chatbot can suggest that the patient consider scheduling an appointment during less busy times or offer information about the hospital's online appointment scheduling system. Personalized suggestions can help patients feel heard and valued, which can improve patient satisfaction and lead to better overall service quality [6].

Analyzing and responding to Google Maps reviews is essential for improving service quality in hospitals. NLP algorithms can be used to collect and analyze reviews, which can provide valuable insights into patient needs and expectations. Using a chatbot to automate responses to reviews can save time and resources while improving the hospital's online reputation. The chatbot can also offer personalized suggestions to patients, which can improve patient satisfaction and lead to better overall service quality. By implementing a chatbot for Google Maps reviews, hospitals can enhance patient satisfaction and stay competitive in today's digital age.

This study contributes to the literature by not only analyzing the online customer reviews by sentiment analysis but also responding to the reviews by a chatbot to identify the problems by voice-of-patient. Patients' feedback is crucial for improving the hospital's level of service quality and should be taken into consideration. The study aims to rapidly take into account the patients' comments and to assist them by replying. The remaining part of the paper is structured as follows. Section 2 summarizes the literature to show the related work. Section 3 explains the methodology applied in this study. Section 4 gives and discusses the results. Finally, Sect. 5 concludes the paper.

2 Literature Review

Research has been conducted on sentiment analysis at various levels of granularity. Some studies have been carried out at the document level [15], where the entire review is categorized, while others have been published at the sentence level [10,11], focusing on determining the polarity (e.g., positive, neutral, negative) of individual sentences based on the semantic information learned from the textual content. More recently, some works have applied sentiment analysis at the phrase level, where the primary focus is on a group of words that often carry a specific idiomatic sense.

Sentiment analysis has been applied in various domains, from travel to healthcare [9,16]. Monitoring customers' online sentiments can be used to give standard or dynamic feedback. Texts generated on social media are very attractive for researchers in this area. [14] mapped sentiments on social media with observations and measurable data. Among the social media channels, Twitter is one of the most used sources for sentiment analysis [1,2]. However Google Reviews can be a valuable tool for analyzing a hospital's performance. Google Reviews is a platform for gathering feedback allowing users to rate and provide comments on various services. The platform identifies important keywords based on frequency to present information more concisely. The reviews can be sorted in four different ways: by relevance, date of posting, the highest rating, and lowest rating.

Sentiment analysis is often combined with other analytical methods to gain a more comprehensive understanding of a particular subject or domain. For example, sentiment analysis can be used alongside natural language processing (NLP) to extract relevant keywords and phrases from a large volume of text data [5]. This can help to identify patterns and trends in customer feedback, social media posts, or news articles. Additionally, sentiment analysis can be used in conjunction with machine learning algorithms to classify text data into different categories based on the sentiment expressed. This can be particularly useful in fields such as marketing, where sentiment analysis can be used to understand consumer preferences and behavior. A combination of machine learning features can be used with lexicon features [9], or perceptron neural networks can be applied to understand the text better [16], and deep learning techniques to detect specific speech [1]. Overall, combining sentiment analysis with other analytical methods can lead to more accurate and nuanced insights into various aspects of human communication and behavior.

This work analyzes the Google reviews of a hospital at the sentence level to determine their positive (olumlu) and negative (olumsuz) sentiments. It also classifies the reviews into three categories: General (Genel), Staff (Personel), and Cleaning (Temizlik). To enhance the usability of this sentiment analysis, a chatbot was developed in this study that provides automated responses to the reviews. It provides standard responses based on these categories, allowing the hospital managers to respond to the reviews more faster for the first contact and addressing the specific concerns and praises mentioned by the reviewer. The chatbot was trained using machine learning algorithms and incorporated various response templates covering various review scenarios.

3 Methodology

Many statistical and machine learning techniques, such as the kNN method, Naive Bayes, Rocchio Algorithm, multivariate regression models, decision trees, Support Vector Machines (SVMs), neural networks, graph partitioning-based approach, and genetic algorithm-based methods, have been developed in the last decade to automatically classify documents [7]. A sentiment analysis and category classification chatbot was developed in this study by using Naive Bayes, Logistic Regression, Support Vector Machines, Random Forest, and Neural Networks.

Naive Bayes is a probability-based algorithm used in machine learning and natural language processing to solve classification problems. It is based on Bayes' Theorem and assumes that when classifying, the features in the data, for example, are independent of each other. Naive Bayes algorithm has high success rates in classification problems and is especially useful for natural language processing problems [13].

Logistic Regression is an algorithm used in machine learning and statistics to solve classification problems. By combining the features in the data, this algorithm determines whether or not the features belong to a class. Based on the values of a set of predictor variables, logistic regression can predict the presence or absence of a characteristic or outcome [12].

Support Vector Machines (SVM) is a machine learning algorithm used in classification and regression analysis. SVM provides high accuracy in linear or nonlinear classification problems and can also work effectively with high-dimensional data [8].

Random Forest is a machine learning algorithm that was created by combining many decision trees. Each decision tree is trained with randomly selected features from different subsets before being combined for use in classification or regression problems [4].

A neural network is a mathematical or computational model based on biological neural networks, or an emulation of a biological neural system. It is made up of an interconnected network of artificial neurons that process data using a connectionist approach to computation. This algorithm can be used to solve a variety of problems, including complex data classification, pattern recognition, estimation, and regression [17].

To begin, the comments made for Izmir Bakircay University Cigli Regional Education Hospital via Google Maps were collected include the people's user names, comments about the hospital, the time of the comments, and the number of stars they gave. Since the text analysis was carried out within the scope of this study, only the comments made by the people about the hospital were used. The comments were then manually tagged in three categories (General, Staff, Cleaning) and two emotions (Positive and Negative). The distribution of the data after labeling is shown in Table 1. Accordingly, there are 115 comments, 26 positive and 86 negative, under the General label. Under the Cleaning label, there are 40 comments, 8 positive and 32 negative. The Staff tag has a total of 273 comments, 101 positive and 172 negative.

Table 1. Labeled Data Set

Category	Positive	Negative	Total
General	29	86	115
Cleaning	8	32	40
Staff	101	172	273
Total	138	290	428

Pandas, NLTK, and Scikit-Learn libraries were used in the Python programming language during the data processing and analysis. First, the preprocessing steps were carried out. All punctuation and links in comments have been cleaned up, all comments have been converted to lowercase, stopwords have been removed, and comments have been lemmatized. After the pre-processing phase was completed and the data were ready for analysis, the subject-based categorization of the comments was executed, followed by the sentiment analysis. A chatbot was created that categorizes newly entered comments based on topic and emotion, employing five algorithms that were discovered to be the most appropriate for the data set and provide the most accurate results within the scope of the study. The success score of each algorithm is compared separately using five algorithms, and the answer of the algorithm with the highest success score was used as the final answer. The process used in the study is shown in Fig. 1.

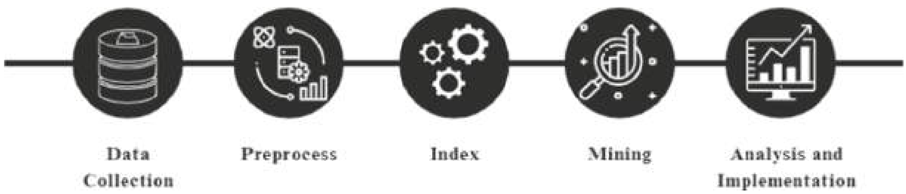


Fig. 1. Text Mining Process

4 Results and Discussion

Subject categorization and sentiment analysis were performed using Naive Bayes, Logistic Regression, Support Vector Machines, Random Forest and Neural Networks algorithms. All classification methods used were compared with the determined metrics. Accuracy, precision, and F1 score are metrics commonly used in evaluating the performance of classification models. Accuracy measures the overall correctness of the model's predictions. Precision measures the proportion of positive predictions that were correct. The F1 score is a weighted harmonic mean of precision and recall. It provides a balance between precision and recall. Recall (also known as sensitivity) is the proportion of positive instances that were correctly identified.

It was aimed to achieve the best success rate and showed the results of classification according to the best success rate of the chatbot. 70% of the tagged data was used as training and 30% as test dataset. Due to the fact that the data set consists entirely of real interpretation data and was obtained from a limited source, there were imbalances in the data set and it was seen as a limitation of this study. The evaluation results of the classification algorithms are shown in Table 2.

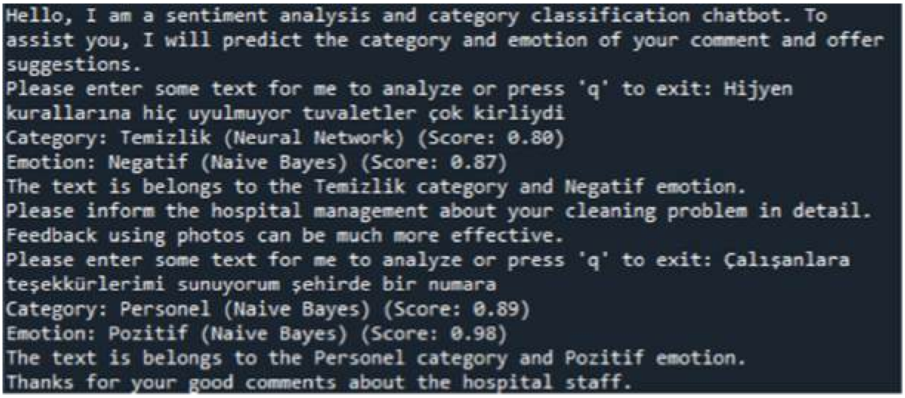
Table 2. Classification Algorithms Evaluation

Classifier	Accuracy	Precision	F1 Score
Naive Bayes	0.79	0.78	0.78
Logistic Regression	0.76	0.78	0.76
Support Vector Machines	0.72	0.75	0.73
Random Forest	0.78	0.77	0.78
Neural Network	0.79	0.80	0.79

According to these results, the algorithms provide relatively similar outcomes, and when all metrics are considered. It was determined that the Neural Network produced the most successful results, albeit with a small difference. Besides the general classification results, the developed chatbot evaluated newly entered sentences using all algorithms and displays the result of the algorithm with the highest success rate to the user. It was designed to guide people according to the category and emotion of their comments. It was expected that this method would result in a more successful chatbot.

Figure 2 shows the chatbot's responses to the sentences entered by the user. The user who made a negative comment about Cleaning was advised that it would be beneficial to report the situation to the hospital management with photographs.

It is necessary to carefully examine and analyze all comments made to organizations that provide services on important issues on behalf of the society, such as hospitals. Managers are frequently unable to check all of the comments one by



```
Hello, I am a sentiment analysis and category classification chatbot. To
assist you, I will predict the category and emotion of your comment and offer
suggestions.
Please enter some text for me to analyze or press 'q' to exit: Hijyen
kurallarına hiç uyulmuyor tuvaletler çok kirliydi
Category: Temizlik (Neural Network) (Score: 0.80)
Emotion: Negatif (Naive Bayes) (Score: 0.87)
The text is belongs to the Temizlik category and Negatif emotion.
Please inform the hospital management about your cleaning problem in detail.
Feedback using photos can be much more effective.
Please enter some text for me to analyze or press 'q' to exit: Çalışanlara
teşekkürlerimi sunuyorum şehirde bir numara
Category: Personel (Naive Bayes) (Score: 0.89)
Emotion: Pozitif (Naive Bayes) (Score: 0.98)
The text is belongs to the Personel category and Pozitif emotion.
Thanks for your good comments about the hospital staff.
```

Fig. 2. Chatbot Responses

one. Therefore, it is seen as an effective way of problem solving that applications such as chatbot both analyze the comments and assist users in their problems.

5 Conclusion and Future Research

The study was conducted to explore how Google Maps reviews for an education and research hospital can be analyzed and responded to using a chatbot. The problem that the study aimed to address was the need for hospitals to improve service quality and enhance their online reputation by analyzing and responding to customer feedback. Google Maps reviews serve as an essential source of feedback in the digital age, and analyzing and responding to them can provide valuable insights into patient needs and expectations. However, responding to reviews can be time-consuming and resource-intensive, which can be challenging for hospitals. The study aimed to address this challenge by proposing the use of a chatbot to automate responses to reviews, saving time and resources while improving patient satisfaction and enhancing the hospital's online reputation.

Within the scope of the study, it was aimed to categorize the comments according to the subject and to perform sentiment analysis in the best way with five different algorithms. According to the results, all algorithms used give similar results. However, in the new comments directed to the chatbot, it was seen that the Naive Bayes and Neural Networks algorithms made the most successful predictions. Since the study was conducted with real user data, the number of data is limited and there are data imbalances between categories. It had been determined that much more successful classification results could be obtained with a more balanced and large data set. The success rate of the designed chatbot was quite good when shorter and clearer sentences are presented, but it decreases in very long sentences due to the small number of data. In addition to the ability to categorize comments from people, the designed chatbot also had the ability to guide people by giving different feedbacks according to their comments.

For future directions, authors can improve the variety of the chatbot's responses. Also, integrating the chatbot into the hospital information system can enable hospital managers to understand patients' expectations to improve not only health but also administrative services.

References

1. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.J.: Sentiment analysis of twitter data. In: Proceedings of the Workshop on Language in Social Media (LSM 2011), pp. 30–38 (2011)
2. Badjatiya, P., Gupta, S., Gupta, M., Varma, V.: Deep learning for hate speech detection in tweets. In: Proceedings of the 26th International Conference on World Wide Web Companion, pp. 759–760 (2017)
3. Bozyigit, F., Dogan, O., Kilinc, D.: Categorization of customer complaints in food industry using machine learning approaches. J. Intell. Syst. Theory Appli. **5**(1), 85–91 (2022)
4. Breiman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)
5. Dang, N.C., Moreno-García, M.N., De la Prieta, F.: Sentiment analysis based on deep learning: a comparative study. Electronics **9**(3), 483 (2020)
6. Deng, X., Yu, Z.: A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. Sustainability **15**(4), 2940 (2023)
7. Deng, X., Li, Y., Weng, J., Zhang, J.: Feature selection for text classification: a review. Multimedia Tools Appli. **78**, 3797–3816 (2019)
8. Gunn, S.R., et al.: Support vector machines for classification and regression. ISIS technical report **14**(1), 5–16 (1998)
9. Harish, B., Kumar, K., Darshan, H.: Sentiment analysis on imdb movie reviews using hybrid feature extraction method. Int. J. Interactive Multimedia Artifi. Intell. **5**(5), 109–114 (2019)
10. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 168–177 (2004)
11. Kim, S.M., Hovy, E.: Determining the sentiment of opinions. In: COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics, pp. 1367–1373 (2004)
12. Lee, S.: Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS. Environ. Manage. **34**, 223–232 (2004)
13. Lewis, D.D., Catlett, J.: Heterogeneous uncertainty sampling for supervised learning. In: Machine Learning Proceedings 1994, pp. 148–156. Elsevier (1994)
14. Neri, F., Aliprandi, C., Capeci, F., Cuadros, M.: Sentiment analysis on social media. In: 2012 IEEE/ACM International Conference on advances in Social Networks Analysis and Mining, pp. 919–926. IEEE (2012)
15. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. arXiv preprint cs/0409058 (2004)
16. Shaukat, Z., Zulfiqar, A.A., Xiao, C., Azeem, M., Mahmood, T.: Sentiment analysis on IMDB using lexicon and neural networks. SN Appli. Sci. **2**, 1–10 (2020)
17. Singh, Y., Chauhan, A.S.: Neural networks in data mining. J. Theor. Appli. Inform. Technol. **5**(1) (2009)
18. Çılgın, C., Ünal, C., Alıcı, S., Akkol, E., Gökşen, Y.: Metin sınıflandırmada yapay sinir ağları ile bitcoin fiyatları ve sosyal medyadaki beklentilerin analizi. Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi **4**(1), 106–126 (2020)

A Novel Decision-Making Framework Based on Interval-Valued Pythagorean Fuzzy AHP and CODAS: An Application to Healthcare Supplier Selection

Irem Otay, Bayram Bulut(✉), Kaan Barutbaş, Ayşe Ecem Arat, Emre Tekin, and Zeynep Şevval Gürz

Department of Industrial Engineering, Istanbul Bilgi University, 34060 Eyüpsultan, İstanbul, Turkey

irem.otay@bilgi.edu.tr, {bayram.bulut,kaan.barutbas,ecem.arat, emre.tekin,sevval.gurz}@bilgiedu.net

Abstract. Especially after the COVID-19 pandemics where market dynamics have changed dramatically, supplier selection has been one of the crucial subjects for almost every industry. After the pandemics, the healthcare industry has played a major role and gained even more importance in all over the world; therefore, selecting the best appropriate supplier has become more critical than ever. In this paper, supplier selection and evaluation problem is analyzed for a specific area of the healthcare industry, i.e. laboratory (lab.) equipments including test tubes and sample containers. Considering the vagueness and imprecise information in the decision-making environment, the problem is handled by considering five main criteria and 20 sub-criteria under fuzzy environment. In the study, a relatively recent extension of the fuzzy sets called as Pythagorean Fuzzy Sets (PFS), which provides more flexibility to the decision makers with membership and non-membership degrees, are utilized. This paper aims to contribute to the literature by proposing a novel fuzzy methodology integrating Analytic Hierarchy Process (AHP) and Combinative Distance-based Assessment (CODAS) method based on interval-valued PFS. Finally, sensitivity analysis is constructed in the paper.

Keywords: Pythagorean Fuzzy AHP · Pythagorean Fuzzy CODAS · Interval-valued Pythagorean Fuzzy Sets · Supplier Selection · Healthcare

1 Introduction

Supplier selection refers to the process including suppliers to find, to evaluate, and to make a contract with the selected supplier/s by evaluating a list of quantitative and/or qualitative attributes which may conflict with each other. By reducing purchasing risks, maximizing overall value to the buyers, and its impacts on cultivating intimacy and long-term relationships with buyers, supplier evaluation and selection provides competitive advantage and enhance corporations' success. Supplier selection is also essential part

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 124–132, 2023.
https://doi.org/10.1007/978-3-031-39777-6_15

of healthcare industry considering the importance of timely and with the right quantity and quality procurements of the medicines, medical kits, or surgical equipment, especially for the cases where it is necessary to intervene patients urgently with the right supply. There are many papers such as Pamucar et al. (2022) and Stević et al. (2020) dealing with healthcare supplier selection problem utilizing multi-criteria decision making (MCDM) methods. However, according to the best knowledge of the authors, this is the first study focusing on supply selection problem for one of the lab. Equipments by proposing an integrated Analytic Hierarchy Process (AHP) integrated with COmbinative Distance-based ASsessment (CODAS) method considering uncertainties. CODAS is a distance-based approach that evaluates the alternatives based on Euclidean and Hamming distances where a degree of closeness to Euclidean distances is determined by a threshold parameter (Ghorabaee et al. 2016). The proposed AHP based CODAS method is designed under Interval-valued Pythagorean fuzzy sets (IVPFSs) are employed to deal with imprecisessness and vaguness by covering a larger domain.

The paper is designed as follows: In the second section, preliminaries are given while in the third section proposed IVPF AHP& CODAS methodology is presented. In Sect. 4, the proposed IVPF methodology is implemented to a real case. In the same section, the results of sensitivity analysis are also displayed. Finally, in Sect. 5 the paper is summarized and concluded.

2 Preliminaries

In Pythagorean fuzzy (PF) sets introduced by Yager (2013), the sum of membership and non-membership degrees can exceed "1"; however, the sum of the squared values is limited by "1". Pythagorean fuzzy set $\tilde{A}$ is presented in Eq. (1) where $\mu_A(x) : X \langle \quad [0, 1]$ and $\nu_A(x) : X \langle \quad [0, 1]$ show the degrees of membership and non-membership of the element x $\rangle$ X to P, respectively (Yager 2016).

$$\begin{gathered}\tilde{\mathrm{A}} = \{< x, \mu_{\tilde{\mathrm{A}}}(x), \nu_{\tilde{\mathrm{A}}}(x) > |x \rangle\ X\} \\ \text{where } 0 \leq \mu_{\tilde{\mathrm{A}}}(x)^2 + \nu_{\tilde{\mathrm{A}}}(x)^2 \leq 1\end{gathered} \tag{1}$$

The hesitancy degree of the Pythagorean fuzzy set $\tilde{A}$ is expressed as $\pi(x) = \sqrt{1 - \mu_{\tilde{\mathrm{A}}}(x)^2 - \nu_{\tilde{\mathrm{A}}}(x)^2}$.

Let $\tilde{A} = \{\langle x, [\mu_L(x), \mu_U(x)], [\nu_L(x), \nu_U(x)]\rangle | x \rangle\ X\}$ be an interval-valued Pythagorean fuzzy numbers (IVPFNs), lower and upper degrees of π (π_L and π_U) can be obtained using Eq. (2) (Boltürk and Kahraman 2019).

$$\pi_U^2 = 1 - (\mu_L^2 + \nu_L^2),\ \mathrm{\pi}_{\mathrm{L}}^2 = 1 - \left(\mathrm{\mu}_{\mathrm{U}}^2 + \nu_{\mathrm{U}}^2\right) \tag{2}$$

Let $\tilde{A}_1$ and $\tilde{A}_2$ be two IVPFNs as represented in Eq. (3).

$$\tilde{A}_1 = [\mu_{\tilde{A}_1}^L, \mu_{\tilde{A}_1}^U], [\nu_{\tilde{A}_1}^L, \nu_{\tilde{A}_1}^U]) \text{ and } \tilde{A}_2 = ([\mu_{\tilde{A}_2}^L, \mu_{\tilde{A}_2}^U], [\nu_{\tilde{A}_2}^L, \nu_{\tilde{A}_2}^U]) \tag{3}$$

Arithmetic operations for $\tilde{A}_1$ and $\tilde{A}_2$ can be listed in the followings (Rahman et al. 2017):

$$\tilde{A}_1 \oplus \tilde{A}_2 = \left[\sqrt{\left(\mu^L_{\tilde{A}_1}\right)^2 + \left(\mu^L_{\tilde{A}_2}\right)^2 - \left(\mu^L_{\tilde{A}_1}\right)^2\left(\mu^L_{\tilde{A}_2}\right)^2}, \sqrt{\left(\mu^U_{\tilde{A}_1}\right)^2 + \left(\mu^U_{\tilde{A}_2}\right)^2 - \left(\mu^U_{\tilde{A}_1}\right)^2\left(\mu^U_{\tilde{A}_2}\right)^2}\right], \left[\nu^L_{\tilde{A}_1}\nu^L_{\tilde{A}_2}, \nu^U_{\tilde{A}_1}\nu^U_{\tilde{A}_2}\right] \tag{4}$$

$$\tilde{A}_1 \otimes \tilde{A}_2 = \left(\left[\mu^L_{\tilde{A}_1}\mu^L_{\tilde{A}_2}, \mu^U_{\tilde{A}_1}\mu^U_{\tilde{A}_2}\right], \left[\begin{array}{c}\sqrt{\left(\nu^L_{\tilde{A}_1}\right)^2 + \left(\nu^L_{\tilde{A}_2}\right)^2 - \left(\nu^L_{\tilde{A}_1}\right)^2\left(\nu^L_{\tilde{A}_2}\right)^2}, \\ \sqrt{\left(\nu^U_{\tilde{A}_1}\right)^2 + \left(\nu^U_{\tilde{A}_2}\right)^2 - \left(\nu^U_{\tilde{A}_1}\right)^2\left(\nu^U_{\tilde{A}_2}\right)^2}\end{array}\right]\right) \tag{5}$$

Let $\tilde{A} = \langle[\mu_L, \mu_U], [\nu_L, \nu_U]\rangle$ be an interval-valued Pythagorean fuzzy number (IVPFN); then, the score function (Sc) is defined as Eq. (6) (Peng and Yang 2016):

$$\mathrm{Sc}\left(\tilde{\mathrm{A}}\right) = \frac{1}{2}\left[(\mu_L)^2 + (\mu_U)^2 - (\nu_L)^2 + (\nu_U)^2\right] \tag{6}$$

3 Proposed IVPF AHP and CODAS Methodology

In this section, the steps of the proposed IVPF AHP & IVPF CODAS methodology are presented briefly (Karasan et al. 2018; Bolturk and Kahraman 2019; Pérez-Domínguez et al. 2021):

Step 1: State fuzzy MCDM problem by identifying the set of criteria ($\mathrm{C_j}$) and alternatives ($\mathrm{A_i}$).
Step 2: Collect the pairwise comparison matrices $R = \left(r_{jt}\right)_{mxm}$ from the experts using linguistic terms given in (Karasan et al. 2019).
Step 3: Obtain the differences matrices $\left(D = \left(d_{ij}\right)_{mxm}\right)$ between lower and upper points of the membership & non-membership degrees based on Eqs. (7–8).

$$d_{ij_L} = \mu^2_{ij_L} - \nu^2_{ij_U} \tag{7}$$

$$d_{ij_U} = \mu^2_{ij_U} - \nu^2_{ij_L} \tag{8}$$

Step 4: Compute interval multiplication matrix $\left(S = \left(s_{ij}\right)_{mxm}\right)$ by means of Eq. (9).

$$s_{ij_L} = \sqrt{1000^{d_{ijL}}}, \quad s_{ij_U} = \sqrt{1000^{d_{ijU}}} \tag{9}$$

Step 5: Calculate indeterminacy degrees $\left(H = \left(h_{ij}\right)_{mxm}\right)$ by Eq. (10).

$$h_{ij} = 1 - \left(\mu^2_{ij_U} - \mu^2_{ij_L}\right) - \left(\nu^2_{ij_U} - \nu^2_{ij_L}\right) \tag{10}$$

Step 6: Derive unnormalized weights $\left(T = \left(\tau_{ij}\right)_{mxm}\right)$ using Eq. (11), and compute the priority weights (Eq. (12)).

$$\tau_{ij} = \left(\frac{s_{ij_L} + s_{ij_U}}{2}\right) h_{ij} \tag{11}$$

$$\omega_i = \frac{\sum_{j=1}^{m} \omega_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{m} \omega_{ij}} \tag{12}$$

Step 7: Obtain the Pythagorean fuzzy weighted normalized matrix.

$$u_{ij} = \omega_j r_{ij} \tag{13}$$

where ω_j is the weight of criteria $(\sum \omega_j = 1)$
Step 8: Compute hesitancy values of π_L and π_U by Eq. (14).

$$\pi_U^2 = 1 - \left(\mu_L^2 + \nu_L^2\right), \pi_L^2 = 1 - \left(\mu_U^2 + \nu_U^2\right) \tag{14}$$

Step 9: Calculate the score values (Eq. (6)) of each judgment in the Pythagorean fuzzy weighted normalized matrix in order to determine Pythagorean Fuzzy Negative Ideal Solutions (PFNIS).
Step 10: Compute Hamming (Eq. (15)) and Euclidean distances (Eq. (16)) from each judgment on the alternatives to PFNIS (Zhang 2017).

$$d_H\left(\tilde{A}_1, \tilde{A}_2\right) = \frac{1}{4}\left(\begin{array}{l} \left|\left(\mu_{\tilde{A}_1}^L\right)^2 - \left(\mu_{\tilde{A}_2}^L\right)^2\right| + \left|\left(\mu_{\tilde{A}_1}^U\right)^2 - \left(\mu_{\tilde{A}_2}^U\right)^2\right| + \left|\left(\nu_{\tilde{A}_1}^L\right)^2 - \left(\nu_{\tilde{A}_2}^L\right)^2\right| \\ + \left|\left(\nu_{\tilde{A}_1}^U\right)^2 - \left(\nu_{\tilde{A}_2}^U\right)^2\right| + \left|\left(\pi_{\tilde{A}_1}^L\right)^2 - \left(\pi_{\tilde{A}_2}^L\right)^2\right| + \left|\left(\pi_{\tilde{A}_1}^U\right)^2 - \left(\pi_{\tilde{A}_2}^U\right)^2\right| \end{array} \right) \tag{15}$$

$$d_E\left(\tilde{A}_1, \tilde{A}_2\right) = \frac{1}{2}\sqrt{\left(\begin{array}{l} \left(\left(\mu_{\tilde{A}_1}^L\right)^2 - \left(\mu_{\tilde{A}_2}^L\right)^2\right)^2 + \left(\left(\mu_{\tilde{A}_1}^U\right)^2 - \left(\mu_{\tilde{A}_2}^U\right)^2\right)^2 + \left(\left(\nu_{\tilde{A}_1}^L\right)^2 - \left(\nu_{\tilde{A}_2}^L\right)^2\right)^2 + \\ \left(\left(\nu_{\tilde{A}_1}^U\right)^2 - \left(\nu_{\tilde{A}_2}^U\right)^2\right)^2 + \left(\left(\pi_{\tilde{A}_1}^L\right)^2 - \left(\pi_{\tilde{A}_2}^L\right)^2\right)^2 + \left(\left(\pi_{\tilde{A}_1}^U\right)^2 - \left(\pi_{\tilde{A}_2}^U\right)^2\right)^2 \end{array} \right)} \tag{16}$$

Step 11: Determine relative assessment matrix based on Eqs. (17–19):

$$Ra = \left[h_{ij}\right]_{mxm} \tag{17}$$

$$h_{ij} = (E_i - E_j) + (\psi(E_i - E_j)x(T_i - T_j)) \tag{18}$$

$$\psi(y) = \begin{cases} 1 if\, |y| \geq \tau \\ 0 if\, |y| < \tau \end{cases} \tag{19}$$

where ψ is a threshold function to identify the equivalence of two options' Euclidean distances.

Step 12: Calculate assessment score (H_i) of each alternative. Then, sort the alternatives according to decreasing order of assessment scores.

$$H_i = \sum_{j=1}^{n} h_{ij} \tag{20}$$

4 Implementation

The proposed IVPF AHP&CODAS methodology is implemented in a real problem of a company providing innovative healthcare solutions and manufacturing medical devices. Using computed tomography, MRI, molecular imaging, X-ray, ultra-sound and imaging CT, the company controls the diagnostic imaging market. It has been noted that there is a problem with the lab. Equipment (blood tubes, urine sample containers, etc.) procurement process.

4.1 Problem Solution

After conducting interviews with the engineers working in the procurement department, data where the engineers reach consensus are used in the analysis. The criteria set is determined based on the meetings with the experts and extensive literature review as illustrated in Table 1. For this study, the top five suppliers are analyzed for a specific product.

Table 1. The set of criteria and sub-criteria.

Criteria	Sub criteria
Cost (C1)	Unit Cost (C11)
	Operational Cost (C12)
	Quantity Discount (C13)
	Payment options (C14)
Company Profile (C2)	Reputation (C21)
	Globalization (C22)
	Past Performance (C23)
	Market Share (C24)
	Communication (C25)
Quality (C3)	Guarantee (C31)
	Return Discount (C32)
	Return Policy (C33)
	Conformance (C34)
Capability (C4)	Prod. Flexibility (C41)
	Multi Item Prod. (C42)
	R&D Studies (C43)
	Functionality (C44)
Service Performance (C5)	Lead Time (C51)
	On Time Delivery (C52)
	Customer Support (C53)

First, linguistic pairwise comparison matrix is collected from the experts utilizing Table 1, as shown in Table 2. In the analysis, consistency ratios are computed using equivalent 1–9 scores for the fuzzy judgments.

Table 2. Linguistic pairwise comparison matrix of main criteria.

	C1	C2	C3	C4	C5
Cost (C1)	EI	HI	AAI	HI	AAI
Company Profile (C2)	LI	EI	LI	BAI	LI
Quality (C3)	BAI	HI	EI	HI	AAI
Capability (C4)	LI	AAI	LI	EI	BAI
Service Performance (C5)	BAI	HI	BAI	AAI	EI

Consistency Ratio: 0.08.

By following the steps of the proposed methodology, the differences matrix between the lower and higher values of membership and non-membership are computed by Eqs. (7–8), and the interval multiplicative matrix is obtained via Eq. (9), as both listed in Table 3.

Table 3. Difference and Interval multiplicative matrices.

Difference matrix					
	C1	C2	C3	C4	C5
C1	(0,0)	(0.30, 0.60)	(0.10, 0.30)	(0.30, 0.60)	(0.10, 0.30)
C2	(−0.60, −0.30)	(0,0)	(−0.60, −0.30)	(−0.30, −0.10)	(−0.6, −0.30)
C3	(−0.30, −0.10)	(0.30, 0.60)	(0,0)	(0.30, 0.60)	(0.10, 0.30)
C4	(−0.60, −0.30)	(0.10, 0.30)	(−0.60, −0.30)	(0,0)	(−0.3, −0.10)
C5	(−0.30, −0.10)	(0.30, 0.60)	(−0.30, −0.10)	(0.10, 0.30)	(0,0)
Interval multiplicative matrix					
	C1	C2	C3	C4	C5
C1	(1,1)	(2.82, 7.94)	(1.41, 2.82)	(2.82, 7.94)	(1.41, 2.82)
C2	(0.13, 0.35)	(1,1)	(0.13, 0.35)	(0.35, 0.71)	(0.13, 0.35)
C3	(0.35, 0.71)	(2.82, 7.94)	(1,1)	(2.82, 7.94)	(1.41, 2.82)
C4	(0.13, 0.35)	(1.41, 2.82)	(0.13, 0.35)	(1,1)	(0.35, 0.71)
C5	(0.35, 0.71)	(2.82, 7.94)	(0.35, 0.71)	(1.41, 2.82)	(1,1)

Next, indeterminacy values have been calculated with Eq. (10); hence, unnormalized weights are computed using Eqs. (11–12) as illustrated in Table 4. Because of the page limitation, only pairwise comparison matrix of criterion C1 is displayed (Table 5). The priority weights of the sub-criteria are obtained as in Table 6.

Table 4. Indeterminacy values, unnormalized and priority weights of the criteria.

The indeterminacy degrees						Unnormalized weights matrix						Priority weights
	C1	C2	C3	C4	C5		C1	C2	C3	C4	C5	
C1	1.00	0.70	0.80	0.70	0.80	C1	1.00	3.77	1.69	3.77	1.69	0.34
C2	0.70	1.00	0.70	0.80	0.70	C2	0.17	1.00	0.17	0.43	0.17	0.06
C3	0.80	0.70	1.00	0.70	0.80	C3	0.43	3.77	1.00	3.77	1.69	0.30
C4	0.70	0.80	0.70	1.00	0.80	C4	0.17	1.69	0.17	1.00	0.43	0.10
C5	0.80	0.70	0.80	0.80	1.00	C5	0.43	3.77	0.43	1.69	1.00	0.21

Table 5. Linguistic pairwise comparison matrix of criterion C1.

	C11	C12	C13	C14
C11	EI	AAI	VHI	HI
C12	BAI	EI	HI	AAI
C13	VLI	LI	EI	LI
C14	LI	BAI	HI	EI

Table 6. Priority weights of sub criteria.

Sub criteria	Weights	Sub criteria	Weights	Sub criteria	Weights	Sub criteria	Weights
C11	0.54	C22	0.06	C32	0.22	C43	0.11
C12	0.23	C23	0.41	C33	0.14	C44	0.22
C13	0.05	C24	0.10	C34	0.58	C51	0.26
C14	0.18	C25	0.31	C41	0.41	C52	0.49
C21	0.11	C31	0.06	C42	0.25	C53	0.26

To evaluate the alternative suppliers, linguistic fuzzy decision matrix is asked from the experts. In Table 7, a part of the decision matrix is presented as a reason of the page limitation.

Herein, the Pythagorean fuzzy weighted normalized matrix is obtained using the weights obtained from IVPF AHP method. Then, score values are computed to identify PFNIS. Hamming and Euclidian distances from each judgment on the alternatives to

Table 7. Pythagorean fuzzy decision matrices.

	A1	A2	A3	A4	A5		A1	A2	A3	A4	A5
C11	VHI	CHI	AAI	AAI	BAI	C21	AAI	EI	AAI	HI	VHI
C12	AAI	AAI	BAI	BAI	AAI	C22	BAI	BAI	BAI	VHI	VHI
C13	EI	AAI	AAI	EI	EI	C23	CHI	BAI	BAI	BAI	HI
C14	AAI	AI	EI	EI	EI	C24	HI	BAI	HI	VHI	VHI
						C25	CHI	BAI	BAI	AAI	VHI

PFNIS are computed. Finally, the relative assessment matrix and assessment scores of the suppliers are derived by means of Eqs. (17–19) as given in Table 8 ($\tau = 0.02$) 0.02).

Table 8. Assessment scores.

Hij	A1		A2		A3		A4		A5
A1-A2	8.52	A2-A1	−8.52	A3-A1	−12.37	A4-A1	−12.16	A5-A1	−12.02
A1-A3	12.37	A2-A3	3.85	A3-A2	−3.85	A4-A2	−3.64	A5-A2	−3.50
A1-A4	12.16	A2-A4	3.64	A3-A4	−0.22	A4-A3	0.22	A5-A3	0.35
A1-A5	12.02	A2-A5	3.50	A3-A5	−0.35	A4-A5	−0.13	A5-A4	0.13

The results indicate that Supplier A1 is the best alternative with the score of 45.07 while Supplier A2 is the second one with 2.48. The rest of the alternatives are ranked as follows: A5 (−15.05) > A4 (−15.71) > A3 (−16.79).

4.2 Sensitivity Analysis

In this study, sensitivity analysis is implemented to five main criteria as changing each criterion weight between "0.1" and "0.9" while keeping the remaining criteria weights equal, and the sum of the weights makes "1". It is observed that the results of the proposed method are robust. Only there are slight differences in the ranking of the alternatives for the changes in the weights of C1, C4 and C5. For instance, for the weights of "0.8" and "0.9" in C4 and C5, the rank is A1>A2>A5>A3>A4 while it is A1>A2>A3>A4>A5 for the weight of C1 equal to "0.9".

5 Conclusion

In light of the global pandemic, the process of selecting an optimal supplier from a list of alternatives has taken on enhanced significance, particularly given the intricate and dynamic nature of contemporary market structures. This study proposes AHP method in conjunction with an IVPF to determine the relative criteria weights. Subsequently, IVPF

CODAS method is utilized to assess five alternative suppliers for the health lab. Equipment. The IVPF AHP and CODAS methods have both yielded meaningful approaches and results. Future research may utilize other fuzzy set extensions i.e. picture fuzzy sets, neutrosophic sets, or circular fuzzy sets, and compare the results with other integrated IVPF MCDM models.

References

1. Bolturk, E., Kahraman, C.: A modified interval-valued Pythagorean fuzzy CODAS method and evaluation of AS/RS technologies. J. Multiple-Valued Logic Soft Comput. **33**, 415–429 (2019)
2. Deschrijver, G., Kerre, E.E.: Triangular norms and related operators in L*-fuzzy set theory. Logical, Algebraic, Analytic and Probabilistic Aspects of Triangular Norms, pp. 231–259 (2005)
3. Ilbahar, E., Karaşan, A., Cebi, S., Kahraman, C.: A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system. Saf. Sci. **103**, 124–136 (2018)
4. Karasan, A., Ilbahar, E., Kahraman, C.: A novel Pythagorean fuzzy AHP and its application to landfill site selection problem. Soft. Comput. **23**(21), 10953–10968 (2019)
5. Keshavarz Ghorabaee, M., Kazimieras Zavadskas, E., Turskis, Z., Antucheviciene, J.: A new combinative distance-based assessment (CODAS) method for multi-criteria decision making. Econom. Comput. Econom. Cybernet. Stud. Res. **50**(3), 25–44 (2016)
6. Peng, X., Yang, Y.: Fundamental properties of interval-valued Pythagorean fuzzy aggregation operators. Int. J. Intell. Syst. **31**(5), 444–487 (2016)
7. Pérez-Domínguez, L., Durán, S.N.A., López, R.R., Perez-Olguín, I.J.C., Luviano-Cruz, D., Gómez, J.A.H.: Assessment urban transport service and Pythagorean fuzzy sets CODAS method: a case of study of Ciudad Juárez. Sustainability **13**(3), 1281 (2021)
8. Rahman, K., Abdullah, S., Shakeel, M., Khan, M.S.A., Ullah, M.: Interval-valued Pythagorean fuzzy geometric aggregation operators and their application to group decision making problem. Cogent Mathematics **4**(1), 1338638 (2017)
9. Yager, R.R.: Generalized orthopair fuzzy sets. IEEE Trans. Fuzzy Syst. **99**, 1–11 (2016)
10. Yager, R.R.: Pythagorean fuzzy subsets. In: 2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), Alberta-Canada, 24–28 June 2013
11. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
12. Zhang, X.: Pythagorean fuzzy clustering analysis: a hierarchical clustering algorithm with the ratio index-based ranking methods. Int. J. Intell. Syst. **33**(9), 1798–1822 (2017)
13. Pamucar, D., Torkayesh, A.E., Biswas, S.: Supplier selection in healthcare supply chain management during the COVID-19 pandemic: a novel fuzzy rough decision-making approach. Ann. Oper. Res. (2022). https://doi.org/10.1007/s10479-022-04529-2
14. Stević, Ž, Pamučar, D., Puška, A., Chatterjee, P.: Sustainable supplier selection in Healthcare Industries using a new MCDM method: measurement of alternatives and ranking according to compromise solution (Marcos). Comput. Ind. Eng. **140**, 106231 (2020)

Machine Learning

Socio-Economic Development Index Estimation for Small Area Urban Neighbourhoods: A Machine Learning Approach

Ulaş Yelkanat(✉) and Ömer Zeybek

Migros Ticaret AŞ R&D Center, Ataşehir, İstanbul, Turkey
ulasy@migros.com.tr

Abstract. Since ancient times, people have lived in regions that improve their quality of life and provide various opportunities to increase their wealth. As a result, small communities in urban areas, called neighborhoods, have unique characteristics, advantages, and disadvantages. Therefore, retailers targeting a high ROI (Rate of Return) for their investments should place their new openings in line with the fundamental characteristics of the neighborhoods. To achieve this, it is necessary to analyze a neighborhood's potential depending on the level of social development or the commercial and public assets located in the area. This study aims to score 32,130 urban neighborhoods, called "mahalle," in the Turkish Local Administrative Units hierarchy using machine learning approaches. We estimated scores using various features related to demographics and points of interest. First, we used official statistics to train our model for the city level. The Turkish Governmental Organizations have been calculating social and economic ratings on the city level since the late 1960 s. Then, we summarized our city-level data and built a regression tree model. Finally, we used this model to score our neighborhood-level data. Using the LightGBM method, we achieved a high r2 score for the test dataset, and then we scored all districts using the model to obtain our scores. When we compared our findings with city-level ratings, we found a strong 81 percent correlation between the two indexes.

Keywords: Machine Learning · LightGBM · Neighborhood Scoring · Spatial Analytics

1 Introduction

The estimation of a Socio-Economic Development Index (SEDI) for small urban areas has been a significant topic in the field of urban planning and development for several decades. A typical development index for small urban areas can be defined as a composite index that quantifies the socio-economic well-being of neighborhoods. These indexes include various dimensions of socio-economic well-being, such as income, education, health, and employment data, and provide a useful tool for policymakers and urban planners to identify areas that require

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 135–141, 2023.
https://doi.org/10.1007/978-3-031-39777-6_16

specific interventions to improve the socio-economic conditions of their residents. However, according to [7], estimating a SEDI at the urban neighborhood level is challenging due to the complex and heterogeneous nature of urban environments, which often exhibit significant variations in socio-economic characteristics across small geographic areas.

Traditional methods of SEDI estimation, such as surveys and censuses, are limited in their ability to provide accurate and timely data at the neighborhood level. Moreover, application of these methods can be expensive and time-consuming, making them impractical for frequent updates. In their studies [5] pointed that, there has been growing interest in using machine learning (ML) approaches to estimate SEDI in urban neighborhoods. ML algorithms can analyze large and diverse datasets to identify patterns and relationships between different socio-economic indicators and SEDI, enabling more accurate and efficient estimation.

The objective of this study is to estimate the Socio-Economic Development Index (SEDI) for 32,130 small area urban neighborhoods in Turkey, utilizing demographic and Point of Interest (POI) data as features. Our approach first involved training a supervised learning model which uses scores estimated in Research on the Socio-Economic Development Ranking of Districts scores [1]. This research estimates a town level development index using various dimensions. Using government data at the district level, we obtained data at a lower level, which is the innovative aspect of our project. To built a sub-town, neighbourhood level development index we summarised our small areas neighbourhood level data into town level and trained a town level regression model using town-level score estimates of SEGE as dependent variable. As Turkish Government uses these predictions to decide regional development projects in sub-city level since 1960 s s this index is a robust measurement for development stage of urban areas.

We trained a LightGBM regression model using features mentioned above and 973 district scores are obtained from the SEGE study and obtained with a correlation of 81%. Then we implemented small are neighbourhood level data for 32,130 areas to this model and estimate scores for them.

The rest of the study is organised as follows: Sect. 2 offers a brief literature review on development of neighborhoods with machine learning. Section 3 presents the conceptual model used in the study. Section 4 summarises the methodology of the study. Finally, Sect. 5 concludes the paper with comments on the output of the model and shares directions for future research.

2 Literature Review

Socio-Economic Development Index (SEDI) is a composite measure of various indicators of social, economic, and environmental development, which can be used to assess the level of development of a particular region or community. In recent years, there has been growing interest in using machine learning (ML) approaches to estimate SEDI in urban neighborhoods, which can provide more accurate and efficient assessments of neighborhood-level socio-economic status for various purposes such as urban planning, policy making, and resource allocation.

The concept and status of poverty vary depending on regions, and its value can signify a lot about a region. The aim of their study [4] was to identify urban poverty in China using image features derived from high-resolution satellite imagery and machine learning approaches, providing empirical evidence of their applicability. Four different machine learning approaches were used in the study, and the best model performance was achieved with a 53% R^2 score. The results of the study revealed poverty levels in regions parallel to their level of development.

Another study [6] aims to go beyond traditional population counting methods and predict socio-economic and demographic characteristics such as age and income in addition to population. Estimates were obtained by combining different machine learning methods and POI and real estate data in administrative areas in Singapore. As a result of this consolidation, it was possible to access the levels of development and various characteristics of the regions from the data obtained.

These studies demonstrate the potential of machine learning approaches for estimating SEDI in neighborhoods. ML approaches can also identify important features associated with SEDI, which can provide insights into the factors that contribute to socio-economic development in urban areas.

3 Conceptual Model

This study examines the relationship between features and socio-economic development in neighborhoods. The data collection process involved obtaining information from various external databases. Once the POI and demographic data were collected, various unsupervised learning methods, such as principal component analysis (PCA), were applied to the data. However, the results of these methods did not accurately reflect the actual indices for many neighborhoods. Additionally, reducing the number of features to minimize the effect of cross-correlations did not improve the results of the analysis. After analyzing the SEGE survey, we decided to use the district-level scores for our study. By aggregating the neighborhood data we had, we transformed it into district (ilçe) data. Subsequently, SEGE scores were added as labels for 973 districts. A Light GBM model employing SEGE scores as a dependent variable showed high performance with more than 98% R^2 scores for both train and test sets.

In the second stage of the study, we implemented our neighborhood level data into an estimated district (ilçe) model and aimed to estimate small urban area (neighborhood level) scores.

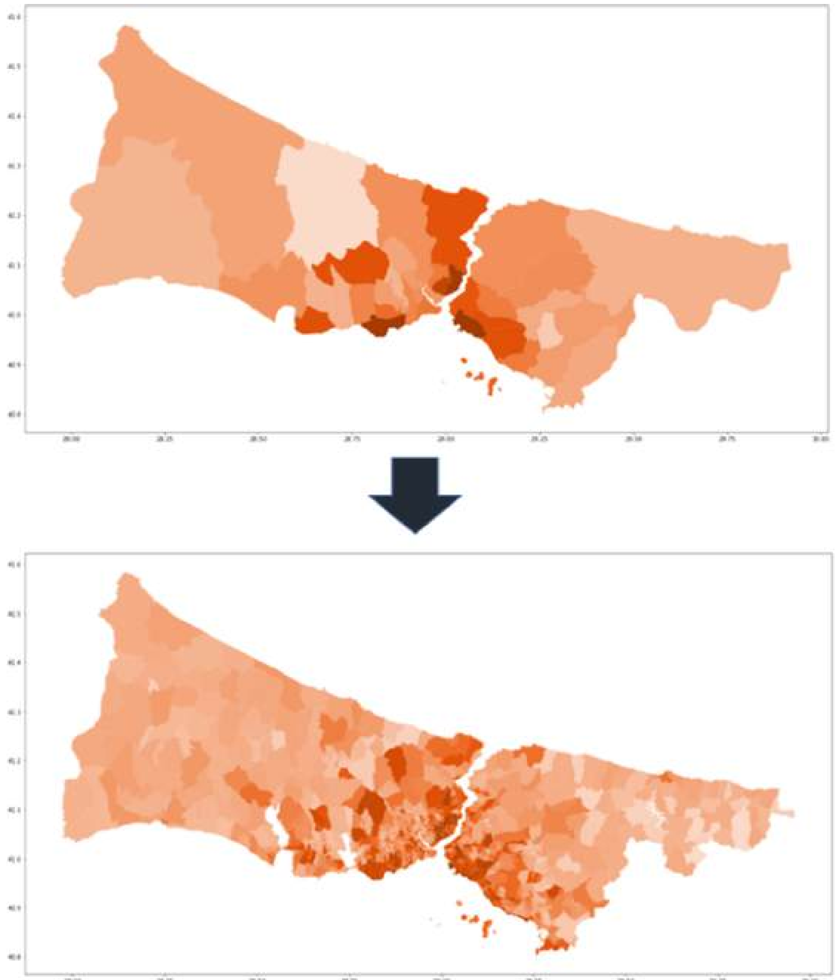


Fig. 1. Comparison between district level average (Top Panel) and small area neighborhood level scores (Bottom Panel)

Our conceptualization could be more clearly presented in Fig. 1. The figure represents two maps of Istanbul, the most crowded city in Turkey, in two dimensions. The first shows district (ilçe) scores, while the second shows small area neighborhood level scores estimated for two different data levels.

4 Methodology

We conceptualized our response model for measuring the socio-economic development index using the Gradient Boosting model family. Gradient boosting [2] is an ensemble method that involves the creation of numerous weak models that are combined to improve overall performance. This study employed the LightGBM method, which is a type of Gradient Boosting model. The LightGBM models, conceptualized by [3], use both regression and classification machine learning tasks. Our project used LightGBM regression. The selected algorithm uses a histogram-based approach where it groups continuous feature values into distinct bins, thereby accelerating the training process. Additionally, the LightGBM algorithm achieves high accuracy by utilizing a leaf-wise split strategy to construct highly complex trees, as opposed to a level-wise approach.

The LightGBM regression uses three decision tree-based boosting types: Gradient Boosting Decision Tree, Multiple Additive Regression Trees, and Random Forest. We employed Gradient Boosting Decision Tree, which is the default parameter of algorithm used in our project. Decision trees are commonly used in machine learning and require less effort in the data preparation stage. They do not require normalization and scaling operations.

First, we aggregated by averaging 124 different demographic and POI features at the district level, which are contained at the neighborhood data level in the database. We then divided the dataset into training and test sets to obtain a measure of accuracy that we confirmed using a 10-fold cross-validation test to ensure consistency of the results. Subsequently, we prepared the dataset to create a model and trained a LightGBM model using SEGE scores estimated by the Turkish Ministry of Industry and Technology as labels.

Table 1. 5 Neighborhoods Highest Score

City	District	Neighborhood	Score
Ankara	Yenimahalle	Beştepe	5332
İstanbul	Şişli	Esentepe	5266
İstanbul	Şişli	Meşrutiyet	5264
İzmir	Bornova	Kazımdirik	5252
Ankara	Çankaya	Üniversiteler	5241

We verified the model's performance by calculating R^2 scores for the training and testing sets separately. After observing successful R^2 scores at the district level, we implemented neighborhood-level observations to the same model. Finally, we estimated the socio-economic development index for 30K neighborhoods scattered in 81 different city centers in Turkey.

The estimated model points out that variables related to the education level, such as the percentage of people with a PhD degree and non-literates, are the two most important features contained in the model. Moreover, the percentage

Table 2. 5 Neighborhoods Lowest Score

City	District	Neighborhood	Score
Şanlıurfa	Bozova	Yaslıca	–1603
Şanlıurfa	Ceylanpınar	Muratlı	–1546
Van	Tuşba	Şemşibey	–1532
Mardin	Derik	Ambarlı	–1531
Diyarbakır	Hazro	İncekavak	–1529

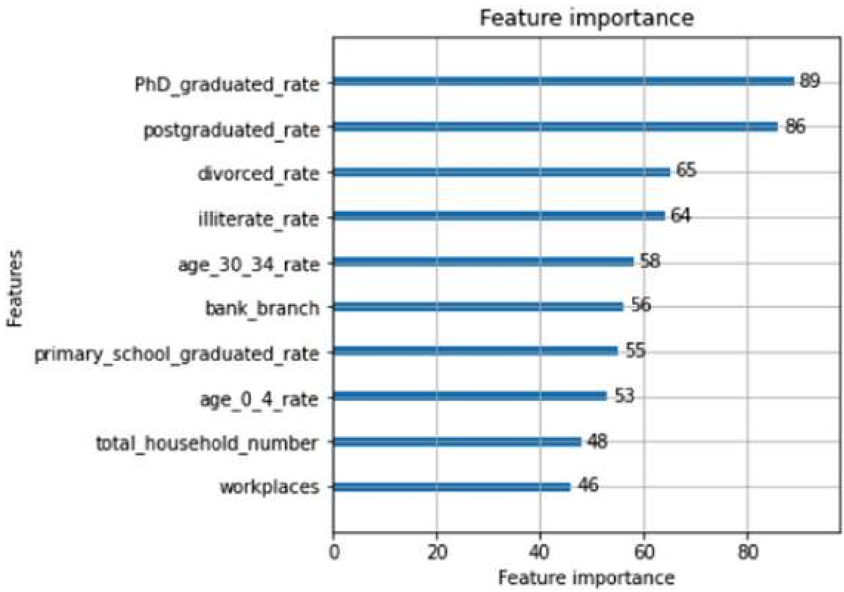


Fig. 2. Feature Importance for Neighborhood Development Index in Turkey

of people living in the observed area between 30–34 years old is also an influencer variable in the model.

We conducted various checks to obtain scores for neighborhoods and verify that the results were both explicable and consistent. While the correlation between district-level averages of neighborhood-level scores and the Ministry's district-level index SEGE is at the 81% level, and the leader and lowest ranking neighborhoods according to the proposed scaling presented in Table 1 and Table 2 is consistent with expectations, we can say that our index is a proper way to describe small area well-being in Turkey. Table 1 shows the top 5 neighborhoods which have highest scores, the neighborhood that has 5332 score called Beştepe is the most developed neighborhood in Turkey according to our study. In Table 2 there are 5 neighborhoods which have lowest scores, the neighborhood called Yaşlıca has the lowest score -1603 according to our study.

5 Conclusion

Our reaseach objective was to score small urban neighborhoods in Turkey according to their level of socio-economic development. Therefore, a large number of effective features are needed to accurately score a large number of neighborhoods. Obtaining these features and their values is crucial during the data preparation stage, which is one of the most important parts of the project. We were able to create a meaningful model using only demographic and POI data. Utilizing SEGE scores during the model creation process further increased the model's accuracy. Finally, the development scores we produced were only prepared for urban areas and were limited to Turkey. To be able to apply the study to different regions or units and generalize it, different approaches may need to be taken.

References

1. Acar, S., Şahin Cinoglu, D., Karagöz, T.: İlçelerin sosyo ekonomik gelişmişlik araştırması [socioeconomic development study of districts]. Technical report, Turkish Ministry of Industry and Technology (2022)
2. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016)
3. Ke, G., et al.: Lightgbm: a highly efficient gradient boosting decision tree. Adv. Neural. Inf. Process. Syst. **30**, 3146–3154 (2017)
4. Li, G., Cai, Z., Qian, Y., Chen, F.: Identifying urban poverty using high-resolution satellite imagery and machine learning approaches: implications for housing inequality. Land **10**(6), 648 (2021)
5. Liu, X., Wang, Y., Li, Y., Wu, J.: Quantifying the spatio-temporal process of township urbanization: a large-scale data-driven approach. ISPRS Int. J. Geo-Infor. **8**(9), 389 (2019)
6. Szarka, N., Biljecki, F.: Population estimation beyond counts-inferring demographic characteristics. PLoS ONE **17**(4), e0266484 (2022). https://doi.org/10.1371/journal.pone.0266484
7. Yakubu, B.I., Hassan, S.M., Asiribo, S.O.: An assessment of spatial variation of land surface characteristics of minna, niger state nigeria for sustainable urbanization using geospatial techniques. Geosfera Indonesia (2018)

Classification of Land Cover Based on Deep Learning Technique and High-Resolution UAV Images

Ching-Lung Fan(✉)

Department of Civil Engineering, The Republic of China Military Academy, No. 1, Weiwu Road, Fengshan, Kaohsiung 830, Taiwan
p93228001@ntu.edu.tw

Abstract. The aim of this study was to investigate the use of Unmanned Aerial Vehicles (UAVs) in combination with convolutional neural networks (CNNs) to classify land cover types. A high-performing classification model was developed by optimizing the CNN architecture and hyperparameters, and training the model with a large dataset. The model's performance was evaluated using four different land cover (LC) classes, with results showing that the accuracy of the model on the training dataset was above 97%. Model F achieved the highest accuracy of 89.4% on the test dataset and had the lowest level of overfitting. The study demonstrates that an effective classification model can be achieved by fine-tuning parameters and is suitable for LC mapping.

Keywords: Convolutional neural network · Land cover · UAV images

1 Introduction

Deep learning is a technique that utilizes mathematical models to extract high-level features from data, leading to improved accuracy in classification and detection tasks. Various neural network models have been proposed; among them, the convolutional neural network (CNN) is the most common. Recently, CNNs have made considerable progress in image recognition and are widely used in land cover (LC) classification with high accuracy [1, 2]. The object-oriented approach is lengthy for image segmentation and image object feature calculation. However, the CNN approach does not require these two processes, thus improving the efficiency of LC classification [3]. Compared with the LC classification method based on pixels and objects, a CNN can be used to extract high-level abstract features from images and yield excellent results in classifying LC.

When applied to high-resolution images, CNNs can overcome the limitations of traditional pixel-based and object-based classifications, making them well-suited for LC classification using high-resolution images captured by Unmanned Aerial Vehicles (UAVs). The approach detailed in this paper focuses on using high-resolution UAV images for LC classification by adjusting the architecture and hyperparameters of the CNN and training and testing the model using a large dataset to construct the most optimized classification model.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 142–148, 2023.
https://doi.org/10.1007/978-3-031-39777-6_17

2 Research Method

2.1 Convolutional Neural Network

Equation (1) displays a very complex formula. A CNN is a deep learning technique that learns features hierarchically from training data. It can adjust and share weights for training data, generalize and optimize parameters, and automatically reduce parameters with high feature differentiation and extraction capabilities [4]. A CNN is made up of one or more layers, such as convolutional, pooling, or fully connected layers, and can classify images or object classes through feature learning. CNNs have an advantage over traditional artificial neural networks (ANNs) in that they add two more convolutional and pooling layers, allowing them to perceive image details rather than just extracting data for computation. Image data typically contains three-dimensional information, such as horizontal, vertical, and color channels, but ANN input must be one-dimensional data, which can result in the loss of important spatial information. The main architecture of a CNN is a series of alternating convolutional and pooling layers that are connected and generalize features into depth and abstract representations. A fully connected layer classifies the final labels.

$$Yi = Wi * X + b \tag{1}$$

2.2 Hyperparameters of CNN

High-resolution images are complex, and therefore, a network with more layers is often used for training. However, as the network deepens, training the model becomes more challenging, and convergence slows. To overcome this, the use of stochastic gradient descent is an effective optimization method when training deep networks [5]. While stochastic gradient descent is simple and efficient, it requires manual adjustments of hyperparameters such as the learning rate, number of epochs, number of hidden layers, activation function choices, and batch size. During the training process, the results from the validation set, such as the validation loss and accuracy, can be monitored to determine if any hyperparameters need to be adjusted [6]. The optimal parameters are determined by validating a set of multiple training networks and evaluating their accuracy. The adjustments of the parameters are grouped into hyperparameters to determine the best structure of the model [7].

Another technique used to improve the training of deep networks is batch normalization (BN). BN standardizes the input data in each layer of the network to a normal distribution (mean of 0 and standard deviation of 1) which reduces the variability of each sample in the batch. This can accelerate deep network training and reduce internal covariate shifting [8]. Additionally, BN makes the model less sensitive to the network parameters, stabilizes network learning by simplifying the process of adjusting parameters, and prevents vanishing gradients during the training process. Other important techniques for training deep networks are dropout and data augmentation. These methods can help improve the training results. Additionally, during the training process, various methods of initializing the parameters can be used to speed up model convergence.

2.3 UAV Images

The research used a DJI Phantom-3 (FC300C focal length: 4 mm) as the UAV, which was equipped with GPS and inertial measurement unit and has automatic navigation and autopilot functions. The UAV was flown at an altitude of 60 m to capture high-resolution images using three channels of RGB and an image resolution of 4,000 pixels x 3,000 pixels. The types of LC in this area include trees and grass (natural LC) and buildings and roads (manmade LC). To train the CNN model, a large amount of image data was needed, so the UAV images were divided into four LC classes: trees, grasslands, buildings, and roads. A total of 800 images were manually labeled and used as the dataset for the four LC classes (Fig. 1).

The 800 images were divided into three sets: training set (400 images, 50%), validation set (200 images, 25%), and test set (200 images, 25%). The training set was used primarily for model training, the validation set for adjusting parameters, and the test set for evaluating model performance. The CNNs were used for image detection, recognition, and classification. A deep network architecture was constructed to match the image dataset, and hyperparameters such as the number and size of network layers, filters, and activation functions were adjusted to improve the performance of the classification model. The main focus of this research was to use CNNs to construct an LC classification model and train and test it on the four types of LCs. Through training the network, the performance of the model was improved.

Building Road Grassland Tree

Fig. 1. UAV images of four types of land cover.

3 Results

This research proposed an architecture for six CNN-based network models (Table 1) and conducted tests to detect four types of LC. All CNN models were designed using convolutional, pooling, and fully connected layers (flat, hidden, and output layers, respectively). The ReLU activation function was applied to the output of all convolutional and fully connected layers, and the softmax function was used in the output layer. First, the architecture of Model A was one convolutional layer and one pooling layer, the number of filters was 32, the kernel size was 5 x 5, and the number of neurons in the two hidden layers were 500 and 1,000 respectively. CNNs are known to have more powerful feature representation capabilities than shallow learning, which can greatly improve the performance of target detection [9]. By doubling the number of filters in each convolutional layer to increase the number of feature maps in the hidden layer, more relevant information can be extracted from the input data [10]. Therefore, Model B was composed of three convolutional layers, three pooling layers, and three fully connected layers, with the number of filters being 32, 64, and 128 respectively and other parameters were the

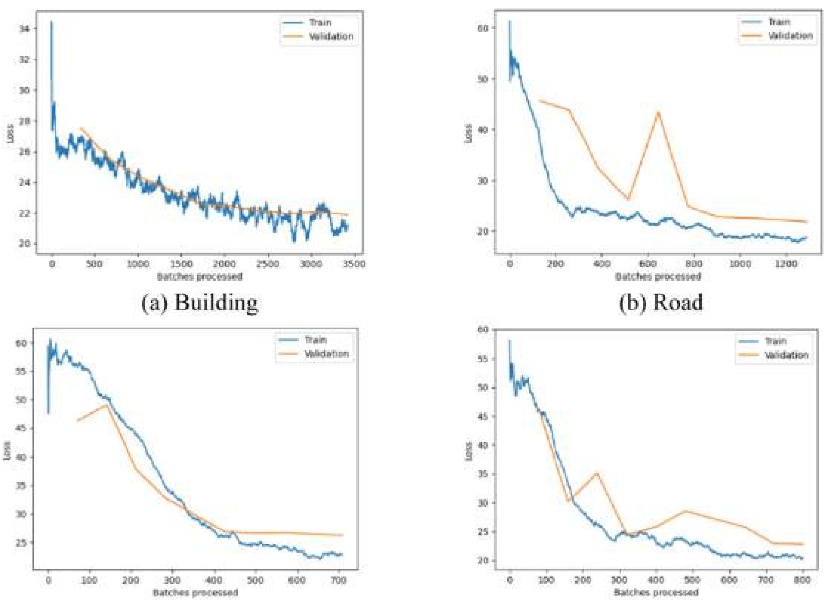
Table 1. Architectures of six CNN models

Model	Layer	Filter	Kernel size	Dropout	Batch normalization	Data augmentation
Model A	Conv1, Maxpool	32	5 × 5	-	-	-
Model B	Conv1, Maxpool Conv2, Maxpool Conv3, Maxpool	32 64 128	5 × 5	-	-	-
Model C	Conv1, Maxpool Conv2, Maxpool Conv3, Maxpool	32 64 128	3 × 3	-	-	-
Model D	Conv1, Maxpool Conv2, Maxpool Conv3, Maxpool	32 64 128	3 × 3	0.5	-	-
Model E	Conv1, Maxpool Conv2, Maxpool Conv3, Maxpool	32 64 128	3 × 3	0.5	BN	-
Model F	Conv1, Maxpool Conv2, Maxpool Conv3, Maxpool	32 64 128	3 × 3	0.5	BN	rescale=1/255, rotation=40 w-shift=0.2 h-shift=0.2 shear=0.2 zoom =0.2

same as Model A. In Model C, the size of all filters was changed from 5 x 5 to 3 x 3 based on the original architecture of Model B.

Overfitting occurs when a learning network performs well on the training set but performs poorly on the validation and test sets. To reduce overfitting in the fully connected layer, the dropout method can be employed. This method reduces the complexity of neural network connections, speeds up model training and improves the generalizability of neural networks. In Model D, a dropout of 0.5 was added to the original architecture of Model C. Model E used batch normalization (BN) to reduce the number of convergences and improve the final performance. Given that the training data was limited, Model F was used to create more training images through data augmentation, which expanded the image data without adding new images. Data augmentation was performed by manually increasing the training dataset which often reduces overfitting.

The accuracy of the six CNN models was over 90%, however, the accuracy of both the validation set and the test set remained under 90%, indicating that the models were overfitting. However, Model F showed the lowest overfitting among all the models and the highest test accuracy of 89.4%. Model F had good detection performance for the four types of LCs. According to the learning process of the model F in Fig. 2, the final loss values of the four types of LCs for training and validation were similar. The curve distribution of training and validation in the learning process of the model tends to be consistent, but the validation curve of the road type fluctuated greatly. This may be due to the fact that roads close to buildings may often be obscured by shadows, which is likely to cause errors. Additionally, the building type had more batches than the other three



(a) Building (b) Road

(c) Grassland (d) Tree

Fig. 2. Loss of Model F for four classes of LC.

categories, which may be due to variations in the shape and size of building samples. Overall, by adjusting the parameters of a CNN, the best classification model can be constructed and can address the requirements of mapping.

4 Conclusion

Adjusting parameters is a critical step in image classification for reducing model overfitting and improving the accuracy of classification. In this study, images were input into the CNN model, and the parameters of the convolutional, pooling, and fully connected layers were adjusted to identify the features of LC. The performance of the model was verified using UAV images. The model was required to have accurate results on the training set, and the test set was required to be generalizable. The results indicated that Model F had the best detection performance for four LC types with an accuracy of 89.4% and a loss of 0.0714. The accuracy of the other five models was also higher than 80.0% and the loss was less than 0.09, however, overfitting was more prevalent in them than in Model F.

This study used UAV images and deep learning methods to construct four LC classification models. High-resolution images and automatic and fast CNN classification methods were used to detect object features in images through extensive training data. Object features effectively solve the problem of classifying high-resolution images. The optimized detection model is valuable because the precise classification results (e.g., area, number, type, location, and size) can be applied in both future research and in areas such as land resource management and planning.

Acknowledgments. The author would like to express their gratitude to Associate Professor Yu-Jen Chung of the Department of Marine Science at the ROC Naval Academy for providing a GPU computer that facilitated the fast execution of image detection tasks in this study.

References

1. Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A.: Deep learning classification of land cover and crop types using remote sensing data. IEEE Geosci. Remote Sens. Lett. **14**(5), 778–782 (2017)
2. Pan, X., Zhao, J.: A central-point-enhanced convolutional neural network for high-resolution remote-sensing image classification. Int. J. Remote Sens. **38**(23), 6554–6581 (2017)
3. Yang, B., et al.: Extraction of road blockage information for the Jiuzhaigou earthquake based on a convolution neural network and very-high-resolution satellite images. Earth Sci. Inf. **13**(1), 115–127 (2020)
4. Wang, Y., Wang, Z.: A survey of recent work on fine-grained image classification techniques. J. Vis. Commun. Image Represent. **59**, 210–214 (2019)
5. LeCun, Y., Bengio, Y.: Hinton G. Deep learning. Nature **521**(7553), 436–444 (2015)
6. Al-Najjar, H.A., et al.: Land cover classification from fused DSM and UAV images using convolutional neural networks. Remote Sensing **11**(12), 1461 (2019)
7. Castagno, J., Atkins, E.: Roof shape classification from LiDAR and satellite image data fusion using supervised learning. Sensors **18**(11), 3960 (2018)

8. Bai, Y., Mas, E., Koshimura, S.: Towards operational satellite-based damage-mapping using u-net convolutional network: a case study of 2011 Tohoku earthquake-tsunami. Remote Sensing **10**(10), 1626 (2018)
9. Cheng, G., Zhou, P., Han, J.: Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE Trans. Geosci. Remote Sens. **54**(12), 7405–7415 (2016)
10. Sharma, A., Liu, X., Yang, X., et al.: A patch-based convolutional neural network for remote sensing image classification. Neural Netw. **95**, 19–28 (2017)

Store Segmentation Using Machine Learning Methods: An Organized Supermarket Chain Case

İskender Ülgen Oğul(✉) and Ömer Zeybek

R&D Department, Migros Ticaret A.Ş, Istanbul, Turkey
iskendero@migros.com.tr

Abstract. Over the last few decades, developing economies have led to a similar increase in per capita income. Rising per capita incomes have allowed consumers to demand a wide range of fast-moving products. Despite the ups and downs of the economy, the expectations that customers have of supermarket chains have not changed. Thus, retail chains have a duty of care towards customer satisfaction In the quest to maintain a consistent level of customer satisfaction and to achieve targeted growth in the customer base, the retail industry needs to follow a well-planned store management process. The need to segment retail stores has therefore become paramount for retail chains. Our company manages six different formats and more than 2700 outlets in all provinces of Turkey. Efficient management of this diversified portfolio requires a comprehensive decision support system. The aim of our study was to segment the stores of a national grocery chain using a machine learning approach. For this purpose, not only the store and store-related customer data was used but also small area demographic indicators, socio-economic and point-of-interest locations around stores was included. After checking the plausibility of the data and enrichment process, we used a k-means clustering algorithm to find the best groupings of the stores. The algorithm suggested 12 different segments for the stores. All segments are profiled and assessed by the sales teams of the company. Finally, to dynamically segment the newly opened stores, we trained a classification model based on the results of the clustering algorithm. The decision support system that was created during the study proves to be a complete segmentation system for retail grocery stores sector.

Keywords: Segmentation · Classification · Clustering · Retail

1 Introduction

The retail industry is characterized by fierce competition and rapidly changing customer preferences. As a result, companies must be in a constant state of innovation and adaptation to new market trends. In this context, store segmentation has emerged as a valuable approach for retail companies seeking to

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 149–156, 2023.
https://doi.org/10.1007/978-3-031-39777-6_18

identify and target specific market segments. Thus, it allows for more effective marketing campaigns, better allocation of resources, efficient process optimization, and improved overall performance. Organisations can gain insights into customer behavior, regional dynamics, competition effect, and store management practices by grouping stores based on common characteristics or performance metrics. This research focuses on store segmentation for a grocery store retail company. It aims to improve the retailer's understanding of store performance, customer behavior, a location analytics and strategic decision making, and contribute to the wider literature on retail management and market segmentation.

Previous store segmentation studies have employed various clustering methods to identify homogeneous groups within the retail landscape. These methods range from hierarchical to partitional algorithms. Each has its strengths and weaknesses.

The rest of the study is organised as follows; Sect. 2 offers a brief literature review, Sect. 3 presents methodology, Sect. 4 explains data preparation and process, Sect. 5 addresses conceptual model, Sect. 6 discusses empirical findings.

2 Literature Review

Different solutions have been employed over various periods to study store segmentation. While small and regionally balanced retailers have adopted the survey based data collection method. Multi-location and multi-regional retailing companies are required to implement advanced data collection techniques.

Gindi et al. focused on segmenting fresh product consumers based on product and store attributes. It used a multi-stage cluster sampling technique and structured questionnaires. Data was collected from 598 respondents from six regions in the Klang Valley. First, factors were identified using exploratory factor analysis and principal component analysis with varimax [7] rotation. Later, cluster analysis was used to group consumers with similar characteristics, resulting in three distinct segments [6].

Another study proposed an improved retail sales forecasting method for convenience stores to improve sales prediction accuracy by leveraging clustering techniques. The methodology calculates the similarity of stores based on the proportion of customer types and extracts the demographic characteristics of customers from the transaction data. Due to insufficient spatial information, demographic factors are studied from transaction data. In the final step, clustering algorithms, k-means, and expectation maximization are employed to segment stores by their similarity [11].

Several other studies have focused on the comparative analysis of clustering in the retail sector using different types of data and clustering methods. In Agarwal's research, four clustering techniques are evaluated: Hierarchical Clustering, Self-Organising Maps, Gaussian Mixture Matrix, and Fuzzy C-Means. Before clustering, data from transactional, product, store, loyalty, and competitive sources is collected and processed. Across all store segmentation solutions, the results show that all techniques are equal. Therefore, it is concluded that several different clustering techniques should be applied, and the most appropriate method should be selected based on the data and characteristics [1].

Many studies have based their data collection processes on survey structure and applied statistical methods, while other studies used their transaction data for feature construction processes. The importance of using demographic and location-based analytics was also highlighted. Comprehensive studies have focused on comparative analyses where different solutions are combined and examined. Our work employs different strategies discussed in the research to achieve robust clustering results for FMCG retail stores.

3 Methodology

In this research, we first explored different clustering techniques during the early-stage testing phase to determine their suitability for the task. Then, we consider factors such as computational efficiency, ease of interpretation, and the ability to generate meaningful and actionable segments. Ultimately, due to its simplicity, efficiency, and effectiveness in generating meaningful and interpretable segments that can inform the company's strategic decisions, the k-means algorithm is selected for the final clustering.

In order to ensure the quality and relevance of the input variables used for store segmentation, our research incorporates various data pre-processing techniques. Principal Component Analysis (PCA) transforms the original variables into a smaller set of uncorrelated components to reduce the dataset's dimension while preserving the most critical information. Not only does this simplify the clustering process by reducing the computational complexity, but it also helps interpret the results by providing a more straightforward representation of the underlying structure in the data. Our work also used Factor Analysis to identify underlying dimensions or factors in the dataset, which helps to understand the relationships between variables and how they contribute to the segmentation process. By reducing the observed variables to a smaller set of latent factors, these techniques allow us to capture the essential patterns in the data, enabling more refined and nuanced segmentation. A correlation analysis is also carried out to examine the linear relationships between the variables, which allows us to eliminate redundant or highly correlated variables from the analysis and ensure the robustness of the segmentation model.

The selection of the optimal number of clusters is crucial in clustering; the elbow method and the silhouette score are widely used for this purpose. The elbow technique plots the sum of the squared intra-cluster distances against the number of clusters, k, and identifies the point at which the rate is most suitable for final clustering. The Silhouette Score measures the clustering quality based on the cohesion within clusters and the separation between clusters, with higher values indicating better quality. The optimal k is chosen based on the highest average Silhouette Score. Combining both methods allows for an informed decision on the optimal number of clusters in our store segmentation model. In order to improve the sustainability of our clustering solution, our work has adopted the clustering technique for the segmentation of new stores.

Through the combination of these advanced analytical techniques, this research aims to provide a comprehensive and insightful store segmentation model for retail store segmentation. The results are expected to improve store performance and customer satisfaction through better decision-making, resource allocation, and marketing strategies. The following sections provide a detailed description of the research questions, methodology, results, and the discussion to provide a complete overview of the store segmentation process.

4 Data Preparation and Pre-processing

Our research consists of several steps: data collection, enhancement, pre-processing, scaling, and dimensional reduction-the quality of the clustering results to is directly affected by all these preliminary stages.

The data collection process consists of three separate stages. The first stage focuses on the generation of store-centered features using transaction sources, this also includes dark stores and online channel sales information. The second stage focuses on the store and the customer-related characteristics, such as the customer lifestyle density for each store. Different from the previous stages, the third stage consists of external data management and spatial information association with stores. All of the data that is collected at these stages is associated with the stores on an individual basis.

The pre-processing stage consists of filling missing values, correlation analysis, dimensional reduction and feature scaling factors. Missing value filling method carried out either by taking the general observation average or, for spatial data, by taking the group average based on location code. Before proceeding dimensional reduction, our research inspected the relationship between the features. Finally, the statistical technique named correlation analysis was used to examine the significant association between variables and the direction of the relationship. As a result highly correlated variables are pruned to achieve accurate clustering results.

Dimensional reduction is an essential step forward. It brings out distinct and robust features to reduce the analysis space while maintaining the quality of the representations. In order to obtain more compact and more densely representative features, principal component analysis is applied. Principal Component Analysis (PCA) is a dimensional reduction technique that reduces the complexity of high-dimensional data while retaining as much variation as possible, making it easier to analyze and visualize [3]. The transformation $T = XW$ maps a data vector $x_{(i)}$ from an original space of p variables to an uncorrelated space of p variables, only keeping the L principal components. Thus gives the $T_L = XW_L$ where the matrix T_L has n rows but only L columns. Formulated as follows, PCA learns a linear transformation.

$$t = W_L^T x, x \in R^p, t \in R^L \tag{1}$$

which the columns of $p \times L$ matrix W_L forms an orthogonal basis for de-correlated features. This score matrix maximises the variance in the original data while minimising the total squared reconstruction error $\|X - X_L\|_2^2$. Our work implements scikit-learn PCA method which incorporates SVD solver based on data characteristics [5].

Through the data preparation steps, we created 245 different features, which include store, customer, POI, and demographic data. After pre-processing steps as described above, our study obtained 176 features with high representation and discriminating power.

5 Conceptual Model

Once the data stages are completed, the clustering stage can be studied to produce the segmentation solution. Getting the optimal number of clusters in an explainable way is the most important problem in clustering studies. The elbow method and silhouette scores are the two most commonly used methods. Before finding the optimal clustering number, all features need to be represented on the same scale, so our work explored different types of scaling factors and discovered that standard scaling is best suited to our problem.

An elbow method is a heuristic approach in which the explained variance is plotted against the number of clusters until an inflection point is reached, which resembles an 'elbow' in the curve. Regardless, the elbow method is criticised for its highly ambiguous behaviour, as the plot does not contain a sharp elbow [9].

Considering this criticism about the elbow, our work also implements silhouette scores to have robust and explainable number of clusters. The silhouette method assesses the clustering quality by evaluating how well each data point fits into the cluster to which it is assigned and how well it differs from the other clusters. It calculates the silhouette coefficient for each data point, ranging from -1 to 1 [10]. High silhouette coefficients indicate good clustering, while low ones indicates the otherwise. As a result, the elbow method indicated 11 clusters, while the silhouette method indicated 12 clusters. Considering the previous discussion, our work decided that 12 clusters would be more suitable for our segmentation strategy.

Clustering approaches are generally based on the k-means algorithm. K-means is an unsupervised machine learning algorithm that divides data into clusters on the basis of similarity [2]. The algorithm initialises K cluster centroids and refines them by minimising the sum of the squares of the distances between the data points and their corresponding centroids within the cluster. K-means is a valuable algorithm for exploratory data analysis because it is computationally efficient, easy to implement, and scales to large datasets.

$$\arg\min S \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 = \arg\min S \sum_{i=1}^{k} |S_i| \mathrm{Var} S_i \tag{2}$$

For a set of observations $(x_1, x_2, ..., x_n)$, where x_i is a d-dimensional real vector, algorithm partitions the n observations into $k(\leq n)$ sets $S = (S_1, S_2, ..., S_k)$ by minimizing the intra-cluster sum of squares.

Setting the iteration size to 1000, our clustering result yielded an average silhouette score of 0.23. In this way, our 2700 stores have been grouped into 12 different segments, each representing the stores' characteristics, the customers and the spatial information.

6 Empirical Findings

6.1 Profilling Segments

The clustering algorithm groups each data point around a particular center in accordance with its relationship to other data points. However, there is no clear indication of the features that are important in the formation of these clusters. Therefore, there is no direct way of naming the segments. To do this, it is possible to extract the underlying meanings using statistical methods named factor analysis.

Factor analysis is a statistical technique used to explore and identify a data set's underlying structure by reducing the data's dimensionality. It uncovers the latent factors or unobserved variables that explain the observed correlations between multiple measures. It helps to simplify the data structure by focusing on the most important factors, which is particularly valuable when dealing with a large number of variables. In the context of clustering, factor analysis can be used as a post-processing step to reduce the dimensionality of the data set so that the key features can be retained for profiling the clustering results [4].

The segment naming process was discussed with participants from different teams by using key variables that were reduced and revealed by factor analysis. At this stage, not only the reduced data, but also the insights provided by different teams were incorporated into the work, to produce a result that all teams could adopt. The final names of the segments were discussed in a large meeting with the stores and regions managers. The goal is to use the insights from the managers to ensure that stores that may be at the intersection of multiple clusters during the clustering phase are aligned with the correct segment.

6.2 Sustainability of the Segments

The segmentation solution was developed using a multi-stage process. A variety of statistical and machine learning methods have been used at each stage. Due to its multi-stage structure, it is very difficult to achieve repeatability and sustainability. Because of these problems, there is a need for a solution that is sensitive to the changes in characteristics that will occur over time and that is easy to explain.

The solution to these problems is derived from classification analysis in the field of machine learning. Clustering analysis revealed a clustering number, a label, for each store that was associated with a group of features that were pruned using statistical techniques. Using the pruned feature set, our work trained a gradient boosted tree classification model called LightGBM [8]. LightGBM is an open-source gradient boosting framework developed by Microsoft. It is designed to be efficient and scalable, handling large datasets and delivering high performance with low memory consumption. It uses a leaf-wise approach with best-first growth, which provides for a more accurate and better-fitting model, particularly for large and complex datasets. The model, trained using the sliding window method with a time interval defined by the team, provides a simple, explainable and accurate result for segmenting newly opened stores.

7 Conclusion

The study we conducted provides a complete segmentation of the retail market. At first, data collected from various sources was enriched using both hand crafting and analytical methods. The enriched data were then filtered according to their representativeness using the PCA method. The filtered data were scaled to the same level using standard scaling. The scaled data were subjected to elbow and silhuette scoring methods to find the optimal number of clusters. Once the approximate number of clusters had been obtained, the k-means clustering algorithm was trained and the relevant cluster labels were obtained for each of the stores. Later, the factor analysis method was used to obtain significant features for each cluster. Using the significant feature and label information, a team from various departments helped with the naming of each segment. In the final stage, the classification algorithm was trained with LGBM algorithm and predictions were obtained at regular intervals to segment new stores for sustainability and to monitor segment changes. Our study provides an end-to-end store segmentation study for the retail sector using a combination of analytical and machine learning techniques.

Future studies could investigate the association between demographic variables and administrative location for a spatial segmentation. Spatial segmentation could help businesses to explore the possible and profitable investment strategies.

References

1. Agarwal, K., Jain, P., Rajnayak, M.A.: Comparative analysis of store clustering techniques in the retail industry. In: International Conference on Data Technologies and Applications (2019)
2. Arthur, D., Vassilvitskii, S.: K-Means++: the advantages of careful seeding. In: ACM-SIAM Symposium on Discrete Algorithms (2007)
3. Karl Pearson, F.R.S.: LIII. on lines and planes of closest fit to systems of points in space. Philos. Mag. Ser. 1 **2**, 559–572 (1901)
4. Gagnon, R.J.: Measurement theory and applications for the social sciences. Meas. Interdisc. Res. Perspect. **17**, 209–210 (2019)

5. Ghriss, A., Girard, R., Thual, A.: Mixtures of probabilistic principal component analysers (1998)
6. Gindi, A.A., Abdullah, A.M., Ismail, M.M., Nawi, N.M.: Segmentation of fresh vegetable shoppers by product and store attributes considered for fresh vegetable purchase in Klang Valley, Malaysia (2015)
7. Kaiser, H.F.: The varimax criterion for analytic rotation in factor analysis. Psychometrika **23**, 187–200 (1958)
8. Ke, G., et al.: LightGBM: a highly efficient gradient boosting decision tree. Adv. Neural. Inf. Process. Syst. **30**, 3146–3154 (2017)
9. Ketchen, D.J., Shook, C.L.: The application of cluster analysis in strategic management research: an analysis and critique. Strateg. Manag. J. **17**, 441–458 (1996)
10. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (1987)
11. Zhang, X., Pei, J., Xiaojun, Y.: Demographic transformation and clustering of transactional data for sales prediction of convenience stores. In: 2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pp. 102–108 (2016)

Demand Forecasting in Pharmaceutical Industry Under Covid-19 Pandemic Conditions by Machine Learning and Time Series Analysis

Irem Tas[1] and Sule Itir Satoglu[2(✉)]

[1] KocDigital Cozumler A.Ş., İstanbul 34700, Turkey
[2] Faculty of Management, Istanbul Technical University, 34369 Istanbul, Turkey
onbaslis@itu.edu.tr

Abstract. As pharmaceutical products carry vital importance for society, demand forecasting of pharmaceuticals is much more critical. A well-designed demand forecasting and planning can prevent pharmaceutical companies from stock-out and high disposal costs of products. However, there is a limited number of studies about demand forecasting in the pharmaceutical industry, especially in pandemic conditions. This article aims to examine this under-researched area and understand the factors that affect the demand for pharmaceuticals significantly in pandemics, and hence perform an accurate demand forecasting. In light of the literature review, the factors affecting the demand for the pharmaceutical are historical sales, price, promotion factors, campaigns, currency rates, market share, and seasonal or epidemic diseases. Since the chosen pharmaceutical product is used in enteric diseases treatments and lockdowns prevent access to public places, the Covid-19 pandemic is thought to be a factor affecting the sales of the selected product. The forecasting methods of Holt-Winter exponential smoothing, multiple linear regression, Artificial Neural Network, and XGBoost were applied. According to the results, XGBoost was determined as the method that gave the best forecasts, and significant factors affecting the demand were determined. This study is the first one in terms of investigating the effects of the Coronavirus pandemic on drug demand.

Keywords: Demand Forecasting · Pharmaceutical Industry · Covid-19 Pandemic · Linear Regression · Holt-Winters Exponential Smoothing · Artificial Neural Network · XG Boost

1 Introduction

In the big data era, there are several sources of demand data within the supply chain. By means of the predictive analytics approaches, the demand data can be exploited, and better forecasts can be made [1]. For many industries, forecasting is imperative for strategic decisions such as facility location, new product initiation in the long term, and aggregate planning in other words sales and operations planning of the manufacturing resources [2], in the medium term. Besides, it is usually needed in the short term,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 157–165, 2023.
https://doi.org/10.1007/978-3-031-39777-6_19

especially to improve the availability of the items' inventory and enhance demand management. Accurate demand forecasting and good demand management can prevent the companies from stock-out and high inventory holding or sometimes disposal costs of the products. So far, traditional time-series analysis techniques were commonly employed for demand forecasting. However, many external factors may affect sales and force us to employ advanced methods related to machine learning, for forecasting [2].

Especially, in the pharmaceutical industry, if the sales occur below the forecasts, the excess stock of the pharmaceuticals may become obsolete, and these may have to be disposed of, as the pharmaceuticals are perishable products. This is an unintended result. On the other hand, if the sales demand exceeds the forecasts, a backlog will incur that causes customer dissatisfaction. Moreover, unmet demands of pharmaceuticals may lead to further health problems which are more difficult to cure. Therefore, accurate demand forecasting is imperative in the pharmaceutical industry.

Many experts believed that the most important factor in demand forecasting of the pharmaceutical industry was historical sales data. However, in most cases, this data is not enough to explain all demand by itself. Candan et al. [3] suggested that external data such as seasonal and epidemic diseases, market shares of the products, and campaigns should be included in demand forecasting models of pharmaceuticals. However, these authors have not examined the effect of any epidemic disease on pharmaceutical demand.

During the Covid-19 pandemic, there is empirical evidence that some problems have arisen in the healthcare products supply chain [4, 5]. Supply chain disruptions occurred and scarcity of some critical medical commodities occurred [6]. Some specific strategies were suggested to enhance supply chain resilience [7]. Therefore, it can be concluded that healthcare products' availability was certainly affected by the Covid 19 pandemic. However, a limited number of studies were conducted to cope with the adverse effects of the pandemic on healthcare products and pharmaceutical products' supply chains.

Besides, our literature review has shown that no single study exists that examined the Covid-19 effect on drug demands. This study aims to understand the kinds of factors that affect the demand for pharmaceuticals, and how these factors explain the demand, especially in Covid-19 pandemic conditions. Hence, better forecasting can be achieved. These are the unique aspects of our study. Thus, in this study, a pharmaceutical product was chosen for the applications, and the state-of-art forecasting methods of Holt-Winter exponential smoothing, multiple linear regression, Artificial Neural Network, and XGBoost were applied. Then, the results were analyzed comprehensively, and significant factors affecting the demand for the specified product as well as the most powerful forecasting technique were determined. To the best of our knowledge, no past single study examined the Covid-19 pandemic effect on drug demands, especially for demand forecasting purposes.

The paper is organized as follows: In the following section, the Literature Review is presented. Later, the Methodology is explained. The demand forecasting techniques that we used during the forecasting are explained in the Methodology section. Later, the methods are applied and the results are explained in the Application Results and Discussion section. Finally, the Conclusion and Recommendations are presented.

2 Methodology

The methodology followed in this study comprises the stages of Business Understanding and Feature Selection, Data Understanding and Collection, Data Pre-Processing, Forecasting Model/Method Applications, and Forecasting Accuracy Evaluation. These stages are also shown in Fig. 1. In the following subsections, these stages will be explained.

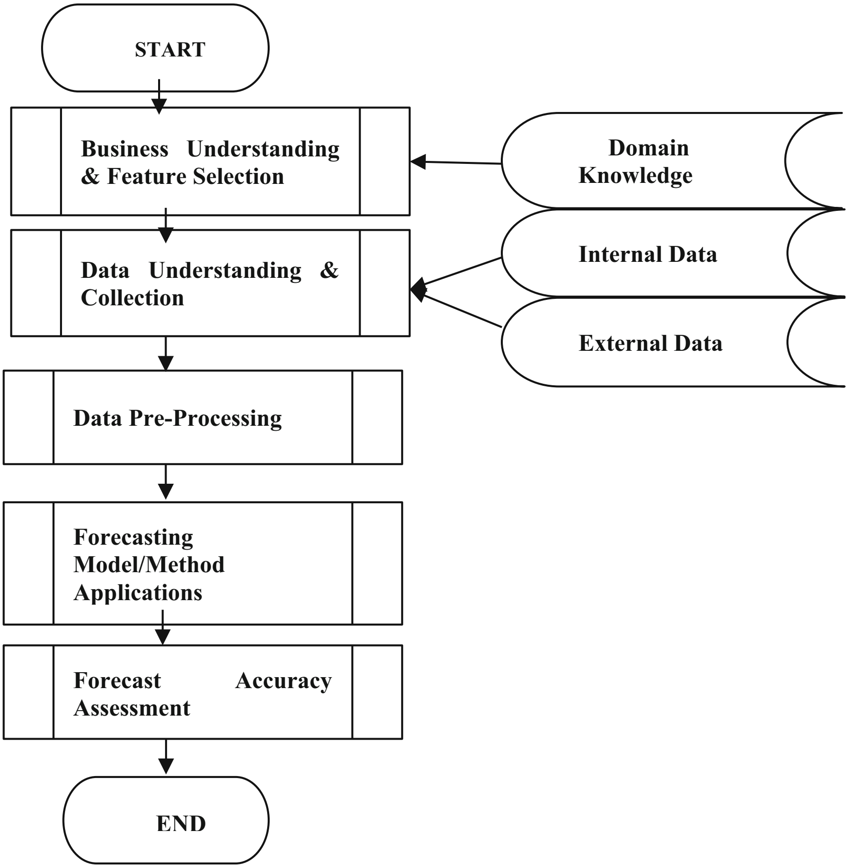


Fig. 1. Methodology of the Study

2.1 Business Understanding and Feature Selection

In demand forecasting models, domain knowledge plays a pivotal role. If the relationships between the demand and the internal and external components are not understood completely, demand cannot be predicted. For the application of demand forecasting techniques, one pharmaceutical product is chosen. Due to confidential issues, this product will be named Product X. This product is a prescription drug. The main use of Product X is enteric disease treatments. The factors affecting the number of patients would also affect the sales of these products. That is why the target users of this product should be identified. Based on the literature review and experts' opinions, data is collected from different sources. The selected features or factors are historical sales, temperature, the Covid-19 pandemic existence (binary variable), the exchange rate of euro/dollar data, and the inflation rate. So, data that pertains to these factors were collected. Especially due to the crisis that emerged with the recent Covid-19 pandemic, many industries have been affected. Measures such as lockdowns have been taken to prevent the pandemic. Since the pharmaceutical product is used in enteric diseases treatments and lockdowns prevent access to public places, the Covid-19 pandemic is thought to be a factor affecting the sales of Product X.

2.2 Data Understanding and Collection

After the business understanding, the data understanding should be gained, too. To gain insights about data and process the data effectively, it should be understood completely. Data understanding involves the exploration and interpretation of data. These steps will contribute to constituting a meaningful whole. The difference between these data is their source. Then, these data will be prepared for data analysis.

2.3 Data Pre-Processing

To process the data in the models, the data should be investigated to ensure whether the data is clean. First, the missing variables are investigated. Several methods are used for missing values based on the columns. These techniques are backfilling, forward filling, and filling with mean values of columns. Then, the statistical information of data is checked, if there was a typo or not. The maximum and minimum values of each column are investigated. Later, the data is normalized, based on Min-Max Normalization. Finally, the dataset is split into two groups: the train set and the test set. The ratio of the train set to the test set is 80:20.

2.4 Forecasting Model/Method Applications

Applied methods/models for forecasting are Holt-Winter Exponential Smoothing, Linear Regression, Artificial Neural Network, and XGBoost. Those techniques and methods will be shortly explained here. Holt-Winter Exponential Smoothing is a traditional time series forecasting method where both trend and seasonality are considered, simultaneously. For a time series data set, where a positive or negative trend and a seasonal pattern are inherent in the demand for the product, this may be a suitable method.

Linear regression is a classical statistical method where the correlation between the dependent variable, and the independent variables is analyzed, and a regression function is reached based on the statistically significant independent variables and the dependent variable. In this study, the dependent variable is the demand (sales) and there are other internal or external variables considered independent variables. Those independent variables are found through the *Business Understanding and Feature Selection* stage**.**

Artificial Neural Network (ANN) is one of the most famous machine learning algorithms used in demand forecasting. It is an algorithm that imitates the learning process of the human brain [8]. By training the algorithm through the training data set, the weights of the input variables are adjusted at each iteration to minimize a non-linear loss function that is usually the mean square error. For detailed explanations of the ANNs, the readers should refer to [9].

XGBoost method, which was invented by Tianqi Chen, stands for extreme gradient boosted trees. As the name implies, it is a version of gradient-boosted trees. According to Tianqi Chen, what makes XGBoost successful is its scalability in all problems. Since this method computes in a parallel and distributed way, this model can be considered as fast [10]. In gradient-boosted trees, every tree is trained, no matter how that tree contributes to the model. On the other hand, XGBoost calculates the gains of each tree, thus it decides whether to continue in that branch or not. If the gain is negative, XGBoost does not continue in that branch, so it saves a lot of time. This is called tree pruning which is very helpful, especially in large datasets.

2.5 Forecast Accuracy Assessment

In this study, the accuracy of the forecasts is assessed based on the *mean absolute percent error* measure. For each forecasting instance, the absolute value of the difference between the actual demand and forecast is divided into the actual sales. These values are summed and divided by the number of instances to reach the "mean" absolute percent error [11]. As it is a percentage, it shows how precisely the forecasting was made.

3 Application Results and Discussion

In the sales data set of Product X, there are 63 observations. These observations were recorded monthly. The sales of Product X which is a drug for enteric diseases show seasonal behaviors. The sales decrease in the winter months and increase in the summer months. Since enteric diseases are contagious, public environments can cause the spread of enteric diseases faster. Moreover, it can be said that dehydration, contaminated foods, and water can cause enteric diseases. Also, the spread of bacteria is getting faster as the temperature increases [12].

The dataset is split into two groups: a train set and a test set. The ratio of the train set to the test set is 80:20. Thus, there are 50 observations in the train test and 13 observations in the test set. All models are applied to the train set and these models are evaluated on the test set. The evaluations are made by the metric of mean absolute percentage error (MAPE). An objective of this study was to predict the sales of Product X. To find out the method that best describes the behavior of the sales, many models/methods were applied. However, the methods employed in this study are Holt-Winter exponential smoothing, multiple linear regression, artificial neural network, and XGBoost. According to the results of all models/methods, XGBoost was found the best-performing forecasting method, or this case. The mean absolute percentage errors of above mentioned models are shown in Table 1. The details of the best-performing forecasting method is also explained below.

Table 1. Overall Model Results.

Method	MAPE
Holt-Winter Exponential Smoothing	%36,4
Linear Regression	%17,3
ANN	%18,3
XGBoost	%11,8

3.1 XGBoost

The inputs included are the binary value indicating if the observation is recorded in July, the binary variables indicating if the observation is recorded in August, December, or February, and the binary value indicating if the observation is pandemic-affected, as well as the sales of periods (t-1), (t-2) and (t-3). Since the demand forecasting problem is a regression problem, the objective is chosen as "linear". The learning rate is a parameter that arranges the balance between overfitting and underfitting. The higher the learning rate, the more the model tends to overfit. As for the learning rate, a 0.2 value is chosen. The maximum depth parameter is chosen as 3. Colsample_bytree is the subsample ratio of columns while constructing each tree. It is chosen as 0.4. Alpha is the L1 regularization term. The higher the alpha is, the less likely the model overfits. The alpha value is chosen as ten. The MAPE score is found as 11.8%. The sales of the past three periods have the greatest feature importance.

Besides, the graph of actual and predicted sales is depicted in Fig. 2. As the features with the strongest importance were selected, the forecasting by the XGBoost was enhanced.

3.2 Discussion

In this application, four methods were applied, which are Holt-Winter exponential smoothing, linear regression, artificial neural network, and XGBoost. The data was shuffled and split into two datasets as train and tests. The size of the test set is 20% of the data and the rest of the data is considered a train set. The models are trained on the train sets and evaluated on the test set. It can be said that XGBoost outperforms other models. Its MAPE value is lower than that of the others. Besides, XGBoost is extremely fast, thus it saves a lot of time. In brief, historical sales, the months of February, July, August, and December, and the Covid-19 pandemic affect the demand for Product X. These inputs explain the demands considerably. However, the temperature, the inflation rate, and the exchange rate do not affect Product X. Having discussed the results of the methods, the final part of this paper gives recommendations about demand forecasting in the pharmaceutical industry.

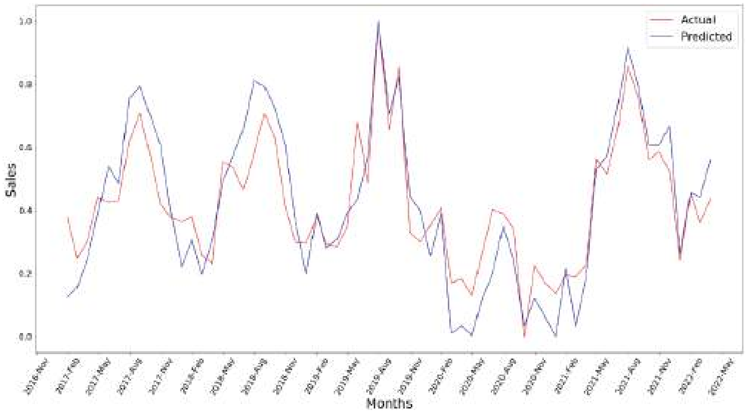


Fig. 2. The Actual Sales versus Predicted Sales Calculated by XGBoost Model.

4 Conclusion and Recommendations

This study set out to understand what kind of factors affect the demand for pharmaceuticals, and how these factors explain the demands. There is a growing body of literature that recognizes the importance of demand forecasting techniques. Yet, there is not sufficient academic research about demand forecasting in the pharmaceutical industry. It is aimed to fill this academic gap with this academic research and application. In this paper, firstly, the literature review is explained to understand the related works. Then, the methodology is explained. Business and data understanding are examined since domain knowledge is critical. After the data collection is made and data is analyzed, the chosen methods which are Holt-Winter exponential smoothing, linear regression, artificial neural network, and XGBoost are applied. The performance of these techniques is compared, and also some managerial implications are proposed as follows:

According to the results of the study, the inputs of the final model are historical sales, calendar information, and Covid-19 inputs. It is found that only internal data does not explain all demands. Thus, external data should be anticipated.

Other pandemic diseases similar to the Covid-19 pandemic that may happen in the future, are believed to affect drugs used in enteric disease treatments. In these types of crises, sales will probably decrease due to lockdown declarations. Thus, pharmaceutical companies can avoid inventory holding costs and possible disposal costs by foreseeing this decline.

After proper demand forecasting, the challenge is to manage demand planning carefully to provide an excellent level of product availability, and optimum inventory holding and backlog costs. After ensuring the forecasts, it is needed to examine the production time, lead time, and other idle times. The products that show seasonal behaviors can show predictable behaviors. To meet the demands of product X, the company should keep products in stock before the peaks. Besides, alternative suppliers can be found to decrease lead times in peak seasons. When the demand for this product declines, the company can postpone purchases to avoid inventory holding costs and possible disposal costs. Continued efforts are needed to make demand planning more manageable and controllable.

Our findings have significant implications for the understanding of how forecasting models are applied and which factors should be input into these models, especially in the pharmaceutical industry. For the pharmaceutical products that have similar characteristics to Product X, this study can be an example. In future studies, applications of these forecasting techniques can be made for other types of pharmaceutical products, as different pharmaceuticals such as cancer drugs, supplements, etc. may be influenced by other factors.

References

1. Boone, T., Ganeshan, R., Jain, A., Sanders, N.R.: Forecasting sales in the supply chain: consumer analytics in the big data era. Int. J. Forecast. **35**(1), 170–180 (2019)
2. Caglayan, N., Satoglu, S.I., Kapukaya, E.: Sales forecasting by artificial neural networks for the apparel retail chain stores. In: the International Conference on Intelligent and Fuzzy Systems, pp. 451–456. Springer, Cham (2019, July)
3. Candan, G., Taskin, M.F., Yazgan, H.: Demand forecasting in the pharmaceutical industry using a neuro-fuzzy approach. J. Manage. Info. Sci. **2**(2), 41–49 (2014)
4. Magableh, G.M.: Supply chains and the COVID-19 pandemic: a comprehensive framework. Eur. Manag. Rev. **18**(3), 363–382 (2021)
5. Raj, A., Mukherjee, A.A., de Sousa Jabbour, A.B.L., Srivastava, S.K.: Supply chain management during and post-COVID-19 pandemic: mitigation strategies and practical lessons learned. J. Bus. Res. **142**, 1125–1139 (2022)
6. Gereffi, G.: What does the COVID-19 pandemic teach us about global value chains? the case of medical supplies. J. Int. Bus. Poli. **3**(3), 287–301 (2020). https://doi.org/10.1057/s42214-020-00062-w
7. Spieske, A., Gebhardt, M., Kopyto, M., Birkel, H.: Improving resilience of the healthcare supply chain in a pandemic: Evidence from Europe during the COVID-19 crisis. J. Purchas. Supp. Manage. 100748 (2022)
8. Khashei, M., Bijari, M.: A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Appl. Soft Comput. **11**(2), 2664–2675 (2011)
9. Alpaydin, E.: Introduction to machine learning. MIT Press (2020)

10. Chen, T., Guestrin, C.: Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)
11. Jacobs, R.F., Berry, W.L., Whybark, D.C., Vollmann, T.E: Manufacturing planning and control for supply chain management. McGraw-Hill Education (2018)
12. Fleury, M., Charron, D.F., Holt, J.D., Allen, O.B., Maarouf, A.R.: A time series analysis of the relationship of ambient temperature and common bacterial enteric infections in two Canadian provinces. Int. J. Biometeorol. **50**(6), 385–391 (2006)

Earthquake Predictions Using Machine Learning and Acoustic Emissions

Ioannis Sioutas[1], Maria Drakaki[1](✉), Vasileios Linardos[1], and Panagiotis Tzionas[2]

[1] Department of Science and Technology, University Center of International Programmes of Studies, International Hellenic University, 14th Km Thessaloniki, Nea Moudania, 570 01 Thessaloniki, Greece
mdrakaki@ihu.gr

[2] Department of Industrial Engineering and Management, International Hellenic University, P.O. Box 141, 57400 Thessaloniki, Greece

Abstract. The impacts from earthquakes can be devastating, therefore accurate earthquake forecasting is of ultimate importance towards effective disaster risk reduction and improved disaster preparedness. Earthquake prediction is the holy grail of earthquake science yet advances in this field have not shown results that will contribute significantly to a reduction of associated disaster losses. In this paper, Machine Learning (ML) and Deep Learning (DL) methods have been employed to forecast the time before a defect fails by using the acoustic signals of a laboratory fault. A Kaggle dataset was used containing experimental laboratory earthquake data. Various algorithms were employed including the CatBoost, Random Forest, Extreme Gradient Boosting, AdaBoost and Convolutional Neural Networks (CNNs). The forecasts consider the instantaneous physical characteristics of the acoustic signal. The best accuracy was achieved by applying rolling forecast using the AdaBoost Regressor algorithm. Promising results have shown the potential of the proposed method. Future research could explore the applicability of the method with continuous signals from seismic waves to provide further insights into earthquake prediction.

Keywords: Earthquake Prediction · Machine Learning · Deep Learning · Laboratory data · Acoustic Emissions

1 Introduction

Earthquakes are costly natural hazards faced by the humans, that occur without an explicit warning and may result in numerous injuries or fatalities due to damage to buildings or other rigid structures [1]. Earthquakes and natural disasters in general, phenomena that occur all over the world, are the subject of daily scientific investigations.

Earthquakes are sudden tremors of the earth's surface caused by the propagation of seismic waves through rocks. These waves result from the sudden release of energy stored within the earth's crust, which occurs when rock masses are temporarily fractured and shifted against each other [13]. Faults, which are narrow regions where rock masses

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 166–172, 2023.
https://doi.org/10.1007/978-3-031-39777-6_20

move relative to each other, are the most common sites of earthquakes. Major tectonic plates that form the earth's crust are bounded by significant fault lines [13]. Prior to the emergence of seismology as a scientific field in the early 1900s, very little was understood about earthquakes. Seismology has since provided insights into a range of questions about seismic activity, including the causes and mechanisms of earthquakes.

Seismologists face a significant challenge in accurately predicting earthquakes to minimize loss of life and economic disruption [1]. Ideally, we would have complete knowledge of the location of all active faults, the timing of their rupture, and the size of the resulting earthquake. Successful earthquake forecasting would provide precise information on the timing, location, and magnitude of forthcoming earthquakes. In the United States, advances in building design, intelligent urban planning, and a better understanding of fault behavior have significantly reduced the risk of earthquakes to human life. Improved comprehension of earthquake mechanisms and the behavior of the Earth will continue to enhance our safety [1].

In the following, literature review is presented next. Methodology is presented next, as well as discussion of the results. Conclusions with directions for future research follow.

2 Literature Review

Linardos et al. [2] analyzed recent studies related to earthquake prediction using machine learning techniques. In the review article the following references were found most relevant to this study. Resch et al. [3] addressed the limitations of traditional methods in identifying hotspot areas and assessing damage caused by natural disasters. The authors combined Latent Dirichlet Allocation (LDA) with temporal and spatial analysis of social media posts to accurately predict earthquakes and generate loss maps, validated by the HAZUS loss model and official earthquake footprint provided by the US Geological Survey. Asim et al. [4] used various ML techniques, including pattern recognition NN, RNN, RF, and linear programming boost ensemble classifier, to predict earthquakes in the Hindukush region of Pakistan. LPBoost ensemble showed the highest prediction accuracy of 65% with an unknown dataset. Chin et al. [5] focused on improving the accuracy of early warning and detection systems for earthquakes, which suffer from false alarm problems. ML-based algorithms, such as SVM, classification tree, and KNN, were experimented to reduce the false alarms rate and increase detection accuracy using the earthquakes related seismic data in Taiwan. Li et al. [6] aimed to mitigate false alerts and noise from earthquake early warning systems by training a generative adversarial network (GAN) to learn the characteristics of first-arrival earthquake P waves, using waveforms recorded in southern California and Japan. An RF classifier was trained with earthquake and noise waveforms, resulting in a high accuracy rate for identifying earthquake P waves. These studies demonstrate the potential of machine learning in improving earthquake prediction and reducing false alarms.

Acoustic emission testing (AE) is a non-destructive method that detects ultrasonic stress waves emitted by materials under stress [7]. This technique involves attaching a small sensor to the material, which converts the waves into electrical signals for analysis on a recording PC. When components are displayed to external stimuli such as stress, high pressure, and temperature, waves are trapped. More energy is released as component damage increases. The intensity, loudness, and activity of acoustic emission are monitored to assess the structural integrity and condition of the component.

In a laboratory experiment that mimics real-world settings, Los Alamos National Laboratory (LANL) researchers found a way to accurately predict slow slip earthquakes (SSEs) in 2017. The group developed software that can recognize and examine quasiperiodic seismic and acoustic data produced by fault movement. They analyzed enormous quantities of data and discovered noise patterns that were once supposed to precede earthquakes [8]. The LANL team recognizes that the properties of laboratory experiments, such as shear stress, differ from natural earthquakes, but work is underway to apply the analysis to the real world to authenticate the results. The method can be applied beyond seismology and can support material failure studies in other fields such as aerospace and energy [8].

To date, earth scientists have mainly relied on catalogs of historical data to characterize fault conditions. These catalogs contain a small fraction of seismic data, some of which is discarded as unwanted noise during analysis. The researchers found that the noisy data are signals that give them information of the state of the fault in a more precise way than catalogs [9, 10]. This study shows that the statistical properties of continuous seismic signals emanating from a laboratory fault zone can predict the timing of future fault events and the frictional state of the fault zone [11].

Acoustic Emission (AE) is considered a form of microearthquake that occurs when a material is subjected to stress and failure [12]. This technique involves the spontaneous release of local strain energy in the material, often resulting from microcracks, which can be detected by sensors on the material's surface. Due to this similarity to seismic methods, Acoustic Emission Testing (AET) is often used to assess the structural integrity of materials. [12].

In a scientific paper on earthquake physics, Ares Rosakis proposed the concept of conducting laboratory earthquakes to observe their growth and development by slowing down their movement [12]. To simulate earthquakes in a laboratory setting, Rosakis and his team used a transparent plastic block called Homalite with properties similar to rocks. They first cut the block in half and then bonded the two parts together under pressure and shear, mimicking the buildup of tectonic pressure along a fault line [12]. Small wire fuses were placed at specific locations beneath the simulated surface to trigger earthquakes.

3 Methodology

Our research focuses on the inference of shear stress, shear displacement, and groove thickness, which involves utilizing continuous seismic signals obtained during shear and relying on statistical data analysis. Our approach employs models that do not consider the past or future signal history, resulting in instantaneous estimations. We employ machine learning techniques that leverage recent advancements in the field, using data collected from controlled laboratory experiments that involve double direct shear on fault material. During these experiments, a drive piston moves at a constant velocity during inter-event time and briefly accelerates during slip, while acoustic emissions (AE) from the shear layer are recorded by accelerometers. The output of our models is the mean absolute error in the estimated properties.

For this project, our training dataset comprises a lengthy time-series of measurements obtained from a laboratory-based simulation of earthquakes. We are also provided with a "test" dataset consisting of segments of similar time-series data, and our task is to predict the occurrence of the next earthquake in the laboratory based on this information. The approach involves analyzing the time-series measurements and utilizing examples of when previous lab earthquakes were recorded to make predictions about future occurrences. Seismic signals were recorded for 157.275 s by utilizing the acoustic emission method and 600 million rows of acoustic data were gathered. The training set consists of two columns:

- Acoustic data which is the seismic signal and
- Time to Failure that corresponds to the time until the laboratory earthquake takes place.

This study relies solely on the analysis of snapshots of continuous seismic data obtained during shear to estimate the time to failure. The task at hand is a regression problem, where the input comprises, the acoustic data captured during the experiment, and the target is the time to failure of the fault. ML and DL models are developed during the training phase using input and target data. During testing, the recorded shear stress is taken as the reference value and is not provided to the model. Despite this, the models are able to accurately predict the timing of laboratory earthquakes. Specifically, we used Random Forest (RF), CatBoostRegressor, AdaBoostRegressor, XGBoost, while we experimented with the architecture of two Neural Networks, a simple Artificial Neural Network (ANN) and a Convolutional Neural Network (CNN).

The data were divided into 4194 bins, each containing 150,000 rows of acoustic data. The objective for each test folder segment ID was to predict the time to failure, i.e., the time (in seconds) from the last segment to the onset of the next laboratory tremor. To enhance learning, several statistical features as well as rolling descriptive statistics were extracted from each bin. The training process involved the following phases.

1. Training on all the extracted features: In this phase, all the statistical features and rolling descriptive statistics extracted from each bin were used to train the model. This approach provides the model with a comprehensive set of features to learn from, but it can also lead to overfitting.
2. Training on the most appropriate ones after feature selection: In this phase, feature selection techniques were used to identify the most relevant features for the model. This approach helps to reduce overfitting and improve training times by removing irrelevant or redundant features. The selected features were then used to train the model.
3. Rolling Forecast: In this phase, the model was trained to predict the time to failure for each bin based on the previous bins. This rolling forecast approach takes into account the time series nature of the data and allows the model to learn from the temporal patterns in the data.
4. Training after Hyperparameter tuning: In this phase, hyperparameter tuning techniques were used to optimize the performance of the model. By tuning these parameters, the model can be optimized for better accuracy and generalization (Tables 1 and 2).

Table 1. Best performances of the machine learning algorithms

Regressor name	MAE score
Random forest	2.53467
CatBoost	2.63198
XGBoost	2.54251
AdaBoost	2.54251

Table 2. Best performances of the deep learning algorithms

Neural Network	MAE score
Artificial NN (feature selection)	2.68692
Convolutional NN (feature sel.)	2.65024
Artificial NN (rolling features)	2.62962
Convolutional NN (rolling feat.)	2.63771

The evaluation of the results of the study was carried out by submitting the research output to the Kaggle Competition page. The performance of the models was assessed based on the Mean Absolute Error (MAE) metric, which is a measure of the average magnitude of the errors in the predictions. The MAE metric returns values on the same scale as the target variable, and there is no universally accepted benchmark for a good MAE score.

Among the models tested, the AdaBoostRegressor demonstrated the best performance, achieving an MAE score of approximately 2.51. This result placed the model among the top 300 submissions on the leaderboard, and it was awarded a bronze medal. The AdaBoostRegressor's ability to minimize the errors in the predictions suggests that it could be a suitable model for similar research studies (Fig. 1).

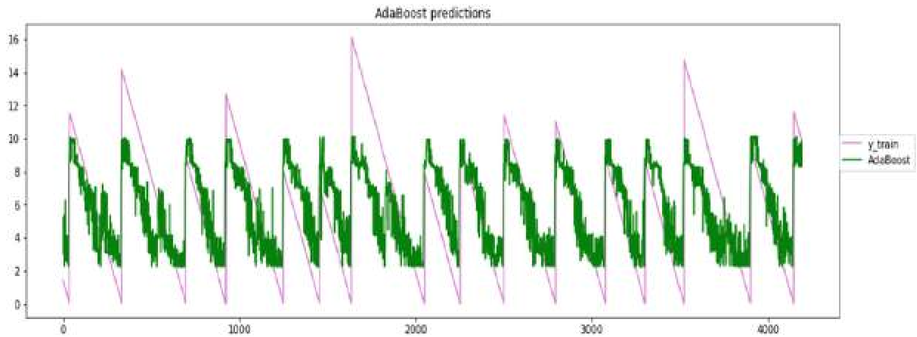


Fig. 1. The plotted Time to Failure predictions of AdaBoost (green) with the ground truth (pink).

4 Conclusions

In conclusion, our study reveals some key findings that contribute to the understanding of laboratory failures and seismic cycles. One of the most significant outcomes of this research is that laboratory failures are not random but occur in a highly predictable manner. Moreover, we found that applying rolling forecast using the AdaBoostRegressor algorithm is the most effective strategy compared to other models in the literature. This highlights the potential for this approach to be applied in practical settings.

Furthermore, our observations demonstrate that key features of laboratory seismic cycles can be inferred from continuous seismic signals emitted by faults. Specifically, the use of an equation of state that relates the strength of the continuous seismic signal to the friction of the fault allowed us to determine the instantaneous state of friction, the critical state of stress, and the position of the fault within the seismic cycle. This finding challenges the traditional seismic catalog approach for analyzing the physical properties of faults, which may overlook crucial information.

To build upon our findings, future research could explore the application of machine learning in other domains beyond laboratory failures. Additionally, investigating the use of continuous signals from seismic waves to better understand Earth's faults and improve earthquake forecasting could be a fruitful avenue for future research. Future studies could investigate the potential application of machine learning algorithms to seismic data analysis on real world data, which could lead to further insights and advancements in earthquake mechanics. Overall, our research suggests that there is much to be explored in the field of seismic cycles and earthquake forecasting, and we hope our findings will stimulate further research in this area.

Funding. This research is co-financed by Greece and the European Union (European Social Fund-SF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014–2020» in the context of the project "Support for Internationalization Actions of the International Hellenic University", (MIS 5154651).

References

1. Moustra, M., Avraamides, M., Christodoulou, C.: Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals. Expert Syst. Appl. **38**(12), 15032–15039 (2011)
2. Linardos, V., Drakaki, M., Tzionas, P., Karnavas, Y.L.: Machine learning in disaster management: recent developments in methods and applications. Mach. Learn. Knowl. Extrac. **4**(2), 446–473 (2022)
3. Resch, B., Usländer, F., Havas, C.: Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. Cartogr. Geogr. Inf. Sci. **45**, 362–376 (2018)
4. Asim, K.M., Martínez-Álvarez, F., Basit, A., Iqbal, T.: Earthquake magnitude prediction in Hindukush region using machine learning techniques. Nat. Hazards **85**(1), 471–486 (2016). https://doi.org/10.1007/s11069-016-2579-3
5. Chin, T.-L., Huang, C.-Y., Shen, S.-H., Tsai, Y.-C., Hu, Y.H., Wu, Y.-M.: Learn to detect: improving the accuracy of earthquake detection. IEEE Trans. Geosci. Remote Sens. **57**, 8867–8878 (2019)

6. Li, Z., Meier, M.A., Hauksson, E., Zhan, Z., Andrews, J.: Machine learning seismic wave discrimination: application to earthquake early warning. Geophys. Res. Lett. **45**, 4773–4779 (2018)
7. Große, C.U.: Acoustic Emission Testing: Basics for Research – Applications in Civil Engineering. Springer (2010)
8. Rouet-Leduc, B.: Los Alamos machine learning discovers patterns that reveal earthquake fault behavior. Science briefs, Los Alamos National Laboratory, P.O. Box 1663 Los Alamos, NM 87545 (March 2018)
9. Tanyuk, O., Davieau, D., South, C., Engels, D.W.: Machine Learning Predicts Aperiodic Laboratory Earthquakes. SMU Data Science Review **2**(2), Article 11 (2019)
10. Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J., Johnson, P.A.: Machine learning predicts laboratory earthquakes. Geophys. Res. Lett. **44**(18), 9276–9282 (2017)
11. Lubbers, N., Bolton, D.C., Mohd-Yusof, J., Marone, C., Barros, K., Johnson, P.A.: Earthquake catalog-based machine learning identification of laboratory fault states and the effects of magnitude of completeness. Geophysical Research Letters **45**(24), (2018)
12. Tal, Y., Rubino, V., Rosakis, A.J., Lapusta, N.: Illuminating the physics of dynamic friction through laboratory earthquakes on thrust faults. Proc. Natl. Acad. Sci. **117**(35), 21095–21100 (2020)
13. Hayes, G., Wald, D.: Earthquake magnitude, energy release, and shaking intensity. Available at: https://earthquake.usgs.gov/learn/topics/mag-intensity/

Machine Learning Prediction of Wind Turbine Energy Production

Osman Emin Erdem, Ahmet Sarucan, and Mehmet Emin Baysal(✉)

Konya Technical University, Selçuklu, Konya, Turkey
mebaysal@ktun.edu.tr

Abstract. This study used statistical data to create predictive models for effective and efficient use of energy generation systems using machine learning algorithms. Four different models predicted wind energy production and compared with actual production data. Model validity and verification were evaluated using mean square error, mean absolute error, and R-squared values. The model constructed with the random forest method is the most successful model.

Keywords: Energy · Machine Learning · Data

1 Introduction

Energy is an essential resource for life's survival and development, and is critical to industry, commerce, agriculture, and households. It is the source of electricity, heat, mechanical power, and other forms of energy needed to sustain human life. Energy resources are divided into natural and manmade. Natural resources include sources such as oil, natural gas, coal, hydroelectric, wind, solar, and geothermal. Sources such as nuclear energy are manmade energy sources. The use of energy resources has been one of the most important technological and economic developments in the history of mankind. However, because of limited energy resources and environmental effects, the demand for sustainable energy sources is growing. Therefore, it is important to invest in renewable energy sources, increase energy efficiency, and reduce dependence on fossil fuels. Hydroelectric, wind, solar, biomass, biogas, and geothermal energy are among the renewable energy sources. They are environmentally friendly and unlimited to use, being renewable. But renewables are also the cheapest form of energy. It is for this reason that efforts are being made all over the world to invest in and increase the use of renewable sources of energy. At this point, it is necessary to know how much energy can be produced to meet the demand. Increasing importance of renewable energy resources makes it necessary to estimate their production. Estimating the amount of renewable energy produced is important for energy planning, making investment decisions, and setting energy policy. Forecasting is the process of determining how much energy is available from energy resources and how much is consumed. In order to meet energy demand, accurate forecasting of renewable energy production is also important. Energy supply and demand can be balanced using forecasts. Additionally,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 173–180, 2023.
https://doi.org/10.1007/978-3-031-39777-6_21

production fluctuations of energy resources can be predicted in advance by forecasting the production amount. As a result, the energy supply can be stabilized. When planning energy investments, forecasting the expected production of energy resources is also very important. These projections guide the amount of renewable energy invested, the payback period, and the costs. All in all, the estimation of renewable energy production is of great importance for planning, investment and decision making processes in the energy sector. It has a significant impact on the sustainability of energy resources.

This study shows that wind turbine energy production predictions can be obtained with machine learning algorithms. These obtained data can be used as input in energy planning and policy determination studies.

The study addresses this important issue with the following structure: After the introductory chapter, the second chapter presents the current literature on the topic. The third chapter describes the application under six subheadings within the framework of methods and materials used. Finally, the fourth chapter is devoted to the conclusions and recommendations.

2 Literature Review

Since there are many studies on the use of forecasting techniques in the literature, the scope of the review is narrowed to the use of machine learning techniques in energy forecasting. Anıl's study used machine learning, deep learning and time series methods to forecast the total production of power plants belonging to five organizations with high geothermal installed capacity for the first week and two weeks of March and April, then calculated imbalance penalties and compared them with the actual values. It concludes that the imbalance penalty obtained as a result of forecasting in five organizations is lower than actual imbalance penalty, optimal and profitable [1]. Inal, Darıca-2 Hydroelectric Power Plant (HEPP) located in Ordu Province, Turkey was analyzed and a neural network model was used to predict the maximum power generation capacity for the next day using the current water level in the reservoir and weather forecasting data. The paper shows that the artificial neural network models are successful in forecasting and can be used in this area [2]. Akpolat's study describes the design, modeling, implementation, and operation of a microgrid in which an independent hybrid power system is installed for an educational and research laboratory. An artificial intelligence-based sensorless control approach is used to solve stability and reliability problems in DC microprocessors. From the results obtained, it is concluded that artificial intelligence-based approaches achieve very small error values, successfully capture the dynamics of the system, allow for a flexible structure with adjustable hyper-parameters, and provide opportunities for practical application [3]. Balsever's study used artificial neural network and support vector machine models to estimate hourly electricity consumption. His analysis showed that artificial neural network predictions were generally more accurate, with about 97% accuracy [4]. Mohamed et al. compared the performance of six forecasting models for electricity consumption in New Zealand. The Harvey model performed best among all models for residential and total electricity consumption. The Harvey logistic model performed best for non-residential consumption [5]. Durğun, used regression analysis and artificial intelligence techniques to forecast Turkey's electricity demand in 2023. To forecast the energy consumption of Turkey, the GDP, population, and weather variable data (humidity, temperature, wind, and rainfall) between 1980 and 2017 were used as the input data of the prediction model [6]. Balsever, used artificial neural network and support vector machine models in his study on estimating the consumption of electrical energy on an hourly basis. As a result of the analyses, it was determined that the predictions of artificial neural networks were generally more accurate, with an accuracy rate of approximately 97% [7]. Sevdim, collected data on the matching of consumer demand from the combination of solar and wind energy with various amounts of energy storage capacity. These data were compared with data from other wind and solar panels in the region. Artificial neural networks were then used to predict the most appropriate location and positioning of wind and solar panels in the region [8]. Based on TurkStat data, Eren used Gray (1,1) model to analyze Turkey annual natural gas consumption and population data to forecast natural gas consumption from 2017 to 2030 [9]. Şenol et al. conducted a study to estimate electricity generation using wind turbine output power as input data using artificial neural network method. It was shown that the predictions made by the model were consistent with the regression curves of the results in the test phase of the wind speed data after training the model using the output power of different types of wind turbines in the training phase [10].

3 Implementation

3.1 Building a Model for Forecasting Wind Turbine Energy Production

Wind energy is an important source of renewable energy that has an important place in the modern era. A wind turbine extracts energy from the wind and converts it into electrical energy. This study aims to estimate the amount of energy produced in wind turbines by using machine learning algorithms in python program.

The dataset used in this project contains two types of variables: independent variables (predictors) and dependent variables (targets). The independent variables include the power generated by the turbine, wind speed, wind theoretical power value, wind direction and wind speed. The dependent variable is the energy produced by a wind turbine. The dataset consists of a set of 5530 observations taken with SCADA software at 10 min intervals during the days of the calendar year 2018 based on data collected from a turbine operating in Turkey and generating electricity. The dataset is taken from Kaggle's repository and the information of the dataset includes the following [11].

- Date/Time (10 min intervals)
- LV ActivePower (kW): Power currently generated by the turbine
- Wind Speed (m/s): Wind speed at the hub height of the turbine (wind speed used by turbines for electricity generation)
- Theoretical_Power_Curve (KWh): The theoretical power produced by the turbine with the wind speed given by the turbine manufacturer
- Wind Direction (°): Wind direction at the hub height of the turbine (wind turbines automatically turn in this direction)

3.2 Procedure of the Machine Learning Algorithms

Step 1: Data Preprocessing: The data set is scaled to avoid any bias or scaling effects. Outlying values are discarded from the data set prior to normalization. A relational graph was created to make sense of the correlation between the variables in the data set. This is shown in Figs. 1–3.

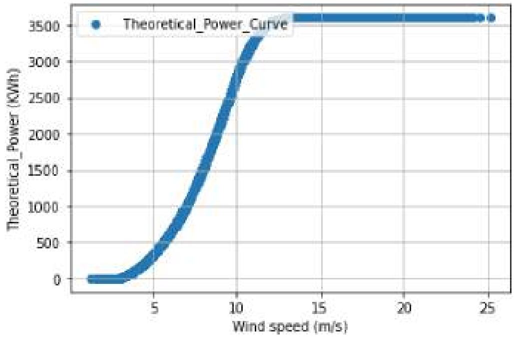


Fig. 1. Relationship between theoretical power curve (KWh) and wind speed (m/s)

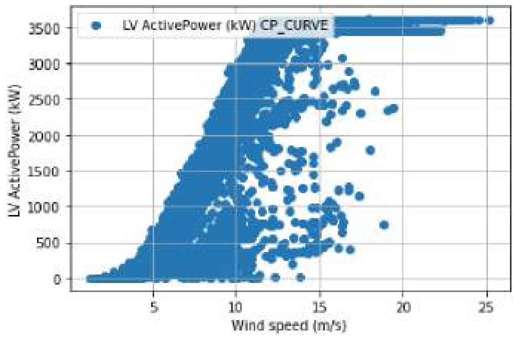


Fig. 2. Relationship between the LV ActivePower (kW) and the wind speed

As shown in Figs. 1–3, the correlation between wind speed (m/s) and power produced and between wind speed (m/s) and theoretical power produced by the wind turbine is positive (+).

Analyzing the observations by month, Fig. 4 shows that January, February, March, September, November, and December have the highest active power range. This is consistent with a higher wind speed during these months as compared to the other months of the year.

Step 2: Create training and test sets of the data using the Pyhton library function.
Step 3: Model Building: Modeling is done using the selected predictors, model, training dataset.
Step 4: Model Evaluation: After the model is created, R-squared, Mean Absolute Error (MAE), Mean Square values are calculated as a result of evaluation on the test dataset.

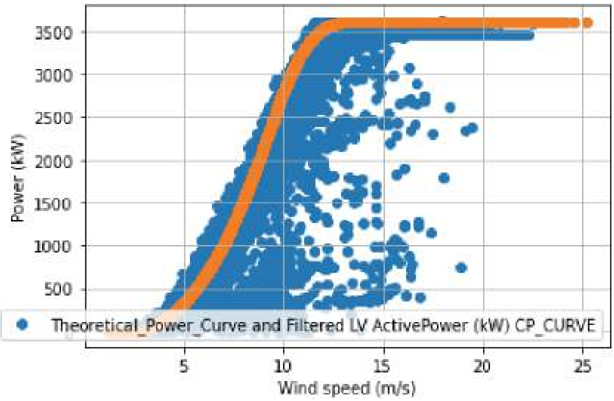


Fig. 3. Relationship between the LV ActivePower (kW) and the wind speed (m/s)

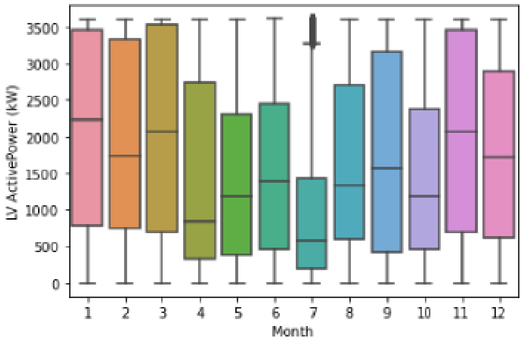


Fig. 4. Wind turbine monthly energy distribution

3.3 Prediction of Wind Turbine Energy Production with Multiple Linear Regression Model

The multiple linear regression model will be used because it is an effective method for estimating the dependent variable. To select variables that have a significant impact on the dependent variable, we need to use the feature selection techniques Stepwise Regression and LASSO Regression.

To evaluate the performance of the multiple regression model, it is necessary to split the data into training and testing sets. To do so, we split the data into training and test sets using the Python library feature. We used 20% of the data to test and 80% to train. The model results are shown in Table1.

The estimation function results from training the model as shown in Eq. 1.

$$\begin{aligned} &LV\ ActivePower(kW) = 50.847803 * Wind\ Speed + 0.30246839* \\ &Wind\ Direction + 0.80591175 * Theoretical\_Power\_Curve- \\ &291.06722981960525 \end{aligned} \tag{1}$$

Table 1. Multiple linear regression model outputs

Coefficients	[50.847803 0.30246839 0.80591175]
Intercept:	−291.06722981960525
Mean Squared Error:	72844
Mean Absolute Error:	137
R-squared:	0.954

3.4 Estimation of the Energy Produced by a Wind Turbine Using a Decision Tree Algorithm

The decision tree algorithm builds a decision tree based on the values of the features and the target variable in a dataset and predicts the target values of new instances by traveling between branches in this tree. The decision tree consists of a set of decision nodes following a tree structure. Each node contains data divided according to a feature value. Each branch leads to the next node or leaf node, which is used to make a prediction.

Evaluating the performance of the decision tree requires splitting the data into training and test sets. To achieve this, we used the Python library function to split the data into training and testing sets. We used 20% for testing and 80% for training. The results are shown in Table 2.

Table 2. The outputs of the decision tree model

Mean Squared Error:	124819
Mean Absolute Error:	152
R-squared:	0.922

3.5 Estimation of the Energy Produced by a Wind Turbine Using a Random Forest Algorithm

The Random Forest Algorithm is an ensemble learning algorithm that is created by combining many decision trees. The Random Forest Algorithm can be used for classification and regression problems. It can be used in two different ways: Random Forest Classifier and Random Forest Regression. This algorithm allows each tree to be trained independently with a random sample and a random set of features. In this way, the trees make different decisions by focusing on different data and a more diverse model is produced.

It is necessary to separate the data into a training set and a test set in order to evaluate the performance of the random forest algorithm. We used the Python library function to split the data into training and test sets. We used 20% of the data for testing and 80% for training. The model outputs are shown in Table 3.

Table 3. Outputs of the Random Forest Model

Mean Squared Error:	61618
Mean Absolute Error:	111
R-squared:	0.962

3.6 Prediction of the Energy Produced by a Wind Turbine Using an Artificial Neural Network Algorithm

An artificial neural network (ANN) is a machine learning technique designed to emulate how the human brain works. Neural networks process data using mathematical operations and algorithms. These operations are performed through connections between artificial neurons organized in various layers.

An artificial neural network algorithm is used with three initial layers, six hidden layers, one final layer, and a linear activation function. We used 20% of the data for the test and 80% of the data for the training. The model outputs are given in Table 4.

Table 4. Output of the artificial neural network model

Mean Squared Error:	167471
Mean Absolute Error:	181
R-squared:	0,903

4 Conclusion

The models were evaluated using the mean squared error, the mean absolute error, and the R-squared as shown in Table 5.

Table 5. Comparison of models performance

	Multiple Linear Reg	Decision Tree	Random Forest	Artificial NN
Mean Squared Error:	72844	124819	61618	167471
Mean Absolute Error:	137	152	111	181
R-squared:	0.954	0.922	0.962	0,903

After the model was created, the test data set was evaluated and R-squared, Mean Squared Error (MSE), and Mean Absolute Error values were calculated. According to the results, the most successful prediction model was the Random Forest method with a higher R-squared value and lower Mean Squared Error and Mean Absolute Error values than the other models.

In future studies, wind turbine energy production forecasts can be found by different demand forecasting methods such as artificial neural networks, metaheuristic algorithms and compared with the results of this study.

References

1. Anıl, Y.: Effects of geothermal production prediction on the energy market with artificial intelligence methods. Bahcesehir University, Ankara, PhD Dissertaion (2022)
2. İnal, S.: Operation of hydroelectric power plants dam reservoirs with artificial intelligence, Master thesis, Ordu University, Ordu (2022)
3. Akpolat, A.N.: Management and control of distributed energy generation systems via artificial intelligence techniques. Marmara University, İstanbul, PhD Dissertaion (2021)
4. Balsever, A.E.: Estimation of energy consumption demand with artificial intelligence methods, Master thesis. İstanbul University, İstanbul (2018)
5. Mohamed, Z., ve Patrick, B.: Forecasting Electricity Consumption: A Comparison of Models for New Zealand. Technological Forecasting and Social Change **72**(8), 1030–1043 (2019)
6. Durğun, S.: Long term estimation of energy demand of Turkey's by artificial intelligent techniques, Master thesis. Necmettin Erbakan University, Konya (2018)
7. Balsever A.E.: Estimation of energy consumption demand with artificial intelligence methods, Master thesis. İstanbul University, İstanbul (2018)
8. Sevdim R.E.: Hybrid of renewable energy resources modelling usage with structural intelligence, Master thesis. İstanbul Gelişim University, İstanbul (2018)
9. Eren, T.: Grey forecasting method in natural gas consumption planning and Turkey implementation, Master thesis. İstanbul Ticaret University, İstanbul (2017)
10. Şenol, Ü.: Wind energy and estimating wind energy potential by artificial neural networks method, Master thesis. Bozok University, Yozgat (2017)
11. Kaggle Homepage, https://www.kaggle.com/datasets, last accessed 24 February 2023

Application of a Fuzzy Based Machine Learning Approach to the Detection of Harmful Algae in Water Monitoring

Stefanie Penzel[1](✉), Mathias Rudolph[1], Helko Borsdorf[2], and Olfa Kanoun[3]

[1] Leipzig University of Applied Science, Karl-Liebknecht-Str. 132, 04277 Leipzig, Germany
Stefanie.penzel@htwk-leipzig.de
[2] Helmholtz Centre for Environmental Research - UFZ, Permoserstr. 15, Leipzig, Germany
[3] University of Technology Chemnitz, Reichenhainer Str. 9, Chemnitz, Germany

Abstract. This paper introduces a new approach for the detection of harmful algae in water monitoring with the application of a fuzzy based machine learning methodic by fuzzily describing the data after feature extraction from the water spectra. The main challenge of this task was to describe the whole transmission and fluorescence spectra to extract the features and build a multidimensional fuzzy pattern classifier that integrates the measurement uncertainties associated with the measurement. The utility and application of such a methodic is to detect with light sensors harmful algae in water monitoring before, for example, algae can pollute the water or toxic algal blooms can cause harm. Using different solutions of a reference substance for alga within chlorophyll a data basis for the learning phase of a classification was generated with the feature extraction. The result is a highly flexible classifier which provides information about unknown spectral data and indicates a possible need for action.

Keywords: Machine learning · Fuzzy logic · Fuzzy pattern classification · Water monitoring

1 Introduction

The increasing occurrence of harmful algae species in waters and the associated negative ecological as well as economic impacts require an innovative approach to detecting and analyzing them. The aim of this paper is to propose a machine learning model methodology with a fuzzy pattern classifier for detection of these harmful algae. Machine learning is already widely used in water quality monitoring and prediction [1]. For assessing the condition of waters and early identification of potential hazards, it requires water monitoring directly in the field. However, especially in on-site analysis, measurement uncertainties may occur, for example due to temperature. In addition, the conditions in the water are in constant change and it is necessary to develop a flexible and realistic methodology that adapts to these boundary conditions. Based on the methods of fuzzy logic according to Zadeh [2], various high-performance approaches are available, by which measurement uncertainties can be integrated and high-dimensional models for

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 181–188, 2023.
https://doi.org/10.1007/978-3-031-39777-6_22

the system state description can be generated. There are already applications of machine learning where fuzzy methods have been integrated [3]. Fuzzy logic has been used for many years for various real-world applications such as monitoring, fault diagnosis, decision support, control, etc. [2]. Most of the methods are based on the fuzzy membership function, which ensures a flexible use, such as neuro-fuzzy systems, fuzzy reinforcement learning [4] and also fuzzy pattern classification [5, 6]. Fuzzy pattern classification refers to the use of fuzzy logic and fuzzy sets in machine learning algorithms for pattern recognition and classification. Fuzzy pattern classification is used in a variety of applications, including, for example, signal processing applications and automation systems [7, 8], pollutant analysis in water [9] or in the field of neural propositions as well as medical diagnostics [5, 10]. In the context of the studies presented here, spectra from dilution series of a reference substance with water were recorded between 300–900 nm on an experimental setup using a combined UV/Vis fluorescence measurement probe. The experimental findings are incorporated into a fuzzy machine learning approach, which uses supervised learning methods to model and record the measurement uncertainties and environmental influences of an on-site analytical procedure for environmental monitoring with the aid of soft computing. In the first step, after a short data preprocessing, the feature extraction of the spectra is performed. In the second step, the learning data resulting from the feature extraction are classified and integrated into a multidimensional class space. The result is a multidimensional fuzzy pattern classifier which, in a subsequent work phase, can utilize new data to generate statements about the current status of the respective water sample. Recommendations for action can then be derived from this. In this paper, after a brief insight into related work in the literature, the used methodology will be presented. Then, the results of the application of this methodology are discussed and finally a short summary is given.

2 Related Work

There are many research approaches and already established methods for water monitoring based on machine learning and/or fuzzy logic. For example, in the article [11] different algorithms based on machine learning to predict algae blooms were generated. The adaptive neuro-fuzzy inference system was named as the best model. The authors of paper [12] describe different modeling approaches for the prediction of water quality parameters. Due to possible uncertainties, an extended wavelet denoising method with Neuro-Fuzzy Inference System is recommended. The article [13] shows a machine learning method for the evaluation and monitoring of water quality characteristics using the example of blue-green algae. In [14] a hybrid model with a combination of data decomposition, fuzzy C-means clustering, and bidirectional gated recurrent unit is presented. In [15] various machine learning models such as multiple linear regression, support vector machine and artificial neural networks are used to predict algal chlorophyll and water clarity. Four machine learning models are investigated in the paper [16] for the assessment of water quality variables for cyanobacterial blue-green algae.

There are many more different approaches for the detection of algae via water monitoring using machine learning and different fuzzy based approaches. The special feature of the approach presented in this work is on the one hand the recording and complex description of the complete water spectra (in contrast to read out the intensity value only at a certain wavelength) and on the other hand the possible integration of measurement uncertainties into the fuzzy pattern classifier using realistic data.

3 Methodology

3.1 Data Generation and Feature Extraction

The data basis is generated at an experimental setup. At this setup, transmission and fluorescence spectra can be recorded in parallel by using electromagnetic radiation. For the transmission spectra a tungsten and deuterium lamp are used as UV/VIS light source. For the fluorescence spectra LEDs at 440 nm and 375 nm are equipped which are placed at an angle of 90 degrees in the experimental setup. The prepared samples are placed in a flow-through cuvette and the spectra are recorded using the spectrometer QMini from Broadcom. For the algae reference substance, diluted solution series were prepared using spinate extract and water. The spinach extract was prepared according to DIN EN 17899 Annex C.3, using water instead of ethanol for the preparation. The Information about the diluted solution series can be found in Table 1 and the recorded spectra are displayed in Fig. 1. In addition, pure water spectra were generated for subsequent further processing of the data. Currently, this experimental setup is being integrated into a submersible probe to record data on-site later.

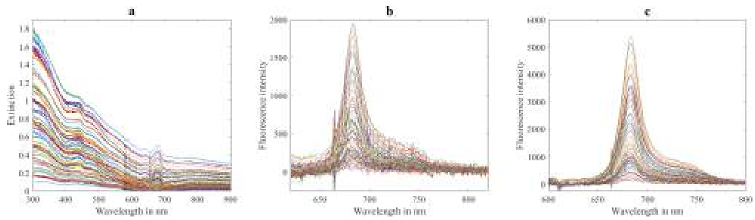


Fig. 1. Recorded spectra of a) extinction spectra with UV/VIS, b) fluorescence spectra with LED 375 nm and c) fluorescence spectra with Led 440 nm.

Table 1. Data recording of the different measuring methods with different light sources.

Measuring method	Lightsource	Solutions	Data points per spectrum
Transmission/Extinction	UV/VIS	(5, 7.5, 10, 15, 20, 22.5, 25, 30, 35, 40, 45, 50) %	2184
Fluorescence	LED 375 nm	(5, 7.5, 10, 15, 20, 22.5, 25, 30, 35, 40, 45, 50) %	808
Fluorescence	LED 440 nm	(5, 7.5, 10, 15, 20, 22.5, 25, 30, 35, 40, 45, 50) %	808

In data pre-processing, the spectra are first filtered with a moving average to minimize noise. Subsequently, the transmission spectra are converted into extinction spectra according to the Beer-Lambert law and the fluorescence or stray light spectra are converted into revised fluorescence spectra by subtracting the water spectrum.

To extract the features, the spectra were described mathematically. For the mathematical description function, a spectrum at 50% solution for each measurement method was taken and calculated by adding Gaussian functions.

$$g_{total,i} = a1 \cdot e^{-\frac{(x-b1)^2}{2 \cdot c1^2}} + a2 \cdot e^{-\frac{(x-b2)^2}{2 \cdot c2^2}} + \ldots + ai \cdot e^{-\frac{(x-bi)^2}{2 \cdot ci^2}} \quad (1)$$

Using this respective overall descriptive function, all other spectral files were fitted using the least squares method. In each case, an R-squared value is output, which gives the degree of possible membership of the respective substance for the spectrum. This was done for all three measurement methods, so that three features per spectrum were extracted for each dilution series value. In Fig. 2, the features for a very low solutions (5%) value and a very high solutions (50%) value are given as examples. The extracted features represent the learning data set for building the classifier. Since there are three features involved, a three-dimensional feature space is required for the classification. As can be seen in the figure, the characteristics differ although they originate from the same solutions values. This can be attributed to the variable substance in the water. Due to this, a fuzzy classification is preferred.

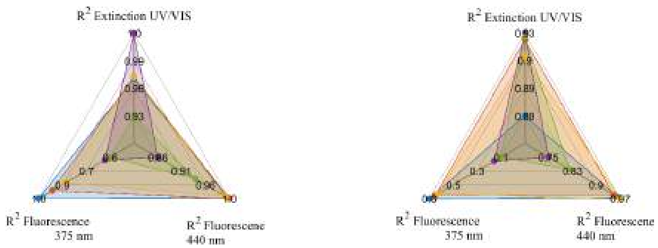


Fig. 2. Example of the three characteristics for a low concentration (dilution series at 5%) on the left and for a high concentration (dilution series at 50%) on the right.

3.2 Structure of the Classification Model

The classification was implemented using a fuzzy approach. As mentioned earlier, the latter was chosen for the application discussed here. This pattern-based classification approach offers the possibility to model mutual dependencies between nature/human and technical variables in a very effective and flexible way. Moreover, expert knowledge can be integrated, and the methodology can work for small data sets as well as very big data sets. Overall, data- and knowledge-based approaches are combined, considering possible fuzziness due to measurement uncertainties.

Fuzzy pattern classification was introduced by Bocklisch and colleagues [17]. The modeling concept is divided into a learning phase and a working phase. In the learning phase, the features are first sharply grouped and then transformed into fuzzy classes using a fuzzy membership function according to AIZERMAN. The simplified one-dimensional formula for this is shown here.

$$\mu(\mathrm{u}) = \frac{a}{1 + \left(\frac{1}{b} - 1\right) \cdot \left(\frac{|u-u_0|}{c}\right)^d} \tag{2}$$

This membership function can be adapted (asymmetrically) with function parameters. The initially one-dimensional functions are then combined into multi-dimensional fuzzy pattern classes. In the working phase unknown working data are read in and assigned to the correct class with a membership of the data points. Due to the flexible modeling and adaptation of different data and applications of the real world, this approach represents an optimal evaluation methodology for the constantly changing water spectra including the numerous occurring measurement uncertainties on-site.

4 Results

Several spectral data with the alga reference substance were acquired as described in Table 1 with UV/VIS, with the LED 440 nm and with the LED 375 nm. These were further processed and features R-squared were extracted from each spectrum. This was used to generate a learning dataset, which was subsequently used in the classification. Table 2 shows a section into the learning dataset after data preprocessing and feature extraction.

Table 2. Excerpt from the learning data set for the classifier construction.

Object number	R^2 extinction UV/VIS	R^2 fluorescence LED 440 nm	R^2 fluorescence LED 375 nm
1	0.886920163	0.968236218	0.629016872
2	0.916924554	0.9804423	0.858470063
3	0.924464757	0.98785643	0.897418698
…	…	…	…
60	0.962591762	0.999458571	0.914236156

Sharp classification is performed using cluster analysis and expert knowledge (see Fig. 3). The meaning of the classes is defined as follows. Class 1: alga reference substance is uncertainly present, class 2: alga reference substance is present and class 3: alga reference substance is certainly present.

The exact procedure with use of the AIZERMAN function is described in [5, 6]. With the parameter c the measurement uncertainties can be considered and integrated. Figure 4 represent the fuzzy pattern classifier. Here, two features are shown in each figure. The entire fuzzy pattern classifier consists of four dimensions.

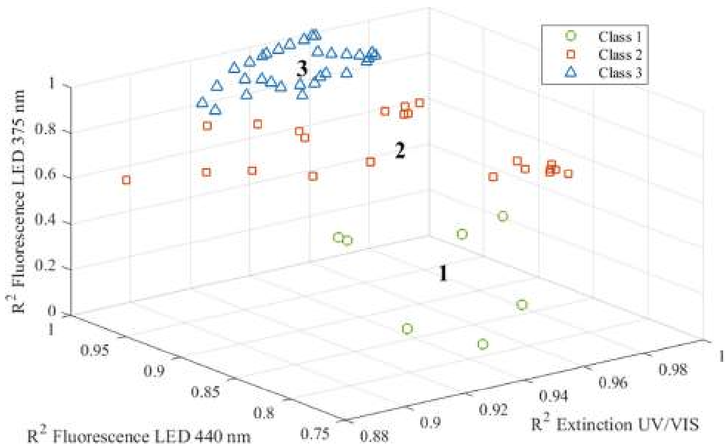


Fig. 3. Sharp classification of the feature using cluster analysis and implemented expert knowledge.

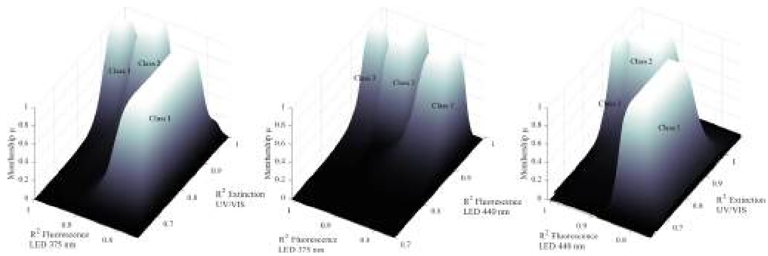


Fig. 4. Representation of the four-dimensional fuzzy pattern classifier divided into three dimensional images.

A working data set was generated for the working phase. For this purpose, solutions with the alga reference substance were prepared again and measured by the experimental set-up with UV/VIS in transmission and with the LEDs at 375 nm and 440 nm in fluorescence. Data pre-processing and feature extraction were performed analogously to the learning dataset. The extracted features are now read into the fuzzy pattern classifier resulting from the learning phase and the membership values for the classes are calculated. The assignment of the features is done by the respective highest membership value to the respective class. The working phase is exemplarily shown in three-dimensional view (only with the x-axis of UV/VIS measurement and the y-axis with LED measurement at 440 nm) in Fig. 5, where the objects of the working phase are displays as blue cross. The calculation is done in the four-dimensional classifier. In Table 3, the membership values of the features from the work phase are assigned to the respective classes.

As shown in Table 3 and Fig. 5, object 1 is assigned to class 2 with a very high affiliation. Here it can be assumed that an algae reference substance is present. In case of an assignment to class 3, as it is calculated for object 2, a pollutant treatment should

Table 3. Membership values of the individual objects from the working data set to the classes.

Object number	Assigned class	Membership to class 1	Membership to class 2	Membership to class 3
1	2	1.665321e-07	9.166185e-01	3.761218e-08
2	3	6.080647e-13	2.911444e-05	8.244349e-01
3	1	6.247900e-01	3.253890e-07	2.116445e-10

be advised. In comparison, object 3 has a rather low membership value to class 1, so that an algae reference substance is only present with uncertainty. Here further water monitoring should be recommended.

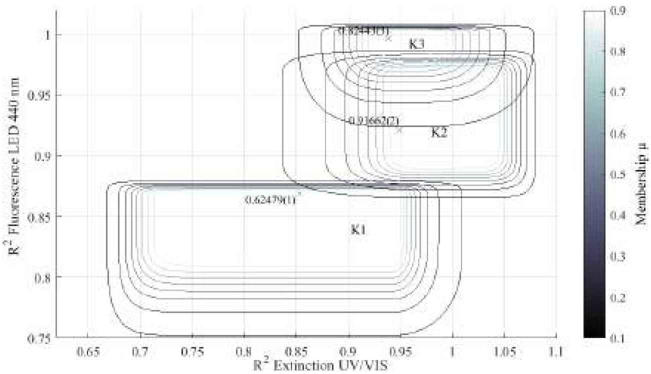


Fig. 5. Graphical representations in the top view of the assignment of test data in the working phase.

5 Conclusion

In this paper, the use of a fuzzy based machine learning approach for the detection of harmful algae in water was investigated. For this purpose, an algorithm was presented which uses data pre-processing and feature extraction of the spectra to create a learning dataset. This represents the basis for a fuzzy pattern classifier. In this classifier, class membership is expressed by the degree of membership in a fuzzy set and not by a binary yes or no decision. The main result of this work is a flexible and robust classification model that can deal with uncertainties and imprecise data, which play a major role in the case of water spectral data, especially for on-site recording. In the created fuzzy pattern classifier, unknown data can be read in after feature extraction and assigned to classes automatically. Thus, a statement about the possible presence of algae substances can be made and actions such as pollutant treatment can be recommended. In future work, even more data will be collected to improve the classifier and the use of on-site recorded data will be tested.

References

1. Zhu, M., et al.: A review of the application of machine learning in water quality evaluation. Eco-Env. Health **1**(2), 107–116 (2022). https://doi.org/10.1016/j.eehl.2022.06.001
2. Zadeh, L.A.: Fuzzy Sets. Inf. Control **8**, 338–353 (1965). https://doi.org/10.1016/S0019-9958(65)90241-X
3. Hüllermeier, E.: Fuzzy methods in machine learning and data mining: status and prospects. Fuzzy Sets Syst. **156**(3), 387–406 (2005). https://doi.org/10.1016/j.fss.2005.05.036
4. Zadeh, L.A., Aliev, R.A.: Fuzzy Logic Theory and Applications Part I and Part II. World Scientific Publishing, New Jersey (2019)
5. Bocklisch, S.F., Bocklisch, F., Beggiato, M., Krems, J.F.: Adaptive fuzzy pattern classification for the online detection of driver lane change intention. Neurocomputing **262**, 148–158 (2017). https://doi.org/10.1016/j.neucom.2017.02.089
6. Hempel, J.A., Bocklisch, S.F.: Fuzzy pattern modelling of data inherent structures based on aggregation of data with heterogeneous fuzziness. In: Romero, G. (ed.) Modelling Simulation and Optimization, pp. 637–655. Intechopen, London (2017)
7. Lohweg, V., Diederichs, C., Müller, D.: Algorithms for hardware-based pattern recognition. EURASIP J. Appl. Signal Process. **2004**, 1912–1920 (2004). https://doi.org/10.1155/S1110865704404247
8. Mönks, U., Lohweg, V., Larsen, H.L.: Aggregation operator based fuzzy pattern classifier design, machine learning in real-time applications (MLRTA 09). In: KI 2009 Workshop, Paderborn, 15th Sep 2009
9. Penzel, S., Rudolph, M., Borsdorf H., Kanoun, O.: Prototypical investigation of the use of fuzzy measurement data in a case study in water analysis. In: Ganzha, M., Maciaszek, L., Paprzycki, M., Ślęzak, D (eds.) Position Papers of the 17th Conference on Computer Science and Intelligence Systems, ACSIS, vol. 31, pp. 27–33 (2022)
10. Bocklisch, F., Hausmann, D.: Multidimensional fuzzy pattern classifier sequences for medical diagnostic reasoning. Appl. Soft Comput. **66**, 297–310 (2018). https://doi.org/10.1016/j.asoc.2018.02.041
11. Ly, Q.V., et al.: Application of Machine Learning for eutrophication analysis and algal bloom prediction in an urban river: a 10-year study of the Han River, South Korea. Sci. Total Env. **797**, 149040 (2021). https://doi.org/10.1016/j.scitotenv.2021.149040
12. Najah, A., et al.: Machine learning methods for better water quality prediction. J. Hydrol. **578**, 124084 (2019). https://doi.org/10.1016/j.jhydrol.2019.124084
13. Saboe, D., et al.: Real-time monitoring and prediction of water quality parameters and algae concentrations using microbial potentiometric sensor signals and machine learning tools. Sci. Total Env. **764**, 142876 (2021). https://doi.org/10.1016/j.scitotenv.2020.142876
14. Yu, J.-W., et al.: Water quality forecasting based on data decomposition, fuzzy clustering and deep learning neural network. Env. Pollut. **303**, 119136 (2022). https://doi.org/10.1016/j.envpol.2022.119136
15. Mamun, M., et al.: Prediction of algal chlorophyll-a and water clarity in monsoon-region reservoir using machine learning approaches. Water **12**(1), 30 (2019). https://doi.org/10.3390/w12010030
16. Heddam, S., Yaseen, Z.M., Falah, M.W., et al.: Cyanobacteria blue-green algae prediction enhancement using hybrid machine learning–based gamma test variable selection and empirical wavelet transform. Environ. Sci. Pollut. Res. **29**, 77157–77187 (2022). https://doi.org/10.1007/s11356-022-21201-1
17. Bocklisch, S.F.: Prozeßanalyse mit unscharfen Verfahren. Verlag Technik, Berlin (1987)

Flight Gate Assignment Problem with Reinforcement Learning

Müge Muhafız Yıldız[1(✉)], Umut Avcı[1], Mustafa Arslan Örnek[1], and Cemalettin Öztürk[2]

[1] Yaşar University, Izmir, Turkey
muge.yildiz@stu.yasar.edu.tr, {umut.avci,arslan.ornek}@yasar.edu.tr
[2] Munster Technical University, Cork, Ireland
Cemalettin.Ozturk@mtu.ie

Abstract. The operation of an airport is a very complex task involving many actors. The primary mission of airport management is to provide sufficient capacity and the best working conditions to all airlines, ground handling, and service provider companies. Flight gate assignment is one of the essential planning problems airport management needs to address, assigning incoming aircraft to the available gates or stands while satisfying operational constraints. Generally, flight arrivals and departures are considered deterministic, and various operational research methods have been applied to solve this combinatorial problem. However, in real-life scenarios, deterministic solutions are generally infeasible because arrival and departure times are uncertain. It is crucial to deal with these uncertainties to create a robust schedule. In this study, we develop a Reinforcement Learning (RL) algorithm to solve the flight gate assignment problem since it is a sequential decision-making method and allows adaptive solutions to address urgent and frequent changes.

Keywords: Flight gate assignment · Reinforcement learning · Robust scheduling

1 Introduction and Background

The aviation industry has experienced significant growth in recent years due to the world's population and standard of living increase. Airports are once considered the most critical infrastructure providers for the industry. Efficient and reliable airport operations are crucial for the viability and further development of individual airports and the air transport industry [6].

The flight gate assignment problem deals with assigning incoming aircraft to the available gates or stands while ensuring several operational constraints. Since this assignment significantly impacts airport traffic and customer satisfaction, it can be considered one of the airports' most important daily decisions. In this study, it is focused on the flight gate assignment problem, which is critical for airlines and whole airway traffic.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 189–196, 2023.
https://doi.org/10.1007/978-3-031-39777-6_23

Flight gate assignment can be formulated as a deterministic or stochastic optimization problem. There are many studies on flight gate assignment problems, and most of them do not consider the stochasticity of flight arrivals and departures. A robust gate assignment helps the decision-makers deal with uncertain events. Hassounah and Steuart [5] studied robust and stochastic flight gate scheduling models and planned buffer times to provide a robust schedule. Yan and Chang [10] and Yan and Huo [11] used fixed buffer time between two consecutive flights of a common gate to absorb uncertainty in flight arrivals. Yan and Tang [12] developed a heuristic approach sensitive to stochastic flight delays. The heuristic approach comprises a stochastic gate assignment model, a real-time assignment rule, and two penalty adjustment methods. Seker and Noyan [9] developed a stochastic programming model which aims to achieve a robust schedule minimizing the number of conflicting flights, and maximizing idle and buffer times. The authors formulated the problem as a large-scale mixed-integer programming problem and resolved it by tabu search. Lim and Wang [8] modeled the gate assignment problem as a stochastic programming model and transformed it into a binary programming model using estimation functions on gate conflicts but surely it did not provide a solution in polynomial time. Accordingly, they proposed a hybrid meta-heuristic combining a tabu search and a local search for resolution using unsupervised estimation functions without knowing the real-time arrival and departure times in advance. Aoun and El Afia [1] modeled the flight gate assignment problem as a Markov Decision Process (MDP) to provide a robust solution that considers possible flight delays. Later, Aoun and El Afia [2] developed the same problem as a Multi-Agent Markov Decision Process by taking the gates as independent agents to cope with stochastic disturbance.

Although studies model the problem as an MDP, to the best of the authors' knowledge, there is no study on applying reinforcement learning to the flight gate assignment problem. In the literature, only Li et al. [7] used RL to solve the gate assignment problem within their tabu search algorithm. They employed an RL algorithm to guide the tabu search toward new promising solutions.

Reinforcement Learning is a form of unsupervised learning method that allows the agent to learn through trial and error, using feedback from its own experiences. The agent interacts with its environment and through this experience, the environment's state changes and the agent gets a reward. The agent uses this feedback from the environment to decide further actions. RL is a convenient technique to solve sequential decision-making problems.

Briefly, there is a gap in the flight gate assignment problem to employ a reinforcement learning algorithm to find a robust schedule as it is allowing decision makers to tackle this dynamic and stochastic combinatorial problem. Thus, in this study, it is focused on applying RL to solve the flight gate assignment problem.

2 Flight Gate Assignment Problem Formulated as an RL-Problem

2.1 Problem Formulation

During a time horizon, many flights are expected to arrive and depart. At each decision point t, operation controllers at the airport (agent) decide to which gate the incoming flights should be assigned or which gate should be discharged by changing the gate's status and the total number of flights assigned to a gate until time t. Once an incoming flight arrived, if all the gates are already occupied, then this flight has to be assigned to the stand. It is assumed that the stand has infinite capacity.

- **State:** A representation of the airport environment with the total number of assigned flights and a set of occupied gates. s_t contains the number of already assigned flights until time t and a list of gates that are occupied at time t.

  $$S = (n_s; O_s) \colon n \in 1..m, O \parallel\ G$$

  where m is the total number of flights expected to arrive during the time horizon and G is the set of all subsets of gates in the airport. Note that G does not include stands. For example, if there are only 2 gates in the airport; Gate 1 and Gate 2, then G will be the following:

  $$G = (\leq \{1\}, \{2\}, \{1, 2\})$$

  Thus, at time t, $s_t = (4; \{1, 2\})$ means that until time t, there are 4 flights assigned to any gate (including stand), and currently, at time t, Gates 1 and 2 are occupied by some flights.
- **Action:** A finite set of actions A of assigning arriving flights to gates and discharging the gates due to departures. A could be the set of all gates plus the stand; one for assigning and one for discharging. For example, if there are only 2 gates in the airport; Gate 1 and Gate 2, then A will be the following:

  $$A = \{G1, G2, S, SD, G1D, G2D\}$$

  Actions $G1, G2$ and S are for assigning an incoming flight to Gate 1, Gate 2, and Stand, respectively. Actions $G1D, G2D$ and SD are for discharging Gate 1, Gate 2, and the stand, respectively.
- **Reward:** Feedback from the environment, the reward after transitioning from state s to state $s^{'}$ with action a. It is aimed to minimize the number of ungated flights by forcing high penalties for assigning an arriving flight to an already

occupied gate or stand. Besides the utility of assigning an arriving flight to an idle gate is gained by giving higher rewards for this action. Additionally, it is aimed to satisfy robustness in the schedule with the aid of reward calculation. In the literature, the main objectives for robustness in flight gate assignment problems were:

- Minimize idle time (Bolat [3])
- Minimize gate conflicts (Lim and Wang [8], Castaing et al. [4], Seker and Noyan [9])

Robustness is principally satisfied with leaving enough slack time between consecutive flights at the same gate. However, leaving a large slack time causes a decrease in gate utilization, and on the other hand, leaving a small slack time may cause gate conflicts in case of possible delays in the arrival and departure of the flights. Thus, it is important to decide when and how much slack time to allow between consecutive flights at the same gate. Step-wise reward calculation is used since a tighter slack time is not useful for robustness and a larger slack time is not efficient for airport operations.

2.2 Reward Recalculation and Step-Wise Reward Function

If an incoming flight is assigned to a gate after another flight, the reward is recalculated in order to force to leave a reasonable slack time between consecutive flights to satisfy both robustness and gate utilization.

$$AT_i : \text{Planned arrival time of flight } i$$

$$DT_i : \text{Planned departure time of flight } i$$

$$DC_j : \text{Latest discharge time of gate} j$$

Once a flight is assigned to a gate, the latest discharge time of its gate will be set as the planned departure time of that flight.

$$DC_j = DT_i$$

Let x_k be the breakpoints in the step-wise reward function where $k = 1 \dots n$ where n is the number of breakpoints. 3 breakpoints are used in our calculation. If the action is assigning flight i to gate j,

$$R = \begin{cases} f(r_1), \; AT_i - DC_j < x_1 \\ f(r_2), \; AT_i - DC_j \geq x_1 \;\&\; AT_i - DC_j < x_2 \\ f(r_3), \; AT_i - DC_j \geq x_2 \end{cases}$$

This is a generic format of step-wise reward calculation. In the algorithm, the functions are defined as follows:

$$f(r_1) = f(r_3) = AT_i - DC_j$$
$$f(r_2) = (AT_i - DC_j)^2$$

3 Q-Learning Algorithm-Training

AT_i : Planned arrival time of flight i
DT_i : Planned departure time of flight i
DC_j : Latest discharge time of gate j

- Set each of the arrival and departure times as the decision points in increasing order
- Generate action set; one assignment and one discharge action for each gate
- Generate all possible states with two elements in each of them:
  - 1^{st} element is the total number of flights that are assigned till the current state
  - 2^{nd} element is the set of busy gates at the current state
- Generate state transition matrix for each state and action couples
- Generate reward matrix for each state and action couples to give higher reward for assigning the flights to the available gates
- Generate the initial Q-table corresponding to the state transition matrix
- Set the initial state as $(0, \trianglelefteq)$
- Start from the first decision point
  - If the decision point is an arrival of a flight
    - * Choose an assignment action with the highest Q-value among the Gates such that $DC_j \leq AT_i$
    - * If $DC_j > 0$, recalculate reward with buffer (buffer $= AT_i - DC_j$))
    - * Set $DC_j = DT_i$
    - * Increase the total number of flights that are assigned
  - If the decision point is a departure of a flight
    - * Set the action as the discharge of the gate that the flight has been assigned to
- Update Q-table
- Update the current state using chosen action through the state transition matrix
- If all flights have not been assigned yet, move on to the next decision point, else repeat starting from the first decision point until the agent learns (40000 times i.e.)

3.1 Application of the Algorithm

In this study, Q-learning algorithm is applied for training on a randomly generated small data set with a size of 6 flights and 3 gates through Python 3.9.7 using Spyder IDE 5.1.5. The algorithm has been run 40000 times to learn. The algorithm is terminated with the condition that the total number of assigned flights achieves 6. An example flight schedule that is worked on is given in Table 1.

Table 1. Training Data

Flight	Arrival Time	Departure Time
1	35	165
2	85	220
3	200	370
4	205	335
5	235	425
6	265	440

After sorting the arrival and departure times, the decision points are obtained. The action set is generated based on the gate information. Initially the latest discharge time DC_j and the latest assigned flight F_j of all gates are set as $j = 1..M$ 0.

Leaving less than 15 min between two consecutive flights at the same gate is risky in case of possible delays and leaving more than 45 min may cause a decrease in the utilization of gates. Therefore, for the step-wise reward calculation, the reward function is adjusted in order to avoid the slack being less than 15 min and more than 45 min. The reward is promoted for the slacks between 15 min and 45 min as follows:

$$R = \begin{cases} AT_i - DC_j, \ AT_i - DC_j < 15 \\ (AT_i - DC_j)^2, \ AT_i - DC_j \geq 15 \ \& \ AT_i - DC_j < 45 \\ AT_i - DC_j, \ AT_i - DC_j \geq 45 \end{cases}$$

For the 1^{st} run, the algorithm starts from the first decision point and increases by one until all the flights are assigned. When all the flights are assigned, it is started from the first decision point and initial state until this loop works for 40000 times.

4 Q-Learning Algorithm-Test Part

In the test algorithm, for the newly generated data set, the agent only follows the Q-table for selecting the action that gives the highest Q-value and moving to the next state by following the state-transition matrix obtained during the training algorithm.

4.1 Application of Algorithm

Any data can be tested with the same size as training data very quickly. You can find the example of a test data set in Table 2.

Table 2. Test Data

Flight	Arrival Time	Departure Time
1	30	190
2	75	265
3	95	255
4	125	285
5	300	500
6	305	495

The Q-table that we obtained at the end of the training algorithm is followed and assignments are completed in less than 1 s (0,18 s). Final flight gate assignments are given in Table 3

Table 3. Assignment Results of Test Data

Flight	Arrival Time	Departure Time	Gate
1	30	190	1
2	75	265	2
3	95	255	3
4	125	285	S
5	300	500	1
6	305	495	2

5 Conclusion

There are many studies on airport resource allocation problems using machine learning, but in the specific area of flight gate assignment, we have yet to find a study that has used machine learning. Hence, there is a gap in the literature for flight gate assignment problems with machine learning techniques. Although the problem may be solved in a very short time through the test algorithm, it takes a longer time to train the agent since state space explodes as the number of gates increases. Hence, we have a time complexity issue for data sets with more than 12 gates. As a future study, we will work on different reinforcement learning algorithms and compare the results gathered with different algorithms. In addition to that, we aim to apply RL algorithms to challenging real-world data.

References

1. Aoun, O., El Afia, A.: Using Markov decision processes to solve stochastic gate assignment problem. In: Proceedings of 2nd IEEE International Conference on Logistics Operations Management, GOL 2014, pp. 42–47 (2014). https://doi.org/10.1109/GOL.2014.6887445
2. Aoun, O., El Afia, A.: Time-dependence in multi-agent MDP applied to gate assignment problem. Int. J. Adv. Comput. Sci. Appl. **9**, 331–340 (2018). https://doi.org/10.14569/IJACSA.2018.090247
3. Bolat, A.: Assigning arriving flights at an airport to the available gates. J. Oper. Res. Soc. **50**, 23–34 (1999). https://doi.org/10.1057/palgrave.jors.2600655
4. Castaing, J., Mukherjee, I., Cohn, A., Hurwitz, L., Nguyen, A., Müller, J.: Reducing airport gate blockage in passenger aviation models and analysis. Comput. Oper. Res. **65**, 189–199 (2014). https://doi.org/10.1016/j.cor.2014.02.011
5. Hassounah, M.I., Steuart, G.N.: Demand for aircraft gates. Transp. Res. Rec. **1423**, 26–33 (1993)
6. Justesen, T.F.: Allocation of Ground Handling Resources at Copenhagen Airport. DTU Management Engineering (2014)
7. Li, M., Hao, J.-K., Wu, Q.: Learning-driven feasible and infeasible Tabu search for airport gate assignment. Eur. J. Oper. Res. (2021). https://doi.org/10.1016/j.ejor.2021.12.019
8. Lim, A., Wang, F.: Robust airport gate assignment. In: Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI 2005, pp. 74–81 (2005). https://doi.org/10.1109/ICTAI.2005.110
9. Şeker, M., Noyan, N.: Stochastic optimization models for the airport gate assignment problem. Transp. Res. Part E-Logist. Transp. Rev. **48**, 438–459 (2012). https://doi.org/10.1016/j.tre.2011.10.008
10. Yan, S., Chang, C.-M.: A network model for gate assignment. J. Adv. Transp. **32**, 176–189 (1998). https://doi.org/10.1002/atr.5670320204
11. Yan, S., Huo, C.-M.: Optimization of multiple objective gate assignments. Transp. Res. Part A Policy Pract. **35**, 413–432 (2001). https://doi.org/10.1016/S0965-8564(99)00065-8
12. Yan, S., Tang, C.: A heuristic approach for airport gate assignments for stochastic flight delays. Eur. J. Oper. Res. **180**, 547–567 (2007). https://doi.org/10.1016/j.ejor.2006.05.002

Clustering AirBnB Listings in Istanbul Using Unsupervised Machine Learning Techniques

Mehmet Ata Özçini(✉), Tuna Yılmaz, and Tolga Kaya

Department of Management Engineering, Istanbul Technical University, 34367 Istanbul, Turkey
{ozcini19,yilmaztun19,kayatolga}@itu.edu.tr

Abstract. The rapid expansion of Airbnb and the increase in the number of listings on the platform have raised concerns about the impact on housing markets and urban dynamics, particularly in regards to housing affordability. This study aims to examine the relationship between Airbnb listings in Istanbul, Turkey, a major tourist destination with a high demand for accommodations and a significant housing affordability problem. Using unsupervised learning techniques, the study clusters and analyzes Airbnb listings in Istanbul to identify patterns and trends in the data. In the study, k-means and fuzzy c-means algorithms are applied to group the data into various numbers of clusters and the results were evaluated using the Silhouette Score and the Dunn Index methods. Results showed that k-means clustering algorithm could be preferred to fuzzy c-means, with the optimal number of clusters being between two and five. Authors selected k-means algorithm with five clusters for further analysis and named the resulting clusters based on their most important characteristics, providing an overview of what to expect from listings within each group, revealing valuable insights for researchers and stakeholders in the housing market.

Keywords: Housing · AirBnB · Machine Learning · Unsupervised Learning · Clustering · Fuzzy C-Means · K-Means

1 Introduction

Airbnb is a platform that enables individuals to rent out their homes or apartments for short-term stays, and has become a major player in the hospitality industry since its launch in 2008. It has disrupted the traditional hotel business model, but has also raised concerns about its impact on housing markets and urban dynamics. Housing affordability is a critical issue in many cities worldwide, and the rise of Airbnb has been linked to rising housing prices and decreased availability of rental units for long-term residents [1].

The study aims to provide insights for policymakers and stakeholders in cities facing challenges related to short-term rentals by examining the relationship between Airbnb listings in a city with a high demand for accommodations and a shortage of affordable housing options. To our knowledge, this is one of the first studies which clusters AirBnB listings in Turkey, particularly in Istanbul. Istanbul is an interesting case study as it is

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 197–205, 2023.
https://doi.org/10.1007/978-3-031-39777-6_24

a major tourist destination but also has a significant housing affordability problem. To achieve this, to provide insights, this study examines the relationship between Airbnb listings in Istanbul, Turkey, using unsupervised learning techniques, such as K-Means and Fuzzy C-Means clustering algorithms. In addition, to assess the resulting clusters and determine the most suitable outcome, Silhouette and Dunn indices are utilized.

The rest of this paper is organized as follows: second section provides a thorough review of pertinent literature. Third section describes the technique used in this project. Fourth section describes the data preparation approach that was used. The study's findings are given in section five. Finally, section six summarizes the results and makes recommendations for further research.

2 Literature Review

Airbnb, a public marketplace, is a meeting point for people looking for accommodation for different purposes and people who want to rent their homes for short or long term. Choudary defines Airbnb as an intermediary that enables trade by connecting hosts and travelers [2]. Customers can access and select listings according to their needs and desires by using filtering methods such as location, size, daily rent, and etc. Airbnb plays an active role in obtaining information about local prices and even in setting prices as a further step, as it has witnessed explosive growth rates and more and more travelers are attracted to the company [3].

For the remaining literature review part, similar prior studies were filtered out and it has been observed that different algorithms and models have been used in previous studies. For example, similar researches were conducted in different metropolitan areas. A study that looked at regional property prices in the boroughs of inner London discovered a significant and modestly positive correlation between the frequency of Airbnb rentals and the boroughs' average house prices per square meter [4]. In the research of Schäfer and Braun, which deals with the removal of the shares of residential apartments from the housing market as a result of short-term rentals for tourists in Berlin and the increasing rents due to this reduction in housing supply, ArcGIS was used and variance analyzes were conducted [5]. An analysis of another touristic city, Barcelona, was carried out by Garcia-López et al. The published article shows that rents rose by an average of 1.9%, transaction prices by 4.6%, advertisement prices by 3.7% as a result of Airbnb activity in local areas and that these rates increased in areas with more dense Airbnb activity [6]. Different methods were used during the research, one of which is clustering, similar to this study. In addition, it is suggested that different clustering techniques such as K-means and C-means clustering can be used to further investigate the impact of increasing use of short-term rental platforms on housing markets [7].

Based on this literature review, it is apparent that while there are studies that have examined Airbnb, there remains a gap in research considering the unsupervised clustering of Airbnb listings in order to uncover distinct patterns and attributes within the data. By conducting such analyses, future researchers may gain a more effective and efficient understanding of how to segment and analyze listings data.

3 Methodology

3.1 K-Means

K-Means is an unsupervised learning algorithm for clustering data into K distinct groups [8]. It is one of the most popular partitioning method of clustering and it was firstly proposed by MacQueen in 1967. The main goal is to improve similarity within clusters while minimizing variation between them. To achieve this, the algorithm follows these steps: first, the number of clusters *(k)* are established and all observations are randomly assigned to them. The cluster centers *(M_i)* are then identified, and the distances between these centers and each data point *($d(M_i, x)^2$)* are calculated. Lastly, each data point is given to the cluster whose center is closest to it [9]. The main formula for this algorithm can be shown as follows.

$$E = \sum_{i=1}^{K} \sum_{p \in C_i}^{n} d(M_i, x)^2 \tag{1}$$

3.2 Fuzzy C-Means

Fuzzy C-Means (FCM) is an extension of the K-Means clustering algorithm that allows for overlapping or fuzzy clusters, rather than strict, discrete clusters. The Fuzzy C-Mean approach finds the position of each cluster through maximum membership defuzzification and neighborhood smoothing methods [10].

Like K-Means, FCM works by iteratively assigning each data point to the cluster with the closest mean. However, in FCM, each data point can belong to more than one cluster, with a degree of membership assigned to each cluster. This degree of membership is represented by a membership value, which ranges from 0 to 1 and indicates the degree to which a data point belongs to a particular cluster.

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m ||x_i - c_j||^2, 1 < m < \infty \tag{2}$$

The U membership matrix is assigned at random to begin the process. The vectors for the centers are then computed in the second stage.

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m x_j}{\sum_{i=1}^{N} u_{ij}^m} \tag{3}$$

Based on the computed cluster centers, the U matrix is updated using the equation given above. The difference between the old and new U matrices are compared, and the procedure is repeated until the difference is less than the stopping condition.

$$u_{ij} = \frac{1}{\sum_{i=1}^{n} \frac{||x_i - c_i||^{2/m-1}}{||x_i - c_k||}} \tag{4}$$

3.3 Silhouette Index

The Silhouette Index is a measure of the quality of a clustering solution. It is used to assess how well the data points in a dataset are assigned to their respective clusters. A high Silhouette Index indicates that the data points are well-separated within their own clusters and are not significantly overlapped with other clusters. Each cluster is depicted by a silhouette, which is determined by comparing its tightness and separation. This silhouette indicates which objects are well within their cluster and which are simply in between clusters [11].

$$\boldsymbol{S}(\boldsymbol{x}) = \frac{\boldsymbol{b}(\boldsymbol{x}) - \boldsymbol{a}(\boldsymbol{x})}{\boldsymbol{max}(\boldsymbol{a}(\boldsymbol{x}), \boldsymbol{b}(\boldsymbol{x}))} \tag{5}$$

3.4 Dunn Index

The Dunn Index is a measure of the compactness and separation of clusters in a dataset. It is used to assess the quality of a clustering solution and to determine the optimal number of clusters for a particular dataset. A high Dunn Index indicates that the clusters in the dataset are well-separated and compact, while a low Dunn Index indicates that the clusters are poorly separated and dispersed. This indicator employs standard inter-set distance and set diameter measurements [12].

$$D = \frac{min_{1<i<j<n} d(i, j)}{max_{1<k<n} d'(k)} \tag{6}$$

4 Data

There are four categorical variables in the dataset, which consists from thirty-six different variables in total, where the other variables are numeric or Boolean. The variables "neighborhood_cleansed, property_type, room_type, host_response_time" are the categorical variables and are encoded before being introduced into the model. Since "host_response_time" has an ordinal relationship, it was encoded via ordinal encoding, while one-hot encoding was applied to the other categorical variables.

Variable	**Explanation**
host_response_time	The time it takes for the host to respond to a customer enquiry
host_response_rate	Host's responsiveness rate to customer requests
host_acceptance_rate	Host response rate to customer requests
host_is_superhost	Whether the host is classed as a "super host"* *For a host to become a Super host, they must meet certain eligibility requirements, including a high average rating from guests, a low cancellation rate, and a high response rate
host_listings_count	Number of listings owned by the host

(*continued*)

(*continued*)

host_total_listings_count	Total number of listings of the host
host_has_profile_pic	Whether there is a profile picture in the host's profile
host_identity_verified	Whether the host's identity has been verified

Variable	**Explanation**
neighbourhood_cleansed	District where the property is located
property_type	Type of property available for renting
room_type	Type of room available for hire
accommodates	Maximum number of people that can be accommodated in the property
bedrooms	Number of bedrooms in the property
beds	Total number of beds in the property
bathrooms	Total number of bathrooms in the property
amenities	Amenities offered in the property
price	Rental fee
minimum_nights	Minimum number of nights of accommodation
maximum_nights	Maximum number of nights of accommodation
has_availability	Whether the property is available
availability_90	Availability of the property in the last 90 days
availability_365	Availability of the property in the last 365 days

Variable	**Explanation**
number_of_reviews	Total number of reviews about the property
number_of_reviews_ltm	Number of reviews about the property in the last 12 months
number_of_reviews_30d	Number of comments about the property in the last 30 days
review_scores_rating	Property rating (values distributed on a 0–5 scale.)
review_scores_accuracy	Rate of whether guests think the information listed by the accommodation is accurate and complete (values distributed on a 0–5 scale.)
review_scores_cleanliness	Rate of cleanliness level in the accommodation unit (values distributed on a 0–5 scale.)
review_scores_checkin	Guest reviews on the check-in process (values distributed on a 0–5 scale.)
review_scores_communication	Guest reviews of the way they communicate with the host and answer their questions about the accommodation (values distributed on a 0–5 scale.)

(*continued*)

(*continued*)

review_scores_location	Reviews of the location of the accommodation (values distributed on a 0–5 scale.)
review_scores_value	Guest reviews on the price-performance ratio of the accommodation (values distributed on a 0–5 scale.)
instant_bookable	Whether the accommodation unit can be booked instantly
calculated_host_listings_count	Number of accommodations owned by a host
reviews_per_month	Average number of monthly reviews received by an accommodation
host_since_year	The year a host joined the platform and started hosting guests

5 An Application and Findings

5.1 K-Means and Fuzzy C-Means Model Building and Comparison

The study employed both k-means and fuzzy c-means clustering algorithms to group the data into various numbers of clusters. The results of these analysis were evaluated using two commonly used metrics: the silhouette score and the Dunn index. The silhouette score measures the similarity of points within and between clusters, with a score of 1 indicating perfect separation and −1 indicating closeness to another cluster. Whereas, the Dunn index measures the compactness and separation of clusters, with a higher score indicating more distinct and compact clusters.

The first approach was to one-hot encode categorical variables with no ordinal relationships. In this approach, both methods displayed a similar pattern, with the highest scores occurring when the number of clusters were between 5 and 6. Moreover, the Dunn index for k-means was consistently higher than that of fuzzy c-means. Similar to the silhouette scores, the highest scores occurred when the number of clusters was between 4 and 7 (Table 1).

Table 1. Silhouette and Dunn Index Scores for the First Approach

Number of Clusters	K-means Silhouette Score	K-means Dunn Score	Fuzzy C-means Silhouette Score	Fuzzy C-means Dunn Score
2	0.125	0.013	−0.036	0.005
3	−0.045	0.004	−0.028	0.005
4	0.086	0.013	−0.177	0.004
5	0.091	0.013	−0.081	0.005
6	0.075	0.007	−0.047	0.003
7	0.063	0.013	−0.192	0.004
8	−0.004	0.004	−0.12	0.004

The second approach was to drop the categorical variables with no ordinal relations from the dataset. This approach was decided due to the immense amount of different dummy variables needed to encode them, as it may lead to worsened model performance due to the curse of dimensionality [13].

In the second approach, the silhouette scores for both k-means and fuzzy c-means methods demonstrated a similar pattern, with the highest scores occurring when the number of clusters were between 2 and 5. In terms of the Dunn index, k-means was consistently higher than that of fuzzy c-means, indicating that the clusters produced by k-means were more compact and distinct, just as it was in the first approach (Table 2).

Table 2. Silhouette and Dunn Index Scores for the Second Approach

Number of Clusters	K-means Silhouette Score	K-means Dunn Score	Fuzzy C-means Silhouette Score	Fuzzy C-means Dunn Score
2	0.149	0.003	−0.04	0.001
3	0.122	0.004	−0.105	0.004
4	0.124	0.004	−0.128	0.004
5	0.139	0.006	−0.083	0.003
6	0.087	0.004	−0.223	0.001
7	0.085	0.002	−0.192	0.001
8	0.082	0.003	−0.152	0.001

Based on both of the approaches' results, it is suggested that k-means may be the more suitable choice for this particular analysis. In addition, it should be stated that, as foreseen, removing the categorical variables with no ordinal relation from the dataset proved to be the better approach as can be seen by the higher calculated scores.

5.2 Further Analysis and Clustering with the Selected Method

For further analysis, second approach and the k-means algorithm with 5 clusters have been selected according to the initial findings. In order to visualize the clusters, a three-dimensional space was created using the Principal Component Analysis (PCA) method due to the high dimensionality of the data. The clusters can be examined in the following figure from two angles (Fig. 1).

To further investigate the characteristics of each cluster, a random forest classifier was applied to evaluate the importance of each feature within each cluster. The clusters were then named based on the most important factors identified, as follows: *"Luxury"* with high number of bathrooms, beds and luxury amenities, *"High-Performing"* with experienced hosts who have quick response times and a high number of recent reviews, *"Highly-Rated"* with positive review scores, *"Convenient"* with prioritized check-ins and communication satisfaction, and lastly *"Immaculate"* with high cleanliness scores, great value, and high availability.

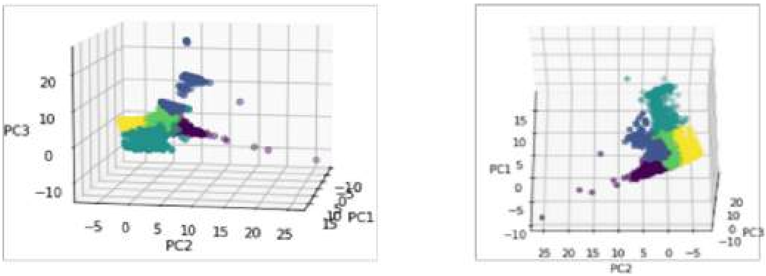


Fig. 1. Visualization of the clusters in a three-dimension space with two viewing angles.

6 Conclusion and Discussion

In summary, this study aimed to investigate the relationship between Airbnb listings in Istanbul. The study applied k-means and fuzzy c-means clustering algorithms to group the data into various numbers of clusters and evaluated the results using the silhouette score and the Dunn index. The study revealed that k-means clustering may perform better than fuzzy c-means in this particular case, with the optimal number of clusters being between 2 and 5, latter being the most optimal. The resulting clusters were named based on their most important characteristics to provide an overview of what to expect from listings within each group, which can be beneficial for both hosts looking to improve their listings and for guests looking for specific types of rentals.

This research can contribute to future studies by providing valuable insights into the relationship between Airbnb listings in Istanbul. It can also aid researchers working on similar datasets and projects by guiding the selection of appropriate clustering methods and identifying patterns and similarities among data. As a suggestion for future studies, this analysis may be expanded by the inclusion of additional clustering algorithms such as GMM (Gaussian Mixture Models) and DBSCAN, while also including different indices for evaluating these said algorithms.

References

1. Koster, H.R.A., van Ommeren, J., Volkhausen, N.: Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles. J. Urban Econ. **124**, 103356 (2021). https://doi.org/10.1016/j.jue.2021.103356
2. Choudary, S.P.: Platform thinking. Medium (n.d.) (2015). https://medium.com/platform-thinking. Retrieved 28 Mar 2023
3. Hati, S.R.H., Balqiah, T.E., Hananto, A., Yuliati, E.: A decade of systematic literature review on Airbnb: the sharing economy from a multiple stakeholder perspective. Heliyon **7**(10), e08222 (2021). https://doi.org/10.1016/j.heliyon.2021.e08222
4. Todd, J., Musah, A., Cheshire, J.: Assessing the impacts of Airbnb listings on London house prices. Env. Plann. B: Urban Analytics City Sci. **49**(1), 206–222 (2022)
5. Schäfer, P., Braun, N.: Misuse through short-term rentals on the Berlin housing market. Int. J. Housing Markets Anal. **9**(2), 287–311 (2016). https://doi.org/10.1108/ijhma-05-2015-0023
6. Garcia-López, M.-À., Jofre-Monseny, J., Martínez-Mazza, R., Segú, M.: Do short-term rental platforms affect housing markets? evidence from airbnb in Barcelona. J. Urban Econ. **119**, 103278 (2020). https://doi.org/10.1016/j.jue.2020.103278

7. Li, X., Liao, J., Gao, T.: Airbnb (air bed and breakfast) listing analysis through machine learning techniques. In: Segall, R.S., Niu, G. (eds.) Biomedical and Business Applications Using Artificial Neural Networks and Machine Learning, pp. 209–232. IGI Global (2022). https://doi.org/10.4018/978-1-7998-8455-2.ch008
8. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297. University of California Press, Berkeley, CA (1967)
9. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning, vol. 112, p. 18. Springer, New York (2013)
10. Chuai-Aree, S., Lursinsap, C., Sophatsathit, P., Siripant, S.: Fuzzy C-mean: a statistical feature classification of text and image segmentation method. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **9**, 661–671 (2001)
11. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (1987). https://doi.org/10.1016/0377-0427(87)90125-7
12. Bezdek, J., Pal, N.: Cluster validation with generalized Dunn's indices. In: 1995 International Conference on Artificial Neural Networks and Expert Systems (ANNES '95), pp. 190–193 (1995). https://doi.org/10.1109/ANNES.1995.499469
13. Bellman, R., Kalaba, R.: On adaptive control processes. IRE Trans. Autom. Control. **4**(2), 1–9 (1959)

Social Media Tri-Domain Analysis for Detection of Potential/Likely Malicious Users

Zafran Khan(✉), Yechan Kim, Yeonsu Seul, and Moongu Jeon

Gwangju Institute of Science and Technology, Gwangju, South Korea
{zafrankhan1830,yechankim}@gm.gist.ac.kr, {comm8123,mgjeon}@gist.ac.kr

Abstract. Online social media(OSM) is at the doorstep of every individual in this era of the global village. According to a recent study by the UN census bureau, 59% world's population is actively using it, and 7 users are added to this figure every second. Users form an unstructured network on these platforms. Thus contents posted by any user would transverse in fractions of seconds. Contents available on social media platform effectively exploits the opinions of the readers. These contents can be defamed to spread misinformation, disinformation, and propaganda. The Spread of such propagandist content in society can adversely lead to fear, uncertainty, panic, or even financial loss in trading markets. Therefore, the detection of those users who spread such propagandistic content is the need of the hour. The users and their interactions with each other can be anticipated as similar to nodes and edges of the graph data structures. The edges are to be non-directed if interactions are two way like Facebook and directed if it is one way like Twitter. In this research, objective is to propose a graph convolution neural (GCN) network-based framework that captures the insights of the online social media unstructured patterns. The user-user request/ response graph is learned by the proposed framework. We suggest using the historic features of users to formulate the user profile. The performance of the proposed model is compared with SVM and LSTM. A series of experiments render the out-performance of the proposed framework on a real-world PHEME dataset. The proposed framework may also be used as an OSINT tool if customized data is available.

Keywords: Social Media Analytic · Graph Convolutional Neural Network · User Profiling · Content profiling

1 Introduction

In present era of digitization and global village use of social media (Facebook, Twitter and Instagram etc.) has been tremendously increased. Thus, any contents posted by an unknown individual travels in fraction of seconds within peoples. Moreover, high ranked officials and public office holders also use social media like twitter to convey their messages to the mob.

Contents available on social media platform can easily exploit the mindset of the readers. Such contents can be defamed to spread misinformation, disinformation, propaganda and rumors etc. Posting of such contents can either be

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 206–214, 2023.
https://doi.org/10.1007/978-3-031-39777-6_25

done by a single individual or by state sponsored agencies aimed to achieve the intended goals. Spread of these malicious contents in the society can lead to panic, uncertainty, fear or financial loss in the trading markets. Earlier detection of such malicious users would safeguard propagation of harmful contents on social media, thus protecting society from being jeopardized against such evils.

Peoples on the social media follow an unstructured patterns in their connection with each other. Such as a person 'A' will be partially or entirely different from that of the person 'B' with respect to friends list, followers, followings, likes/ dislikes, activities, and check-ins. Therefore, detecting a user with malign intents is difficult in such an unstructured network of copious users. Manual identification of such users is undoubtedly a tedious and laborious job. Moreover, available datasets in the academia literature are mostly based on the identification of rumors detection itself and not the spreaders of these rumors. This leads to the non availability of labeled data that can be used for Artificial Intelligence/ machine learning models training.

Social media users and their interactions with each other, can be anticipated as nodes and edges of the graph data structures. In this research study, core objective is to propose a graph convolution neural network (GCN) based framework that concurrently learn the user profile, contents profile and unstructured flow of information in social network. The profiling is done on the bases on textual tweets, retweets, followings, followers and the ego-network indicating his interactions in the social network. Based on these learned features, trained framework predicts a user to be classified as suspected or likely to be malicious or not. Moreover, in order to have labeled dataset for training of framework, we transformed the available PHEME dataset into malicious or non-malicious users.

2 Literature Review

Existence of rumors, fake news or defamed contents over the social media networks is of no surprise in today's era of digitization. For years, academia researchers have been focusing on to detect such fake contents aimed to safeguarded the society. Similar to rumors detection, identification of probable person behind such defamed information is the emerging research trend in the academia as well as industry. However, absence of accurately labeled data-sets in this domain makes it nearly impossible. In order to reach-out to the user, different attributes of social media and shared contents are required to be explored. We present recent literature review of both the domains of malicious user's detection and rumor detection.

2.1 Rumor Detection

In order to detect rumorous contents, researchers have put-in their best and evaluated the performance of their work. The authors in [1] implemented fake news detection in two steps. Initially they converted the unstructured data to a structured format and then applied various supervised learning algorithms by text mining methods. Deep neural network based approach is suggested in [2] that detects the fake news based on the contents semantics, their context and related temporal information. A framework which is independent of platform source and language for propaganda detection is suggested in [3]. It is entirely based on text features extraction techniques. The proposed framework was examined on three different language groups with optimal results. GNN based model with enhanced textual contents representation is proposed in [4] for early detection of fake news. The textual contents representation is achieved by integration of semantic relation and sequential ordering of textual contents. [5] implemented deep learning based feature extraction, CNN classifier and a learning framework to detect rumor at the earliest to stop it spreading in the society. The authors tested their proposed framework on the dataset generated from twitter and Weibo.

2.2 Propagandist Detection

Similar to rumor detection another domain to be explored is the detection of the user behind these contents. Law enforcement agencies in a country would prefer to reach out to the user with malicious intention. Though limited, but researchers have also explored in this domain as well wherein, we have vast room of improvements. In this regards, [6] examined the # of followings and # of followers of a user to establish that whether the user possesses malintent or not. Similarly, [7] analyzed user's screen name, parameters of email and user's profile name for malicious user identification. Exploitation of multi-modal information is proposed in [8] for detection of malicious individuals in the online social media platform. The author's focus on the concept of "*who follows whom*". The proposed framework is to catch the dual patterns of followers in which one is caused by zombie followers and other is the in-consistent behaviour that is due to other user's. Diffusion of information flow in twitter social network is studied by [9]. It is stated that such diffusion pattern indicates about the veracity of available data. The diffusion pattern of social media contents was modeled as directed graphs depicting the user's interactions. As stated earlier that very few labeled datasets are available in this domain, thus [10] enlisted various real-world bench mark datasets that can be used for rumor detection or can be modified for rumor spreader detection. [11] analyzed the regular behavioral pattern of frequent or habitual mobile users' for detecting the fake profiles.

Our proposed framework is based on the exploration of user's profile, profile of his shared contents, and the essence of user's centered ego-network. Amalgamation of these three aspects and utilization of recent deep learning approaches render better performance in ascertaining the malicious user.

3 Proposed Framework

In this section, the proposed framework is explained in detail. It is appraised that malicious contents, fake news, or rumor detection is being explored by various researchers in academia. Fake news detection can be done by analyzing the contents shared by users on various social media platforms like Twitter, Facebook, Instagram, etc. In this regard, various benchmark datasets are already available for experimentation. On the contrary, detection of potential/ likely propagandists in a social media network does not involve the shared content analysis only, but also requires the study of user-profiles and a thorough analysis of the social network. Moreover, the non-availability of labeled dataset(for malicious user detection) in academia makes it further difficult. Thus to undertake the proposed task, a labeled dataset is required. This study has been carried out on the dataset that is converted from rumor and non-rumor dataset named PHEME. The conversion of dataset was carried out on various heuristic assumptions and threshold values. After data transformation, it has been exploited in three different domains to evaluate the efficacy of the proposed framework. The tri-domain analysis of the data is, user's profiling, shared content profiling, and the user-centered ego-network in which the user interacted with each other. Accordingly, the flow of the proposed methodology can be segregated into three different sub-tasks i.e., data transformation, feature extraction, and user identification. This segregation is depicted in Fig. 1.

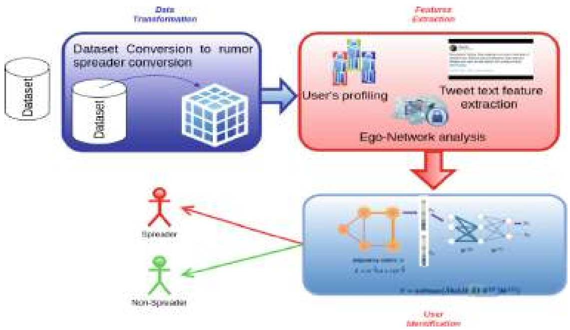


Fig. 1. Proposed Framework

3.1 Data Transformation

In order to examine the performance of the proposed model, we PHEME dataset transformed into the malicious users dataset. The reason for selecting this dataset is because it contains all the details that are required for malicious user detection in an online social media network. Primarily the subject dataset is for rumor and non-rumor detection. The dataset contains tweets from source user as well as from reply user. Here the source user is the one who initially posted a content, whereas, the reply user is the one who responded to the source user. The tweets have been categorized as rumors and non-rumors. Out of the provided features in dataset, We used user id, favorites count, followers count and account is verified or not for user profiling. Similarly, for content analysis, we considered the text of the tweet shared by the user. Lastly, we also considered the information flow pattern or the user-user interaction pattern that is formed among the user's while interacting with each other. It is generally understood that frequent interactions would occur between close friends and vice versa.

3.2 Feature Extraction

After the data transformation, its the time to extract requisite features out of it. So hereinafter, the step of feature extraction comes into play its role. For feature extraction, we can examine any social media data in three different domains i.e. social media tri-domain analysis. The three domains involve analysis of user's profile, analysis of contents shared by users' and the user's centred ego-network analysis. The first two sub-tasks (user's profile and content's profile analysis) requires the enriched features extractions. The user's centered ego-network analysis along with the features extracted would lead to a clue that the user can be classified as propagandist or not.

User's Profile: For the user's profile analysis, we took lead from the very famous saying that a man is known by the company he keeps. For this we considered to look for the no of followers and the no of favorites of source user. Moreover, in order to check the validity of the source user account, the verification of the user's account is considered. The verified user account means that such account is active, authentic and notable. moreover, during data transformation, we formulated the degree matrix of the user-user network wherein high weight is assigned to the edge of the source user and follower who have frequent interaction history. This reflects the user inclination towards a particular followers and vice-versa.

Content's Profile: For the contents profiling, we considered the tweet text that is shared by each user's. We generated embedding's of the tweets using pretrained Glove language model [12]. It captures global contextual information in a text corpus by calculating word-word co-occurrence matrix. GloVe extract the semantic of words present in the tweets and represents the textual tweets in the numerical values that are to be processed by proposed framework.

User's Classification: Identification of all malicious users requires exploitation of social media in all domains i.e. user's profile, shared contents, and the unstructured flow of information. Machine learning models need to be trained on data that should follow a regular format which does not holds true in the case of social media network. Therefore, conventional neural networks are unable to be trained on this unstructured data of social media. Nonetheless, they can be used for user classification on the bases of the user's profile or the content analysis only. However, this would leave behind the user-centered ego-network unexplored which carries rich information regarding user-to-user interactions and contents dissipation in social media platforms. In order to simulate such unstructured flow of information, we assume that the user's and their inter-connections within the social media platforms mimic the graph data structure. Users can be considered as nodes of a graph and interconnections among different users look like edges of the graph. Thus we need a mechanism to be trained on this unstructured graph data. For this graph convolutional network (GCN) [13] is an adequate choice that would considers the user and their connectivity as nodes and edges respectively. GCN takes two matrices as input, one the nodes feature matrix and the other is the graph adjaceny matrix.

4 Implementation and Results

This section covers the practical implementation details that includes both hardware configuration and software packages. Moreover, it also narrates the outcomes of the proposed framework.

4.1 Practical Implementation

We implemented the proposed framework on a desktop system installed with Ubuntu 18.04 LTS bionic, having 16 GB RAM. The system is fitted with AMD®Ryzen 7 3700x 8-core processor and NVIDIA GeForce RTX 2080 Ti Graphic card. Besides the hardware details, software packages used in the implementation includes Tensorflow, Keras, Spektral, Simpletransformers, NLTK, Networkx and Gensim. In order to avoid over-fitting and to cater for class imbalance problem, we used K-fold cross validation. The proposed model is trained while using 'binary cross entropy' as a loss function and RMSProp as an optimizer with a learning rate set to $1e^{-6}$. We fabricated the proposed framework with GraphConv layers with l_2 as kernel regularizer and dropout ratio set to 0.25.

4.2 Results

The users in the transformed data-set fall onto two categories, either malicious user or non-malicious uses. Thus this is the case of bi-classification. In order to give a consolidated and comprehensive performance evaluation of the proposed framework, we presented the f1-score and area under the curve (AUC) as our evaluation protocol. F1-score can be assumed to be the hormonic mean of precision and recall thus giving the overall health or overall performance of the proposed framework. Similarly AUC indicates the extent or capability of separability of the model. To the best of our knowledge, our work is the first of its kind for malicious user detection. Therefore, to have some baseline for comparative analysis, we implemented SVM and LSTM as well. Table 1 tabulates the f1-measure and AUC-ROC values of the proposed model compared to the two baselines. Moreover, for each of the event in the PHEME dataset, we plot the AUC-ROC curve for graphical analysis. It can be easily observed from Table 1 and Fig. 2 that our proposed framework outperforms then the SVM and LSTM. In Fig. 2 horizontal axis indicated the false positive rate while vertical axis represent the true positive rate. As part of application, It is pertinent to mention that detection of such misinformation spreaders can be used as open source intelligence (OSINT) as well. Based on the availability of customized dataset of any organization, it can be used as OSINT aimed to find the propagandist in the wide range of users.

Table 1. Performance Comparison

Dataset		Ferguson	Germenwing crach	Ottawa	Charlie Hebdo	Sydney Siege
F1-score	SVM	0.782	0.572	0.578	0.853	0.751
	LSTM	0.676	0.546	0.559	0.778	0.638
	Ours	0.783	0.709	0.655	0.864	0.690
AUC-ROC	SVM	0.652	0.552	0.566	0.600	0.618
	LSTM	0.550	0.540	0.550	0.570	0.540
	Ours	0.660	0.720	0.680	0.690	0.660

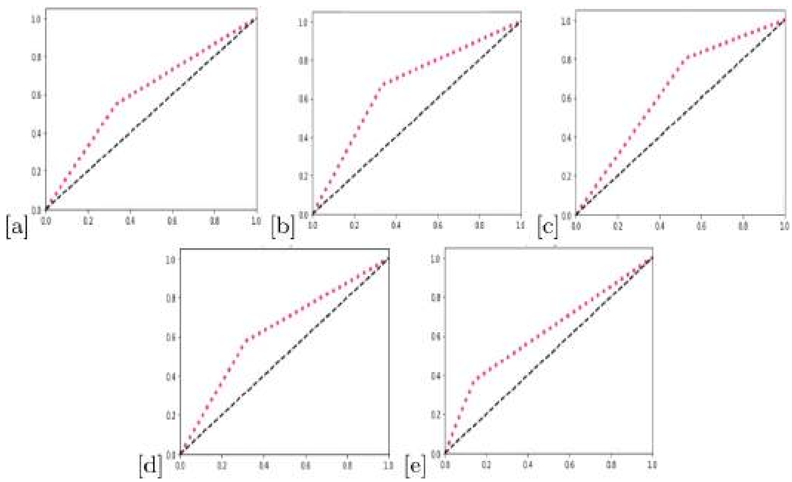


Fig. 2. Performance of Proposed model with respect to ROC-AUC (a): Ferguson, (b): Sydney, (c): Ottawa, (d): German, (e): Cherlie.

5 Conclusion

Social media is an open platform where anyone can post any contents with a pre-determined aim. It is being used as an easily available platform for disinformation campaigns by anti-state agents or state sponsored organizations.

The detection of malicious individual is the need of hour and this work aims to detect such malicious users. We proposed a graph convolutional neural network based framework that exploits the social media platform in three different domains. The performance of the model is tested on real-time PHEME dataset. As part of future work, we would try to transform our dataset into multi class problems and utilize the graph attention frameworks to improve its accuracy further. Moreover, researcher can also utilize this idea as an open source intelligence tool in real-time twitter analysis.

Acknowledgements. This work was supported by GIST-MIT Research Collaboration grant funded by the GIST in 20xx and the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2014-3-00077).

References

1. Ozbay, F.A., Altas, B.: Fake news detection within online social media using supervised artificial intelligence algorithms. Physica A **540**(Article no. 123174), 0378–4371 (2020)
2. Kaliyar, R.K., Goswami, A., Narang, P., Sinha, S.: FNDNet-a deep convolutional neural network for fake news detection. Cogn. Syst. Res. **61**, 32–44 (2020)
3. Faustini, P.H.A., Covões, T.F.: Fake news detection in multiple platforms and languages. Expert Syst. Appl. **158** Article no. 113503 (2020)
4. Wang, Y., Wang, L., Yang, Y., Lian, T.: SemSeq4FD: integrating global semantic relationship and local sequential order to enhance text representation for fake news detection. Expert Syst. Appl. **166**, Article no. 114090 (2021)
5. Liu, Y., Wu, Y.: FNED: a deep network for fake news early detection on social media. ACM Trans. Inf. Syst. **38**, 1–33 (2020)
6. Benevenuto, F., Magno, G., Rodrigues, T., Almeida, V.: Detecting spammers on Twitter. In: Collaboration, Electronic Messaging, Anti-abuse and Spam Conference (CEAS), vol. 6, pp. 9–12 (2010)
7. Thomas, K., McCoy, D., Grier, C., Kolcz, A., Paxson, V.: Trafficking fraudulent accounts: the role of the underground market in Twitter spam and abuse. In: Presented as Part of the 22nd fUSENIXg Security Symposium (fUSENIXg Security 2013), pp. 195–210 (2013)
8. Jiang, M., Cui, P., Beutel, A., Faloutsos, C., Yang, S.: Detecting suspicious following behavior in multimillion-node social networks. In: Proceedings of the 23rd International Conference on World Wide Web, pp. 305–306 (2014)
9. Rosenfeld, N., Szanto, A., Parkes, D.C.: A Kernel of truth: determining rumor veracity on Twitter by diffusion pattern alone. In: Proceedings of the Web Conference 2020, pp. 1018–1028 (2020)
10. Li, Q., Zhang, Q., Si, L., Liu, Y.: Rumor detection on social media: datasets, methods and opportunities. In: Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Sensorship, Disinformation, and Propoganda, China, pp. 66–75 (2019)
11. Bertini, F., Sharma, R., Iannì, A., Montesi, D.: Profile resolution across multilayer networks through smartphone camera fingerprint. In: Proceedings of the 19th International Database Engineering & Applications Symposium, pp. 23–32 (2015)

12. Pennington, J., Socher, R., Manning, C.D.: GloVe: global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Qatar, pp. 1532–1543 (2014)
13. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: ICLR (2018)

3-D Face Reconstruction Method Using Deep Learning Based Simulated Annealing

Fei Fei Chen[1], Bing Guan[2], Sumi Kim[3], and Jaeho Choi[1](✉)

[1] Department of Electronics Engineering, JBNU, Jeonju, Republic of Korea
wave@jbnu.ac.kr
[2] School of CSE, Huizhou University, Huizhou, China
[3] Seoyoung University, Gwangju, Korea

Abstract. Face reconstruction becomes more accurate and popular in the age of artificial intelligence and deep learning. Particularly, 3-D face reconstruction can be one of the key technologies in meta-world and virtual reality applications. In recent years, studies on realistic 3-D face reconstruction draw much attention among the researchers. For realizing accurate shape and facial textures, the neural network models for deep learning are under investigation and several schemes are present-ed. In this paper, a deep learning based simulated annealing algorithm is proposed for 3-D face reconstruction. Face labeling, feature extraction, and 3-D reconstruction are three major elements investigated in this study. A set of computer simulation is performed by using the CelebFaces Attributes and Labeled Faces in the Wild data sets. The system performance is evaluated in terms of reconstruction accuracy and the results show us that the proposed method can be a successful alternative for providing accurate and robust 3-D face reconstruction.

Keywords: 3-D face reconstruction · Deep learning · Simulated annealing · Face labeling · Feature extraction · Face textures

1 Introduction

The 3D deformation model (3DMM) of Blanz and Vetter [1] plays a crucial role in 3D face reconstruction. The implementation of 3DMM has inspired researchers. More and more researchers have become interested in 3D face reconstruction [2]. 3D face reconstruction techniques have become increasingly important with the concept of meta-space. The impact of COVID-19 in recent years, the experience of researchers or teachers using AR and VR tells us that 3D face reconstruction techniques can provide effective experiences in areas such as education, communication and meetings [3]. Unrealistic 3D faces pose a danger to the development of metaverse such as terrorists recruit in second life's virtual communities [4]. Therefore, it is very important to recover the authenticity of the 3D shape of a face from a 2D picture.

In recent years, due to the increasing data sets and the development of deep learning, 3D face reconstruction based on deep learning has become more and more popular. Among them, identity, expression, texture, pose and lighting are important elements of

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 215–221, 2023.
https://doi.org/10.1007/978-3-031-39777-6_26

3D face reconstruction. Deng [5] has pointed out that expression and texture are the key factors for the realism of 3D face reconstruction. However, most of the current work focuses on the shape of the face reconstruction and is not very concerned with the realistic expression and texture. In this paper, we enhance the 3D face reconstruction based on simulated annealing algorithm for deep learning, and strengthen the weight of expression and texture. The paper is organized as follows. Section 2 provides the related work. The dataset and preprocessing work are described in Sect. 3. The proposed deep learning model is described in Sect. 4. In Sect. 5, the performance of the proposed model is verified by computer simulation and the results are evaluated. Finally, conclusions are made in Sect. 6.

2 Related Work

Tuan [6] has proposed the generative renderer and inverse rendering to reduce the domain offset from the real image without involving the scene lighting and face texture, thus improving the realism and similarity of face reconstruction results. Sanyal [7] has proposed RingNet to introduce a ring structure to automatically detect 2D facial features. Wojciech [8] has optimized the reconstruction methods for non-metric and metric benchmarks. The goal of all these methods is to obtain more realistic 3D face models. In this paper, the texture and texture of 3D face reconstruction are optimized by using ResNet50, Confidence-Net, and simulated annealing algorithm.

2.1 R-Net Architecture

Deng [5] has used the ResNet-50 network to regress these coefficients by modifying the last fully connected layer to 239 neurons. The unknown vectors to be predicted include identity, expression, texture, pose, and lighting for a total of 239 parameters; $\alpha = 80, \beta = 80, \delta = 64, \gamma = 9, p = 6$:

$$x = (\alpha, \beta, \delta, \gamma, p) \infty R\ 239 \tag{1}$$

Here in this work, the same model of 239 neurons for the fully connected layer is maintained. Regression is performed with parameters containing 239 unknown vectors. The weights of textures is increased with the addition of simulated annealing.

Simulated annealing (SA) was first proposed by Dowsland [9]. The purpose of training CNNs using SA is to meet certain accuracy requirements to minimize the approximation error accuracy and network complexity metrics. In this paper the simulated annealing algorithm is used for neural networks with two termination conditions [10]. The first is that the iteration reaches a specified maximum iteration, and the second is that the C-Net fitness function is less than a specific constant. Although the computation time increases, for each training round the classification error of this method is smaller than that of the original CNN. The flow chart of this method is shown in Fig. 1.

R-Net is used to predict higher-order information, such as face pose and illumination, and C-Net will predict the features of higher-order information with confidence. In this paper, we fuse the important parameters of simulated annealing with those of R-Net. The fused simulated annealed R-Net will try to keep the realism of the texture.

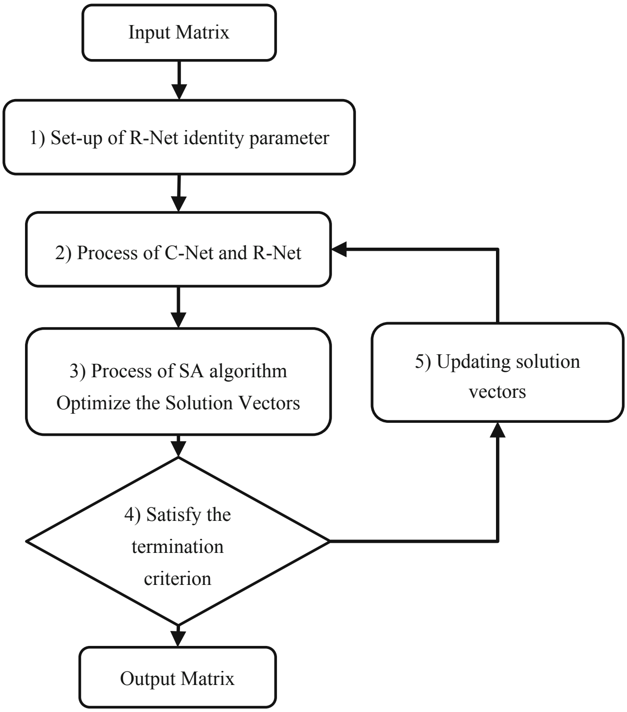


Fig. 1. Flowchart of SA algorithm for optimal R-Net.

2.2 C-Net Architecture

Since the R-Net parameters with simulated annealing added are the same as the R-net proposed by Deng [5], the C-Net as feature map detection will remain consistent. Three convolutional layers with a core of 3 and 256 channels are applied. Finally, we apply the sigmoid function for confidence prediction. The confidence of the feature map is positive.

3 Datasets and Pre-processing

In order to train the neural network proposed in the paper, the labeled data of the dataset must be preprocessed with feature points. We use some publicly available facial datasets such as CelebFaces Attributes dataset (CelebA) [11] and LFW (labeled faces in the wild) face dataset [12]. 68-point detector and 5-point facial landmark detector are used. Table 1 shows that each dataset is divided into 80% on training and 20% on testing.

Table 1. Training and testing split dataset details.

Dataset	Number of images	Training	Testing
CelebA	202599	80%	20%
LFW	13233	80%	20%

3.1 Datasets

CelebA: CelebA is a large dataset of face attributes; the images in CelebA cover large pose variations and background clutter. CelebA has a large amount of diversity face images.

LFW: The LFW database, which is mainly collected from the Internet rather than the laboratory, contains more than 13,000 face images. Each image is identified by the name of the corresponding person, of which 1,680 people correspond to more than one image. he face images provided are derived from natural scenes in life, which makes recognition more difficult, especially due to multiple poses, lighting, expressions, age, occlusions, and other factors that can cause even the same person's photo to vary greatly.

3.2 Pre-processing

To train a model with custom images, 5 facial landmarks of each image are needed in advance for an image pre-alignment process. We use MTCNN to detect these landmark files. The files contain 5 facial landmarks of shape 5x2 and should have the same name as their corresponding images. 68 landmarks and skin attention masks are generated based on images of 5 landmark.

The MTCNN [13] pipeline is divided into the following steps. (1) Obtain an image pyramid by resizing the image. (2) The image pyramid is input to Pnet and a large number of candidate images are obtained. (3) The filtered candidate images are fine-tuned by Rnet. (4) After Rnet takes out a large number of candidate images, the images are input to Onet and the exact bbox coordinates and landmark coordinates are output.

4 SAR-Net Models

The 3D face model using 3DMM [1] is represented in shape and texture as:

$$S = S(\alpha, \beta) = \overline{S} + B_{id}\alpha + B_{exp}\beta; T = T(\delta) = \overline{T} + B_t\delta \tag{2}$$

where $\overline{S}$ and $\overline{T}$ are the average face shape and texture of the popular Basel face model from 2009, which also contains B_{id} and B_t expressions. Using the expression library B_{exp} is constructed.

The input image matrix will be fitted as the feature parameters of C-Net and used as the input parameters of simulated annealing for vector optimization until the termination criterion is satisfied. Shapes and textures will be regressed in R-Net. 3DMM adds a loss function to the model coefficients in order to prevent texture degradation as follows:

$$L_{coef}(x) = \omega_\alpha \alpha^2 + \omega_\beta \beta^2 + \omega_\gamma \delta^2 \quad (3)$$

In the vector for generating 3D faces, α is the identity coefficient, β is the expression coefficient, and δ is the texture coefficient. Since our network incorporates the simulated annealing algorithm in the texture coefficients, the expression in Eq. 3 is modified to Eq. 4 as follows (Fig. 2):

$$L_{coef}(x) = \omega_\alpha \alpha^2 + \omega_\beta \beta^2 + \omega_\gamma SA(\delta^2) \quad (4)$$

The texture coefficient in the feature matrix is calculated as a judging term in the simulated annealing algorithm. The final texture coefficients are optimized as a matrix that meets the termination criteria.

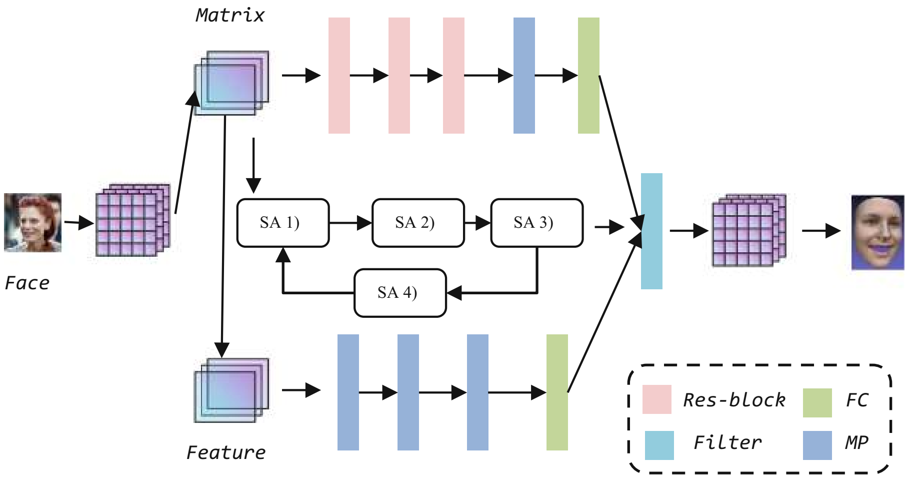


Fig. 2. R-Net and C-Net with SA Neural Network Structure.

5 Experiments

The pre-processed image size is 224 × 224. The same size is used for the input data in this paper and the weight parameters of ImageNet are used. Iterative training has been performed with the initialization parameters of Deng [5] and the simulated annealing algorithm is added as an optimizer for specific parameters (texture). The accuracy of texture has been improved while maintaining the efficiency of other loss functions, as shown in Table 2.

Without considering the underlying information such as photometric similarity, we have compared the publicly available test data with the private face data. The results are shown in Fig. 3.

Fig. 3. The texture representation of our 3D reconstructed model compared to that of Dung [5]. The more prominent and consistent results are obtained by the propose model.

Our model is more prominent in the mouth part and the texture is more visible. The simulated annealing algorithm is effective in the calculation of the prominence parameter.

Table 2. Comparison of parametric optimization algorithms with simulated annealing under the same dataset MICC [14] and Face warehouse [15].

Losses				MICC	Facewarehouse
L_{photo}	L_{lan}	L_{per}	$L_{SA(coef)}$		
			√	1.90 ± 0.56	2.83 ± 0.86
		√		1.87 ± 0.43	2.70 ± 0.65
√	√			1.80 ± 0.52	2.17 ± 0.65
	√	√		1.71 ± 0.43	2.11 ± 0.48
√	√	√		1.67 ± 0.50	1.81 ± 0.50

6 Conclusion

Based on the hybridized loss function proposed by Deng, we have presented a face texture reconstruction method based on simulated annealing CNN. This method uses image information for texture enhancement learning. It is demonstrated experimentally that the simulated annealing CNN method optimizes the face texture while maintaining accuracy and robustness. A loss function to prevent face texture degradation is also included in R-Net. Expression and lighting are also important parameters for face reconstruction. In future work we will optimize several important parameters of the face and believe that 3D faces will be more realistic and exact.

References

1. Blanz, V., Thomas, V.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques (1999)
2. Egger, B., Smith, W.A., Tewari, A., Wuhrer, S., Zollhoefer, M., Beeler, T.: 3D morphable face models—past, present, and future. ACM Trans. Graph. **39**(5), 1–38 (2020)
3. Bolter, J.D., Engberg, M., MacIntyre, B.: Reality media: Augmented and Virtual Reality. The MIT Press (2021)
4. Boellstorff, T.: Coming of Age in Second Life: An Anthropologist Explores the Virtually Human. Princeton University Press, Princeton (2015)
5. Deng, Y., Yang, J., Xu, S., Chen, D., Jia, Y.: Accurate 3d face reconstruction with weakly-supervised learning: from single image to image set. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (2019)
6. Tuan T., Hassner, T., Masi, I., Medioni, G.: Regressing robust and discriminative 3D morphable models with a very deep neural network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
7. Sanyal, S., Bolkart, T., Feng, H., Black, M.: Learning to regress 3D face shape and expression from an image without 3D supervision. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019)
8. Zielonka, W., Bolkart, T., Thies, J.: Towards metrical reconstruction of human faces. In: Avidan, S., Brostow, G., Cissé, M., Maria Farinella, G., Hassner, T. (eds.) Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIII, pp. 250–269. Springer Nature Switzerland, Cham (2022). https://doi.org/10.1007/978-3-031-19778-9_15
9. Dowsland, K.A., Thompson, J.M.: Simulated annealing. In: Rozenberg, G., Bäck, T., Kok, J.N. (eds.) Handbook of Natural Computing, pp. 1623–1655. Springer Berlin Heidelberg, Berlin, Heidelberg (2012)
10. Gülcü, A., Kuş, Z.: Multi-objective simulated annealing for hyper-parameter optimization in convolutional neural networks. PeerJ. Comput. Sci. **7**, e338 (2021)
11. Liu, Z., Luo, P., Wang, X., Tang, X.: Large-scale celebfaces attributes (celeba) dataset. Retrieved 11 Aug 2018
12. Huang, G., Mattar, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: a database for studying face recognition in unconstrained environments. In: Workshop on faces in real-life images: detection, alignment, and recognition (2008)
13. Zhang, K., Zhang, Z., Li, Z., Qiao, Y.: Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process. Lett. **23**(10), 1499–1503 (2016)
14. Bagdanov, A. D., Del Bimbo, A., Masi, I.: The florence 2d/3d hybrid face dataset. In: Proceedings of the 2011 joint ACM workshop on Human gesture and behavior understanding (2011)
15. Cao, C., Weng, Y., Zhou, S., Tong, Y., Zhou, K.: Face warehouse: a 3d facial expression database for visual computing. IEEE Trans. Visual Comput. Graph. **20**(3), 413–425 (2013)

Intelligent Unsupervised Defect Detection of Rail Surface via Generative Adversarial Networks

İlhan Aydın(✉) and Mehmet Sevi

Computer Engineering Department, Fırat University, Elazığ, Turkey
iaydin@firat.edu.tr

Abstract. One of the most important factors affecting the safety of the train during the operation of the railway system is rail surface defects. Therefore, detecting these defects at an early stage is very important for rail safety. Defect detection approaches using supervised learning are based on the use of healthy and defective data in training and the evaluation of the trained model during testing. In this study, an adversarial generative network-based unsupervised defect detection system is proposed to detect defects in a situation where only healthy rail surface data is available. The performance of the proposed method was also evaluated with rail detection and image enhancement studies. The proposed approach has been tested with healthy and defective data taken in different light conditions, and defects have been detected with an accuracy rate of over 98%.

Keywords: Railway · Defect Detection · Unsupervised Deep Learning · Adversarial Generative Networks

1 Introduction

Rails are an important component of the infrastructure in railway transportation, and periodic inspection is very important for passenger and train safety. Rails used in freight transportation as well as passenger transportation are exposed to different types of defects due to different working environments, heavy load-carrying situations, and increased stress [1]. If rail surface defects are not detected at an early stage, it is possible to cause serious accidents by creating certain risks. Therefore, visual inspection of rail surfaces at certain periods is important for rail safety [2].

Detection of traditional rail defects is done by an expert walking along the railway. However, this method is impractical because it occupies the railroad and depends on expert knowledge. Cracks on the rail surface were detected with acoustic signals [2, 3]. Although cracks on the rail surface can be detected with this sensor, the location information cannot be determined exactly and there is a cost of installation at certain intervals. Eddy currents were used to detect cracks in the rails [4]. However, since eddy currents cause even small cracks to be perceived as very large cracks, the rate of generating false alarms is high [5].

Recently, with the developments in high-speed cameras and computer vision technology, many studies have been carried out in the field of vision-based inspection of

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 222–229, 2023.
https://doi.org/10.1007/978-3-031-39777-6_27

rail components. Zhang et al. [1] used features extracted from DenseNet, local binary patterns, and gray-level co-occurrence matrix to classify rail surface defects. With the proposed approach, 5 different rail surface defects are classified. Liu et al. [6] performed pyramid feature extraction with a lightweight convolutional neural network to detect rail surface defects. The proposed approach segments the rail surface defects. Li et al. [7] proposed a hybrid object detection method to detect different types of rail surface defects. The proposed method gave better results than known methods such as Yolo and Faster RCNN. Aydin et al. [8] proposed a hybrid method to classify rail surface defects using two lightweight networks, MobilenetV2 and SqueezeNet. Tu et al. [9] proposed a YOLACT-based segmentation network to detect rail components and used a lightweight image classification approach for defect detection of rails from images taken from the camera. To determine the rail surface defects, segmentation was performed with a GrubCut-based approach and the locations of the defects were determined in the image segmented with YOLOv2 [10].

In this study, an adversarial-generative network-based approach is proposed for the detection of rail surface defects in a situation where only healthy rail samples are available. The proposed approach does not require any kind of data labeling or the creation of ground truths. An effective method is also used to correct for noise and light conditions in low-resolution images. The proposed approach uses only healthy rail images for defect detection. Defects identified as anomalies can be classified separately according to their degree of anomaly.

2 Unsupervised Defect Detection Based on Generative Adversarial Networks

The proposed approach for the detection of rail surface defects consists of three stages. First, image enhancement is applied to solve the problems related to noise and light in the collected rail images. Then, the rail surface region is extracted from the image. The extracted rail surface is given to the Generative Adversarial Network (GAN) and the training process is carried out. The block diagram of the proposed approach is given in Fig. 1.

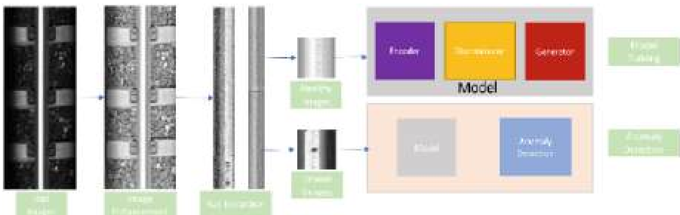


Fig. 1. Flowchart of the proposed method.

In Fig. 1, an image enhancement approach is applied based on a multi-scale Retinex algorithm [11]. Then, the rail surface region is extracted from the image. A model is constructed by giving the healthy rail surface images obtained to GAN, and anomaly detection is done by giving images that were not used in the training to the model obtained during the test phase.

2.1 Multiscale Retinex Algorithm for Image Enhancement

The Retinex algorithm has been developed for images that do not neglect gray image variances. If the reflection in each channel is similar in an RGB image, it can be said that the image provides the gray image variances and is suitable for the Retinex algorithm. The Retinex equation is given in Eq. (1).

$$R_i(x, y) = \log I_i(x, y) - \log[F(x, y) * I_i(x, y)] \tag{1}$$

In Eq. (1), $I_i(x, y)$ represents the color in the relevant pixel for each channel. The pixel obtained in the relevant channel with the Retinex method is represented with $R_i(x, y)$. $F(x, y)$ in Eq. (1) is calculated as in Eq. (2).

$$F(x, y) = e^{-r^2/c^2} \tag{2}$$

The Gaussian function in Eq. (2) achieves a better resolution at a large value. It is generally recommended to choose the c value between 50 and 100 for reasonable image quality. However, the Retinex algorithm causes gray tones to appear in gray regions in all or part of the image [12]. The multi-scale Retinex algorithm proposes a new color restoration to overcome these shortcomings. Multiscale color restoration can be given by Eq. (3).

$$R_{MSi} = \sum_{n=1}^{N} w_n R_{ni} \tag{3}$$

The N and R_{ni} values in Eq. (3) are the number of scales and the Retinex function at the relevant scale, respectively. According to the $R_i(x, y)$ in the original Retinex algorithm, the F(x,y) function is calculated as in Eq. (4).

$$F_n(x, y) = ke^{-r^2/c^2} \tag{4}$$

In Eq. (4), k is the scaling factor. It has been proven experimentally that using three scales is sufficient. The original Retinex algorithm creates an image with more gray values in the image. This situation causes the rail and ballast to be similar during the rail extraction stage and the rail extraction algorithm generates false alarms.

2.2 Rail Extraction Algorithm

For defect detection, the region containing the rail surface should be obtained from the rail images taken from the measuring train. For this purpose, it is necessary to create the data set by obtaining the surface part of the rail. This will be done by extracting the rail from the complex background image. For rail extraction, the left and right boundaries of the rail must be determined. In the rail images taken from under the measuring train, the rail surface remains in the middle. After determining the left and right boundaries of the rail in Fig. 2, the region containing the rail surface is cropped. This step will be used to construct the dataset for the contentious generator network.

```
Algorithm: RailExtractor(I)
Input: An enhanced rail Image I
Output: Cropped rail surface image CI
w, h←width and height of I
G←Convert I to Grayscale image
t←otsu thresholding(G)
R←G>t
H←zeros(1,w)
for i in w
    H(i)←H(i)+sum(R(:,i))
endfor
D←zeros(1,w-1)
for i in w
    D←|H(i)-H(i-1)|
endfor
Edge_min←max(D(1:w/2))
Edge_max←max(D(w/2+1:w)
CI←I(edge_min:edge_max,1:h)
```

Fig. 2. Rail extraction algorithm.

2.3 Defect Detection of Rail Surface Using Generative Adversarial Networks

After detecting the rail surfaces, a fast-GAN-based approach is used to detect surface defects [13]. This method is GAN training on normal images and encoder training on the trained GAN model. After the training phase, the images are given to the trained model, and an anomaly score is obtained for the images. First, a model that maps the images to the hidden area is trained by giving healthy images to the generating networks. The encoder part consists of two basic architectures. These are z-image-z (ziz) and image-z-image (izi) encoder training. First, the GAN training maps G(z) = z → x. Then, E(x) = x → z mapping is done on the encoder side. During the training of the ziz architecture, the samples taken in the z space are matched to the image space with the generator G, and the encoder E is trained to match each sample in the z space. In the training phase, the error between the input z samples and the reconstructed samples is minimized as in Eq. (5).

$$L_{ziz}(z) = \frac{1}{d}\|z - E(G(z))\|^2 \tag{5}$$

After the encoder is trained with the generated images, the izi architecture is used to train with real images. In training, matching from real images to hidden z-coding is done with a trainable encoder. On the other hand, the mapping to the z-image space is done with the standard generator. The loss between the input image x and the structured image is calculated as in Eq. (6).

$$L_{izi}(x) = \frac{1}{n}\|x - G(E(x))\|^2 \tag{6}$$

The distance used in Eq. (4) is the sum of the pixel-based residuals and the value n is the number of pixels in the image. The general architecture of the system is given in Fig. 3.

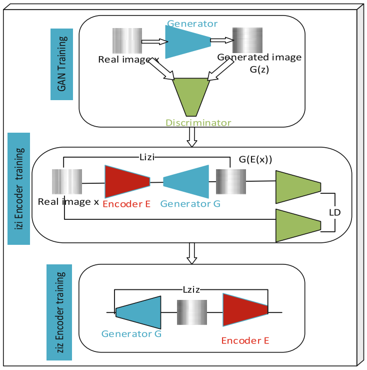


Fig. 3. Training stage of Fast-GAN-based anomaly detection.

In Fig. 3, images are produced from real images at the generator stage with GAN training. Then the izi encoder is trained to minimize the loss between the real image and the generated image. In the last step, the loss between randomly sampled images and generated images in z space is minimized.

In the detection of anomalies in the image, the deviation from the query images and the images produced with GAN is measured. Anomaly scoring is done in a similar way to the structure used in coder training.

In addition to the receiver operating characteristic (ROC) and area under ROC (AUC) metric, the performance measurement of the proposed defect detection system was used, as well as metrics such as precision, sensitivity, and F1 score.

3 Experimental Results

The proposed method for defect detection was applied to different datasets. The first dataset was taken from a measurement train operated by the Railway research and Technology Center (DATEM). The dataset is robust and consists of collapses and joint defects. Some examples from the data set are given in Table 1.

Table 1. The characteristics of used dataset.

Defect Type	Healthy	Squat	Joint
Number of Defects	400	100	100

The image enhancement step is applied to create the data set given in Table 1. Then, the rail surface is extracted from the image with the rail extraction algorithm. The image preprocessing and rail extraction stages are shown in Fig. 5.

Figure 4a shows the ROC curve. The ROC curve is over 0.99 and the rail defects are correctly modeled. In Fig. 4b, the histogram of anomaly scores is shown for solid and defective examples.

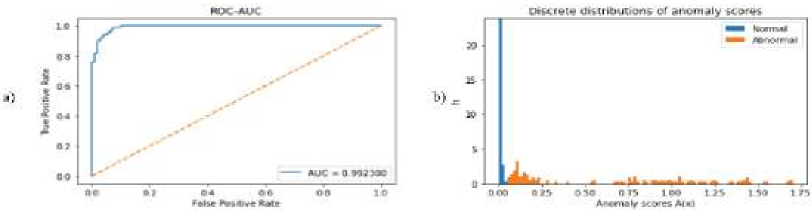
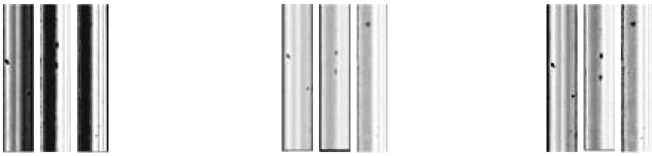


Fig. 4. a) ROC curve for test images. b) The histogram of anomaly scores for healthy and defective images.

As shown in Fig. 4b, it is seen that healthy and defective data can be distinguished correctly. The second data set consists of Rail Surface Defect Detection (RSDD-I) data set. This data set consists of a total of 67 images and the images have noises because of lighting conditions. Since there was no healthy rail image in this data set, healthy rail images in the previous data set were used in training. However, since the data set is very noisy, the noise should be removed by pre-processing. In the data set, both Retinex and Multi-Scale Retinex were applied to remove defective areas and eliminate other noise. Figure 5 gives some original images with pre-processed images.



a) Original b) Retinex c) Multi scale retinex

Fig. 5. The histogram of anomaly scores for healthy and defective images.

As seen in Fig. 5, light-related problems are also seen as a defect, as the original dataset is a very noisy image due to light problems. In the image enhanced with Retinex, the gray level is higher, and the defects are suppressed. In the multi-scale Retinex algorithm, it is seen that both the noises are removed, and the defective regions are more prominent. ROC curves for Retinex and multiscale Retinex are given in Fig. 6.

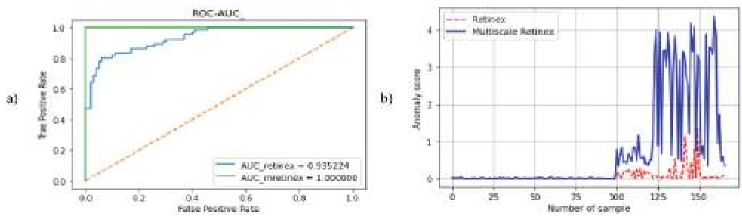


Fig. 6. a) ROC curves for two methods for RSDD dataset. b) Anomaly scores of two methods.

As shown in Fig. 6a, it is seen that the ROC curve is better in the multi-scale Retinex algorithm. Anomaly scores for two image enhancement methods are given in Fig. 6b. In

Fig. 6b, the first 100 samples consist of the health condition and the anomaly scores are low. The next 67 examples show the anomaly scores of the defective rail images. While the anomaly scores for some samples in the Retinex algorithm are close to the healthy state, it is seen that the defective states are fully distinguished in the multi-scale Retinex algorithm. Comparison results with the methods suggested in the literature are given in Table 2.

Table 2. Comparison results with other studies.

Study	Method	Dataset	P	R	F1
[6]	Deep learning-based segmentation	RSDD-I	0.86	0.83	0.83
[10]	Morphological image processing	RSDD-I	0.81	0.82	0.78
[14]	Gaussian mixture models and curvature filter	Author's dataset	0.92	0.89	0.79
[15]	Gaussian mixture models and faster RCNN	RSDD-I	0.95	0.96	0.95
This Study	**Unsupervised adversarial generative networks**	**RSDD-I**	**1.0**	**1.0**	**1.0**
		Author's dataset	**0.98**	**0.99**	**0.98**

The methods in Table 2 generally used segmentation-based techniques. The proposed technique needs only healthy state images during the training phase, but other works generally require ground truth images. In addition, a more accurate result is obtained as the noise in the image is removed because of the proposed preprocessing.

4 Conclusions

In this study, an approach to detecting rail surface defects is proposed to detect anomalies only when healthy data is available. In the proposed approach, the problems caused by lighting while taking rail images are solved with the multi-scale retinex image enhancement method. The proposed approach has been applied to two different data sets and it has been shown experimentally that it gives very good results in defect detection. In the experimental results, it has been shown experimentally that the multi-scale Retinex algorithm is more successful both in highlighting the defects and in removing noise. Overall, future research endeavors should focus on overcoming the current limitations, refining existing methodologies, and exploring innovative approaches to ensure more accurate, efficient, and reliable detection of rail surface defects.

Acknowledgments. This Work Was Supported by The FUBAP (Fırat University Scientific Research Projects Unit) Under Grant No: ADEB.2022.02.

References

1. Zhuang, L., Qi, H., Zhang, Z.: The automatic rail surface multi-flaw identification based on a deep learning powered framework. IEEE Trans. Intell. Transp. Syst. **23**(8), 12133–12143 (2021)
2. Chang, Y., Zhang, X., Lin, C., Liu, J., Shen, Y.: An efficient method for wheel-flattened defects detection based on acoustic emission technique. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **69**(2), 843–853 (2021)
3. Wagner, A., Nash, A., Michelberger, F., Grossberger, H., Lancaster, G.: The effectiveness of distributed acoustic sensing (DAS) for broken rail detection. Energies **16**(1), 522 (2023)
4. Zhu, J., et al.: Characterization of rolling contact fatigue cracks in rails by eddy current pulsed thermography. IEEE Trans. Industr. Inf. **17**(4), 2307–2315 (2020)
5. Rajamäki, J., Vippola, M., Nurmikolu, A., Viitala, T.: Limitations of eddy current inspection in railway rail evaluation. Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit **232**(1), 121–129 (2018)
6. Liu, Y., Xiao, H., Xu, J., Zhao, J.: A rail surface defect detection method based on pyramid feature and lightweight convolutional neural network. IEEE Trans. Instrum. Meas. **71**, 1–10 (2022)
7. Li, H., Wang, F., Liu, J., Song, H., Hou, Z., Dai, P.: Ensemble model for rail surface defects detection. PLoS ONE **17**(5), e0268518 (2022)
8. Aydin, I., Akin, E., Karakose, M.: Defect classification based on deep features for railway tracks in sustainable transportation. Appl. Soft Comput. **111**, 107706 (2021)
9. Tu, Z., Wu, S., Kang, G., Lin, J.: Real-time defect detection of track components: considering class imbalance and subtle difference between classes. IEEE Trans. Instrum. Meas. **70**, 1–12 (2021)
10. Nieniewski, M.: Morphological detection and extraction of rail surface defects. IEEE Trans. Instrum. Meas. **69**(9), 6870–6879 (2020)
11. Liu, S., Long, W., He, L., Li, Y., Ding, W.: Retinex-based fast algorithm for low-light image enhancement. Entropy **23**(6), 746 (2021)
12. Parthasarathy, S., Sankaran, P.: An automated multi scale retinex with color restoration for image enhancement. In: 2012 National Conference on Communications (NCC), pp. 1–5. IEEE (2012)
13. Schlegl, T., Seeböck, P., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U.: f-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks. Med. Image Anal. **54**, 30–44 (2019)
14. Zhang, H., Jin, X., Wu, Q.J., Wang, Y., He, Z., Yang, Y.: Automatic visual detection system of railway surface defects with curvature filter and improved Gaussian mixture model. IEEE Trans. Instrum. Meas. **67**(7), 1593–1608 (2018)
15. Jin, X., et al.: DM-RIS: deep multimodel rail inspection system with improved MRF-GMM and CNN. IEEE Trans. Instrum. Meas. **69**(4), 1051–1065 (2019)

Intelligent Classification of Defective Rails Through GAN and Ensemble Predictive Model

Mehmet Sevi(✉) and İlhan Aydın

Computer Engineering Department, Fırat University, Elazığ, Turkey
m.sevi@alparslan.edu.tr

Abstract. Detection of defects on the rail is an important part of rail inspection. Therefore, in recent years, automatic inspection systems are needed to inspect railway tracks quickly and accurately. Methods such as deep learning are used in automatic control systems. Deep learning methods need a lot of data sets for training. A traditional convolutional neural network cannot learn features with a small data set. Although it is easy to create a dataset of solid rail components for the training process, it is quite difficult to create a dataset of defective rail components. To create this type of dataset, it is necessary to collect images from hundreds of kilometers of railways. Therefore, in this study, images of artificial defective rail components were created using GAN, and the data set was reproduced. The quality and variety of images produced by the GAN architecture directly affect the performance of the proposed model. For this reason, the performance of the GAN model was evaluated with evaluation metrics. Then, the imperfections of the rail components with the neural network based on ensemble prediction were examined. In the study, there are three rail classes as healthy, cracked, and hole. The performances of the new images produced in each class were calculated. The basic idea of the study is to increase individual model performances by making ensemble predictions with models created as a result of training with different deep-learning methods. In the experimental results, an accuracy rate of 99.44% was obtained for the proposed method and its advantage over traditional methods was demonstrated.

Keywords: Railway · Defect Classification · Deep Learning · Adversarial Generative Networks

1 Introduction

In order to maintain the continuity of railway transportation, the maintenance of railway components must be done regularly. Therefore, the examination of defects in railway lines has become more and more important and essential. If problems in railway components are not detected at an early stage, they can cause serious accidents. Rails, one of these components, are one of the most important components that make up the railway. There are a total of two rails on the railway line, the right and left rails. Rails are components that allow wagons or trains to move comfortably on them. The cracking of the rails over time or the formation of gaps on them endanger transportation safety.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 230–237, 2023.
https://doi.org/10.1007/978-3-031-39777-6_28

Therefore, the detection of defects on the rail surfaces is extremely important. In recent years, with the developments in deep learning technologies, robust and highly accurate image classification architectures have been developed. In this study, 2 different deep learning models classified the dataset containing rail images according to the defects they contain. ResNet50 and InceptionV3 deep learning models were selected for training. An ensemble technique is proposed in which multiple models contribute equally to a combined prediction by averaging the deep learning models after training two different deep learning models. The mean ensemble model combines predictions from multiple trained models. The two models selected for this purpose, ResNet50, and InceptionV3, were chosen for their smaller size and faster processing compared to other deep learning models. The experimental results demonstrate that the proposed method achieves improved performance in classifying rail surface defects compared to using a single deep learning model. It is difficult to reach defective rail images due to their scarcity. For this reason, the rail images with defects in the dataset were augmented using a contentious generator network (GAN). Thanks to GAN, the number of defective images has been increased and diversity has been provided.

There is a need for techniques that will automatically inspect the rails, shorten the inspection time and reduce maintenance costs in order to increase the safety of railway transportation [1]. Deep learning models make a difference in many areas [2, 3]. There are several studies on railway maintenance operations in the literature. Gibert et al. Using a Deep Convolutional Neural Network, they classified the rails according to 10 material types from which they are made. They achieved a success rate of 93% [4]. Ye et al. They have proposed a system that classifies defects on the rail surface based on 3D [5]. Several geometric sequences were extracted from the 3D model of the defects in the rail surface and a distinguishable pattern was defined for each type of defect. Successful tests have been carried out with multi-class classifiers. Many studies in the literature are on deep learning-based automatic maintenance systems to facilitate the maintenance of the railway line. Rail identification and measurement are essential for safety, comfort, and maintenance. Therefore, the rail profile diagnostic system can be used to detect faults in railways [6]. Faghih-Roohi et al. proposed using convolutional neural networks as a suitable technique for feature learning to classify defects on the rail. Different network architectures characterized by different sizes and activation functions have been tested on the dataset. In this way, the efficiency of the proposed deep convolutional neural network for detection and classification has been proven. The proposed architecture achieved a validation rate of 92% on the dataset consisting of 5 different defects [7].

Alvarenga et al. introduced an embedded system that utilizes eddy current-based detection and localization of rail faults [8]. They also proposed a novel approach that involves analyzing wavelet transforms using a convolutional neural network to interpret eddy current signals. This embedded system effectively detects and classifies various types of anomalies, thereby facilitating the optimization of railway maintenance plans. The researchers conducted field tests where rail anomalies were categorized into three distinct classes. The results show a classification efficiency of 98%.

The purpose of the work is to develop a method for automatically classifying defects on railway rails using deep learning models. The originality lies in proposing an ensemble technique that combines the predictions of two different deep learning models (ResNet50 and InceptionV3) to improve classification accuracy. Additionally, the study addresses the scarcity of defective rail images by augmenting the dataset using a generative adversarial network (GAN) to increase the number and diversity of defective images. The evaluation metrics were used to assess the performance of the GAN model, which directly impacts the effectiveness of the proposed model based on the quality and diversity of generated images.

In the second part of the article, the proposed method, data set, and preprocessing steps are mentioned. Description of the dataset containing rail images and the need for increasing the number of defective images using GAN are mentioned. The paper then outlines the structure of the ensemble model, which combines the features of ResNet50 and InceptionV3 models to achieve higher accuracy in classifying rail surface defects. In the final part, the results of the proposed method are presented and compared to those of individual deep-learning models.

2 Defect Classification Based on Generative Adversarial Networks and Ensemble Method

The data set used in the study consists of 1838 rail images. Of these, 416 contain cracked rail surfaces, 492 healthy rail surfaces, and 930 hole rail surfaces. It is very difficult to reach the defective rail image on the actively used railway lines. In auto-matic maintenance systems developed for railway maintenance, the defect as well as the healthy component should be found in the images of the component. In this study, the GAN algorithm was used for this. GAN algorithmic architecture consists of two different neural networks. These; are the producer and discriminator. These networks operate in a contentious manner. While the generative network aims to simulate the real images in the training set of the noise vector that it receives as input, the splitter network tries to distinguish between synthetically generated images (fake) and real training images [9]. Fig. 1 shows the traditional GAN architecture. In the GAN architecture, the goal is to map the noise space to the actual data distribution. Training continues until the discriminator's guesses cannot distinguish between what is real and what is fake. The quality and variety of images produced by the GAN architecture directly affect its recommended performance. Some metrics in the literature measure the performance of the GAN architecture. This performance analysis includes the Visual Turing Test (VTT), Frechet Inception Distance (FID), and Number of Statistically Different Bins (NDB). VTT is the voting of artificial data and real data by experts. It is a time-con-suming process. FID and NDB are automatic performance metrics. The FID metric is as in Eq. (1).

$$FID(x, g) = \|\mu_x - \mu_g\|_2^2 + Tr(C_x + C_g - 2(C_x C_g)^{1/2} \tag{1}$$

In (1), (μ_x, C_x) and (μ_g, C_g) show the mean and covariance values of the real and fake images, respectively [9]. The closer the fake image is to the real image, the closer the FID value is to 0. NDB was proposed by Richardson and Weiss to quantify the sample diversity and mode collapse produced [10]. Given two sets of samples from the same distribution, the number of samples in a given bin should be the same up to the sampling noise. The closer the NDB value is to 0, the more successful it is.

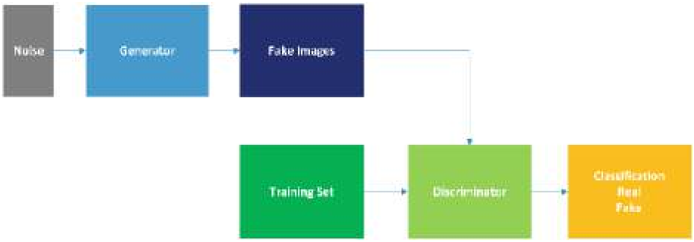


Fig. 1. GAN architecture.

After data generation with GAN, the expanded data set was made ready for classification. For training, 1500 rail images of each class were produced with GAN. In the proposed method, defective and healthy rail images in the data set are classified. ResNet50 and InceptionV3, which are transfer learning architectures, were used in this study. First of all, ResNet50 and InceptionV3 deep learning models were tested on the dataset alone.

The ResNet architecture, introduced by He Kaiming in 2015, revolutionized deep learning models by enabling increased depth without sacrificing performance. Compared to other ResNet models, ResNet50, with its 50 layers, strikes a balance between depth and complexity. It comprises 48 convolution layers, along with a Max Pooling layer and an Average Pooling layer [11]. The ResNet architecture enables the utilization of multiple layers in convolutional neural networks (CNNs), employing activation functions such as Softmax and ReLU. ResNet achieves this by incorporating residual values and blocks into the model, enhancing its overall capabilities. It has now been used to help reduce loss, maintain information gain, and improve performance at the block stage [12]. ResNet is created by adding residuals and residual blocks to the model. It prevents gradient loss by adding a different value to a system in both layers [13]. This stage also improves the value learning error from the two previous layers, even if the weight is zero in one case [13]. Thus, the network is trained more quickly. An example schematic of the ResNet50 architecture is shown in Fig. 2.

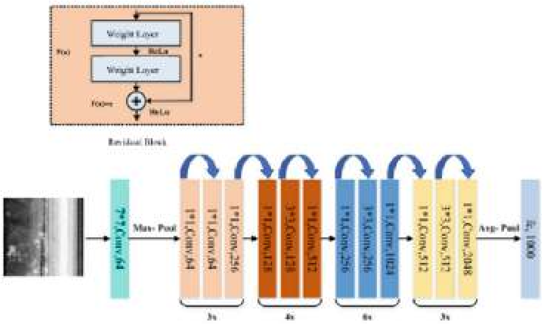


Fig. 2. ResNet50 architecture.

InceptionV3 is a deep learning model developed by Google. In 2014, it proved successful by training around 1.2 million images in the ImageNet competition [14]. The Inceptionv3 model has more layers than a simple convolutional neural network model. In this architecture, the convolution and pooling layers overlap in the sequential structure. Therefore, it showed a very strong performance in terms of memory. Unlike other models, the Inceptionv3 architecture has a layer called the Inception layer. This layer combines the input from the previous layer into a single vector after convolution and merging with different filter sizes. Thanks to this process, the inner layers recognize the shape despite the size differences in the objects by using the appropriate size filter [15]. Instead of using 5 × 5 and 7 × 7 size filters, two or three 3 × 3 size filters are used.

The study proposes a model that automatically classifies rail images according to the defect they contain, based on ensemble decision, using transfer learning-based deep neural networks (ResNet50 and InceptionV3). The schematic of the proposed model is shown in Fig. 3.

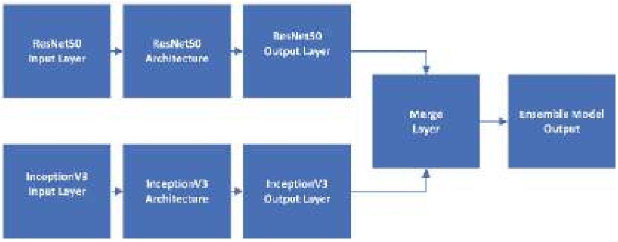


Fig. 3. Proposed method.

ResNet50 and InceptionV3 models primarily performed training operations separately. In the experimental studies for both models, ReLU is used in the convolution layers and softmax is used in the output layer. Since multi-class classification is needed in the fully connected layer in deep neural network architectures, the softmax function is used in both models. The trained ResNet50 and InceptionV3 models were combined by merging their respective model weights, and the ensemble voting technique was employed for making final predictions.

The proposed method has demonstrated its superiority over conventional and deep neural network-based classification approaches. The developed ensemble model effectively categorizes rail images into three classes: healthy, cracked, and hole. The effectiveness of the proposed defect detection system based on deep learning is assessed by analyzing rail component images obtained from a measurement sequence.

3 Experimental Results

Healthy, broken, hole rail images produced with GAN were used for training. The goal here is to improve accuracy by using more samples of defective rails for training. VTT, NBD, and FID metrics were used to evaluate the quality of images produced with GAN. When we evaluate the samples produced with the human eye (VTT), it is seen that the images produced in terms of diversity and quality are at a satisfactory level.

Table 1. Comparison results with other studies.

Class/Metric	NBD	FID
Healthy	8	103.750
Cracked	9	120.504
Hole	2	74.359

Table 1 shows the performance of the GAN architecture by class. The closer the NBD and FID metrics are to 0, the more successful the GAN model is. The GAN model was more successful in terms of variety and quality than other classes in creating a rail image with holes on its surface.

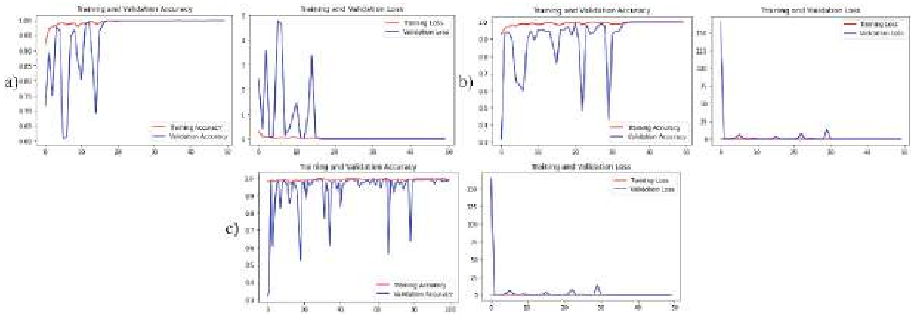


Fig. 4. a) ResNet50 accuracy and loss graph. b) InceptionV3 accuracy and loss graph. c) Ensemble model accuracy and loss graph.

Figure 4a shows the accuracy and loss graphs of the ResNet50 model. The ResNet50 model achieved an accuracy rate of 0.9889 during validation. A checkpoint parameter was defined during the training, where we monitor the accuracy rate after each validation step. Using this, a copy of the best-performing model in the training process was registered for ensemble prediction. Figure 4b shows the accuracy and loss graphs of the InceptionV3 model. The InceptionV3 model also achieved an accuracy rate of 0.9889 during validation. The best performance of the model during the training process was recorded for ensemble prediction. Figure 4c shows the accuracy and loss graph of the ensemble model. The ensemble predictive model also achieved an accuracy rate of 0.9944 in the validation step. The ensemble model has outperformed the models that make individual predictions.

4 Conclusions

The timely identification of defects in train rails is of paramount importance for ensuring human safety and optimizing economic performance. This study introduces a novel deep network model based on CNN to address this issue. The focus of this research is on the classification of defects on the rail surface. The proposed approach encompasses image replication, deep learning models, and an ensemble decision-based classification technique. Based on the comparison results presented in Table 1, it can be observed that the proposed model achieves competitive performance in classifying different rail defects. The evaluation metrics, NBD and FID, indicate that the GAN model used in generating images exhibits satisfactory quality and diversity. Lastly, Table 2 demonstrates the accuracy rates of the individual deep learning models, ResNet50 and InceptionV3, as well as the ensemble model. The ensemble model achieves the highest accuracy rate, outperforming both individual models. These results collectively support the effectiveness of the proposed ensemble model in accurately classifying rail surface defects. The proposed approach achieved a classification success of 99.44%, which is better than the currently reported methods in the literature, as can be seen in Table 2. Overall, the findings highlight the potential of deep learning models and the proposed ensemble approach in enhancing the safety and efficiency of railway transportation through automated defect detection. Based on the results obtained in this study, several suggestions for future research can be proposed. Firstly, although the proposed ensemble model demonstrated improved performance in classifying rail surface defects, further investigations can be conducted to explore other deep learning architectures or combination techniques to potentially achieve even higher accuracy rates. Additionally, GAN method expanding the dataset to include a larger number of defective rail images would provide a more comprehensive and diverse training set, enabling the models to better generalize and improve their performance.

Table 2. Comparison of the proposed ensemble model with studies in the literature.

Reference	Model	Accuracy Rate
[16]	CNN	91.19%
[17]	ResNet	95.94%
[17]	VGGNet	92.91%
[18]	CNN	97.00%
The Proposed Model	Ensemble Model	99.44%

Acknowledgments. This Work Was Supported by The FUBAP (Fırat University Scientific Research Projects Unit) Under Grant No: ADEB.2022.02.

References

1. Li, Q., Ren, S.: A visual detection system for rail surface defects. IEEE Trans. Syst., Man, Cybern., Part C (Appl. Rev.) **42**(6), 1531–1542 (2012)
2. Çetintaş, D., Tuncer, T.: Determining the type of document read using eye movement properties by hybrid CNN method. Traitement du Signal **39**(4), 1099–1108 (2022)

3. Berman, D.S., Buczak, A.L., Chavis, J.S., Corbett, C.L.: A survey of deep learning methods for cyber security. Information **10**(4), 122 (2019)
4. Giben, X., Patel, V.M., Chellappa, R.: Material classification and semantic segmentation of railway track images with deep convolutional neural networks. In: 2015 IEEE International Conference on Image Processing (ICIP), pp. 621–625. IEEE (2015)
5. Ye, J., Stewart, E., Zhang, D., Chen, Q., Roberts, C.: Method for automatic railway track surface defect classification and evaluation using a laser-based 3D model. IET Image Proc. **14**(12), 2701–2710 (2020)
6. Karaduman, G., Karaköse, M., Akin, E.: Experimental fuzzy diagnosis algorithm based on image processing for rail profile measurement. In: Proceedings of 15th International Conference Mechatronika, pp. 1–6. IEEE (2012)
7. Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R., De Schutter, B.: Deep convolutional neural networks for detection of rail surface defects. In: 2016 International Joint Conference on Neural Networks (IJCNN), pp. 2584–2589. IEEE (2016)
8. Alvarenga, T.A., Carvalho, A.L., Honorio, L.M., Cerqueira, A.S., Filho, L.M., Nobrega, R.A.: Detection and classification system for rail surface defects based on Eddy current. Sensors **21**(23), 7937 (2021)
9. Çelik, G., Talu, M.F.: Çekişmeli üretken ağ modellerinin görüntü üretme performanslarının incelenmesi. Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi **22**(1), 181–192 (2019)
10. Borji, A.: Pros and cons of gan evaluation measures. Comput. Vis. Image Underst. **179**, 41–65 (2019)
11. Mascarenhas, S., Agarwal, M.: A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification. In: 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), vol. 1, pp. 96–99. IEEE (2021)
12. Reddy, A.S.B., Juliet, D.S.: Transfer learning with ResNet-50 for malaria cell-image classification. In: 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0945–0949. IEEE (2019).
13. Serin, Z.: Meme kanserinin histopatolojik görüntüler üzerinde derin sinir ağları kullanılarak bilgisayar destekli otomatik tespiti (Master's thesis, Bilecik Şeyh Edebali Üniversitesi, Fen Bilimleri Enstitüsü) (2020)
14. Erdem, E., Aydın, T.: Göğüs Kanseri Histopatolojik Görüntü Sınıflandırması. Bilişim Teknolojileri Dergisi **14**(1), 87–94 (2021)
15. Bilginer, O., Tunga, B., Demirer, R.M.: Deri lezyonlarının evrişimsel yapay sinir ağları ile sınıflandırılması. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi **28**(2), 208–214 (2022)
16. Tripathi, M.: Analysis of convolutional neural network based image classification techniques. J. Innovative Image Process. (JIIP) **3**(02), 100–117 (2021)
17. Shang, L., Yang, Q., Wang, J., Li, S., Lei, W.: Detection of rail surface defects based on CNN image recognition and classification. In: 2018 20th International Conference on Advanced Communication Technology (ICACT), pp. 45–51. IEEE (2018)
18. Gasparini, R., et al.: Anomaly detection, localization and classification for railway inspection. In: 2020 25th International Conference on Pattern Recognition (ICPR), pp. 3419–3426. IEEE (2021)

A Case of Customer Segmentation in the Saving and Finance Industry

Fidan Khalilbayli[1(✉)] and Zekeriya Köse[2]

[1] Computer Engineering Department, Istanbul Technical University, Istanbul, Turkey
khalilbayli15@itu.edu.tr

[2] IT Application Solutions Division, Eminevim, Istanbul, Turkey
Zekeriya.Kose@eminevim.com.tr

Abstract. Businesses in the finance and investment sector are continually looking for different ways to gain a competitive advantage. One of the most effective ways to achieve this is through customer segmentation. The aim of the conducted study is to implement segmentation on demographic and behavioral customer data of the company Eminevim. After the pre-processing stage, the optimum cluster number is decided using the Elbow method. The performance of different clustering techniques is compared using the SI score where K-means outperformed.

Keywords: Customer segmentation · Data mining · Saving and finance sector

1 Introduction

With the increasing competition where customers' needs and expectations continuously change over time, it is of utmost importance to gain customer insight in order to build and maintain long-term relationships. Customer segmentation which can be defined in simple terms as the process of dividing customers into smaller distinct groups based on their common characteristics, behavior, and other relevant factors, is an essential tool for creating targeted marketing campaigns, tailoring products and services to meet specific customer needs, and also improve customer retention rates.

Some of the benefits of segmenting customers can be given as improving customer experience by better understanding their needs and developing marketing strategies that resonate with each group to increase customer satisfaction and loyalty. Also, in terms of source allocation, by focusing efforts on the most profitable customer segments, companies can maximize their return on investment and improve their bottom line. Segmenting the customers can also lead to the emergence of new service and product opportunities.

Apart from numerous benefits, certain challenges are associated with this approach. One of the most important challenge is the need for accurate and up-to-date customer data, which might be problematic in cases where customers

Supported by Eminevim.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 238–245, 2023.
https://doi.org/10.1007/978-3-031-39777-6_29

are reluctant to share personal information. Another example of a possible challenge is the need for effective communication and collaboration across different departments within the business.

In this study, the segmentation is implemented on the dataset obtained from the customers of Eminevim which is a Turkish saving and financing services company of Emin Group that provides an innovative financing solution known as a "household participation system" or "household savings system" to customers who might not otherwise have access to traditional mortgage options. Essentially Eminevim's model enables customers to purchase their homes and vehicles by paying monthly installments. The main difference between Eminevim and other finance companies located in Turkey is that the underlying profit-sharing model of Eminevim is based on non-interest-based transactions. Instead of charging interest on loans, the company operates by pooling the savings of participants and using the funds to distribute to participants according to their level of investment.

Eminevim has been growing rapidly in recent years, and as of 2021, it had over 2 million participants and had helped approximately 250,000 individuals to become homeowners. The company is headquartered in Istanbul and has branches throughout Turkey, as well as expanding its operations in other countries.

To best of our knowledge, the topic of customer segmentation which is conducted by utilizing both demographic and behavioral data has not been exploited in this field.

The rest of the paper is organized as follows: In Sect. 2, a detailed investigation of the conducted existing work in the field of segmentation has been presented, followed by Sect. 3 which describes the utilized dataset and the implemented pre-processing methods. In Sect. 4, the methods and their respective results are reported, respectively. In the final section, the study is summarized and suggestions regarding future work are given.

2 Literature Review

Segmentation of customers based on their characteristics has gained significant attention, especially in the last decade, thus there has been extensive scholarly work in this field. Numerous researchers have focused on different aspects of this topic which can be listed as the utilized methods, used criteria for performance measurements, and also the efficiency of segmentation in the marketing context.

A broad range of features for segmenting purposes have been utilized in the literature that can be categorized under three main subjects which are demographic, life cycle, and behavior [17]. The demographic information of the customers is usually obtained by using questionnaires during the registration process and includes features such as ethnicity, age, gender, income, marital status, family size, education, birthplace, etc. [21]. The features regarding the life cycle divide the duration of the customer in the company into different stages such as newly joined, long-standing, recently lost, lost after a short period of time, etc. which enables the development of different strategies for each group to ensure maximum retention. The most frequently utilized features in the segmentation studies are regarding the behavioral patterns and habits [19].

Selecting the optimal number of clusters to partition the data is tedious work and either can be determined through domain knowledge or several methods such as elbow methods, and silhouette score. An incorrect cluster count can lead to uninformative results in terms of segmentation. Automatic clustering techniques do not require a predefined cluster number, and have been widely investigated by using vastly different methods which can be listed as genetic [6], bacterial evolution [7], particle swarm optimization [13] algorithms, etc. Another study proposed the hybrid approach which consisted of artificial bee colony optimization and K-means algorithm which outperformed other automatic clustering techniques in the benchmark datasets [11].

Apart from automatic clustering techniques, various unsupervised machine learning methods are commonly applied for segmentation purposes. K-means is one of the most utilized algorithms on this subject [18]. In [2], the performance of K-Means and Fuzzy C-Means algorithms were compared on a wholesale customer's data in terms of cluster integrity and computational time which concluded that the latter algorithm is a more decisive way for segmentation. By using RFM score obtained from the customer's internet banking transactions, K-Means, and K-Medoids methods were implemented where the former achieved significantly smaller score both in the intra-cluster distance and Davies-Bouldin index [3].

The usage of ensemble learning for clustering purposes was exploited in the study [12] where the outcomes of three different machine learning models namely Random Forest, Gradient Boosting, and k-Nearest Neighbors have combined and a precision score of 76% was achieved on the customer behavior dataset obtained from UCI ML.

To make the cluster assignment more interpretable to the end user, several studies focused on the explainability part of clustering by implementing decision-based tree visualization [4,10].

3 Dataset

The foundation of the segmentation models relies on customer data. To achieve desired outcomes and make informed decisions, it is crucial to work with accurate data. In this particular case, the dataset is obtained from the customers of the company Eminevim registered between 2020 and 2022. The data of the customers who cancelled their contract and left the company were removed from the dataset. The demographic information as well as the behavioral data of the customers were analyzed to generate appropriate segments. The list of the utilized features is provided in Table 1. Here, the organization amount refers to the total amount the client will receive on the determined date according to the agreement, whereas the organization fee indicates the amount that the company gains and is estimated based on the organization amount and installment count. Installment count and amount are predetermined fixed values and indicated in the contract, but paid installments can differ according to the financial status of the customer during each month. The amount of paid organization fee is also calculated since it can be paid in monthly installments.

Table 1. Feature List

Feature
Age
Gender
Marital Status
Registration District
Organization Amount
Organization Fee
Down Payment Amount
Installment Count
Installment Amount
Average Paid Organization Fee
Average Paid Installment

Pre-processing is an essential step in segmentation since it prepares the data for clustering analysis, which is usually extremely sensitive to the input data. So, the quality of the input data significantly affects the quality of the clustering results. Pre-processing involves several tasks such as data cleaning, dimensionality reduction, data normalization, and transformation which can enhance the performance of the clustering.

Clustering requires a reliable dataset free from any noise or outliers which can be ensured by removing irrelevant data, correcting errors, and handling missing data. In some of the utilized features, irrelevant data is detected by using domain knowledge. For instance, during the years 2020–2022, it was not possible to sign a contract with the company if the client was a minor. Thus, clients younger than 18 were removed from the dataset. The same process was implemented on customers older than 100. Another usage of domain knowledge can be observed in inspecting irrelevant organization amounts since each campaign has a fixed limit on this amount.

For detecting and removing outliers in installment amount, and paid installment amount, the IsolationForest algorithm is utilized which is scalable and effective in large datasets [8]. This method relies on creating random decision trees, where each tree is built by selecting a random subset of features and splitting the data at random until each point is isolated [20]. The contamination hyperparameter which estimates the number of outliers is set to 0.1.

To transform the monetary behavioral data into normal distribution, Box-Cox statistical technique is utilized which is calculated as raising the variable Y to a power λ to transform the data, and calculating the logarithm of the result if λ is equal to zero. Here, λ is chosen as the value that maximizes the log-likelihood function [16] (Eq. (1)).

$$Y_i^{(\lambda)} = \begin{cases} \frac{Y_i^{\lambda}-1}{\lambda} & \text{if } \lambda \lesseqqgtr 0 \\ ln(Y_i) & \text{else} \end{cases} \tag{1}$$

To assure equal contribution of all features to the clustering analysis, the monetary behavioral data is standardized using min-max normalization by scaling the values to a range between 0 and 1 [14].

To deal with the missing data problem which was observed only in the demographic features (approximately 0.0625% of all the dataset), data imputation based on monetary behavioral data is conducted. At first, the organization amount distribution is categorized according to its histogram, then the customers with missing age value are assigned with the mode age of the corresponding organization amount category. The other demographic features were imputed in the same manner based on different behavioral features. Also, the customers which did not possess any demographic features were removed from the dataset and all of the nominal features were categorized.

Principal Component Analysis (PCA) is implemented to reduce the dimensionality of the dataset while retaining most of the important information. The operation of PCA is as follows [9]:

1. For each feature calculate the mean (μ) and standard deviation (σ) and apply the following formula:
$$X_{new} = \frac{X - \mu}{\sigma} \tag{2}$$
2. Compute the covariance matrix for Step 1.
3. Calculate the eigenvalues and eigenvectors of Step 2.
4. Sort eigenvalues and the respective eigenvectors in descending order.
5. Select top k eigenvalues.
6. Transform the original dataset.

4 Method

After the pre-processing step, the optimal number of clusters is selected by using the elbow method. At first, the clustering algorithm is implemented for different numbers of clusters. It was selected to be between 1 and 10 for this case. For each cluster count, the sum of the squared distances between each data point and its cluster centroid (WCSS) is calculated [5] and the optimal number of clusters is selected as the number at the elbow point which is where the WCSS value levels off. This value corresponds to 3 in our dataset (Fig. 1).

Once the cluster number is determined, several unsupervised machine learning models were implemented on the dataset namely, K-means, K-medoids, K-means++, Fuzzy C-Means, and Mini Batch K-means and their Silhouette (SI) coefficient were compared and K-means algorithm outperformed. The visualization of the clusters using K-means can be observed in Fig. 2. The SI coefficient

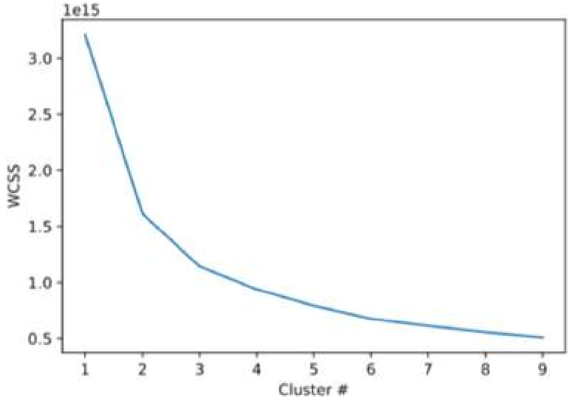


Fig. 1. Elbow Method Result

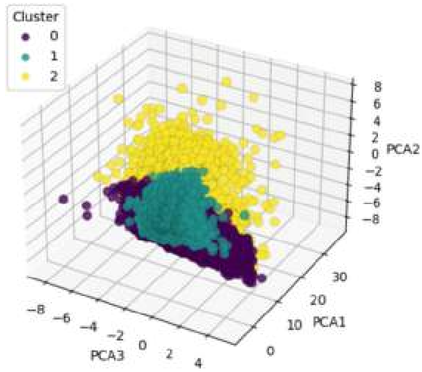


Fig. 2. Clustering Result

basically defines how well a data point fits into its assigned cluster, as well as how well it is separated from other clusters [1]. Thus, the quality of clustering is assessed by both separation and cohesion [15]. The calculation is as in Eq. (3) where a_i indicates the average distance between a data point and the rest of the data points in the same cluster, whereas b_i is the average distance between a data point and all other points in the nearest cluster.

$$SI_i = \begin{cases} 1 - \frac{a_i}{b_i} & \text{if } a_i < b_i \\ 0 & \text{if } a_i = b_i \\ \frac{b_i}{a_i} - 1 & \text{if } a_i > b_i \end{cases} \tag{3}$$

The main steps of the K-means algorithm can be described as [10]:

1. Choose the number of clusters k.
2. Assign initial centroids to each of the k clusters from randomly selected data points.

3. Calculate the distance between each data point and each centroid. Assign data points to the cluster with the nearest centroid.
4. Measure the average of all data points for each cluster and assign the average as the new centroid.
5. Repeat steps 3 and 4 until no significant change is observed in the centroids or pre-defined iteration count has been reached.

5 Conclusion

While investigating the distribution of each feature in each cluster distinct differences between clusters are observed. In terms of monetary behavior, C2 has the highest values in all features and also has the highest married and male rates, whereas C0 has the lowest results in monetary behaviour and has the least amount of married clients.

In this study, customer segmentation based on demographic and behavioral data obtained from Eminevim is conducted. Various data pre-processing steps are implemented and Elbow method is used to achieve the ideal cluster number. K-means is selected as the best performing algorithm in terms of SI score and the clusters are visualized.

As for the future work, the usage of not only partitioning-based but also density, hierarchical and grid based clustering techniques will be implemented. Also, clustering will be exploited on the time series payment features with dynamic time warping.

References

1. Acar, S., Köroğlu, F., Duyuler, B., Kaya, T., Özcan, T.: Customer segmentation using RFM model and clustering methods in online retail industry. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds.) INFUS 2021. LNNS, vol. 307, pp. 69–77. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-85626-7_9
2. Afrin, F., Al-Amin, M., Tabassum, M.: Comparative performance of using PCA with K-means and fuzzy C means clustering for customer segmentation. Int. J. Sci. Technol. Res. **4**, 70–74 (2015)
3. Aryuni, M., Madyatmadja, E.D., Miranda, E.: Customer segmentation in XYZ bank using K-means and K-medoids clustering. In: 2018 International Conference on Information Management and Technology (ICIMTech), pp. 412–416. IEEE (2018)
4. Bertsimas, D., Orfanoudaki, A., Wiberg, H.: Interpretable clustering: an optimization approach. Mach. Learn. **110**, 89–138 (2021)
5. Bholowalia, P., Kumar, A.: EBK-means: a clustering technique based on elbow method and K-means in WSN. Int. J. Comput. Appl. **105**(9), 17–24 (2014)
6. Chang, D.X., Zhang, X.D., Zheng, C.W., Zhang, D.M.: A robust dynamic niching genetic algorithm with niche migration for automatic clustering problem. Pattern Recogn. **43**(4), 1346–1360 (2010)

7. Das, S., Chowdhury, A., Abraham, A.: A bacterial evolutionary algorithm for automatic data clustering. In: 2009 IEEE Congress on Evolutionary Computation, pp. 2403–2410. IEEE (2009)
8. Hariri, S., Kind, M.C., Brunner, R.J.: Extended isolation forest. IEEE Trans. Knowl. Data Eng. **33**(4), 1479–1489 (2019)
9. Karamizadeh, S., Abdullah, S.M., Manaf, A.A., Zamani, M., Hooman, A.: An overview of principal component analysis. J. Signal Inf. Process. **4**(3B), 173 (2013)
10. Khan, R.H., Dofadar, D.F., Alam, M.G.R.: Explainable customer segmentation using K-means clustering. In: 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp. 0639–0643. IEEE (2021)
11. Kuo, R.J., Zulvia, F.E.: Automatic clustering using an improved artificial bee colony optimization for customer segmentation. Knowl. Inf. Syst. **57**(2), 331–357 (2018)
12. Mustafa, S.M.N., Akhtar, A., Noronha, J.T.P., Salman, M., Baig, M.A.: Customer segmentation using machine learning techniques. In: 2023 International Multidisciplinary Conference in Emerging Research Trends (IMCERT), vol. 1, pp. 1–7. IEEE (2023)
13. Omran, M.G., Salman, A., Engelbrecht, A.P.: Dynamic clustering using particle swarm optimization with application in image segmentation. Pattern Anal. Appl. **8**, 332–344 (2006)
14. Patro, S., Sahu, K.K.: Normalization: a preprocessing stage. arXiv preprint arXiv:1503.06462 (2015)
15. Petrovic, S.: A comparison between the Silhouette index and the Davies-Bouldin index in labelling IDS clusters. In: Proceedings of the 11th Nordic Workshop of Secure IT Systems, vol. 2006, pp. 53–64. Citeseer (2006)
16. Poirier, D.J.: The use of the box-cox transformation in limited dependent variable models. J. Am. Stat. Assoc. **73**(362), 284–287 (1978)
17. Romdhane, L.B., Fadhel, N., Ayeb, B.: An efficient approach for building customer profiles from business data. Expert Syst. Appl. **37**(2), 1573–1585 (2010)
18. Shen, B.: E-commerce customer segmentation via unsupervised machine learning. In: The 2nd International Conference on Computing and Data Science, pp. 1–7 (2021)
19. Tabianan, K., Velu, S., Ravi, V.: K-means clustering approach for intelligent customer segmentation using customer purchase behavior data. Sustainability **14**(12) (2022)
20. Tokovarov, M., Karczmarek, P.: A probabilistic generalization of isolation forest. Inf. Sci. **584**, 433–449 (2022)
21. Wu, Z., Jin, L., Zhao, J., Jing, L., Chen, L.: Research on segmenting E-commerce customer through an improved K-medoids clustering algorithm. Comput. Intell. Neurosci. **2022**, 9930613 (2022)

Intelligent Network Monitoring System Using an ISP Central Points of Presence

Yousef Alkhanafseh[1], Mahsun Altın[1(✉)], Altan Çakır[2], Esra Karabıyık[1,2], Ekrem Yıldız[1,2], and Sencan Akyüz[1,2]

[1] Istanbul Technical University, 34467 Sarıyer, Turkey
{alkhanafseh15,altinma21}@itu.edu.tr,
{yousef.alkhanafseh,mahsun.altin}@turk.net
[2] Turknet İletiŞim Hizmetleri, 34394 Şişli, Turkey
{altan.cakir,esra.karabiyik,ekrem.yildiz,sencan.akyuz}@itu.edu.tr

Abstract. The proliferation of both internet usage and users have been remarkably increased due to certain situations that influenced face-to-face communications, which in turn have created high pressure on Internet Service Providers (ISPs). This research mainly aims to boost ISP services by conducting near real-time analysis for customer's behavior movements based on their score of central Points of Presence (POP). In addition, this study focuses on establishing special Recurrent Artificial Intelligence (RNN) architecture to make daily sales predictions based on various central POPs. The process utilizes different RNN architectures, Long Short Time Memory (LSTM) and Gated Recurrent Unit (GRU), and compares them in order to make smart scoring measurements for customers' high-dimensional data. As a result, it can be concluded that LSTM architecture has achieved much better Mean squared Error (MSE) than GRU architecture. LSTM outperforms GRU in forecasting less sensitive outliers, with an average Mean Absolute Error (MAE) of 1.354 for LSTM and 1.554 for GRU. Additionally, LSTM performs better in forecasting outliers, with an average MSE of 3.592 compared to GRU's average of 4.8. Thereafter, the obtained results are merged over private Application Programming Interface (API) and monitored over smart reports. Eventually, the outcomes of this research can be summarized in providing several benefits for customers such as increasing internet performance, reaching promised speed, and shortening activation times. ISP-related benefits such as gaining reputation, promoting sales, and reducing customers' negative support tickets can be achieved as well.

Keywords: Internet · Telecommunication · Data Analysis · Network Monitoring · Smart Reports · POP Scoring

1 Introduction

The advent of new technologies and the increasing the accessibility of the internet have led to increase the number of people that can access to the online world

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 246–254, 2023.
https://doi.org/10.1007/978-3-031-39777-6_30

which in turn have created networks congestion, bad internet connections, and more frequent service outages [1]. Therefore, it has put pressure on ISPs to ensure that their networks can handle the increased traffic and provide reliable service to their customers. The need for near real-time analysis of customer behavior movements based on their score of central POP is becoming increasingly important in today's network business environment. A central POP is a key location where the network connects to the internet [2]. POPs Scoring technique can be accomplished using different sources of data such as data of internet package sales, customer activation, surveys, customer support tickets, cancellation requests, internet speed measurements, and delivered speed parameters. These data-sets offer important insights into customers preferences, satisfaction, faced issues, and their internet connection consistency. Thus, it paves the way for promoting sales, making improvements, and taking smart decisions to enhance the overall quality of internet service. For instance, if a large number of customers are experiencing slow network speeds or connectivity issues, businesses need to be able to identify and address these issues as quickly as possible to prevent customer churn [3]. The objective and originality of this paper can be stated as building special RNN architectures from scratch and pick the best one in order to develop smart forecasting model that can be used for all POPs. The main used models are LSTM and GRU and their results are evaluated based on two different metrics which are MSE and MAE. The obtained results will be merged using private Application Programming Interface (API) and will be monitored through smart reports. The rest of the paper is organized as follows. Section 2 introduces the related work. Section 3 describes the general architecture of the used NN system. Section 4 exploits the internal structure of our special LSTM and GRU models, respectively. Section 5 explains the implementation details of RNN to forecast sales. Section 6 discusses and compares the obtained results from both models and then states the best one. Finally, Sect. 6 gives a brief conclusion.

2 Related Works

Several studies have been conducted to explore the feasibility and effectiveness of intelligent network monitoring system using ISP central points. In this section, it is supposed to review the related works that have been published in the literature and discuss their contributions to the field. One research has conducted a study on the same filed utilizing a longitudinal database of internet infrastructure development to examine city accessibility to the commercial internet [4]. Their research investigated fiber-optic backbone points of presence (POP) established by commercial Internet service providers, finding that larger metropolitan areas maintain dominant shares of telecom infrastructure, but several mid-sized metros are emerging as important centers for telecommunication interconnection. This study provides valuable insights into internet accessibility. In addition, another study has made a comparative analysis of a developed web-based application named ISP-Perf with test mobile systems (TEMs) in a mobile broadband measurement environment [5]. Their research evaluated the quality of service (QoS) metrics such as upload and download speeds, as well as the latency of 3G MTN network in Nigeria, and recommended ISP-Perf for measuring network performance due to its low error margin. Also, a methodology for customer churn prediction (CCP) using machine learning techniques has been implemented in [6]. Their research utilized data pre-processing, feature analysis, feature selection, predictive models (logistic regression, naive bayes, support vector machine, random forest, decision trees, boosting, and ensemble techniques), and K-fold cross-validation for hyper-parameter tuning. Besides, it is proposed an efficient and accurate sales forecasting model using machine learning, and demonstrates its performance on a Walmart retail goods dataset [7]. Their study lights the way for predicting the sales and forecasting it in different sectors. Lastly, a framework for near-real time monitoring server reboots in a large-scale internet service environment, using multiple data infrastructures and machine learning-based anomaly detection has been presented in [8]. A research presents LSTM and GRU models that incorporates customer heterogeneity into model parameters and attention mechanisms to extract trends from customer purchase history, resulting in improved sales forecasting accuracy [9].

3 System Architecture

The main architecture of POPs based multivariate sales forecasting is clearly summarized in Fig. 1. Fundamentally, it consists of seven stages which are data collecting, data manipulation, data splitting, neural network preparing, neural network optimizing, and results visualization. The first level involved collecting data from multiple tables in different databases and merging them into one data file using Extract, Transform, Load (ETL) processes. The Extract phase involves identifying and collecting the necessary data, Transform involves cleaning and preparing the data for analysis, and Load involves inserting the data into the destination system. Also, it is used for near real-time monitoring to control the network. The next level is, principally, responsible about performing data manipulation such as data imputation to compensate the missed observations, data grouping to get the desired shape of the data, data scaling to unit observations scales between 0 and 1, and multi co-linearity to detect related features. After that, the stage of data splitting has been implemented to give three data-sets which are train, validation, and test with the ratios of %80, %10 and %10, respectively. After achieving these steps, data becomes ready to be fed to RNN model. The special RNN model of this project has been completely built using TensorFlow and Python Programming Language. The whole details of the model will be entirely discussed in the next section. Following, established RNN models are precisely fine tuned by optimizing their hyper-parameters such as batch size, number of hidden layers, neurons, learning rate, regularization, etc. The model's performance is then evaluated based on MSE and MAE metrics. Eventually, obtained model is used for forecasting POPs sales.

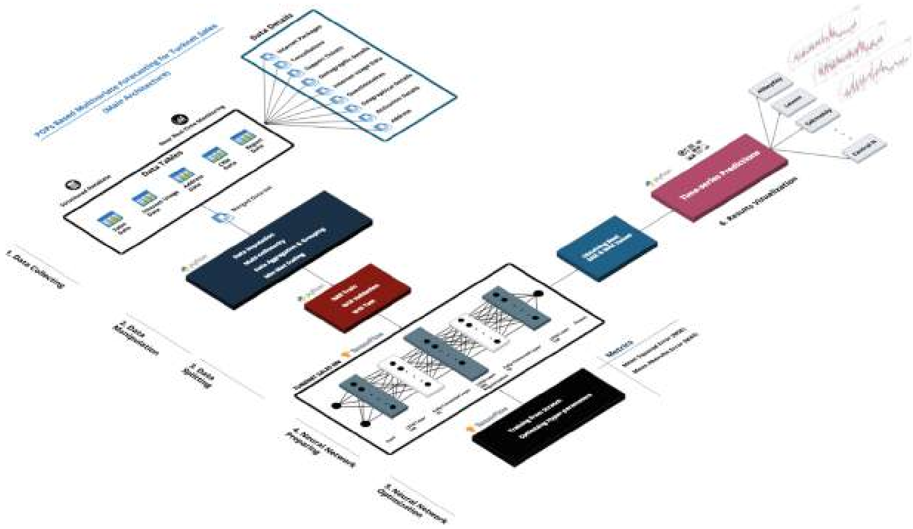


Fig. 1. TurkNet POP Analytic System Architecture.

4 Model

Inspired from the huge learning capacity of Artificial Neural Networks (ANNs), the model of this project is entirely built using one type of ANN which is RNN. However, one model of RNN should be carefully selected. During this project, two RNN models, LSTM and GRU, are particularly compared based on two different metrics which are MSE and MAE. The design of their networks are the same which is obviously mentioned in Fig. 2. The architecture of the two models is the same, the only difference is that LSTM layers are replaced with GRU layers. As it is clear from Fig. 2 that the architecture consists of multiple inputs and one outputs so that it is supposed to do multivariate LSTM/GRU forecasting. In addition it mainly consists of three LSTM/GRU layers with 128, 256, and 128 neurons, respectively. It is crucial to notice that L1 regularization is applied on the activity and L1L2 regularization is applied on the kernel of the middle LSTM/GRU layer in order to prevent over-fitting. Moreover, it has two fully-connected layers with 32 and 16 neurons which are connect LSTM/GRU layers with each others. It is important to note that the design of this architecture is specially designed based on several experiences.

The parameters, weights of connections between layers, of both models, LSTM and GRU, are stated in Tables 1 and 2, respectively. In summery, the established LSTM and GRU architecture have total parameters of 445,617 and 337,841, respectively. The difference between the total parameters comes from the difference between the architectures of both LSTM [10] and GRU [11] models.

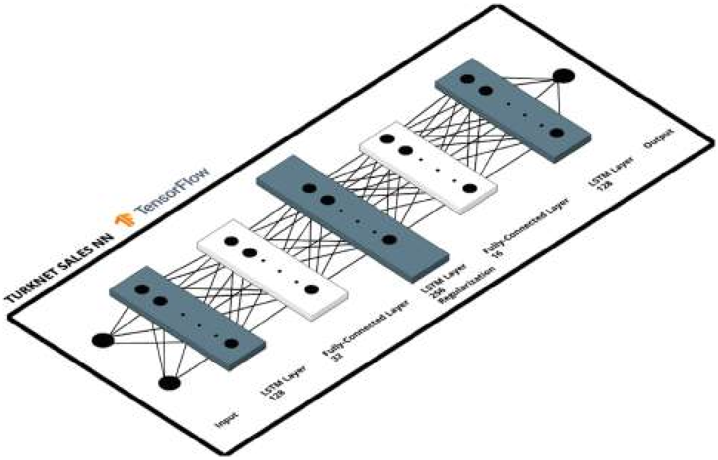


Fig. 2. Special Neural Network Architecture.

Table 1. LSTM Architecture

Layer type	Output Shape	Parameters
lstm_1	(None, 5, 128)	67072
Dense_1	(None, 5, 32)	4128
lstm_2	(None, 5, 256)	295936
Dense_2	(None, 5, 16)	4112
lstm_3	(None, 128)	74240
Dense_3	(None, 1)	129

Table 2. GRU Architecture

Layer type	Output Shape	Parameters
gru_1	(None, 5, 128)	50688
dense_1	(None, 5, 32)	4128
gru_2	(None, 5, 256)	222720
dense_2	(None, 5, 16)	4112
gru_3	(None, 128)	56064
dense_3	(None, 1)	129

5 Implementation Details

It is common that before establishing any Artificial intelligence model some operations and changes should be done on data-set called data manipulation in order to enhance the performance of the model. In this project, four different operations were implemented on the data-set. The first one is data imputation which has been implemented using mean value to compensate missed values. Moreover, values of observations must be in the same range so that metrics can be expressed well. Accordingly, min max scalar is used to scale the data-set between 0 and 1. Furthermore, multi co-linearity is executed to detect related and unrelated features so that related features can be dropped. Thus, the performance of the model is significantly improved as it is fed by independent features. When there are similar features, RNN model can have difficulty in updating weights and biases at each iteration. Consequently, it is detected that four types of support tickets have the most serious impact on sales. The other features are somehow dependent so that it is decided to drop them. Lastly, selected features have been aggregated into one feature and the data is grouped based on different POPs and date-times. As a result, the input data can be considered as a matrix with

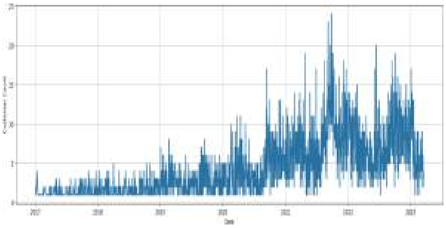


Fig. 3. Sales Distribution.

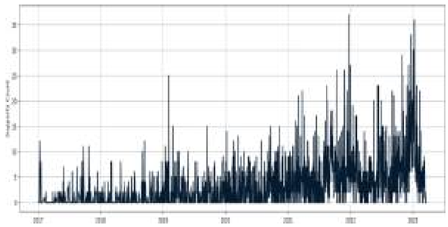


Fig. 4. Support Tickets Distribution.

Table 3. Models' Hyper-parameters.

Hyper-Parameter	Value
Past days	7
Batch size	32
LSTM_1 layer	128
Dense_1 layer	32
LSTM_2 layer	256
Dense_2 layer	16
LSTM_3 layer	128
Learning rate	0.00005
L1 Regularization	0.045
L2 Regularization	0.04
Optimizer	Adam
Epochs	60
Loss	MSE

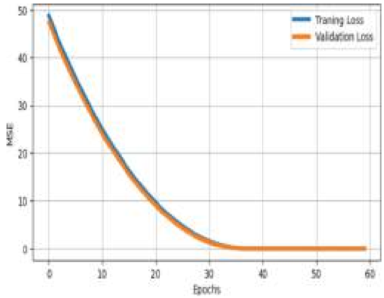


Fig. 5. Models' Train and Validation losses.

two columns. A sample of input data is represented in Fig. 3 and 4. Based on several experiments, it is assumed that five past days are the most appropriate number to forecast one future day. Then, data is split into three parts which are train, validation, and test with the ratios of 0.8, 0.1, and 0.1, respectively.

At the beginning, only one LSTM or GRU layer was not sufficient to give rational results and it suffered from under-fitting and vanishing gradients problem. Subsequently, the RNN architecture has been developed as the one that is mentioned in the previous section. Nevertheless, this architecture experienced over-fitting problem. So as to overcome this problem and to be able to get accurate results, the hyper-parameters of the RNN architecture, see Table 3, are extremely optimized. Regularization algorithms of Lasso Regression (L1) and combined L1 and Ridged Regression (L2) are added only to the middle LSTM/GRU layer in order to penalize the RNN architecture from learning small details which in turn helps in reducing over-fitting problem. The resulted graph of train and validation losses can be clearly seen in Fig. 5. It is obvious that both train and validation losses are decreasing at the same time and approximately at the same level.

6 Results

The results that are obtained from constructed RNN models, LSTM and GRU, are totally discussed in this section. When data is divided into train, validation, and test the resulted maximum observations per POP are 1057, 131, and 131, respectively. As a result, it can be understand that the model learning part is done using information of 1057 days and its learning process has been validated using information of 131 and it is tested using information of 131 days. Tables 4 and 5 show the obtained results of MSE, MAE, and %100 correct day predictions of five different POPs from both LSTM and GRU models, respectively. It can be claimed that LSTM Architecture performs better than GRU architecture for all the available POPs. For instance, while forecasting the sales of Levent POP, LSTM architecture was able to get MAE of 1.67, number of error predicted customers sales for each day. On the other hand, GRU architecture has contributed to MAE of 1.91. In addition, in the second POP example, Alibeykoy, GRU architecture outperforms LSTM architecture in number of %100 predicted days. But, its performance for the other days is worse than the performance of LSTM architecture. Figure 6 and Fig. 7 display the predictions of LSTM and GRU architectures for the same POP, Küçükçekmece. It is obvious from the figures that LSTM architecture is much better than GRU architecture in detecting average and high frequency sales.

Table 4. Metric Results of LSTM Model

POP	MSE	MAE	Predictions
Levent	4.87	1.67	38
Alibeyköy	4.96	1.69	31
Küçükçekmece	5.8	1.85	33
Çekmeköy	1.29	0.79	54
Esenler	1.04	0.77	42

Table 5. Metric Results of GRU Model

POP	MSE	MAE	Predictions
Levent	6.21	1.91	32
Alibeyköy	6.51	1.87	34
Küçükçekmece	8.2	2.15	31
Çekmeköy	1.73	0.96	45
Esenler	1.4	0.88	38

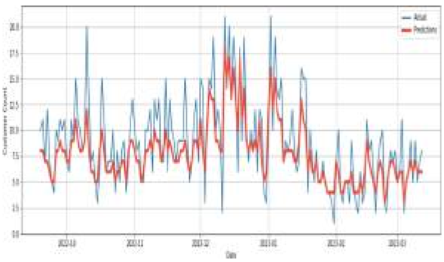


Fig. 6. LSTM Result (Küçükçekmece).

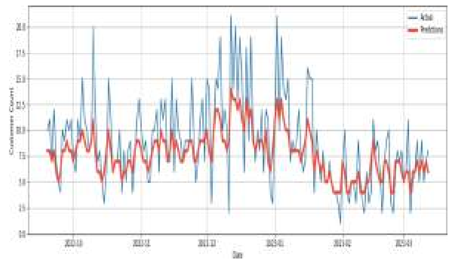


Fig. 7. GRU Result (Küçükçekmece).

7 Conclusion

In conclusion, this study primarily utilized specially designed RNN architectures with LSTM and GRU layers. The study successfully demonstrated the effectiveness of LSTM and GRU models in forecasting daily sales for POPs. However, even though both models were accurate, it is important to note that the RNN architecture with LSTM layers outperformed the one with GRU layers based on MSE and MAE metrics. It is emphasized that LSTM has shown to be superior to GRU in forecasting all points that are less sensitive to outliers, with an average MAE of 1.354, while GRU has an average of 1.554. Furthermore, even in forecasting outliers, LSTM has performed better, with an average MSE of 3.592 compared to GRU's average of 4.8. ISPs can use the established model to improve their sales forecasting and strategic decision-making processes. Furthermore, the results of our study have broader implications for the ISP industry beyond sales forecasting. By implementing these models, ISPs can improve their overall service quality and customer satisfaction by increasing internet performance, reaching promised speeds, and shortening activation times. Additionally, ISPs can benefit from improved reputation, increased sales, and reduced negative support tickets. For future research, it is recommended to explore the potential of incorporating additional features such as activation and cancellation numbers related to POPs as inputs to the RNN architecture. This may enhance the model's performance and generate more accurate predictions. Overall, this study provides a strong foundation for future research in this area and underscores the potential benefits of deep learning models for the ISP industry.

References

1. Shim, J.P., Warkentin, M., Courtney, J.F., Power, D.J., Sharda, R., Carlsson, C.: Past, present, and future of decision support technology. Decis. Support Syst. **33**(2), 111–126 (2002)
2. Chamberland, S.: Point of presence design in Internet protocol networks with performance guarantees. Comput. Oper. Res. **32**(12), 3247–3264 (2005)
3. Hadden, J., Tiwari, A., Roy, R., Ruta, D.: Computer assisted customer churn management: state-of-the-art and future trends. Comput. Oper. Res. **34**(10), 2902–2917 (2007)
4. Grubesic, T.H., O'Kelly, M.E.: Using points of presence to measure accessibility to the commercial Internet. Prof. Geogr. **54**(2), 259–278 (2002)
5. Orakwue, S.I., Al-Khafaji, H.M., Rathi, A.: Comparative analysis of ISP-Perf and TEMs in mobile broadband QoS metrics measurement. In: 2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT), pp. 1–3. IEEE, January 2023
6. Lalwani, P., Mishra, M.K., Chadha, J.S., Sethi, P.: Customer churn prediction system: a machine learning approach. Computing **104**, 1–24 (2022)
7. Shilong, Z.: Machine learning model for sales forecasting by using XGBoost. In: 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE), pp. 480–483. IEEE, January 2021
8. Lin, F., et al.: Near-realtime server reboot monitoring and root cause analysis in a large-scale system. In: 2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks-Supplemental Volume (DSN-S), pp. 37–40. IEEE, June 2021
9. Kaneko, Y.: Customer-base sequential data analysis: an application of attentive neural networks to sales forecasting. In: 2019 International Conference on Data Mining Workshops (ICDMW), pp. 349–355. IEEE, November 2019

10. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
11. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling (2014). arXiv preprint arXiv:1412.3555

Estimating the Success Likelihood of Variant Price Promotion Schemes: A Machine Learning Approach

Burak Kilercik[1,2] and Ömer Zeybek[1,2](✉)

[1] Migros, İstanbul, Turkey
{burakkil,omerze}@migros.com.tr
[2] R&D Department, Migros Ticaret AŞ, Ataşehir, 34735 İstanbul, Turkey

Abstract. Price promotions are an essential marketing tool for acquiring new customers and retaining existing ones while providing them with a personalized store experience. It is also a powerful tool for sustaining main topics such as the consistent increase in return on investment and customer loyalty. This study took an in-depth look at predicting successful promotions. It examined the transition from traditional rule-based evaluation methods to explainable machine learning methods in identifying successful promotions. In this research, firstly, the problem of defining a successful past promotion (on a large scale) was addressed, which is still an argumentative topic in the sector. Since the subject of promotion is inherently erratic, it was aimed to define a successful promotion by formalizing an objective Key Performance Indicator. Based on this definition, two target metrics were created using profitability and quantity. Decisive features shaped around these two metrics were created with historical data. Consequently, two distinct classification models were estimated for both target definitions, the first assessing a significant increase in the product's profitability and the second considering any remarkable growth in sales quantities. Finally, promotion success is inferred by blending probabilities estimated through two different algorithms. The study revealed that the success of price promotions, which are erratic in nature, can be predicted with high confidence with up-to-date methods and a blended interpretation strategy, most of which can be generalized to other sectors.

Keywords: Machine learning · Product promotion · Successful promotion estimation

1 Introduction

Price promotions are a popular marketing strategy used by retailers to stimulate consumers to purchase their products or services. These kinds of promotions are also called temporary price reductions (TPR) which contain temporary deals or discounts on products. Promotion has four main objectives; providing information to consumers and others, increasing demand, differentiating a product or

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 255–261, 2023.
https://doi.org/10.1007/978-3-031-39777-6_31

category, stabilising sales and accentuating product's value. Although measuring the effectiveness of price promotions is crucial for businesses, it is also difficult to assess real impact created through promotions.

Assessing the effectiveness of price promotions has been a longstanding research interest among researchers working in marketing domain. In recent years, the focus has intensified, as marketers seek to leverage promotions to boost sales and cultivate customer loyalty. Promotions are recognized as a crucial component of the marketing mix, as they inform, persuade, and remind consumers with the aim of influencing their opinions or eliciting a response.

This paper presents a novel approach to classification of successful promotion campaigns. Our study provides new insights into research area by employing hi-performance machine learning algorithms.

2 Literature Review

Early studies on sales promotions were based on notional mathematical models, as the potential of scanner data had not yet been recognized. However, with the advent of more advanced data collection methods, it has become possible to evaluate the quantitative effects of promotions. Once calculation method for the net profit is well studied, researchers began to work on studies aiming to classify successful price promotions.

Identifying successful price promotions is a prominent research area in the field of marketing science. With the emergence of advanced quantitative methods and improvements in computing technology, the number of studies employing machine learning methods to assess promotions has increased significantly. By analyzing large datasets and considering a range of factors such as promotion type and timing, researchers have been able to identify factors that contribute to successful promotions. Such insights have important implications for marketers seeking to optimize their promotion strategies and achieve greater returns on investment.

In their study using an Indian retailer's data [4] they employed a decision tree model to predict which promotions are more effective. The study found that promotion type and timing are the most important features in increasing the probability of success. Another study by [6] used neural networks for identifying of successful price promotions. They found that promotion type and promotion duration are the most significant factors affecting the success of promotions. A comparative study by [5] examines effectiveness of four classification methods such as Logistic Regression, Naive Bayes, Decision Tree, and Random Forest in an assessment of promotional assessment. They found that all Random Forrest approach is the most effective one. Another study by [2] aimed to develop a classification model that would identify the impact of promotional campaigns based on their effects on customer life time value.

3 Conceptual Model

3.1 Overview

The conceptual model proposed in this research serves as a framework for predicting the success of variant price promotion schemes by integrating explainable machine learning techniques. The model considers two key performance indicators: profitability and sales quantity, as well as decisive features derived from historical data. It aims to provide a generalizable approach for determining the likelihood of success for price promotions.

3.2 Data Collection and Preprocessing

To develop and test the proposed conceptual model, we collected related historical price promotion and sales data from a large Turkish retailer's database. The data covers a period from January 2021 to November 2022 and includes promotions from various product categories. To ensure the validity of our analysis, we applied the following data restrictions and preprocessing steps:

- Only promotions with a minimum discount of 10

These data boundaries and preprocessing steps were applied consistently to all promotions included in the analysis. The resulting dataset consisted of 5,000 unique promotions from ten different product categories. By applying these data restrictions and preprocessing steps, we aimed to ensure that the data used in our analysis was representative, unbiased, and reliable for the purpose of developing and testing the proposed conceptual model.

3.3 Data Collection and Preprocessing

To develop and test the proposed conceptual model, we collected historical promotional data from a large retail chain in the United States. The data covers a period from January 1, 2021 to November 30, 2022 and includes promotions from various product categories. To ensure the validity of our analysis, we applied the following data restrictions and preprocessing steps:

- Sales of tobacco and alcoholic beverages were excluded from the analysis, as these products are not subject of promotions according to legislation.
- Promotions with personalized discounts based on customer or store preferences were excluded from the analysis, as these promotions may not be directly comparable to other promotions that are available to all customers and stores.

These data boundaries and preprocessing steps were applied consistently to all promotions included in the analysis. The resulting dataset consists of more than 135,000 unique promotional products, 12,000 unique promotional campaigns, and 24,000 unique products. By applying these data restrictions and preprocessing steps, we aimed to ensure that the data used in our analysis was representative, unbiased, and reliable for the purpose of developing and testing the proposed conceptual model.

3.4 Formation of Classifier Target Value

In machine learning, every classification model requires a classifier target value, which is extracted from past values to act as a discrimination rule between desired and undesired conditions. Then, a classification model is built using this classifier as a target variable and past values as model features. In our study, we chose to use two different classifier target values to create two different models.

The calculation idea behind the two key performance metrics is as follows: we first calculate the profitability and sales quantity of each promotion using historical data. Then, we use these two metrics to define two different target variables. The first target variable indicates whether a promotion is profitable or not, and the second target variable indicates whether a promotion leads to increased sales or not.

We then use these two target variables to build two separate classification models, each using the historical promotion data as model features. The resulting models can then be used to predict the likelihood of success for new promotions, based on their profitability and sales quantity.

Profitability. It reflects the net gain or loss resulting from a promotional campaign. To calculate the labels derived from this metric, the following formula is used:

$$\frac{\text{IntraP Profitability}}{\frac{\text{PostP Profitability}+\text{PreP Profitability}}{2}} \tag{1}$$

The pre and post promotion time periods are assessed as follows:

- Pre Promotion (PreP) period start date: *promotion start date* $-$ *promotion duration* -1
- Pre Promotion (PreP) period end date: *promotion start date* -1
- Post Promotion (PostP) period start date: *promotion end date* $+$ *promotion duration* $+1$
- Post Promotion (PostP) period end date: *promotion end date* $+1$

From the values obtained using the formula above, promotions exceeding the pre-determined thresholds given by category managers are marked as successful, while the others are marked as failed promotions.

Sales Quantity. An another essential metric for assessing promotion success, as it indicates the volume of products sold during the promotion period. To calculate the labels derived from this metric, the following formula is used:

$$\frac{\text{IntraPQuantity}}{\frac{\text{PostPQuantity}+\text{PrePQuantity}}{2}} \tag{2}$$

The pre and post promotion time periods are assessed as follows:

- **Pre Promotion (PreP) Start Date:**
  promotion start date − *promotion duration* −1
- **Pre Promotion (PreP) End Date:**
  promotion start date −1
- **Post Promotion (PostP) Start Date:**
  promotion end date + *promotion duration* +1
- **Post Promotion (PostP) End Date:**
  promotion end date +1

From the values obtained using the formula above, promotions exceeding the determined thresholds are marked as successful, while the others are marked as failed promotions.

3.5 Feature Extraction

Using historical data, decisive features are extracted to be used as input for the machine learning models. These features may include, but are not limited to:

- **Promotion duration:** The length of the promotion period.
- **Promotion type:** The specific type of price reduction or deal offered, such as percentage discounts or buy-one-get-one-free deals.
- **Product category:** The classification of the product being promoted.
- **Product expensiveness level:** The relative costliness of the product being promoted.
- **Seasonality:** The time of year the promotion takes place, accounting for seasonal trends and fluctuations in consumer behavior.
- **Past promotion performance:** The historical performance of similar promotions on the same or related products.
- **Demographics:** The target consumer segment for the promotion, including age, gender, income, and location.
- **Promotion depth:** The magnitude of the price reduction or discount offered.

$$\frac{\text{Distributed Discount During Promotion}}{\text{Distributed Discount During Promotion} + \text{Total Sales}} \tag{3}$$

- **Cannibalism:** The degree to which the promotion of one product affects the sales of other products within the same category. The same formula for promotion depth (shown above) is applied to other products within the same category to assess the cannibalism effect.

4 Method and Empirical Findings

4.1 Methodology

The machine learning approach used in this study employs the Light Gradient Boosting Machine (LightGBM) classifier, a popular gradient boosting framework that was developed by [3]. It offers high performance and efficiency in classification problems. The LightGBM classifier is utilized for both target metrics: profitability and sales quantity in order to allow for a transparent and interpretable understanding of the factors influencing promotion success.

The LightGBM classifier is chosen for its ability to handle large datasets, its efficiency in training on high-dimensional data, and its capacity to handle categorical features. Additionally, it provides high accuracy and interpretability, which are essential for understanding the factors that contribute to promotion success.

For each target metric, a separate LightGBM model is trained and tuned to optimize the performance in predicting the success likelihood of a promotion. The models are evaluated using cross-validation, and their hyperparameters are fine-tuned to achieve the best results.

In hyperparameter optimization phase, the Optuna library is employed, which is an efficient and flexible framework for hyperparameter tuning [1]. Optuna facilitates the optimization process by automatically searching for the best hyperparameter configurations, taking into account the specific characteristics of the dataset and the LightGBM model.

4.2 Model Evaluation and Validation

The performance of the classification models is evaluated using standard metrics, such as precision, recall, F1 score, and area under the receiver operating characteristic (ROC) curve. These metrics provide a comprehensive assessment of the model's ability to accurately predict the success likelihood of a promotion based on profitability and sales quantity.

To ensure the reliability and generalizability of the model, cross-validation techniques are employed. By partitioning the data into multiple training and testing sets, the models are tested on various subsets of the data to evaluate their performance and prevent overfitting. This process enables the model to adapt to different market conditions and customer behaviors, ensuring its applicability across a range of promotional scenarios.

In this study, the F1 score is chosen as the evaluation metric to assess the performance of the predictions made by the LightGBM models. The F1 score is a suitable choice because it combines both precision and recall into a single metric, providing a balanced view of the model's performance. This is particularly important when dealing with imbalanced datasets or when both false positives and false negatives are of concern.

The achieved F1 scores for the trained models with test data corresponds to November 2022 in this research are 0.86 for sales quantity and 0.80 for profitability. Thee confusion matrices for both models can be seen from Fig. 1.

These high F1 scores indicate that the models are capable of predicting the success likelihood of variant price promotion schemes with a high degree of accuracy, providing valuable insights for businesses in optimizing their promotional strategies (Fig. 2).

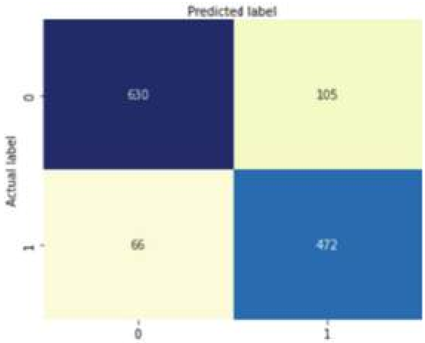


Fig. 1. Confusion Matrix for Quantity Model (Test Data)

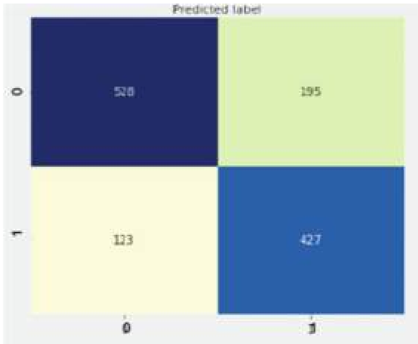


Fig. 2. Confusion Matrix for Profit Model (Test Data)

5 Conclusions

According to our model product's sales amount prior to promotion term, ratio of loyalty cardholders among the customers purchased for promoted items, discount rates are most influential variables determining success of any promotion. Moreover regarding on generalisable test date validations our model proved we are able to identify success conditions with over 0.80 accuracy with F1 score.

Once the machine learning models have been validated and optimized, they can be deployed for real-world applications. Retailers and businesses can use the models to predict the success likelihood of their proposed price promotions and refine their marketing strategies accordingly. By leveraging the insights provided by the models, businesses can optimize their promotional campaigns for maximum profitability and sales quantity, ultimately improving customer acquisition, retention, and overall business performance.

References

1. Akiba, T., Sano, S., Yanase, T., Ohta, D., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631 (2019)
2. Gyulai, T., Molnár, B.: Classification of promotional campaigns based on their impact on customer lifetime value. J. Bus. Res. **89**, 336–342 (2018)
3. Ke, G., et al.: LightGBM: a highly efficient gradient boosting decision tree. Adv. Neural. Inf. Process. Syst. **30**, 3146–3154 (2017)
4. Kumar, N.R., Kumar, P.S.: A decision tree approach to classify effective sales promotion. Int. J. Bus. Manag. Invention **3**(8), 01–05 (2014)
5. Narasimhan, S., Udupi, R., Shankar, R.: A comparative study of classification models for predicting the success of marketing promotions. J. Retail. Consum. Serv. **32**, 1–11 (2016)
6. Yee, P.M., Rau, P.P.: Identifying successful sales promotions: a neural network approach. J. Retail. Consum. Serv. **27**, 33–41 (2015)

Classifying the Success of Transfers Made by Turkish Super-League Teams Using Advanced Machine Learning Techniques

Serhan Berke Erden(✉), Ünal Berke Kaymak, and Tolga Kaya

Department of Management Engineering, Istanbul Technical University, Sarıyer, 34367 Istanbul, Turkey
{erdens18,kaymaku18,kayatolga}@itu.edu.tr

Abstract. It is a matter of controversy whether the transfers made in the football industry are efficient or not. The aim of the study is to explore the efficiency of transfers made in the football industry using machine learning techniques. In this context, a methodology to model the success of transfers based on Turkish Super League data is suggested. In the modelling processes, the data of the transfers taken from the Tranfermarkt website were used. The target variable is created as binary and the classification problem is the consideration. Accordingly, the data of 16 teams and 2261 players in total were analysed using advanced machine learning methods. Results reveal that transfers of young and homegrown players are relatively more efficient compare to those of the others.

Keywords: Machine Learning · Football · Transfer Efficiency · Turkish Super League · Classification

1 Introduction

Football has become the focus of attention of many people. As Paramio, Buraimo, and Campos explain, the sport of football began to become popular in England in 1885. From this date on, players started to be employed for money and transferred with testimonial fees [1]. Football teams can add players to their teams by covering the sales price of them. However, not every player can show the expected efficiency. In the market, it is seen that football teams fall into financial difficulties as a result of costly expenditures not getting performance compensation. It is an undeniable fact that the clubs in the Turkish league are worse in the list of teams that have financial problems compared to Europe. According to Taştan's research in 2021, Turkish league teams took the last place in the debt ranking of European clubs with a loss of €263 millions. Major teams Fenerbahçe, Beşiktaş, Galatasaray and Trabzonspor spent €195.445.000 on 120 players between 2014 and 2017 [2]. In order to prevent these debts, financial fair play (FFP) rules are followed and different solutions are explored.

In this study, it was desired to learn whether the teams in the Turkish league will be successful in their future transfers by using machine learning techniques. The main

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 262–268, 2023.
https://doi.org/10.1007/978-3-031-39777-6_32

purpose here is to provide minimum loss in player transfers for clubs by providing suggestions based on the predictions. The target variable to be studied is the change between the market value of the football player at the time of the transfer and the value reached at the end of the following season. If the market value has appreciated during the season, the team has made a successful transfer and vice versa. As a result of various methods, it was aimed to contribute to the teams by providing maximum efficiency.

The paper is structured into five sections after the Introduction, which include literature review, methodology, data, findings and conclusion. The Literature review provides an overview of the previous research on transfer efficiency in football, while the methodology explains the research approach and data sources used. The data section offers an in-depth analysis of the players in the Turkish Super League, and findings section presents the study results. Finally, conclusion section summarizes the main findings and suggests recommendations for future research.

2 Literature Review

Machine learning is a field of computer science that allows machines to learn by increasing performance without human intervention, wants to achieve prediction and optimization results, and is directly related to artificial intelligence [3]. This state of art is very popular today and is needed in almost every field. Kumar, Kaur and Singh state that machine learning is widely used in many areas such as the algorithm that shows the most suitable way for the person in the travel industry, studies that measure people's emotions and stress level in social networks, and cancer diagnosis in medicine [4]. Furthermore, Baboota and Kaur have made progress with feature engineering and machine learning algorithms on English Premier League data, and they have signed a study that tries to predict the match results [5]. Although the studies on the English League are more intense due to the fact that the English league is extremely popular in the world, some researches have also been carried out in the Turkish League. For instance, there is a study has been carried out by Tümer and Akyıldız, which tries to predict the league rankings of the teams in Turkish league. It is achieved by using various machine learning procedures on the Turkish Super League data between the years 2015 and 2018 [6]. As a result, machine learning is used in so many fields, football is one of them. These examples will shed light on new football analyses and will pave the way for a study that will enable clubs to make their transfers rationally and classify their efficiency.

3 Methodology

Various methods have been developed thanks to improved computers and the expansion of many machine learning algorithms. The classification methods may have different mathematical formulations from regression problems within supervised techniques. In this study, 15 different classification methods were applied to data.

LGBM (Light Gradient Boosting Machine) is a fast and scalable gradient boosting framework that uses tree-based algorithms. Massaoudi, Abu-Rub, Refaat, Chihi and Oueslati explain that the complexity of samples with large gradient values is obtained and preserved by a method called Gradient-based One Side Sampling (GOSS). This is achieved from random sample to ensure distribution of the data. Massaoudi, Abu-Rub, Refaat, Chihi and Oueslati explain that the complexity of samples with large gradient values is obtained and preserved by a method called Gradient-based One Side Sampling (GOSS). This is achieved from random sample to ensure distribution of the data.

$$T_i = -\frac{\lambda L(y_i, f_{t-1}(X_i))}{\lambda f_{t-1}(x_i)}, i = 1, \ldots m \tag{1}$$

L, y and f are loss functions. Leaf Wise Tree Growth (LWTG) method is applied in LightGBM. In order to decrease the loss, the leaf with the most delta loss is chosen, which leads to increase in LWTG. The targeted function is given as Obj(t) and equals to

$$Obj(t) = L(t) + \pi(t) + c \tag{2}$$

Here, $\pi(t)$ and $L(t)$ correspond to regular and loss, t is sampling time and c is any parameter. If the method is complex, it is penalized by the $\pi(t)$ function. In the $L(t)$ function, the difference between the predicted value and the actual value is calculated. This calculation is like below

$$L(t) = \sum_{n=1}^{n} \left(y_i(t) - (\hat{y})_i(t)\right)^2 \tag{3}$$

where y corresponds to the true value and $\hat{y}$ is the estimated [7]. Moreover, CatBoost is a gradient boosting technique that employs decision trees as its foundation and has the ability to handle categorical features directly, bypassing the need for preprocessing. Massaoudi, Abu-Rub, Refaat, Chihi and Oueslati also mentioned that it works permutation-based without any manipulation on categorical variables and uses Target Based Statistics (TBS) method for this. x_{υ_p} can be calculated by assuming $\upsilon = (\upsilon_1, \ldots, \upsilon_n)$ as

$$\frac{\sum_{j=1}^{p-1} \left[x_{\upsilon_j k} = x_{\upsilon_p k}\right] Y_{\upsilon_j} + a.P}{\sum_{j=1}^{p-1} \left[x_{\upsilon_j k} = x_{\upsilon_p k}\right] + a} \tag{4}$$

Here, if $\left[x_{j,k} = x_{i,k}\right]$ is 1 when $x_{j,k} = x_{i,k}$; if not, it takes the value 0. P is the prior value and is its weight is a when greater than 0. This a can solve the problem of incoming variables with a low probability [7]. In addition, XGBoost is a decision tree-based ensemble algorithm that uses gradient boosting for automatic model optimization. Furthermore, AdaBoost is a technique that combines multiple simple classifiers to create a more accurate classifier by adjusting the weights of wrongly classified observations.

4 Data

In order to have reliable and accurate data, information was obtained from tranfermarkt.com [8]. According to Wikipedia, it is a German-based website founded in 2000 by Matthias Seidel and currently 51% owned by Axel Springer. On the site, there are many football-related information such as football teams, players, stadium information, etc. Furthermore, there are also information such as match scores, player transfer information, estimated transfer fee and statistical data [9].

In order to question the success of the clubs in the transfers of the players, the data of the teams that took part in the Turkish Super League at least 4 times in the last 5 seasons (2017–2022) were consulted and the dataset was created (Table 1).

Table 1. Data Dictionary

#	Variable	Type	Definition
1	Name	character	Name and surname of football player
2	Club	character	Football team that transferred that season
3	Season	integer	Season in which he is transferred and will play immediately after
4	Position	character	Original position in the squad
5	Age	integer	Age of the player in mentioned season
6	Nationality	character	Nationality of player
7	Former_Club	character	Football club he played before the transfer
8	Transfer_Fee	float	Money paid to the former club when the transfer took place
9	Goals	integer	The number of goals player scored during the season
10	Assists	integer	The number of assists player made during the season
11	Match_Played	integer	Number of matches played in the mentioned season in the club to which he was transferred
12	Total_Time	integer	Total game play duration of the player in the mentioned season in the club
13	YC_cnt	integer	Number of yellow cards the player has seen in the mentioned season
14	YRC_cnt	integer	Number of cards that the player received that season turned from a yellow to a red card
15	RC_cnt	integer	Number of red cards directly seen by the player that season
16	PPG	float	Points Per Game, which is an indicator of success performance per match
17	Success	binary	Success of the transfer by club: 1 successful, 0 unsuccessful transfer

The dependent variable is binary and created by the difference between the market value of the player at the time of the transfer and at the end of the season after the transfer. Accordingly, if there is an increase in value, it takes 1 and indicates a successful transfer for the club; if there is a decrease or not change, the transfer is unsuccessful and indicated by 0. When everything is created, a total of 2261 player information from 16 teams has been reached. Among these players, there may be players who have taken part more than once due to the transfer of the same player in different seasons. All manipulations, feature engineering and algorithm developments on the data were carried out in the Python environment. When the dependent variable efficiency is examined, it is seen that there are 747 missing values. When the reason was investigated, it was understood that there were non-integers in the market values of the players who were excluded from the squad, who were thought to be young and on loan, and whose interest in football was approaching the end of their career. According to first analysis of data, 73.1% of transfers were unsuccessful while remaining was successful.

5 Findings

15 different classification algorithms were used for the data set and as a result, accuracy and F1 scores were obtained (Table 2).

Table 2. Models as well as Accuracy and F1 Scores

Models	Accuracy	F1 Score
LGBM Classifier	0.76	0.75
XGB Classifier	0.77	0.76
Random Forest Classifier	0.78	0.73
Decision Tree Classifier	0.67	0.67
Ada Boost Classifier	0.8	0.87
Bagging Classifier	0.77	0.75
Catboost Classifier	0.77	0.85
Logistic Regression	0.72	0.72
Linear SVC	0.62	0.65
SGDC Classifier	0.72	0.63
Linear Discriminant Analysis	0.65	0.67
K Neighbors Classifier	0.72	0.62
Quadratic Discriminant Analysis	0.35	0.34
Dummy Classifier	0.74	0.64
SVC	0.71	0.63

When the results are evaluated, XGBoost, CatBoost, AdaBoost, LightGBM, Bagging and Random Forest are the most successful among them. Considering these methods, 10-fold cross validation was applied (Fig. 1).

Finally, the age of the football player is the most important variable affecting the transfer efficiency in the algorithm results. Accordingly, the younger the player, the more successful the transfer of the team can be said. The excess of the total time played, Turkish player transfers, player productivity per match, are other variables that significantly affect the success of the transfer. In addition, teams have made more successful transfers in recent seasons (Fig. 2).

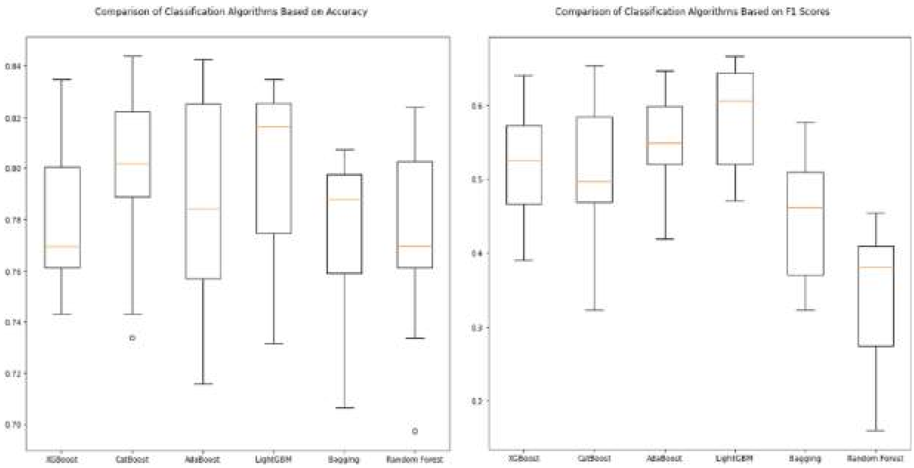


Fig. 1. 10-Fold Cross Validation for 6 Best Models Based on Accuracy and F1 Score

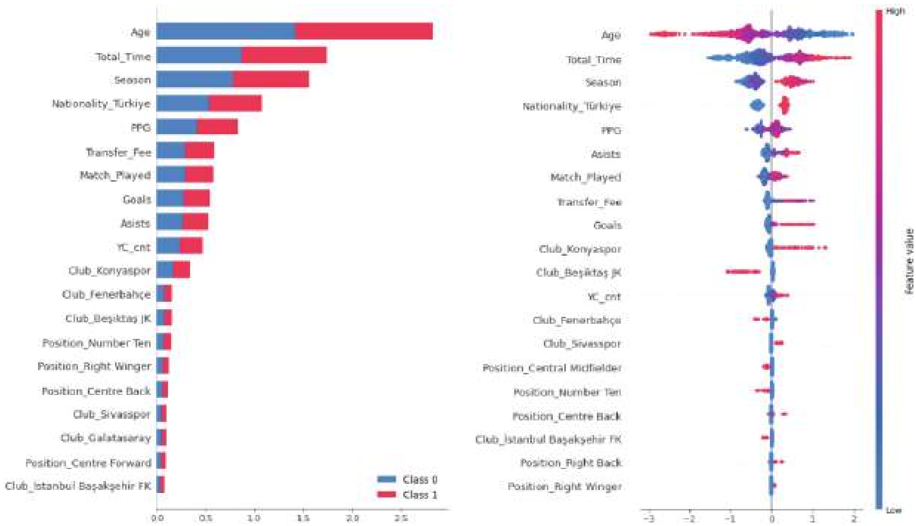


Fig. 2. Feature Valuation

6 Conclusion

This study shows LightGBM to be the most effective method for analyzing football transfer success. Successful transfer maps have been created for Turkish clubs, with recent seasons showing more success and increased earnings from Turkish transfers. The study identifies Beşiktaş and Fenerbahçe as making the most ineffective transfers, while Konyaspor and Sivasspor perform well. Additionally, number 10 transfers tend to be unsuccessful, while right-back transfers are effective.

For further studies, larger datasets, continuous auditing, hyperparameter tuning, and the use of neural network methods like Gradient Descent Algorithm are recommended. Overall, this study is open to improvement with no major obstacles.

References

1. Paramio, J.L., Buraimo, B., Campos, C.: From modern to postmodern: the development of football stadia in Europe. Sport Soc. **11**(5), 517–534 (2008)
2. Taştan, H.Ş: The effects of the number of foreign players and market values of the super league clubs on the sports success of the clubs. Int. J. Sport Culture Sci. **9**(2), 268–281 (2021)
3. Bi, Q., Goodman, K.E., Kaminsky, J., Lessler, J.: What is machine learning? A primer for the epidemiologist. Am. J. Epidemiol. **188**(12), 2222–2239 (2019)
4. Kumar, Y., Kaur, K., Singh, G.: Machine learning aspects and its applications towards different research areas. In: 2020 International Conference on Computation, Automation and Knowledge Management, pp. 150–156. IEEE (2020)
5. Baboota, R., Kaur, H.: Predictive analysis and modelling football results using machine learning approach for English premier league. Int. J. Forecast. **35**(2), 741–755 (2019)
6. Tümer, A.E., et al.: Prediction of soccer clubs' league rankings by machine learning methods: the case of Turkish super league. In: Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 17543371221140492 (2022)
7. Massaoudi, M., Abu-Rub, H., Refaat, S.S., Chihi, I., Oueslati, F.S.: An effective ensemble learning approach-based grid stability assessment and classification. In: 2021 Kansas Power and Energy Conference, pp. 1–6. IEEE (2021)
8. Transfermarkt. (2022, November 7). Transfermarkt Football Database. Transfermarkt.com. https://www.transfermarkt.com/
9. Wikipedia contributors. (2023, January 10). Transfermarkt. In Wikipedia. https://en.wikipedia.org/wiki/Transfermarkt

Workforce Prediction with Machine Learning Algorithms for an FMCG Retailer

Furkan Oral(✉), Mustafa Kaan Aslan, Buse Mert, and Birol Yüceoğlu

Migros Ticaret A.Ş. R&D Center, Ataşehir, İstanbul, Turkey
foral@migros.com.tr

Abstract. In FMCG retailing, store employees have an essential role in performing store operations (such as receiving goods, replenishing shelves) and helping customers make purchase decisions. It is estimated that 6 percent of customers leave the store without making any purchases when they do not receive the service they need during shopping. An important reason for this problem is the lack of sufficient employees in the store. By contrast, having more employees than necessary negatively affects the profitability of the store. This study aims to provide an estimation of monthly workforce level for each store of an FMCG retail chain using machine learning approaches. Due to the nature of the problem the labels are provided by actual workforce levels and can be faulty because of previous poor planning. In order to avoid this problem, we select a sample of stores by eliminating outliers according to their efficiency levels. In this study, we use both actual and forecasted sales, store properties and previous workforce information. We approach this case as a regression problem and use machine learning algorithm named LightGBM. We present our computational results.

Keywords: Workforce Management · Machine Learning · Retail Analytics

1 Introduction

There are certain financial factors that are the basis of retail stores. One of them is the cost of products to be sold. In order to solve this problem, it is necessary to minimise the Cost of Goods Sold (COGS) and for this purpose supply chain optimisation is applied. Another cost factor is operational costs. Employee hiring, training and employment costs constitute an important part of operational costs. We decided to carry out this study in order to optimise these costs on a store basis. Some situations arise when retail store workhours are not optimized. These are called undersizing and oversizing. Undersizing a retail stores' workhour level causes customer dissatisfaction and labor dissatisfaction due to inability to adequately assist the customer during the shopping process,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 269–274, 2023.
https://doi.org/10.1007/978-3-031-39777-6_33

make suggestions, and the need to constantly work at a very high level of performance. In retail stores, especially FMCG stores, routine and crucial tasks have to be done by store employees such as taking products from supplier into store, product transportation between backroom to shelves, detecting goods misplaced by customers, helping customers when they make a purchase decision if needed [4]. In general, retail store managers consider the sales level and store budget while allocating the number of employees required. These staffing rules ignore the effects of inventory levels. Thus, errors occur in determining the number of employees for a retailing store. Some surveys indicate the 33% of customers who have experienced poor services in stores could not find any solution, at last 6% of all customers could not buy anything due to the stated problem [2] . This study aims to provide an estimation of optimum workload level needed by employees monthly per each store by considering sales, number of goods sold and physical attributes of retail stores. The data has a 24-month history. This approach ensures whether the workload level change is due to inflation. In retail stores, fingerprinting is used to record the time spent at work by employees. Employees have their fingerprints scanned separately at the beginning and end of their working hours. We try to estimate the optimum value of the working hours for the stores. However, since the recorded working hours of the store are based on the number of employees determined by the human resources department, we try to avoid bias when using this data as a label. To avoid this, the productivity statistics of the stores are taken into account. In this paper, our main goal is to find the best-performing algorithm for this specific regression problem. After trying several algorithms, we choose LightGBM. and for the error metric we choose Mean Absolute Percentage Error(MAPE).

The organization of the paper is as follows: Sect. 2 includes our literature search of related subject. Section 3 introduces how approached to the problem. Section 4 exhibits our prediction performances. Section 5 includes suggestions for further researchers about related subject and conclusion of the paper.

2 Literature Review

In retail industry, companies have to utilize their expenses to compete other competitors and also to continue their existance in the market. The performans of a retail store determined by how correctly conducted routine tasks such as shelving, reaching customer in store and inventory tasks [5]. On the other hand, nowadays online retail shoppings suppress the size of sales of physical retail store. A report has published by US Census states that the retail market share of online retailers noticeably increasing. In 2019, online retail is growing faster than traditional retail and has an 11.1% share of the retail market [1]. That situation indicates how important physical retailers have to use their resources efficiently. The assigned workforce level of each store plays a critical role. Considering the expenses of retail stores, the first factor that stands out apart from the costs of goods sold is the expenses of retail store employees. A research indicates that hire, train and employing costs 10–20% of sales and more than 50% of operating costs [5]. Undersizing or oversizing of store workforce comes with the different circumstances but it affects negatively store performance either way. Undersizing results in customer dissatisfaction and labor dissatisfaction. The long-term consequence is a worsening of the store's financial situation. In contrast, oversizing leads conformance, service quality and profit increases at first but when sales predicted in a unrealistic that situation results with downsizing and the morale of the rest of employees will be decreased due to downsizing. When a 168-store retail organization optimized its workforce, its annual revenue increased by 4.5% [3]. Briefly, optimizing retail workforce is challenged task because both to perform the necessary tasks in the best way possible and to do so at the least cost must be reconsidered. A brief explanation is visualised about the optimal workforce for retail stores in Fig. 1 [5].

Fig. 1. Retail workforce sizing how effects the retail store performances

3 Methodology

This study aims to predict optimum workforce level for each retail store which is in specific store formats. We use past shopping activies for 24 months, employee data, item category data and sales area of retail stores as features. We use these to create regression problem. We collect past shopping activities which is in specific store formats and within the metrics which will be introduced later. We use PySpark, Hadoop technology to reach these datasets. Shopping activity data is collected on a daily basis. Then we convert this data into monthly data. As we convert this data into monthly data, we collect it according to the job descriptions we will mention later. Instead of using past shopping activities as features of regressor, we create features by dividing the shopping activities of each month in order to obtain both the sales changes between the years and the sales changes within the season. Due to feature importance results, we take out some of these ratios. Additionally, we add formats of retail stores as a categorical feature to regressor.

Table 1. Detailed information of employee data

Feature	Description
month_period	Indicate which month personnel make this shift
personnel_id	Indicate personnel id
store_number	Indicate in which store the personnel make this shift
job_desc	Indicate the job description of personnel
shift_hour	Indicate how much time spend on shift by a personnel

Employees scan their fingers on the sensor at the beginning and end of the shift. In this way, it is possible to measure which employee has worked how much time in which store as seen in Table 1. This allows us to monitor and analyze the time employees spend on the job and the work they do. Thus, the workforce predictions we make for the stores are more accurate.

Table 2. Detailed information of past shopping activity data

Feature	Description
sum_quantity_lag1	Total number of items sold in a store one month ago
sum_quantity_lag2	Total number of items sold in a store two month ago
sum_quantity_lag1/lag2	Total number of items sold in a store one month ago/two month ago
sum_quantity_lag1/lag13	Total number of items sold in a store one month ago/thirteen month ago
sum_quantity_lag13/lag25	Total number of items sold in a store one month ago/twenty-five month ago

As stated in [6] yearly, periodical, seasonal effects of the forecast periods are very important. In order to obtain yearly and monthly seasonal effects of shopping activities for each retail store, we create these features as shown in Table 2.

Table 3. Detailed information of item-category data

Feature	Description
category	Indicate the sold item belongs which category
job_desc	Indicates which job deals with the mentioned category

As shown above in Table 3. Specific job groups determine the sales of the categories for which they are responsible. For example, butchers are responsible for the sale of meat products. This provides us with the ability to produce analyses and forecasts for each group separately. The label for this study is the total monthly working hours of the stores. As previously mentioned, the working hours can be biased. In order to eliminate this bias, we created a metric by considering the efficiency of the stores. Through this metric, we plan to build the model by eliminating very efficent and very unefficent stores from the training data. After the preprocessings, we created 45 features into two parts. First part is past shopping activities of stores and the second one is physical attributes of stores. We create the training and testing periods as follows. If the forecast is to be produced for the current month, forecasts for the previous month are produced and reflected in current day's forecast with updated data (Table 4).

Table 4. Summary of model data and algorithm structure

Target	2020 - May	2020 June	2020 - May	2022 - June	2022 - July
		Training Period		Label	
Store A			Testing Period		Label
		Training Period		Label	
Store B			Testing Period		Label
		Training Period		Label	
Store C			Testing Period		Label

As shown in Table 2. We assume that the predictions will be produced for July 2022 period. For each store, training dataset covers the data between May 2020 and May 2022. For training section workforce of June 2022 period will be considered as label. Furthermore, in testing section dateset will cover the data between June 2020 and June 2022. The model creates future predictions for 2022 July according to this dataset.

Fig. 2. Comparison of workhours predicted by model and actual workhour values

4 Conclusion

The error values we consider most important when evaluating the results are those in the testing data. Since the label data is biased, we did not take an approach where training and validation errors are evaluated in a common structure to determine model success. Instead, the predictions made on the testing dataset are evaluated in terms of both metric and business insight.

Each month is located on the x-axis (e.g. 202203 represents the 3rd month of the year 2022.) The Actual and Prediction values shown in the Figure include the total hours of operation of all retail stores for these months. Considering MAPE as an error metric, the highest MAPE value occurs in January and March, with a ratio of 11%. In the remaining months, the maximum MAPE value is observed as 6%.

As mentioned earlier in the paper, the working hour data used as the label of the model is biased and it would be wrong to use traditional methods to determine the success of the model. In addition, it is also important to have a minimum bias in the training of the model. In future research on this subject, a solution can be found to minimize or eliminate this problem.

References

1. Bureau, U.C.: 2020 census. U.S, Department of Commerce (2020)
2. Chuang, H.H.C., Oliva, R., Perdikaki, O.: Traffic-based labor planning in retail stores. Prod. Oper. Manag. **25**(1), 96–113 (2016)
3. Fisher, M., Gallino, S., Netessine, S.: Setting retail staffing levels: a methodology validated with implementation. Manuf. Serv. Oper. Manag. **23**(6), 1562–1579 (2021)
4. Fisher, M., Krishnan, J., Netessine, S.: Retail store execution: an empirical study. Available at SSRN 2319839 (2006)
5. Madhani, P.M.: Enhancing retail stores' performance: managing compensation costs with optimal retail workforce sizing. IUP J. Bus. Strat. **18**(1), 1–15 (2021)
6. Serengil, S.I., Ozpinar, A.: Workforce optimization for bank operation centers: a machine learning approach (2017)

Using LSTM as Intelligent Machine Learning Method to Forecast the Annual Average Relative Humidity: A Case Study for Konya, Turkiye

Abubakar Mayanja[1,2], Şule Eryürük[3](✉), and Kağan Eryürük[4]

[1] Graduate School in Science, Department of Computer Engineering, Selcuk University, Konya, Turkey

[2] Vocational School of Commerce and Industry, Department of Computer Programming, KTO Karatay University, Konya, Turkey

[3] Department of Industrial Engineering, KTO Karatay University, Konya, Turkey
sule.eryuruk@karatay.edu.tr

[4] Department of Civil Engineering, Necmettin Erbakan University, Konya, Turkey

Abstract. Nowadays, one of the most demanding areas of Artificial Intelligence application is to forecast the weather condition as rapid and accurate as possible. The ratio of air vapor pressure to saturation vapor pressure is defined as relative humidity which is represented by a percentage of how much moisture the atmosphere can hold at the same temperature and pressure. The relative humidity has a significant impact on climate change and global warming which have had a negative impact on agriculture, economy, food security, and public health. In addition, accurate weather information and humidity forecasts are frequently required to warn of natural disasters result from unexpected changes in climatic conditions such as relative humidity. Therefore, recent decades have seen a rise in interest in this field of study. The aim of this study is to use Long Short-Term Memory (LSTM) models to forecast the annual relative humidity in Konya, Turkey between 1931 and 2010 by utilizing the annual average amounts of rain, sunlight, wind speed, and temperature. This paper represents a novel approach to predict annual relative humidity in Konya, which has not been previously explored in this location and between the given dates. As a result, the findings were able to forecast the monthly average humidity for next five years in advance with a mean square error of 0.1705% within 150 episodes.

Keywords: Recurrent Neural Network · LSTM · Root Mean Square Error · Relative Humidity

1 Introduction

The continuous rise in global temperatures that began in the twenty-first century remains a major concern for humanity. As a result, climate change is becoming increasingly important. The primary causes of climate change are increasing sun's radiation and global warming caused by the deterioration of greenhouse effect [1, 2] Higher temperatures,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 275–282, 2023.
https://doi.org/10.1007/978-3-031-39777-6_34

depending on climate change, alter weather patterns and disrupt the natural balance of nature. This situation has a negative impact on many aspects of human life, including agriculture, the economy, food security, and public health [3, 4].

The amount of moisture in the air in comparison to what the air can hold at that temperature is referred to as relative humidity (RH) and it is expressed as a percentage [5]. Relative humidity is a weather component that has a direct effect on crucial matters like human health, hydrological studies, agriculture, and irrigation planning [6, 7]. Extreme RH values can harm human health by causing diseases like colds and flu, nosebleeds, vomiting, asthma attacks, and allergies, and high RH values can lead to a rise in precipitation, which could also negatively impact the economy [8–10]. Besides, accurate weather forecasts and humidity forecasts have become more and more critical in the prevention of environmental hazards caused by dramatic alterations in the climate [11]. The studies of weather forecasting are getting important day by day [12, 13]. In recent years, a lot of futuristic heuristics have been used to make sure that predicted weather results are as timely and accurate as possible. Many technological tools and data mining techniques, such as Naive Bayes and the Chi-square test, have been used with excellent results [14]. One of the major drawbacks of conventional machine learning techniques is the use of hand-designed feature extraction methods that do not bring out better results with some complex data sets [15]. With the improvement of deep learning, which is a branch of machine learning, that uses neural networks which play a big role in making it possible for mathematical models to acquire knowledge about data in a both structured and unstructured manner [16], more accurate results can be drawn for complex datasets. This can be categorized into three branches, i.e., supervised learning, Hybrid, and non-supervised learning as demonstrated in Fig. 1 below.

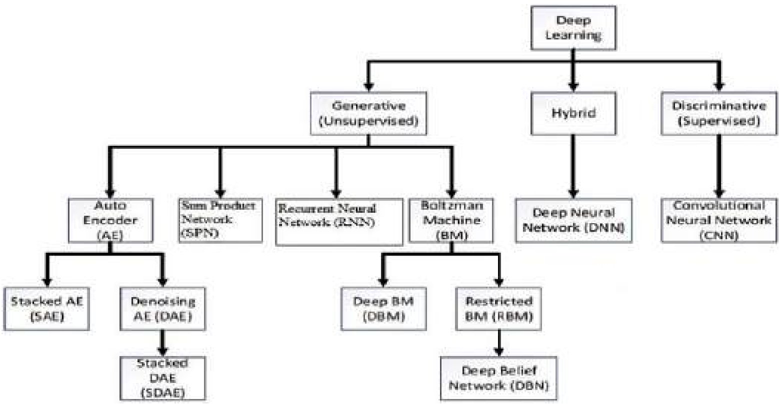


Fig. 1. Anatomy of Deep Learning [17]

A branch of supervised learning called Recurrent neural networks (RNN) which was first proposed by Hochreiter in 1997 [18] and also explained in detail by many researchers [19, 20]. The goal of this paper is to use Long Short-Term Memory (LSTM) models to forecast annual relative humidity in Konya, Turkey. The annual average amounts of rain, sunlight, wind speed, and temperature were used to predict relative humidity be-tween

1931 and 2010. The novelty of this study is that to predict annual relative humid-ity in Konya, which is classified as the second-degree arid region in Turkey and has 12.2% of the total arable land in Turkey [21].

The remaining part of the paper is structured as follow; Introduction provides a summary of related works in the field of humidity forecasting using machine learning techniques. Materials and method section describes the dataset used in this study, including data preprocessing, feature engineering and the LSTM models used to forecast annual relative humidity. Experimental Results section presents the experimental setup and results, followed by a discussion of the findings covering summary of the research's findings and future directions.

2 Materials and Methods

2.1 Long Short-term Memory (LSTM)

RNN encompasses a variety of techniques such as recursive neural networks, Boltzmann machine, gated recurrent etc. [14] RNN might have a vanishing gradient problem. A technique called Long short-term memory (LSTM) which plays a role in learning and refining sequential data by solving this problem including gates that can control the flow of information with long-term interdependence. So, in this study was used [4]. The LSTM consists of three gates which are the input, output, and forget gates. The input and output play the role of loading and unloading data from the network respectively while the forget gate acts as a lever that standardizes the weights and creation of long-term memory functions. Additionally, the gates are controlled by a weight matrix called recurrent weight matrix. The nature of LSTM can be diagrammatically represented in Fig. 2.

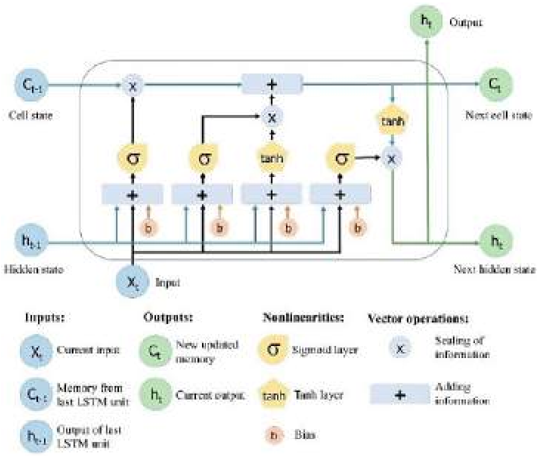


Fig. 2. LSTM neural network [21]

The uniqueness of the LSTM network in continuous use of the network output as the new input in the same network is also shown in the Fig. 2 using the tanh element-wise multiplication.

Figure 3 represents a single LSTM cycle that is repeated depending on the number of time steps configured. The other important point to note is that neural network weights and biases are not descendants of the time steps.

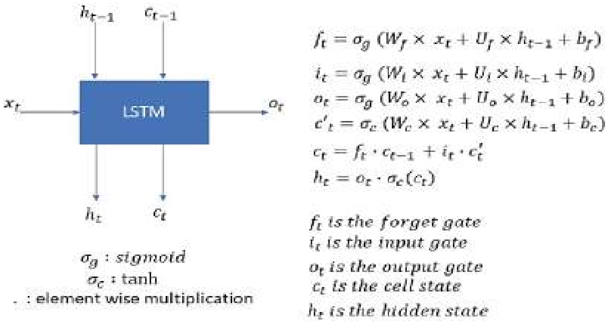


Fig. 3. Expansion of the Long Short-Term Memory [22].

A comma-separated dataset obtained from the Konya weather forecast station was used in this study. The dataset was composed of 1200 rows of data but after the data cleaning process, a total of 960 rows of average monthly rain, temperature, relative humidity, wind speed, and temperature were adopted.

The original dataset file was composed of various weather feature-related information such as snow-covered days, evaporation, snow depth, etc., but for experimental results in forecasting humidity, four major elements were considered, i.e., the amount of rain measured in millimeter (mm), speed of wind in meters per second (m/s), temperature (degree Celsius, °C) and humidity in grams of water per cubic centimeter (g/cm^3).

Different combinations of the above factors team up to come up with an equation of relative humidity below.

$$RH = (Pw)/(Ps).100 \tag{1}$$

where RH denotes the relative humidity, Pw denotes the density of water vapor and Ps denotes the density of water vapor at saturation.

2.2 Data Preparation and Model

Preparation starts with loading and converting the Microsoft Excel spreadsheet file into a data frame workbook. From that point, data frames are converted into separate data arrays which are scaled to a range of 0 and 1 for proper use in the RNN model. The dataset was divided into 90% percent training and 10% test set. Then, data is processed by the LSTM model. The performance of the model output is calculated as mean square error. Finally training and validation loss of results are displayed for the experimental result. The graphics of the input parameters are illustrated and given at Fig. 4 a to d.

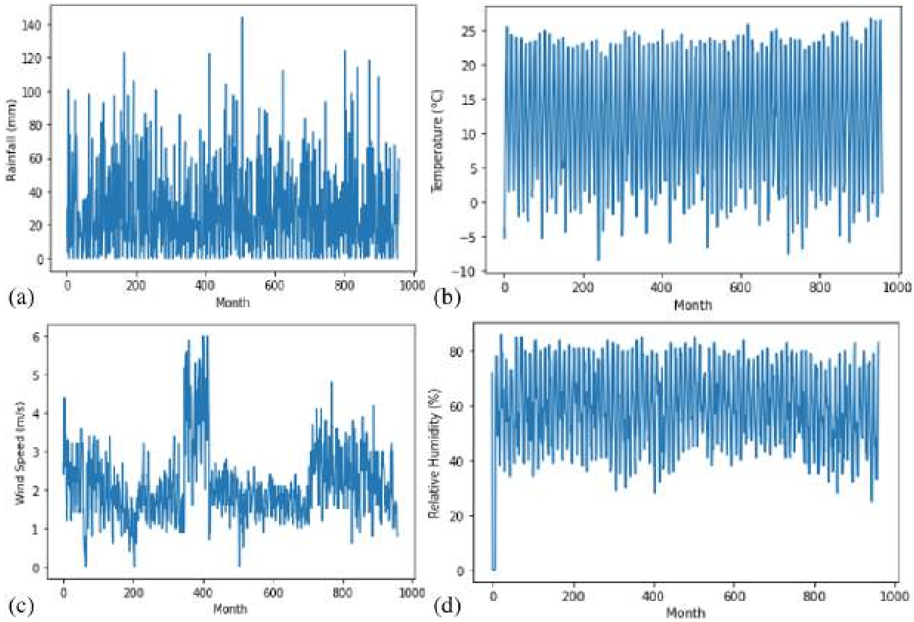


Fig. 4. Time Series Graphs for Rainfall (a), Temperature (b), Wind Speed (c) and Relative Humidity (d) between 1931 and 2010.

It is then passed through a three-layered RNN sequential model that takes a feed-in of a 1 by 4 matrix and a single output as shown below after which the root mean square error (RMSE) is computed.

$$[\text{Humidity, Rain, Temperature, Wind}] = [\text{Humidity(next 5 years)}]$$

The details of the neural network structure showing both input and output is in Fig. 5. Given below.

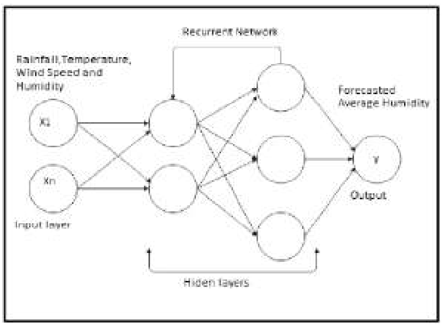


Fig. 5. The RNN structure

As demonstrated in Fig. 5, the network is a multivariate series LSTM that takes more than an argument (Average rainfall, temperature, wind speed, and humidity) to predict a single output (Average humidity). During the training process, the network was set to use data from the previous four months as input to predict the values in the fifth month. It then takes a single-step slide for the next prediction until the end. As an example, from input [m_1, m_2, m_3, m_4], output [m6] is generated. So on, input [$m_{n\text{-}4}$, $m_{n\text{-}3}$, $m_{n\text{-}2}$, $m_{n\text{-}1}$] creates output of [m_n] where $m_{(1,2,3\ldots n)}$ is the monthly data fed in and predicted out of the LSTM network.

3 Experimental Results

In this section, the experimental results after 150 episodes run is discussed. Both the training and validation loss were examined and graphical results showed a Mean Squared error of 0.1076 and 0.1705 for training and testing data respectively as demonstrated in Fig. 6 (a, b) given below.

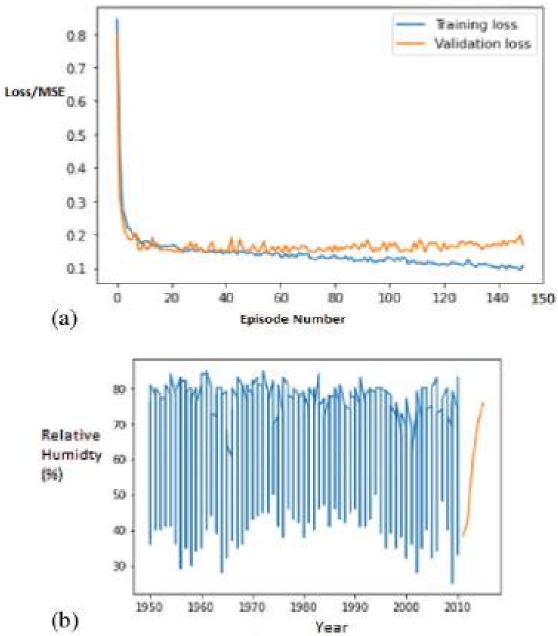


Fig. 6. Training and Validation Loss (a) and Five-years prediction result (b)

Beyond training and testing, the network was used to predict relative humidity values 5 years ahead by selecting a date range of 1950 to 2010 as graphically demonstrated in Fig. 6(b) given below. From the figures above, it is seen that the gap between the two mean squared error values is minimal though fluctuating from the beginning to the last episode of the line graph. This shows that the network learned significantly. Additionally, the 5-year prediction gave promising results that it gives a pattern almost like that of the original data in Fig. 5(d).

4 Discussion and Conclusion

The current study utilized Long Short-Term Memory (LSTM) models to predict the annual relative humidity in Konya, Turkey from 1931 to 2010 by using the annual average amounts of rain, sunlight, wind speed, and temperature. The results of the study demonstrated that the use of monthly average amounts of these variables as input parameters was a novel approach to predicting annual relative humidity in this location for the given dates. The study's findings showed that the LSTM model could predict the monthly average humidity for the next five years with a mean square error of 0.1705% within 150 episodes. By accurately forecasting relative humidity, decision-makers can prepare for natural disasters resulting from unexpected changes in climatic conditions. Additionally, predicting relative humidity can also aid in developing sustainable agricultural practices, mitigating the effects of global warming and climate change. The study's approach can be extended to other locations and time periods, further advancing the field of weather prediction using different machine learning algorithms. The study results provide a solid foundation for future studies on this topic. Future research may explore the impact of incorporating additional parameters such as air quality and oceanic data in predicting humidity. Additionally, the study could be extended to include more locations and more extended periods. In conclusion, the current study successfully predicted the annual relative humidity in Konya, Turkey from 1931 to 2010 using Long Short-Term Memory (LSTM) models. The study's novel approach, utilizing monthly average amounts of rain, sunlight, wind speed, and temperature as input parameters, was effective in predicting relative humidity accurately. The findings have significant implications for predicting climate change and natural disasters while also having practical applications for agriculture, public health, and the economy. This study provides a solid foundation for future studies on weather prediction using machine learning techniques.

References

1. Griggs, D.J., Noguer, M.: Climate Change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Weather **57**(8), 267–269 (2002)
2. Asadi, M., Karami, M.: Modeling of relative humidity trends in Iran. Model. Earth Syst. Env. **8**(1), 1035–1045 (2021)
3. Kogo, B.K., Kumar, L., Koech, R.: Climate change and variability in Kenya: a review of impacts on agriculture and food security. Environ. Dev. Sustain. **23**(1), 23–43 (2020)
4. Pasquini, L., van Aardenne, L., Godsmark, C.N., Lee, J., Jack, C.: Emerging climate change-related public health challenges in Africa: a case study of the heat-health vulnerability of informal settlement residents in Dar es Salaam, Tanzania. Sci. Total Env. **747**, 141355 (2020)
5. Ghadiri, M., Marjani, A., Mohammadinia, S., Shokri, M.: Machine learning approaches for accurate prediction of relative humidity based on temperature and wet-bulb depression. Preprints, 2020020075 (2020)
6. Davis, R.E., McGregor, G.R., Enfield, K.B.: Humidity: a review and primer on atmospheric moisture and human health. Env. Re. **144**, 106–116 (2016)
7. Gunawardhana, L.N., Al-Rawas, G.A., Kazama, S.: An alternative method for predicting relative humidity for climate change studies. Meteorol. Appl. **24**(4), 551–559 (2017)

8. Silveira, A.: Problems of modern urban drainage in developing countries. Water Sci. Technol. **45**(7), 31–40 (2002)
9. Falagas, M.E., et al.: Effect of meteorological variables on the incidence of respiratory tract infections. Respir. Med. **102**(5), 733–737 (2008)
10. Zhang, Ds., Zhang, X., Ouyang, Yh., et al.: Incidence of allergic rhinitis and meteorological variables: non-linear correlation and non-linear regression analysis based on Yunqi theory of Chinese medicine. Chin. J. Integr. Med. (2016). https://doi.org/10.1007/s11655-016-2588-9
11. Flora, M.L., Potvin, C.K., Skinner, P.S., Handler, S., McGovern, A.: Using machine learning to generate storm-scale probabilistic guidance of severe weather hazards in the warn-on-forecast system. Mon. Weather Rev. **149**, 1535–1557 (2021)
12. Holtslag, A.A.M., De Bruijn, E.I.F., Pan, H.-L.: A high resolution air mass transformation model for short-range weather forecasting. Mon. Weather Rev. **118**(8), 1561–1575 (1990)
13. Fathi, M., Haghi Kashani, M., Jameii, S.M., Mahdipour, E.: Big data analytics in weather forecasting: a systematic review. Arch. Computat. Methods Eng. **29**, 1247–1275 (2022)
14. Du, K.-L., Swamy, M.N.S.: Recurrent Neural Networks. In: Neural Networks and Statistical Learning, pp. 337–353. Springer, London (2014). https://doi.org/10.1007/978-1-4471-5571-3_11
15. Wani, M.A., Bhat, F.A., Afzal, S., Khan, A.I.: Advances in Deep Learning, Studies in Big Data. Springer, Singapore (2020)
16. Mathew, A., Amudha, P., Sivakumari, S.: Deep learning techniques: an overview. In: Hassanien, A.E., Bhatnagar, R., Darwish, A. (eds.) AMLTA 2020. AISC, vol. 1141, pp. 599–608. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-3383-9_54
17. Selvaganapathy, S., Nivaashini, M., Natarajan, H.P.: Deep belief network based detection and categorization of malicious URLs. Inform. Secur. J.: A Global Perspect. **27**(3), 145–161 (2018)
18. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
19. Zhong, G., Lin, X., Kang, C., Huang, K.: Long Short-Term Attention. arXiv:1810.12752v2 (2019)
20. Giunchiglia, E., Nemchenko, A., van der Schaar, M.: RNN-SURV: a deep recurrent model for survival analysis. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I. (eds.) ICANN 2018. LNCS, vol. 11141, pp. 23–32. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01424-7_3
21. Asaad, M.N., Eryürük, Ş, Eryürük, K.: Forecasting of streamflow and comparison of artificial intelligence methods: a case study for Meram stream in Konya, Turkey. Sustainability **14**(10), 6319 (2022)
22. https://towardsai.net/p/machine-learning/tutorial-on-lstm-a-computational-perspective-f3417442c2cd

Multi-criteria Decision Making

Assessment of Sustainable Energy Resources with Hesitant Fuzzy Linguistic AHP-MULTIMOORA Methods

Esin Mukul, Merve Güler(✉), and Gülçin Büyüközkan

Industrial Engineering Department, Galatasaray University, Ciragan Cad., 36 Ortaköy, Istanbul 34349, Turkey
{emukul,mguler,gbuyukozkan}@gsu.edu.tr

Abstract. Building a sustainable development path in energy, supporting the transition to a low-carbon energy system, and promoting clean energy are among the critical components for sustainable energy and energy efficiency. Sustainable energy has particular importance in diversifying resources, meeting energy demand with greener options, and making better use of domestic resources to achieve flexible energy systems. This study aims to assess sustainable energy resources with Hesitant Fuzzy Linguistic (HFL) Multi-Criteria Decision Making (MCDM) methods. The Hesitant Fuzzy Linguistic Term Sets (HFLTS) approach is used to convey Decision Makers' (DMs') evaluations by addressing the effort of expressing concepts through crisp numbers and ambiguity. The weights of the assessment criteria are computed with the HFL Analytic Hierarchy Process (AHP) method, whereas the sustainable energy resources are evaluated with the updated form of HFL Multi-Objective Optimization on the basis of Ratio Analysis (MULTIMOORA) method. A case study for this evaluation problem is provided to illustrate the power of the proposed methodology. Finally, the application results are given, and future remarks are presented.

Keywords: AHP · Hesitant Fuzzy Linguistic Term Sets · MCDM · MULTIMOORA · Sustainable Energy

1 Introduction

In many countries, great importance is attached to sustainable energy and energy efficiency to create a sustainable development route, support the transition to a low-carbon energy system, and spread clean energy. Sustainable energy investments and energy efficiency studies have long been on Turkey's agenda as vital components of green growth. Sustainable energy research, bolstered by the mobilization of hydroelectric power plants in the mid-2000s, has accumulated wealth alongside various types of power plants throughout time. In the previous ten years, Turkey has achieved significant development in this field [1–3]. In order to evaluate Turkey's potential in the electricity and natural gas sectors in the most effective, competitive, and efficient way, there are some requirements that must be met in the field of sustainable energy. These requirements are

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 285–292, 2023.
https://doi.org/10.1007/978-3-031-39777-6_35

resource optimization, predictable investment climate, strong grid infrastructure, and environmental protection throughout the value chain [3]. Sustainable energy provides a much cleaner and safer environment than non-renewable energies; so that both tomorrow and future generations can be prepared for a clean and reliable world. The characteristics of sustainable energy can be listed as follows; being accessible, harmonious, flexible and fair.

This study uses an integrated Hesitant Fuzzy Linguistic (HFL) Multi-Criteria Decision-Making (MCDM) approach to assess the sustainable energy resource alternatives. To overcome hesitation in decision-makers' (DMs) decisions, Hesitant Fuzzy Linguistic Term Set (HFLTS) technique proposed by Rodriguez et al. [4] is used. Considering that it has a hierarchical structure, the HFL Analytic Hierarchy Process (AHP) method is used [5]. The updated form of HFL Multi-Objective Optimization on the Basis of Ratio Analysis (MOORA) method named as HFL MULTIMOORA method is used since the method is the most robust system of multiple objective optimization [6]. The criteria weights are calculated using HFL AHP, and HFL MULTIMOORA is utilized to assess the sustainable energy resource options.

In recent years, the usage of sustainable energy resources has grown intensely, therefore some of the previous studies have examined the subject with various MCDM methods [7–11]. Oberschmidt et al. [7] evaluated sustainable energy alternatives in Pakistan with Preference Ranking Organisation Method for Enrichment Evaluations (PROMETHEE) method. Streimikiene et al. [8] evaluated sustainable electricity production technologies with Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) and MOORA and MULTIMOORA methods. Luthra et al. [9] recognized the adoption barriers of renewable and sustainable technologies in India with AHP method. Ghenai et al. [10] evaluated the renewable energy systems according to sustainability indicators by using an integrated Step-wise Weight Assessment Ratio Analysis/Additive Ratio Assessment (SWARA/ARAS) methodology. Nsafon et al. [11] integrated AHP-VIšekriterijumsko KOmpromisno Rangiranj (VIKOR) methods for sustainable energy planning in Africa. As a result, the primary contribution of this paper is the first use of the HFL AHP and HFL MULTIMOORA methods in the assessment of sustainable energy resources.

The paper is structured in the manner described below. The research methodology is described in the following section. A presentation of the case study for research methodology is included in the third section. The final section includes a summary of the overall observations and suggestions for future research.

2 Research Methodology

The proposed HFL AHP-HFL MULTIMOORA methodology includes 3 steps as seen in Fig. 1.

Hesitant Fuzzy Sets (HFS) were proposed by Torra [12] in 2010. A set of HFLTS were used to create the HFLTS approach, which was first introduced in 2012 [4]. In this study, the fuzzy envelope technique put forward by Liu and Rodriguez [13] is used to convert HFLTS into fuzzy numbers. For more information on the HFLTS technique, please refer to [4, 13]. The steps of HFL AHP and HFL MULTIMOORA methods are explained in Fig. 2 [5, 6, 14, 15].

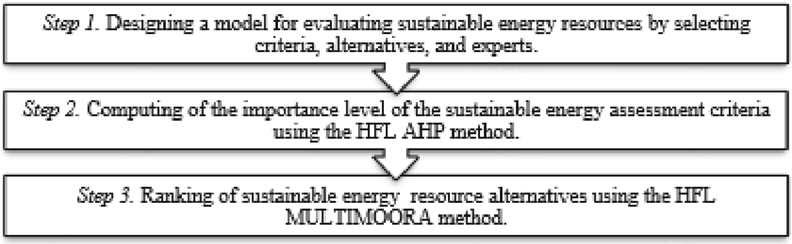


Fig. 1. The framework of the proposed research methodology.

3 Case Study

Globally, the climate crisis is a serious problem that affects the energy sector as well as all sectors. Many developments such as the increasing pressure on natural resources, extreme weather events, and forest fires in various parts of the world reveal that countries facing the climate crisis also need to radically change to reach their short, medium, and long-term goals. Sustainable resources play a key role in achieving these goals, reducing emissions and increasing energy efficiency globally [3, 18].

This study proposes a model for assessing sustainable energy resources that was developed after evaluating academic articles, industry reports, and expert opinions. Table 1 is an illustration of this model. The proposed criteria are divided into four main criteria:

- *The economic criteria (C1)* aim to meet the increasing share of renewable and sustainable energy by providing a market environment in which electricity is economically accessible, reducing investment and maintenance costs.
- *The environmental criteria (C2)* aim to reduce the environmental impact of energy production in order to attain the objective of clean and sustainable energy. This goal is reached by selecting the most appropriate sustainable energy source based on its varied emissions and environmental implications.
- *The sociopolitical criteria (C3)* aim to dramatically enhance sustainable energy supplies by popularizing digital technology. This may be accomplished through making sustainable energy sources more socially and politically attractive, as well as by eliminating poverty, contributing to communities, and adjusting to national energy policies.
- *Technical criteria (C4)* aim to deliver dependable and modern energy services by focusing on mature but novel energy technologies that produce energy in a reliable, safe, and efficient manner using local know-how.

The sustainable energy resource alternatives are determined based reports and academic papers [2, 5, 16–18]. These alternatives are as following: Hydro (A1), Solar (A2), Geothermal (A3), Wind (A4), and Biogas (A5).

Three DMs evaluate this proposed model. All three DMs have appropriate knowledge and expertise in the field of sustainable and renewable energy. The weights of the DMs are considered equal. DM1 has 14 years of private sector experience in the development of sustainability indicators for different sustainable energy sources projects. DM2

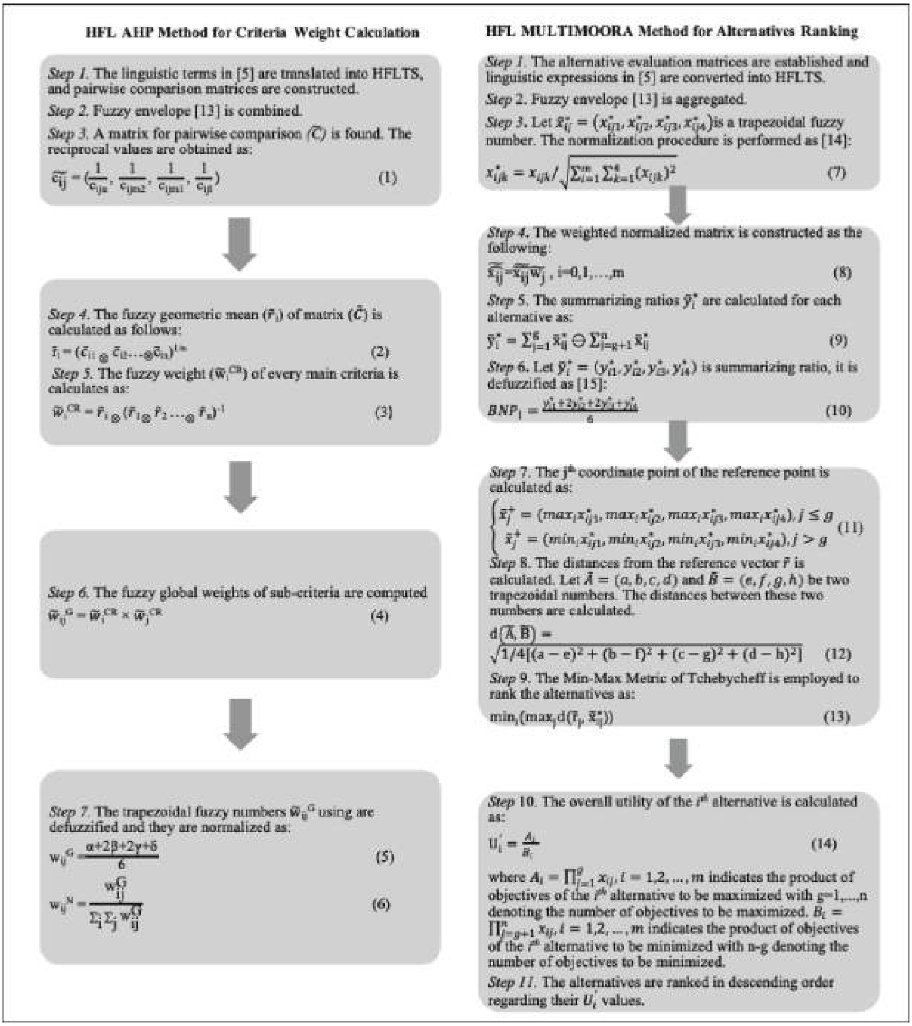


Fig. 2. The flowchart of the proposed research methodology.

has been conducting academic studies, lectures, industrial research, and consultancy on sustainable energy. DM3 has consulting and project management experience in the development of renewable energy sources projects in Turkey.

The HFL AHP method is used in the first phase to compute the criteria weights. Table 2 displays the evaluations made by the DMs with linguistic expressions. These expressions provided are converted into trapezoidal fuzzy numbers. Then, the HFL AHP method's stages are applied with Eqs. (1)–(4). Then, defuzzification is realized using (5), and then the values are normalized with Eq. (6). Table 3 shows the final criteria weights.

Table 1. The assessment criteria for sustainable energy resources.

Main Criteria	Sub Criteria
C1. Economic	C11. Affordability of electricity [5, 18, 19]
	C12. Investment cost [5, 16, 17]
	C13. Operation and Maintenance cost [5, 16–18]
	C14. Return on Investment [2, 3, 16, 18]
C2. Environmental	C21.Land use and ecologic footprint [5, 17, 18]
	C22.Climate change [1, 5]
	C23.Greenhouse gas emissions [1, 16]
	C24.Ecologic impacts [1, 5, 16, 18]
C3. Sociopolitical	C31. Policy compatibility [5, 16, 18]
	C32. Job creation [5, 16–18]
	C33. Poverty reduction and welfare [5, 16, 17]
	C34. Social acceptability [5, 16–18]
C4. Technical	C41. Maturity and reliability [16, 17, 19]
	C42. Technical feasibility and safety [3, 16]
	C43. Know-how and maintainability [1, 5, 19]
	C44. Modernity and efficiency [3, 17, 19]

Table 2. The DMs' evaluations for the main criteria.

	C1	C2	C3	C4
C1	EE	Btw. ESLI and ELI	Btw. ELI and EHI	Btw. ESLI and ELI
C2		EE	Btw. EHI and WHI	Btw. WHI and ESHI
C3			EE	Btw. EHI and WHI
C4				EE

The most appropriate criterion is found as "C21. Land use and ecologic footprint", the second important one is "C24. Ecologic impacts," and the third-ranked criterion is "C23. Greenhouse gas emissions".

In the second step, the alternatives are assessed by using HFL MULTIMOORA method. First, experts evaluated the alternatives regarding criteria by using the linguistic scale given in [5]. Table 4 shows the evaluations made by the DMs. Then, by using Eqs. (7)–(14), the steps of the HFL MULTIMOORA technique are applied. Table 5 displays the ranking lists.

Table 3. The weights of the criteria.

Sub-Criteria	Global Weights				Defuzzified Weights	Normalized Weights
C11	0.003	0.021	0.074	0.869	0.177	0.052
C12	0.001	0.016	0.043	0.511	0.105	0.031
C13	0.001	0.012	0.033	0.389	0.080	0.023
C14	0.002	0.028	0.057	0.832	0.167	0.049
C21	0.009	0.078	0.333	2.575	0.568	0.166
C22	0.003	0.018	0.065	0.610	0.130	0.038
C23	0.006	0.059	0.222	1.722	0.382	0.112
C24	0.009	0.089	0.275	2.413	0.525	0.153
C31	0.007	0.042	0.197	1.191	0.279	0.082
C32	0.001	0.009	0.029	0.282	0.060	0.018
C33	0.007	0.055	0.173	1.295	0.293	0.086
C34	0.001	0.008	0.029	0.282	0.060	0.017
C41	0.001	0.013	0.030	0.385	0.079	0.023
C42	0.005	0.049	0.111	1.084	0.235	0.069
C43	0.005	0.049	0.111	1.084	0.235	0.069
C44	0.001	0.009	0.020	0.221	0.047	0.014
					3.421	1.000

Table 4. The DMs' evaluations for the alternatives.

	C11	C12	C13	C14
A1	Btw. ESHI an VHI	At least VHI	Btw. ESHI an VHI	At least ESHI
A2	Btw. WHI an ESHI	Btw. EHI and WHI	At least ESHI	Btw. WHI an ESHI
A3	Btw. EHI and WHI	Btw. WHI an ESHI	Btw. EHI and WHI	At least ESHI
A4	Btw. ESHI an VHI	At least VHI	Btw. ESHI an VHI	Btw. WHI an ESHI
A5	Btw. EHI and WHI	Btw. WHI an ESHI	Btw. EHI and WHI	Btw. WHI an ESHI

The most suitable alternative is found as Wind (A4), and the second most appropriate alternative is found as Solar (A2), according to the fuzzy reference point approach and the fuzzy ratio system ranking. However, Solar (A2) is found as the most suitable alternative according to the fuzzy full multiplicative form ranking, and Hydro (A1) is found as the second most suitable alternative.

Table 5. The results of the HFL MULTIMOORA.

	The Fuzzy Reference Point Approach		The Fuzzy Ratio System Ranking		The Fuzzy Full Multiplicative Form Ranking	
	max $d(r_j,x_{ij})$	Ranking	BNP_i	Ranking	BNP_i	Ranking
A1	11.346	5	0.893	3	1.61588E-28	2
A2	10.884	2	2.234	2	2.28797E-28	1
A3	11.346	5	0.726	4	9.40884E-36	4
A4	2.126	1	5.822	1	1.30231E-35	3
A5	11.058	3	0.377	5	1.77858E-39	5

4 Conclusion

The aim of this study was assessing the sustainable energy resource alternatives with an integrated HFL MCDM approach. HFL AHP-HFL MULTIMOORA methodology is used for the first time in sustainable energy evaluation. HFL AHP technique is utilized for calculating the sustainable energy assessment criteria weights. HFL MULTIMOORA method is utilized for ranking the sustainable energy resource alternatives. A new framework for evaluating sustainable energy resources using HFL MCDM approaches is proposed in this paper as a contribution to the field.

At the end of the case study, the most important criteria is found as "C21. Land use and ecologic footprint", and the first ranked alternative is determined as A4; which is "Wind". It will make a significant contribution to the installed capacity of wind energy in a short time, by paving the way for capacity increase investments, which is one of the most economical and rapid solutions for maximizing the use of wind energy potential, and bringing them into the system as soon as possible.

Future studies may look at models with greater detail. On the other hand, the framework might be extended to include further fuzzy set expansions.

Acknowledgments. The authors gratefully acknowledge the assistance of industry professionals. This work has been supported by the Scientific Research Projects Commission of Galatasaray University under grant number FOA-2021–1059 and FOA-2023–1181.

References

1. Jacobson, M.Z., Delucchi, M.A.: A path to sustainable energy by 2030. Sci. Am. **301**(5), 58–65 (2009)
2. Paris Agreement: https://unfccc.int/process-and-meetings/the-paris-agreement/what-is-the-paris-agreement. Last accessed 2 Dec 2021
3. Sandalkhan, B., Bolukbasi, S., Selcuk, F.: Sürdürülebilir Gelecek icin Sürdürülebilir Enerji – Kısa ve Orta Vadeli Öneriler (Sustainable Energy for Sustainable Future – Short and Medium Term Recommendations). TUSIAD & BCG (2018)

4. Rodríguez, R.M., Martinez, L., Herrera, F.: Hesitant fuzzy linguistic term sets for decision making. IEEE Trans. Fuzzy Syst. **20**(1), 109–119 (2012)
5. Büyüközkan, G., Karabulut, Y., Mukul, E.: A novel renewable energy selection model for United Nations' sustainable development goals. Energy **165**, 290–302 (2018)
6. Büyüközkan, G., Güler, M.: A combined hesitant fuzzy MCDM approach for supply chain analytics tool evaluation. Applied Soft Computing, 112 (2021)
7. Oberschmidt, J., Geldermann, J., Ludwig, J., Schmehl, M.: Modified PROMETHEE approach for assessing energy technologies. Int. J. Energy Sect. Manag. **4**, 183–212 (2010)
8. Streimikiene, D., Balezentis, T., Krisciukaitienė, I., Balezentis, A.: Prioritizing sustainable electricity production technologies: MCDM approach. Renew. Sustain. Energy Rev. **16**, 3302–3311 (2012)
9. Luthra, S., Kumar, S., Garg, D., Haleem, A.: Barriers to renewable/sustainable energy technologies adoption: Indian perspective. Renew. Sustain. Energy Rev. **41**, 762–776 (2015)
10. Ghenai, C., Albawab, M., Bettayeb, M.: Sustainability indicators for renewable energy systems using multi-criteria decision-making model and extended SWARA/ARAS hybrid method. Renew. Energy **146**, 580–597 (2020)
11. Nsafon, B.E.K., Butu, H.M., Owolabi, A.B., Roh, J.W., Suh, D., Huh, J.S.: Integrating multi-criteria analysis with PDCA cycle for sustainable energy planning in Africa: application to hybrid mini-grid system in Cameroon. Sustain. Energy Technol. Assess. **37**, 100628 (2020)
12. Torra, V.: Hesitant fuzzy sets. Int. J. Intell. Syst. **25**(6), 529–539 (2010)
13. Liu, H., Rodríguez, R.M.: A fuzzy envelope for hesitant fuzzy linguistic term set and its application to multicriteria decision making. Inf. Sci. **258**, 220–238 (2014)
14. De Liu, P., Guan, Z.L.: Hybrid multiple attribute decision making method based on 2-tuple. Kongzhi Yu Juece/Control Decis. **24**, 3716–3724 (2009)
15. Erdoğan, M., Kaya, İ: A combined fuzzy approach to determine the best region for a nuclear power plant in Turkey. Appl. Soft Comput. **39**, 84–93 (2016). https://doi.org/10.1016/j.asoc.2015.11.013
16. Büyüközkan, G., Güleryüz, S.: Evaluation of renewable energy resources in Turkey using an integrated MCDM approach with linguistic interval fuzzy preference relations. Energy **123**, 149–163 (2017)
17. Büyüközkan, G., Karabulut, Y., Güler, M.: Strategic renewable energy source selection for turkey with hesitant fuzzy MCDM method. In: Kahraman, C., Kayakutlu, G. (eds.) Energy Management—Collective and Computational Intelligence with Theory and Applications. SSDC, vol. 149, pp. 229–250. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75690-5_11
18. WEC Energy Group: Pathway to a Clean Energy Future (2022)
19. Büyüközkan, G., Güleryüz, S.: A new GDM based AHP framework with linguistic interval fuzzy preference relations for renewable energy planning. J. Intell. Fuzzy Syst. **27**(6), 3181–3195 (2014)

The ELECTRE Method for Hesitant Fuzzy Multi-Attribute Decision-Making in Emergency Evacuation

Evgeniya Gerasimenko and Alexander Bozhenyuk(✉)

Southern Federal University, Taganrog, Russia
egerasimenko@sfedu.ru, avb002@yandex.ru

Abstract. Decision-making is becoming extremely complex these days. It requires much specific knowledge about the attributes, which can be difficult in real-time tasks. Emergency evacuation is considered as a valuable strategy during the disaster preparedness stage of evacuation to save lives and health of potential victims. In recent years, emergency evacuation problems have used fuzzy logic and soft computing approaches to support effective decision-making, providing robust and low-cost solution. Modern extensions of fuzzy logic such as hesitant and intuitionistic fuzzy sets are applied for the tasks of decision-making by incorporating experts in the problem, which have different assessments because of hesitation without providing a common result to declare their preferences. In addition, it is hard for experts to get the common decision, because decision-makers have different expertise and background, and they cannot easily persuade each other. Therefore, hesitant fuzzy sets that provide the set of possible membership degrees instead of the one common degree can be used to solve the problem. In the paper, hesitant fuzzy ELECTRE method will be used to address the problem of emergency evacuation in fuzzy conditions in order to o find the best evacuation strategy. Various emergency evacuation alternatives can be compared on the attributes. Furthermore, different classes of hesitant fuzzy concordance (discordance) set are considered as the hesitant fuzzy concordance (discordance) set and the weak hesitant fuzzy concordance (discordance) set. The best alternative has the highest score or the lowest deviation degree when the alternatives have the same score. A case study is conducted to illustrate the proposed algorithm.

Keywords: Intermediate storage · Hesitant fuzzy elements · Evacuation

1 Introduction and Literature Review

On October 9, 1963, one of the largest accidents in the history of hydraulic engineering construction occurred in Italy, which claimed the lives of, around 3 thousand people. In seven minutes, the water caused the following destruction: the wave destroyed five villages in the Piave valley, and also subjected several other settlements to serious destruction. The break of the dam in Brumadinho occurred on January 25, 2019, due to catastrophic damage to the tailings at the iron ore mine in the city of Brumadinho in the

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 293–302, 2023.
https://doi.org/10.1007/978-3-031-39777-6_36

Brazilian state of Minas Gerais., 40 dead and at least 300 missing are known, of which 200 are employees of the plant and 100 are residents of neighboring villages. Sayano-Shushenskaya power station accident in Russia was an industrial man-made disaster that occurred on August 17, 2009. As a result of the accident, 75 people died, and serious damage was caused to the equipment and premises of the station. The consequences of the accident affected the ecological situation in the water area adjacent to the power station, the social and economic spheres of the region.

To cope with dam-breaks, authorities could initiate preventive evacuation as an effective tool to save lives of potential victims by shifting people out from the dangerous area to safe destinations. However, mass evacuation management is a very complex tool that requires the coordination of many governments, authorities, and members of the civil society. Moreover, the process of mass evacuation is not very common, since it is a risky measure, costly in time, money that can expose people to risk of injuries and loss or stealing of goods or properties [1]. Therefore, it is necessary to make an effective decision and to define the best evacuation strategy in case of a dam-break. In this paper, multi-attribute decision-making (MADM) is used to define necessity of evacuation while a dam-break. To evaluate the necessity of evacuation and its type, the direct and indirect criteria will be applied such as flood severity, rescue capabilities, the level of understanding degree of flood severity, etc. However, these criteria are hard to express using exact values due to uncertainty and fuzziness inherent into their nature. Moreover, in real-time problems it is not always possible to make decisions without hesitation and doubts by providing the specific preference. Therefore, hesitant fuzzy sets [2] will be incorporated in the process of DM for this purpose. The ELECTRE method [3] is widely used in decision-making to outrank the alternatives. Author [4] were the first who introduced it. MADM tasks in hesitant fuzzy conditions were solved by Chen et al. (Chen et al.[5])who developed an ELECTRE I method based on hesitant fuzzy theory. Concepts of hesitant fuzzy concordance set, and hesitant fuzzy discordance set underlie the proposed method to determine the preferable alternative. Vague environment was considered in ELECTRE method proposed in [6, 7]. Wu and Chen [7] considered intuitionistic fuzzy ELECTRE method to solve MADM problems. In order to avoid drawbacks of aggregation operation, authors [8, 9 used additive value functions to aggregate the average sum of performances. Torra and Narukawa [2] showed that introducing hesitant fuzzy elements allows avoiding aggregation. Chen et al. [5] elaborated the ELECTRE I method according to hesitant fuzzy environments to consider the differences in experts' opinions. The objective of the present research is to apply HF ELECTRE I method to rank the evacuation alternatives in case of a dam-break. The main contribution of the paper is determining the "necessity to evacuate" as the degree of potential need for evacuation in a chosen area based on HF ELECTRE I method.

The paper has the following structure. In Sect. 2, we observe the methodology of hesitant fuzzy ELECTRE method and the algorithm for finding the best evacuation strategy. Section 3 observes a case study of the proposed method. Finally, Sect. 4 concludes the paper and gives topics for future research.

2 Methodology

Let us consider basic definitions underlying the proposed method. Definitions of hesitant fuzzy sets, score and deviation degree are given in [10]. ELECTRE method relies in four binary relations, the concepts of preference, concordance and discordance [3]. Binary relations on the set $X$are given in [3].

Let a set of evacuation attributes that influence evacuation is denoted as $X = \{x_1, x_2, \ldots, x_n\}$. HFS A_i on X is defined as $A_i = \{< x_j, h_{A_i}(x_j) > | x_j \in X\}$, where $h_{A_i}(x_j) = \{\chi | \chi \in h_{A_i}(x_j), 0 \le \chi \le 1\}$, $(i = 1, 2, \ldots n)$, $(j = 1, 2, \ldots m)$ stands for all membership degrees of the i-th alternative A_i under the j-th attribute x_j. The hesitant fuzzy concordance set $Z_{c_{ij}}$ for each pair of alternatives A_i and A_j is the sum of such attributes where the performance of A_i is superior to A_j (Eq. 1):

$$Z_{c_{ij}} = \left\{k|s\left(h_{ik}\right) \ge s(h_{jk}\right) \, and \, \tilde{\sigma}\left(h_{ik}\right) < \tilde{\sigma}\left(h_{jk}\right)\right\}. \tag{1}$$

The weak hesitant fuzzy concordance set $Z_{c'_{ij}}$ is defined as Eq. 2:

$$Z_{c'_{ij}} = \left\{k|s\left(h_{ik}\right) \ge s(h_{jk}\right) and \, \tilde{\sigma}\left(h_{ik}\right) \ge \tilde{\sigma}\left(h_{jk}\right)\right\} \tag{2}$$

Two concordance sets differ in the values of deviation functions. The lower deviation the higher consistency degree, which is why $Z_{c_{ij}}$ is more concordant than $Z_{c'_{ij}}$.

Furthermore, the hesitant fuzzy discordance set $Z_{d_{ij}}$, which consists of attributes for which A_i is inferior to A_j as shown in Eq. 3.

$$Z_{d_{ij}} = \left\{k|s\left(h_{ik}\right) < s(h_{jk}\right) and \, \tilde{\sigma}\left(h_{ik}\right) \ge \tilde{\sigma}\left(h_{jk}\right)\right\} \tag{3}$$

In the case when $s\left(h_{ik}\right) < s\left(h_{jk}\right)$ and $\tilde{\sigma}\left(h_{ik}\right) < \tilde{\sigma}\left(h_{jk}\right)$, we apply the weak discordance set $Z_{d'_{ij}}$ (Eq. 4):

$$Z_{d'_{ij}} = \left\{k|s\left(h_{ik}\right) < s(h_{jk}\right) and \, \tilde{\sigma}\left(h_{ik}\right) < \tilde{\sigma}\left(h_{jk}\right)\right\} \tag{4}$$

Therefore, $Z_{d_{ij}}$ is more discordant than $Z_{d'_{ij}}$.

To find the hesitant fuzzy concordance index, the sum of the attribute weights in the hesitant fuzzy concordance sets should be divided by all attributes.

The concordance index c_{ij} of A_i and A_j is defined as Eq. 5:

$$c_{ij} = \omega_C \sum_{k \in Z_{c_{ij}}} w_k + \omega_{C'} \sum_{k \in Z_{c'}{}_{ij}} w_k \tag{5}$$

where w_k is the weight of criterion and $\sum_{k=1}^{n} w_k$=1.

ω_C and $\omega_{C'}$ are the weights of the hesitant fuzzy concordance and the weak concordance sets depending on the values of attributes. The element c_{ij}, $0 \le c_{ij} \le 1$ shows the relative importance of A_i with respect to A_j. A large value of c_{ij} indicates that the alternative A_k is superior to the alternative A_j.

The hesitant fuzzy discordance index indicates the relative difference of A_i towards terms of discordance attributes. The discordance index is defined by Eq. 6:

$$d_{ij} = \frac{\max\limits_{k=Z_{d_{ij}} \cup Z_{d'_{ij}}} \left\{ \omega_{\overline{D}} \times d\left(w_k h_{ik}, w_k h_{jk}\right), \omega_{\overline{D'}} \times d\left(w_k h_{ik}, w_k h_{jk}\right) \right\}}{\max\limits_{k=\{1,2,\ldots n\}} d\left(w_k h_{ik}, w_k h_{jk}\right)} \tag{6}$$

In the formula, $\omega_{\overline{D}}$ and $\omega_{\overline{D'}}$ are weights of the hesitant fuzzy discordance and the weak discordance set, respectively. These values depend on experts' preferences; d $\left(w_k h_{ik}, w_k h_{jk}\right)$ is a distance measure found by a Hamming distance.

The hesitant fuzzy concordance dominance matrix us found based on the cut-level of hesitant fuzzy concordance indices (Eq. 7):

$$\overline{c} = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{c_{ij}}{n(n-1)} \tag{7}$$

Based on the concordance level, the concordance dominance matrix is filled with elements by the rule: $\begin{cases} f_{ij} = 1, \ if \ c_{1j} \geq \overline{c}; \\ f_{ij} = 0, \ if \ c_{1j} < \overline{c}. \end{cases}$

Similarly, find the hesitant fuzzy discordance matrix based on the values of discordance level $\overline{d}$ (average of the elements in the hesitant fuzzy discordance matrix (Eq. 8):

$$\overline{d} = \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1, j \neq i}^{n} \frac{d_{ij}}{n(n-1)} \tag{8}$$

Based on the discordance level, the discordance dominance matrix is filled with elements by the rule: $\begin{cases} q_{ij} = 1, \ if \ d_{1j} \leq \overline{d}; \\ q_{ij} = 0, \ if \ d_{1j} > \overline{d}. \end{cases}$

The aggregation dominance matrix P is constructed based on the elements of the matrices F and Q as: $P = F \otimes Q$ whose elements $p_{ij} = f_{ij} q_{ij}$.

Finally, the outranking relations to give recommendations concerning the results are introduced [3].

Present the emergency decision-making in hesitant fuzzy conditions for the maximum flow evacuation with intermediate storage at nodes. Algorithm for HF-ELECTRE I is summarized in [3, 5].

3 Case Study

Apply the presented method to the task of finding the optimal pre-evacuation strategy. The main goal is to find the necessity to evacuate potential victims in case of a dam-break. The term "necessity to evacuate" will show the degree of potential need for evacuation in a chosen area. The higher it is, the more urgent and mass it should be. The necessity of evacuation thus will be evaluated in percent: evacuation necessity will be determined as the level of potential need for evacuation, relatively to the priority objective of people's security [1]. This parameter varies from 0 to 100% (Table1).

Table 1. The necessity of evacuation based on linguistic assessments and percentage.

Evacuation alternative	Description	Necessity to evacuate	
		Linguistic assessment	Percent %
No evacuation	No additional actions are needed	Very low	(0,15)
Alertness	Precaution of the dam-break risk including the preparation for evacuation	Low	(15, 30)
Partially recommended evacuation	Partial evacuation of the most vulnerable population groups (disabled people.)	Medium Low	(30, 45)
Recommended evacuation	Partial evacuation of the rest part of vulnerable population groups (people who were injured, sick etc.)	Medium	(45,60)
Mild evacuation	A half of people are in danger, almost full evacuation is needed	Medium High	(60,70)
Evacuation	Full evacuation is needed, the majority of people are in danger	High	(70,80)
Urgent evacuation	Order of full evacuation; citizens must urgently leave dangerous areas	Extremely high	(85,100)

Suppose a dam-break has happened, and authorities have decided to invite experts in order to evaluate the necessity of evacuation in case of a dam-break by ranking the possible evacuation alternatives. Experts were chosen based on their expertise and knowledge to order the choose the best evacuation option [10, 11]. The main goal was to assess the evacuation alternatives regarding the attributes. Experts provide their evaluation as hesitant fuzzy values. Seven evacuation strategies A_i, $(i = 1, 2 \ldots, 7)$ are available (see Table 1). In order to make decision what alternative is the most effective, experts evaluate alternatives regarding seven attributes X_j, $(j = 1, 2, ..7)$. Direct criteria are flood severity, warning time, rescue capability, the level of understanding degree of flood severity. Indirect criteria include temperature at the dam-break time, distance between population at risk and the dam site, and reliability of the dam.

Step 1. A group of experts reconcile their evaluations as hesitant fuzzy numbers, which leads to the hesitant fuzzy decision matrix (Table 2).

Experts define the relative weights of the hesitant fuzzy concordance sets (weak hesitant fuzzy concordance sets) and the hesitant fuzzy discordance sets (weak discordance sets): $(\omega_C, \omega_{C'}, \omega_{\overline{D}}, \omega_{\overline{D'}})^T = \left(1, \frac{2}{3}, 1, \frac{2}{3}\right)^T$. The weight vector of the attributes is given as $w = (0.22, 0.15, 0.12, 0.11, 0.13, 0.14, 0.13)$.

Step 2. Find the score and the deviation degree of each alternative regarding attributes in Tables 3 and 4.

Table 2. Hesitant fuzzy decision matrix

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
A_1	{0.2, 0.4}	{0.4,0.6,0.7}	{0.6,0.8}	{0.4, 0.5, 0.7}	{0.2,0.3,0.7}	{0.3,0.6,0.8}	{0.5,0.6}
A_2	{0.7,0.8}	{0.3,0.5,0.7}	{0.1,0.2, 0.4}	{0.5,0.6, 0.8}	{0.5,0.6,0.8,0.9}	{0.1,0.3,0.4}	{0.3,0.4,0.6}
A_3	{0.7,0.9}	{0.2,0.4,0.5}	{0.3,0.4,	{0.4,0.6,0.7, 0.8}	{0.5,0.7}	{0.2,0.5,0.6}	{0.8,0.9}
A_4	{0.4,0.5, 0.6,0.7}	{0.1,0.3,0.4}	{0.4,0.6, 0.7,0.9}	{0.2,0.3}	{0.4,0.5,0.6}	{0.4,0.6,0.8}	{0.7,0.8}
A_5	{0.3,0.5}	{0.5,0.7,0.8}	{0.6,0.7}	{0.3,0.4,0.5, 0.7}	{0.8,0.9}	{0.4,0.6}	{0.2,0.3,0.5,0.7}
A_6	{0.5,0.7, 0.8}	{0.3,0.5}	{0.6,0.7,0.9}	{0.2,0.3, 0.4}	{0.7,0.8}	{0.3,0.5,0.6,0.8}	{0.4, 0.7,0.8}
A_7	{0.6,0.8, 0.9}	{0.4,0.8}	{0.7,0.9}	{0.3,0.5,0.6, 0.8}	{0.5,0.6,0.8}	{0.6,0.8}	{0.1,0.3}

Table 3. Scores of each alternative regarding attributes

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
A_1	0.300	0.567	0.700	0.533	0.400	0.567	0.550
A_2	0.750	0.500	0.233	0.633	0.700	0.267	0.433
A_3	0.800	0.367	0.350	0.625	0.600	0.433	0.850
A_4	0.550	0.267	0.650	0.250	0.500	0.600	0.750
A_5	0.400	0.667	0.650	0.475	0.850	0.500	0.425
A_6	0.667	0.400	0.733	0.300	0.750	0.550	0.633
A_7	0.767	0.600	0.800	0.550	0.633	0.700	0.200

Step 3. Find the hesitant fuzzy concordance set Z_C and the weak concordance set $Z_{C'}$.

$$Z_C = \begin{pmatrix} - & \{2,3,7\} & \{2\} & \{2,3\} & \{4,7\} & - & \{7\} \\ \{1,5\} & - & \{4\} & \{1\} & \{1,4,7\} & \{1\} & \{4\} \\ \{5\} & \{3,7\} & - & \{1,2,7\} & \{7\} & \{1,7\} & \{1,4,7\} \\ \{5,6\} & \{7\} & \{6\} & - & \{7\} & \{6,7\} & \{7\} \\ \{5\} & \{2,3,5,6\} & \{3,5,6\} & \{5\} & - & \{5\} & \{2,5\} \\ \{5\} & \{5\} & \{2,5\} & \{2,3,5\} & \{7\} & - & \{5\} \\ \{5,6\} & \{3,6\} & \{6\} & \{3,6\} & - & \{3,6\} & - \end{pmatrix}$$

Table 4. Deviations of each alternative regarding attributes

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
A_1	0.100	0.125	0.100	0.125	0.216	0.205	0.050
A_2	0.050	0.163	0.125	0.125	0.158	0.125	0.125
A_3	0.100	0.125	0.050	0.148	0.100	0.170	0.050
A_4	0.112	0.125	0.180	0.050	0.082	0.163	0.050
A_5	0.1	0.125	0.050	0.148	0.050	0.100	0.192
A_6	0.125	0.100	0.125	0.082	0.050	0.180	0.170
A_7	0.125	0.200	0.100	0.180	0.125	0.100	0.100

$$
Z_{C'} = \begin{pmatrix}
- & \{6\} & \{3,6\} & \{4\} & \{3,6\} & \{2,4,6\} & - \\
\{4\} & - & \{2,5\} & \{2,4,5\} & - & \{2,4\} & \{5,7\} \\
\{1,4,7\} & \{1,6\} & - & \{4,5\} & \{1,4\} & \{4\} & - \\
\{1,7\} & \{3,6\} & \{3\} & - & \{1,3,6\} & - & - \\
\{1,2\} & - & \{2\} & \{2,4\} & - & \{2,4\} & \{7\} \\
\{1,3,7\} & \{3,6,7\} & \{3,6\} & \{1,4\} & \{1,3,6\} & - & \{7\} \\
\{1,2,3,4\} & \{1,2\} & \{2,3,5\} & \{1,2,4,5\} & \{1,3,4,6\} & \{1,2,4\} & -
\end{pmatrix}
$$

Step 4. Define the hesitant fuzzy discordance set $Z_{\overline{D}}$ and the weak discordance set $Z_{\overline{D'}}$.

$$
Z_{\overline{D}} = \begin{pmatrix}
- & 1,5 & 5,7 & 5,6 & 1,5 & 5 & 3,5,6 \\
2,3,7 & - & 3,7 & 7 & 2,3,5,6 & 5 & 3,6 \\
2 & 4 & - & 6 & 3,5,6 & 2,5 & 6 \\
2,3 & 1 & 1,2,7 & - & 2,5 & 2,3,5 & 3,6 \\
4,7 & 1,4,7 & 4,7 & 7 & - & 7 & - \\
- & 1 & 1,7 & 6,7 & 5 & - & 1,3,6 \\
7 & 4 & 1,4,7 & 7 & 2,5 & 5 & -
\end{pmatrix}
$$

$$
Z_{\overline{D'}} = \begin{pmatrix}
- & 4 & 1,4 & 1,7 & 2 & 1,3,7 & 1,2,4 \\
6 & - & 1,6 & 3,6 & - & 3,6,7 & 1,2 \\
3,6 & 2,5 & - & 3 & 2 & 3,6 & 2,3,5 \\
4 & 2,4,5 & 4,5 & - & 4 & 1,4 & 1,2,4,5 \\
3,6 & - & 1 & 1,3,6 & - & 1,3,6 & 1,3,4,6 \\
2,4,6, & 2,4 & 4 & - & 2,4 & - & 2,4 \\
- & 5,7 & - & - & 7 & 7 & -
\end{pmatrix}
$$

Step 5. Construct the hesitant fuzzy concordance matrix C based on hesitant fuzzy concordance indexes

$$C = (c_{ij})_{7\times 7} = \begin{pmatrix} - & 0.493 & 0.323 & 0.343 & 0.413 & 0.267 & 0.130 \\ 0.423 & - & 0.297 & 0.480 & 0.460 & 0.393 & 0.283 \\ 0.437 & 0.490 & - & 0.660 & 0.350 & 0.423 & 0.460 \\ 0.503 & 0.303 & 0.220 & - & 0.450 & 0.270 & 0.130 \\ 0.377 & 0.540 & 0.490 & 0.303 & - & 0.303 & 0.367 \\ 0.443 & 0.390 & 0.453 & 0.620 & 0.450 & - & 0.217 \\ 0.670 & 0.507 & 0.407 & 0.667 & 0.393 & 0.580 & - \end{pmatrix}$$

Step 6. Construct the hesitant fuzzy discordance matrix $\overline{D}$ based on hesitant fuzzy discordance indexes

$$\overline{D} = (d_{ij})_{7\times 7} = \begin{pmatrix} - & 1 & 0.667 & 0.652 & 1 & 0.667 & 0.667 \\ 0.606 & - & 1 & 0.875 & 0.675 & 0.667 & 1 \\ 0.273 & 0.237 & - & 0.354 & 0.409 & 0.667 & 0.497 \\ 1 & 1 & 1 & - & 1 & 1 & 0.559 \\ 0.429 & 1 & 0.667 & 0.758 & - & 0.667 & 0.667 \\ 0.233 & 0.407 & 0.833 & 0.500 & 0.455 & - & 0.538 \\ 0.477 & 0.241 & 1 & 1 & 0.414 & 0.667 & - \end{pmatrix}$$

Step 7. Find the hesitant fuzzy concordance level $\overline{c}$ and construct the concordance dominance matrix F, respectively: $\overline{c} = 0.409$.

Step 8. Find the hesitant fuzzy discordance level $\overline{d}$ and construct the discordance dominance matrix Q, respectively:

$$F = \begin{pmatrix} - & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & - & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & - & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & - & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & - & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & - & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & - \end{pmatrix} Q = \begin{pmatrix} - & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & - & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & - & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & - & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & - & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & - & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & - \end{pmatrix}$$

Step 9. The aggregation dominance matrix P is determined as:

$$P = \begin{pmatrix} - & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & - & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & - & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & - & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & - & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & - & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & - \end{pmatrix}$$

Step 10. Aggregation dominance matrix gives the following results: A_2 is preferred to A_1 and A_5; A_3 is preferred to A_1, A_2, A_4, A_6, A_7; A_5 is preferred to A_3; A_6 is preferred to A_1, A_4, A_5; A_7 is preferred to A_1, A_2, A_6. Therefore, A_3 is the best alternative. In evacuation terms, it means that partial evacuation of the most vulnerable population groups in case of a dam-break (e.g. disabled people) is required.

4 Conclusion and Future Study

Despite the durability of dams, they may not withstand due to engineering errors or natural disasters such as rainfalls, which leads to severe disasters. In these situations, preventive evacuation may be needed as an effective tool that allows saving lives of potential victims by shifting them out from the dangerous area to safe destinations. On the other hand, mass evacuation management is connected with a high cost, risk of injuries and loss or stealing of goods or properties. Therefore, it is necessary to make an effective evacuation decision when a dam-break occurs. We suggest the following evacuation strategies: no evacuation, alertness, partially recommended evacuation recommended evacuation, mild evacuation, evacuation, urgent evacuation. To evaluate the necessity of evacuation and its type, direct and indirect criteria are incorporated into the task of decision-making. Among direct criteria there are: flood severity, warning time, rescue capability, and the level of understanding degree of flood severity. Indirect criteria include temperature at the dam-break time, distance between population at risk and the dam site, reliability of the dam. These aforementioned criteria are hard to define using exact values due to uncertainty and fuzziness inherent into their nature. Experts can be hesitative and be unsure regarding the specific evaluation assessment in the form of a membership degree thus giving their evaluation as hesitant fuzzy values. Therefore, in this paper, decision-makers will provide the evaluation of evacuation alternatives in the case of a dam-break using hesitant fuzzy values. The process of decision-making relies on the hesitant fuzzy ELECTRE I method in order to choose the optimal evacuation strategy. A case study of evaluating the evacuation alternatives in the case of a dam-break and finding the best one is conducted. A decision graph is presented to visualize the results. As a part of our future research, we intend to apply the extensions of hesitant fuzzy ELECTRE method and hesitant fuzzy PROMETHEE method to compare the results and give recommendations.

Acknowledgments. The research was funded by the Russian Science Foundation project No. 22–71-10121, https://rscf.ru/en/project/22-71-10121/ implemented by the Southern Federal University.

References

1. Jia, X., Morel, G., Martell Flores, H., Hissel, F.: Fuzzy logic-based decision support for mass evacuations of cities prone to coastal or river floods. Environ. Model. Softw. **85**, 1–10 (2016)
2. Torra, V., Narukawa, Y.: On hesitant fuzzy sets and decision. In: The 18th IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, pp. 1378–1382 (2009)
3. Xu, Z.: Hesitant Fuzzy Sets Theory. Studies in Fuzziness and Soft Computing, vol. 314 (2014)
4. Benayoun, R., Roy, B., Sussman, B.: ELECTRE: Une méthode pour guider le choix en présence de points de vue multiples. Note de Travail 49, SEMA-METRA International, Direction Scientifique (1966)

5. Chen, N., Xu, Z.S., Xia, M.M.: The ELECTRE I multi-criteria decision-making method based on hesitant fuzzy sets. Int. J. Inf. Technol. Decis. Mak. **14**(3), 621–657 (2015)
6. Hatami-Marbini, A., Tavana, M.: An extension of the ELECTRE I method for group decision-making under a fuzzy environment. Omega **39**(4), 373–386 (2011)
7. Wu, M.C., Chen, T.Y.: The ELECTRE multicriteria analysis approach based on Atanassov's intuitionistic fuzzy sets. Expert Syst. Appl. **38**, 12318–12327 (2011)
8. Kim, S.H., Ahn, B.S.: Interactive group decision making procedure under incomplete information. Eur. J. Oper. Res. **116**(3), 498–507 (1999)
9. Kim, S.H., Choi, C.H., Kim, J.K.: An interactive procedure for multiple attribute group decision making with incomplete information: Range-based approach. Eur. J. Oper. Res. **118**(1), 139–152 (1999)
10. Gerasimenko, E., Kureichik, V.: Hesitant fuzzy emergency decision-making for the maximum flow finding with intermediate storage at nodes. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds.) INFUS 2021. LNNS, vol. 307, pp. 705–713. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-85626-7_82
11. Gerasimenko, E., Kureichik, V.: Minimum cost lexicographic evacuation flow finding in intuitionistic fuzzy networks. J. Intell. Fuzzy Syst. **42**(1), 251–263 (2022)

An Integrated Fuzzy Decision-Making Approach to Emergency Evacuation Based on HF ELECTRE

Evgeniya Gerasimenko[1], Alexander Bozhenyuk[1](✉), and Sergey Gorbachev[2]

[1] Southern Federal University, Nekrasovsky Street 44, Taganrog 347922, Russia
egerasimenko@sfedu.ru, avb002@yandex.ru

[2] School of Artificial Intelligence, Chongqing University of Education, 1, Chongjiao Road, Nanshan Street, Nan'an District, 400065 Chongqing, China
gorbachev@cque.edu.cn

Abstract. Evacuation during emergencies relies on transporting the maximum number of potentially affected people from hazard areas to the safe destinations are becoming more and more popular. However, hard computing approaches to solving these problems have limited capabilities due to such factors as complexity of the problems, environmental factors, human behavior. Uncertainty and fuzziness incorporated into emergency evacuation strategies require fuzzy emergency problem statements and decision-making. Another problem researchers tackle during emergency response is transporting potential victims by the path of the minimum time in a dynamic network, whose arc traversal time can vary because of blocking of hallways with fire. This article integrates flow approaches in transportation networks and hesitant fuzzy decision-making to move the maximum possible number of potential victims to the safe destination in shortest time. On order to determine the priority list in which the aggrieved must be transported to the sink, tools of multi-attribute decision making are used. Hesitant fuzzy ELECTRE method is applied find the ranks of intermediate nodes for transportation the potential victims. The goal of the designed method is to shift the potential victims from the dangerous source nodes to the safe sinks which have various levels of priority based on different attributes such as level of reachability, capacity of destination nodes, reliability and affordability. A case study illustrates the designed algorithm.

Keywords: Multi-attribute decision-making · fuzzy ELECTRE method · dynamic network

1 Introduction

Various disasters such as floods, hurricanes, tsunamis, earthquakes, and wildfires, which often spontaneously occurred, lead to severe damage to life, property and society. To remove the effects of such emergencies, countries all over the world pay great attention to an emergency evacuation.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 303–310, 2023.
https://doi.org/10.1007/978-3-031-39777-6_37

One of the urgent tasks of emergency evacuation is devoted to evacuation from buildings. In this regard, a building under evacuation is presented as a network woth nodes and arcs, where nodes are rooms or lobbies, and detect connection such as corridors, hallways, and stairways. In this paper, we will consider the task of transporting the maximum possible number of potential victims from the hazard area to the safe destination.

Despite their relevance and urgent type, emergency evacuation problems are difficult to be solved accurately. Arc capacities and arc costs (times) can only be evaluated but not accurately defined in exact values. Multi-attribute decision-making (MADM) considers the task of making effective decisions with the highest level of crtainty from a set of alternatives. However, experts cannot always give exact assessments because of the high level of hesitation and uncertainty while making decisions. In this regard, many valuable tools have been suggested to simulate uncertainty while decision-making. Fuzzy sets proposed by L. Zade [1] indicated an expert's uncertainty in the form of a membership functionб that shows the level of belongingness of the element to the set. Later, different extensions of fuzzy sets were developed to express experts' hesitations regarding specific values. Among these extensions there are type-2 fuzzy sets, fuzzy multisets, intuitionistic fuzzy sets, intuitionistic soft fuzzy sets, linguistic arguments, and hesitant fuzzy sets [2]. Hesitant fuzzy sets appear when a decision-maker has some possible values of attributes and is not sure what to choose so using a set of possible membership degrees to assess the attribute [3]. In this paper, experts' evaluations will be presented as fuzzy hesitant numbers for ranking the shelters for evacuation. Various tasks of MADM were introduced in crisp and fuzzy conditions such as TOPSIS, the maximizing deviation method, the ELECTRE, the VIKOR, and the PROMETHEE method [3].

The ELECTRE finds its application in decision-making to outrank the alternatives. It was first proposed in [4] as so-called ELECTRE-I. Later, its extensions, ELECTRE Iv and ELECTRE, were developed to rank the problems by authors in [5], which are called current version of the ELECTRE methods. ELECTRE method consists in finding the best alternative in a group decision making tasks [6]. Authors [7, 8] considered ELECTRE in uncertain and fuzzy conditions due to uncertainty of the surroundings. Torra and Narukawa [9] showed that introducing hesitant fuzzy elements allows avoiding aggregation. Chen et al. [10] considered the ELECTRE I method in fuzzy conditions to consider the differences in experts' opinions. Later, authors [10] proposed a hesitant fuzzy ELECTRE I method for MADM problems in hesitant fuzzy conditions. The present paper illustrates HF ELECTRE I method in order to rank the shelters for evacuation so that to transport the maximum possible number of potential victims.

The paper is structured as follows. In Sect. 2, we overview the main concepts and algorithm, which underlies the paper. Section 4 illustrates a case study. Section 4 gives the concludes.

2 A Proposed Integrated Method of Emergency Decision-Making Based on HF ELECTRE for Finding the Maximum Number of Aggrieved

2.1 Ranking the Shelters for Evacuation Based on HF ELECTREE

Definitions of hesitant fuzzy sets, score, and deviation degree are given in [11]. ELECTRE method relies in four binary relations, the concepts of preference, concordance and discordance [3]. Binary relations on the set X are given in [3].

Present the HF EELECTRE method [3] for finding the priority order of shelters for evacuation [11].

Step 1. Find the hesitant fuzzy decision matrix along with importance vector $w = (w_1, w_2, \ldots, w_n)^T$ for the attributes, and the relative weight vector $\omega = (\omega_C, \omega_{C'}, \omega_{\overline{D}}, \omega_{\overline{D'}})^T$ for various classes of hesitant fuzzy concordance sets (HFC sets) and hesitant fuzzy discordance sets (HFD sets).

Step 2. Find the matrixes with scores and deviations of each evacuation alternative on the attributes as HFEs.

Step 3. Define HFC sets and weak HFC sets.

Step 4. Find HFD sets and weak HFD sets.

Step 5. Find the HFC matrix based on HFC indexes.

Step 6. Calculate the HFD matrix based on the HFD indices.

Step 7. Find the concordance dominance matrix.

Step 8. Construct the discordance dominance matrix.

Step 9. Find the aggregation dominance matrix.

Step 10. Present a decision graph and select the preferable alternative, exit.

Finally, the outranking relations to give recommendations concerning the results are introduced [3]: if $p_{ij} = 1$, then A_i is strictly preferred to A_j or A_{ih} is weakly preferred to A_j. If $p_{ij} = 1$ *and* $p_{ji} = 1$, then A_i is indifferent to A_j. If $p_{ij} = 0$ *and* $p_{ji} = 0$, then A_i is incomparable to A_j.

2.2 Emergency Evacuation Based on the Maximum Dynamic Flow Finding

Equations (1)–(3) describe a mathematical model for the maximum dynamic evacuation flow as a fuzzy transportation network $\tilde{G} = \left(X, \tilde{A}\right)$, where X is the set of nodes, and $\tilde{A}$ is the set of nodes arcs.

The fuzzy flow $\tilde{\xi}_{ij}(\theta)$ travelling by the arc $\left(x_i, x_j\right)$ satisfies the capacity constraint (3). Flow conservation is given in Eq. (2). The goal is to determine the maximum flow from the source to the sinks considering priority order of intermediate shelters for evacuation (1). Intermediate nodes ranked according to expert' evaluation by hesitant fuzzy ELECTREE method described above.

$$\sum_{\theta=0}^{T} val\left(S^k\right) \to max, \theta \in T, \tag{1}$$

$$\sum_{\theta=\tau_{ij}}^{T} \sum_{(i,j)\in \tilde{I}_{x_i}} \tilde{\xi}_{ij}\left(\theta - \tau_{ij}\right) = \sum_{\theta=0}^{T} \sum_{(i,j)\in \tilde{O}_{x_i}} \tilde{\xi}_{ij}(\theta), \forall x_i \notin \{s, d\}, \tag{2}$$

$$0 \le \tilde{\xi}_{ij}(\theta) \le \tilde{u}_{ij}(\theta), \forall \left(x_i, x_j\right) \in \tilde{E}, \theta \in T. \tag{3}$$

Present the algorithm for emergency evacuation of the maximum number of aggrieved after finding the priority order of shelters [12].

Step 1. Transform the initial dynamic network $\tilde{G}$ into a time-spaced network $\tilde{G}^*$ based on copying each node and arc at the specific time period.

Step 2. Pass the flow along the augmenting paths in the residual network $\tilde{G}^{*r}$.

Step 2.1. The If $\tilde{\xi}^{*r}\left(x_i^*, x_j^*, \theta, \vartheta\right) < \tilde{u}^*\left(x_i^*, x_j^*, \theta, \vartheta\right)$ in $\tilde{G}_e^{*r}$, then $\tilde{u}^{*r}\left(x_i^*, x_j^*, \theta, \vartheta\right) = \tilde{u}^*\left(x_i^*, x_j^*, \theta, \vartheta\right) - \tilde{\xi}^*\left(x_i^*, x_j^*, \theta, \vartheta\right)$. If $\tilde{\xi}^*\left(x_i^*, x_j^*, \theta, \vartheta\right) > \tilde{0}$, then $\tilde{u}^{*r}\left(x_j^*, x_i^*, \vartheta, \theta\right) = \tilde{\xi}^*\left(x_i^*, x_j^*, \theta, \vartheta\right)$.

2.2. If the path exists, move to the step 2.3

2.3 If a path to the sink does not exist, turn to step 2.4 as the maximum flow to the destination t is found.

Step 3. Pass the flow $\tilde{\sigma}^* = min[\tilde{u}^{*r}\left(x_i^*, x_j^*, \theta, \vartheta\right)]$, turn to the step 2.5.

Step 4. Find the augmenting paths from the intermediate shelters to the sink T according to the priority order of nodes based on hesitant fuzzy ELECTREE method. The most preferable shelter has the highest priority; then there is the intermediate shelter x_i with the priority highest among others and so on.

4.1 If a path exists, move back to the step 2.3

4.2 If there is no path, the maximum flow to the sink t is found, move to step 2.6

Step 5. Transform the evacuation flows: 1) for arcs joining $(x_j^{*\mu r}, \vartheta)$ and $(x_i^{*\mu r}, \theta)$, decrease the flow value $\tilde{\xi}^{\mu}\left(x_i^*, x_j^*, \theta, \vartheta\right)$ by the value $\tilde{\sigma}^{*\mu}$. The total flow is $\tilde{\xi}^{\mu}\left(x_i^*, x_j^*\theta, \vartheta\right) - \tilde{\sigma}^{*\mu}$. Move back to the step 2.2. 2) for arcs joining $(x_i^{*\mu r}, \theta)$ and $(x_j^{*\mu r}, \vartheta)$, increase the flow value $\tilde{\xi}^{\mu}\left(x_i^*, x_j^*, \theta, \vartheta\right)$ by the value $\tilde{\sigma}^{*\mu}$. Total flow value is $\tilde{\xi}^{\mu}\left(x_i^*, x_j^*, \theta, \vartheta\right) + \tilde{\sigma}^{*\mu}$ and return to the step 2.2

Step 6. Delete dummy terminals. Return to the initial network.

3 Case Study

Simulate the emergency evacuation of potential victims in the fuzzy network considering the priority list of intermediate nodes. To find the order of intermediate shelters, apply a method given in Sect. 2.1. A building for evacuation is given in Fig. 1, where s is the hazard area (source node); exits 4, 5, 6, 7 are the safe locations (sink nodes). A corresponding graph that represents the building plan is shown in Fig. 2. An artificial sink t is introduced to determine the flow in the given network. The goal of the evacuation is to shift the maximum number of evacuees from the source s to the shelter t in 4 time periods considering the order of intermediate shelters for evacuation.

To assess the priority list of terminals, a decision-maker rank them towards four attributes: the level of reachability (B_1), reliability (security) (B_2), capacity (B_3), and affordability of transportation (B_4). Expert's evaluations are given in Table 1 as fuzzy hesitant values.

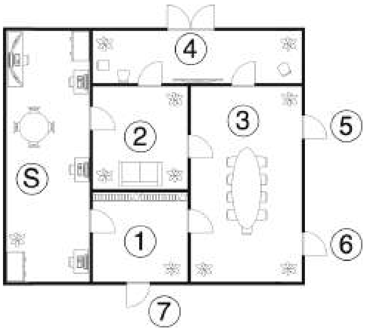


Fig. 1. A building for evacuation.

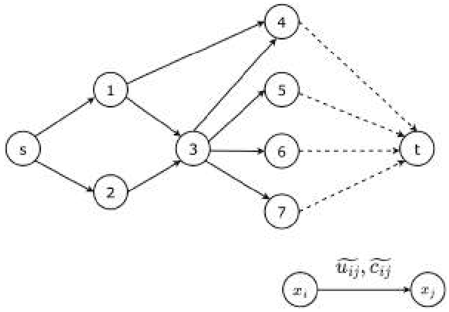


Fig. 2. Network presentation of the graph Fig. 1.

Table 1. Hesitant fuzzy decision matrix of the DM.

	B_1	B_2	B_3	B_4
x_1	(0.4, 0.5)	(0.3, 0.5, 0.7)	(0.6, 0.7)	(0.4, 0.5, 0.6)
x_2	(0.2, 0.3, 0.4)	(0.7, 0.8)	(0.3, 0.5, 0.6)	(0.5, 0.6)
x_3	(0.7, 0.8)	(0.3, 0.5, 0.6, 0.7)	(0.5, 0.8)	(0.4, 0.6)
x_4	(0.4, 0.5, 0.6)	(0.3, 0.5, 0.7)	(0.8, 0.9)	(0.2,0.3,0.5)

Decision-makers find the relative weights of HFC, and weak HFC sets along with HFD, and weak HFD sets: $(\omega_C, \omega_{C'}, \omega_{\overline{D}}, \omega_{\overline{D'}})^T = \left(1, \frac{3}{4}, 1, \frac{3}{4}\right)^T$. The weight vector of criteria is given as $w = (0.23, 0.27, 0.35, 0.25)$.

According to the main steps of the algorithm from Sect. 2.1, we finally turn to the aggregation dominance matrix P: $P = \begin{pmatrix} - & 0 & 0 & 0 \\ 0 & - & 0 & 0 \\ 0 & 0 & - & 0 \\ 1 & 1 & 0 & - \end{pmatrix}$.

From the above matrix, it follows that the shelter 7 is preferred to the shelters 4 and 5. The shelter 7 is indifferent to 6. Exits 4 and 5 are indifferent. It means that the order of transportation is 7, 6, 4, 5 or 7, 6, 5, 4 (Fig. 3).

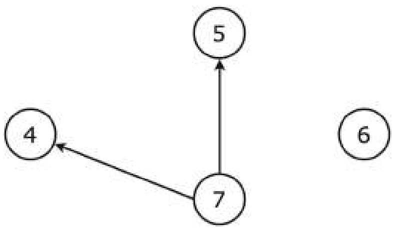


Fig. 3. Decision graph of the evacuation example.

Draw a time-expanded static graph regarding the graph in Fig. 2, as shown in Fig. 4.

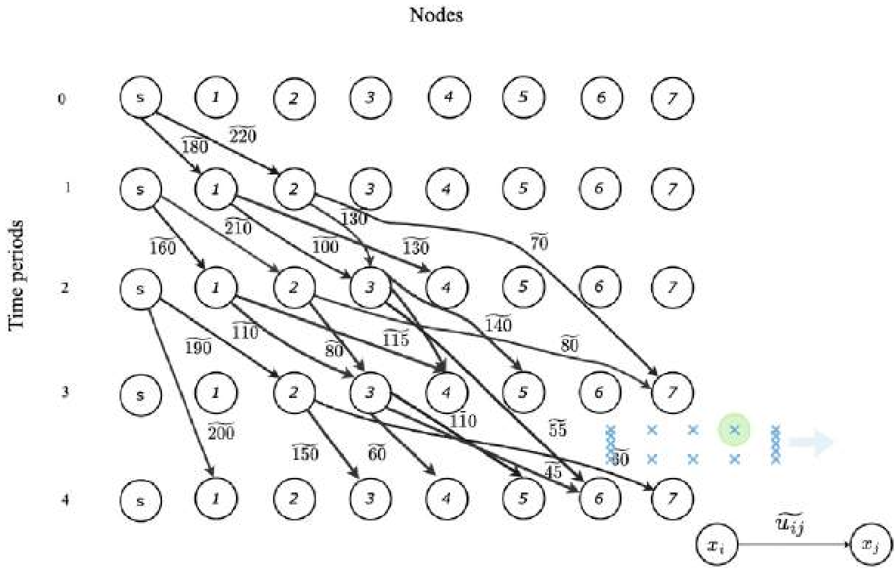


Fig. 4. Time-expanded network of the dynamic network in Fig. 2

Therefore, the final flow with storage at nodes is $\widetilde{760}$ units (Table 2).

Table 2. Flow distribution with intermediate storage.

The path	*Departure node*	*Flow value*
$s_0 \to 2_1 \to 7_3 \to T$	7	$\widetilde{760}$
$s_1 \to 2_2 \to 7_3 \to T$	7	$\widetilde{80}$
$s_2 \to 2_3 \to 7_4 \to T$	7	$\widetilde{60}$
$s_0 \to 1_1 \to 3_2 \to 6_4 \to T$	6	$\widetilde{55}$
$s_1 \to 1_2 \to 3_3 \to 6_4 \to T$	6	$\widetilde{45}$
$s_0 \to 2_1 \to 3_2 \to 5_3 \to T$	5	$\widetilde{130}$
$s_1 \to 2_2 \to 3_3 \to 5_4 \to T$	5	$\widetilde{80}$
$s_0 \to 1_1 \to 3_2 \to 5_3 \to T$	5	$\widetilde{10}$
$s_1 \to 1_2 \to 3_3 \to 5_4 \to T$	5	$\widetilde{30}$
$s_0 \to 1_1 \to 4_2 \to T$	4	$\widetilde{115}$
$s_1 \to 1_2 \to 4_3 \to T$	4	$\widetilde{85}$
Total flow		$\widetilde{760}$

The corresponding flow regarding the maximum evacuation flow is shown in the network (Fig. 5).

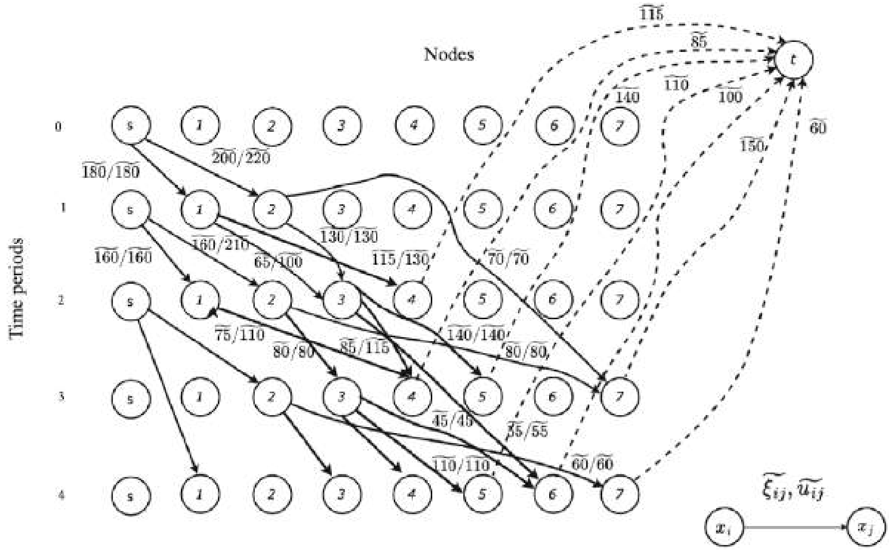


Fig. 5. Network with augmenting paths and a total flow.

4 Conclusion

Emergency evacuation as the significant task of evacuation relates to building evacuation. This task consists in transporting the maximum number of potential victims as quickly as possible to the safe shelter. However, complexity of evacuation tasks and fuzzy environment influences emergency evacuation problems so that they are considered in fuzzy conditions using hesitant fuzzy sets. Several membership degrees instead of one can be utilized by a decision-maker while giving the evaluation. These degrees indicate experts' hesitation and uncertainty regarding the specific value. In the paper, MADM method considers the task of making the effective evacuation order with the highest level of satisfaction from a set of evacuation alternatives. In particular, the order of intermediate shelters for evacuation is found based on hesitant fuzzy ELECTRE method, which allows to consider experts' hesitation and insurance. This article integrates flow approaches in dynamic transportation networks and hesitant fuzzy decision-making to convey the maximum possible number of potential aggrieved in shortest time to the safe destination. Dynamic network handles time-varying arc capacities and parameters of time to consider covering of hallways with fire. The pivotal role of the illustrated algorithm is to convey the maximum number of aggrieved from the dangerous source nodes to the safe sinks which have various levels of priority based on different attributes such as level of reachability, capacity of destination nodes, reliability and total expenses. In the future, the extensions of hesitant fuzzy ELECTRE method and hesitant fuzzy PROMETHEE method will be applied to compare the results and give recommendations.

Acknowledgments. The research was funded by the Russian Science Foundation Project No. 22–71-10121, https://rscf.ru/en/project/22-71-10121/ implemented by the Southern Federal University.

References

1. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
2. Xu, Z., Cai, X.: Intuitionistic Fuzzy Information Aggregation: Theory and Applications. Science Press, Springer (2012)
3. Hu, Z.: Hesitant Fuzzy Sets Theory. Studies in Fuzziness and Soft Computing (2014)
4. Benayoun, R., Roy, B., Sussman, B.: ELECTRE: Une méthode pour guider le choix en présence de points de vue multiples. Note de Travail 49, SEMA-METRA International, Direction Scientifique (1966)
5. Roy, B.: The outranking approach and the foundation of ELECTRE methods. Theor. Decis. **31**, 49–73 (1991)
6. Kaya, T., Kahraman, C.: An integrated fuzzy AHP–ELECTRE methodology for environmental impact assessment. Expert Syst. Appl. **38**, 8553–8562 (2011)
7. Wu, M.C., Chen, T.Y.: The ELECTRE multicriteria analysis approach based on Atanassov's intuitionistic fuzzy sets. Expert Syst. Appl. **38**, 12318–12327 (2011)
8. Vahdani, B., Jabbari, A.H.K., Roshanaei, V., Zandieh, M.: Extension of the ELECTRE method for decision-making problems with interval weights and data. Int. J. Adv. Manufact. Technol. **50**, 793–800 (2010)
9. Torra, V., Narukawa, Y.: On hesitant fuzzy sets and decision. In: The 18th IEEE International Conference on Fuzzy Systems, Jeju Island, Korea, pp. 1378–1382 (2009)
10. Chen, N., Xu, Z.S., Xia, M.M.: The ELECTRE I multi-criteria decision making method based on hesitant fuzzy sets. Int. J. Inf. Technol. Decis. Mak. **14**(3), 621–657 (2015)
11. Gerasimenko, E., Kureichik, V.: Hesitant Fuzzy emergency decision-making for the maximum flow finding with intermediate storage at nodes. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds.) INFUS 2021. LNNS, vol. 307, pp. 705–713. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-85626-7_82
12. Kacprzyk, J., Bozhenyuk, A., Gerasimenko, E.: Maximum dynamic evacuation modelling in networks in fuzzy conditions with partial lane reversal. LNNS, vol. 610, pp. 16–24 (2023). https://doi.org/10.1007/978-3-031-25252-5_8

Evaluation of MENA Countries in Views of SDG6 Targets Using a Hybrid MCDM Approach

Serhat Aydın[1](✉), Ahmet Aktas[2], Mehmet Kabak[3], and Cengiz Kahraman[4]

[1] National Defence University, 34149 Istanbul, Turkey
saydin3@hho.msu.edu.tr
[2] University of Turkish Aeronautical Association, 06790 Ankara, Turkey
[3] Gazi University, 06570 Ankara, Turkey
[4] Istanbul Technical University, 34367 Istanbul, Turkey

Abstract. United Nations formulated Sustainable Development Goals (SDG) in 2015, which are the collection of a set of objectives to provide a world to people with peace and prosperity now and in the future. 17 goals are identified as achievable targets in 2030 and they contain elimination of poverty, hunger, inequalities and improvement of health, education, environment, security and economic conditions around the world. One of SDGs is defined as clean water and sanitation and known as SDG6. Clean water and sanitation is essential for health and well-being of people and evaluation of countries considering several indicators of SDG6 simultaneously should be done. Especially in MENA countries, clean water and sanitation is a popular topic and the main aim of this study is to develop a multi-criteria evaluation model for these countries in views of SDG6 indicators. To do so, a hybrid approach based on Interval Valued Intuitionistic Fuzzy Analytic Hierarchy Process (IVIF-AHP) and Technique of Order Preference by Similarity to Ideal Solution (TOPSIS) methods is proposed. IVIF-AHP is used to determine indicator weight values and TOPSIS is used to obtain the ranking of countries. Evaluation of MENA nations is conducted to show the viability of the suggested approach. The results of the study could be useful for sustainable management of water and sanitation systems in these countries.

Keywords: SDG6 · AHP · Interval Valued Intuitionistic Fuzzy Sets · TOPSIS · MENA Countries

1 Introduction

The Earth can no longer support our population levels. On Earth now, there are more than 8,000,000,000 individuals. The first billion people were not counted in the world's population until the early 1800s. Every 12 to 15 years, we now add a billion. Therefore, some measures must be taken due to the rapid population growth in the world.

The Sustainable Development Goals (SDGs) were established by the United Nations in 2015 with the intention of removing poverty, preserving the natural world, and guaranteeing that everyone lives in tranquility and abundance by the year 2030. SDGs seek

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 311–318, 2023.
https://doi.org/10.1007/978-3-031-39777-6_38

to alter the course of global affairs [1]. They operate as a catalyst for action in the fight against poverty and inequality, environmental protection, and ensuring that everyone may live in calm, equity, and affluence.

The United Nations outlines 17 Goals with 169 goals. According to the 17 SDGs, sustainable development must strike a balance between social, economic, and environmental sustainability, and decisions made in one area will affect outcomes in other areas. Governments have decided to give priority to those who are falling behind the most while progress is being made. The SDGs seek to end bias against women and girls, AIDS, hunger, and poverty [2]. The SDGs are " affordable and clean energy; no poverty; climate action; decent work and economic growth; sustainable cities and communities; reduced inequalities; good health and well-being; gender equality; life on land; clean water and sanitation; life below waterzero hunger; industry, innovation and infrastructure; responsible consumption and production; quality education; peace, justice, and strong institutions; and partnerships for the goals".

One of SDGs is defined as clean water and sanitation and known as SDG6. SDG 6 states that water and sanitation are available and are managed sustainably for all people. Significant indicators for this goal are the proportions of the population who have access to securely managed sanitation and utilize safely managed drinking water. Water scarcity affects more than 40% of people, a stunning ratio that is anticipated to climb as temperatures rise. Every continent is being impacted by declining drinking water resources, even though 2.1 billion people have improved the quality of their water since 1990. By 2030, we must invest in sufficient infrastructure, offer sanitary facilities, and promote good hygiene to ensure that everyone has access to safe, cheap drinking water.

Multi-Criteria Decision Making (MCDM) provides efficient decision-making in situations where selecting the best option is quite challenging [3]. The main objective is to discover various applications and processes and to suggest methods that can be used to choose the best alternatives quickly and accurately. Multi-criteria decision-making has been utilized in numerous sectors.

Because many real measurements are unclear and variable, as well as because human reasoning is complex, it is difficult to utilize accurate numerical values to evaluate criteria and alternatives in multi-criteria decision-making. Experts must make fuzzy judgments at this point since many real-world issues need it. The premise of fuzzy logic [4] is that people make decisions based on ambiguous and non-numerical data. To deal with ambiguity, fuzziness, and imprecision, fuzzy logic is a helpful technique. As a result, numerous applications integrate MCDM and fuzzy logic.

In this research a hybrid approach based on Interval Valued Intuitionistic Fuzzy Analytic Hierarchy Process (IVIF-AHP) and Technique of Order Preference by Similarity to Ideal Solution (TOPSIS) methods is proposed. The proposed method is utilized to evaluate Middle East and North Africa (MENA) countries according to the SDG6 criteria. In the literature some works have been realized by researchers on MENA countries; [5–7]. The difference between this study and other studies is that this paper proposed a fuzzy hybrid approach using IVIF-AHP and TOPSIS method to evaluate MENA countries first time in the literature.

The remaining sections are arranged as follows: Sect. 2 presents the preliminaries of interval valued intuitionistic fuzzy sets and the steps of the proposed algorithm.

Application is given Sect. 3. Section 4 offers a conclusion at the end.

2 Preliminaries and the Proposed Approach

Basic definitions related to intuitionistic and interval valued intuitionistic fuzzy sets are given in this section as follows:

Definition 1: Intuitionistic Fuzzy Set ($\dot{I}$) of a fixed set X can be shown as in Eq. (1):

$$\tilde{I} = \sum x, I\left(\mu_{\tilde{I}}(x), v_{\tilde{I}}(x)\right); x \in X \sum \tag{1}$$

where $\mu_{\tilde{I}}(x) : X \rightarrow [0, 1]$ and $v_{\tilde{I}}(x) : X \rightarrow [0, 1]$ shows the membership and non-membership degree of the element $x \in X$, respectively. The degree of indeterminacy can be calculated by Eq. (2):

$$\pi_{\tilde{I}}(x) = 1 - \mu_{\tilde{I}}(x) - v_{\tilde{I}}(x) \tag{2}$$

Definition 2: An interval valued intuitionistic fuzzy set (IVIF) ($\tilde{I}$) can be defined as follows:

$$\tilde{I} = \left[\,\right]x, \ \ \mu_{\tilde{I}}^{-}(x), \mu_{\tilde{I}}^{+}(x)\Big], \ \ v_{\tilde{I}}^{-}(x), v_{\tilde{I}}^{+}(x)\Big]\Big\rangle; x \in X\Big\} \tag{3}$$

where $0 \leq \mu_{\tilde{I}}^{-}(x) \leq 1$, $0 \leq \mu_{\tilde{I}}^{+}(x) \leq 1$, $0 \leq v_{\tilde{I}}^{-}(x) \leq 1$, $0 \leq v_{\tilde{I}}^{+}(x) \leq 1$, $\mu_{\tilde{I}}^{-}(x) \leq \mu_{\tilde{I}}^{+}(x)$, $v_{\tilde{I}}^{-}(x) \leq v_{\tilde{I}}^{+}(x)$, $0 \leq \mu_{\tilde{I}}^{+}(x) + v_{\tilde{I}}^{+}(x) \leq 1$ and $0 \leq \mu_{\tilde{I}}^{-}(x) + v_{\tilde{I}}^{-}(x) \leq 1$.

Definition 3: Some arithmetic operations of two IVIF numbers $\widetilde{I_1} = \ \mu_{\widetilde{I_1}}^{-}, \mu_{\widetilde{I_1}}^{+}\Big], \ v_{\widetilde{I_1}}^{-}, v_{\widetilde{I_1}}^{+}\Big]$ and $\widetilde{I_2} = \ \mu_{\widetilde{I_2}}^{-}, \mu_{\widetilde{I_2}}^{+}\Big], \ v_{\widetilde{I_2}}^{-}, v_{\widetilde{I_2}}^{+}\Big]$ can be given as follows:

$$\widetilde{I_1} \oplus \widetilde{I_2} = \ \mu_{\widetilde{I_1}}^{-} + \mu_{\widetilde{I_2}}^{-} - \mu_{\widetilde{I_1}}^{-}\mu_{\widetilde{I_2}}^{-}, \mu_{\widetilde{I_1}}^{+} + \mu_{\widetilde{I_2}}^{+} - \mu_{\widetilde{I_1}}^{+}\mu_{\widetilde{I_2}}^{+}\Big], \ \ v_{\widetilde{I_1}}^{-}v_{\widetilde{I_2}}^{-}, v_{\widetilde{I_1}}^{+}v_{\widetilde{I_2}}^{+}\Big] \tag{4}$$

$$\widetilde{I_1} \otimes \widetilde{I_2} = \ \mu_{\widetilde{I_1}}^{-}\mu_{\widetilde{I_2}}^{-}, \mu_{\widetilde{I_1}}^{+}\mu_{\widetilde{I_2}}^{+}\Big], \ \ v_{\widetilde{I_1}}^{-} + v_{\widetilde{I_2}}^{-} - v_{\widetilde{I_1}}^{-}v_{\widetilde{I_2}}^{-}, v_{\widetilde{I_1}}^{+} + v_{\widetilde{I_2}}^{+} - v_{\widetilde{I_1}}^{+}v_{\widetilde{I_2}}^{+}\Big] \tag{5}$$

$$\lambda\widetilde{I_1} = \left[1 - \left(1 - \mu_{\widetilde{I_1}}^{-}\right)^{\lambda}, 1 - \left(1 - \mu_{\widetilde{I_1}}^{+}\right)^{\lambda}\right], \left[\left(v_{\widetilde{I_1}}^{-}\right)^{\lambda}, \left(v_{\widetilde{I_1}}^{+}\right)^{\lambda}\right] \tag{6}$$

$$(\widetilde{I_1})^{\lambda} = \left[\left(\mu_{\widetilde{I_1}}^{-}\right)^{\lambda}, \left(\mu_{\widetilde{I_1}}^{+}\right)^{\lambda}\right], \left[1 - \left(1 - v_{\widetilde{I_1}}^{-}\right)^{\lambda}, 1 - \left(1 - v_{\widetilde{I_1}}^{+}\right)^{\lambda}\right] \tag{7}$$

$$\frac{\widetilde{I_1}}{\widetilde{I_2}} = \ \min\left(\mu_{\widetilde{I_1}}^{-}, \mu_{\widetilde{I_2}}^{-}\right), \min\left(\mu_{\widetilde{I_1}}^{+}, \mu_{\widetilde{I_2}}^{+}\right)\Big], \ \ \max\left(v_{\widetilde{I_1}}^{-}, v_{\widetilde{I_2}}^{-}\right), \max\left(v_{\widetilde{I_1}}^{+}, v_{\widetilde{I_2}}^{+}\right)\Big] \tag{8}$$

Definition 4: Score function for an IVIF number $S\left(\widetilde{I_1}\right)$ can be calculated by Eq. (9):

$$S\left(\widetilde{I_1}\right) = \frac{\mu_{\widetilde{I_1}}^{-} + \mu_{\widetilde{I_1}}^{+} - v_{\widetilde{I_1}}^{-} - v_{\widetilde{I_1}}^{+}}{2} \tag{9}$$

Definition 5: Accuracy function of $\widetilde{I_1}$ is denoted by $H\left(\widetilde{I_1}\right)$ can be expressed by Eq. (10):

$$H\left(\widetilde{I_1}\right) = \frac{\mu_{\widetilde{I_1}}^{-} + \mu_{\widetilde{I_1}}^{+} + v_{\widetilde{I_1}}^{-} + v_{\widetilde{I_1}}^{+}}{2} \tag{10}$$

Definition 6: Comparison of two IVIF numbers can be done by the following rules:

$$if\ S\left(\widetilde{I_1}\right) < S\left(\widetilde{I_2}\right);\ then\ \widetilde{I_1} < \widetilde{I_2} \tag{11}$$

$$ifS\left(\widetilde{I_1}\right) = S\left(\widetilde{I_2}\right);\ then$$

$$ifH\left(\widetilde{I_1}\right) < H\left(\widetilde{I_2}\right);\ then\widetilde{I_1} < \widetilde{I_2}$$

$$ifH\left(\widetilde{I_1}\right) = H\left(\widetilde{I_2}\right);\ then\widetilde{I_1} = \widetilde{I_2} \tag{12}$$

Definition 7: Defuzzification of an IVIF number can be done by Eq. (13):

$$D\left(\widetilde{I_1}\right) = \frac{\mu_{\widetilde{I_1}}^{-} + \mu_{\widetilde{I_1}}^{+} + \left(1 - v_{\widetilde{I_1}}^{-}\right) + \left(1 - v_{\widetilde{I_1}}^{+}\right) + \mu_{\widetilde{I_1}}^{-}\mu_{\widetilde{I_1}}^{+} - \sqrt{\left(1 - v_{\widetilde{I_1}}^{-}\right)\left(1 - v_{\widetilde{I_1}}^{+}\right)}}{4} \tag{13}$$

Based on the given preliminary information, a hybrid method based on Interval Valued Intuitionistic Fuzzy Analytic Hierarchy Process (IVIF-AHP) and Technique of Order Preference by Similarity to Ideal Solution (TOPSIS) methods was presented in this study for evaluation of MENA countries in views of SDG6 targets. IVIF-AHP approach was utilized for the calculation of importance degree values of effective factors on SDG6 targets. After obtaining the importance degree of these factors, TOPSIS method was applied to evaluate countries. The suggested model's steps are shown in Fig. 1. Due to the page limitations in proceedings, details for steps of the methods can be seen in Kahraman et al. [8] and Kabak et al. [9].

3 Evaluation of MENA Countries using Proposed Model

Application steps of the evaluation procedure of MENA countries in views of SDG6 targets are presented in this section as follows:

Step 1.1: Evaluation of some of the MENA countries with respect to SDG6 targets is the main concern of this study. Water quality is a critical issue for human being, and water is a critical issue for people in MENA countries.

Step 1.2: Eight countries in the region (Algeria, Bahrain, Iraq, Jordan, Kuwait, Morocco, Qatar, and Tunisia) were evaluated in terms of seven criteria (drinking water

Phase 1: Problem Definition

- Step 1.1: Problem Statement
- Step 1.2: Definition of criteria and alternatives

Phase 2: IVIF-AHP

- Step 2.1: Construction of pairwise comparison matrices for criteria evaluation
- Step 2.2: Formation of score judgement matrices
- Step 2.3: Calculation of interval exponential matrices
- Step 2.4: Calculation of priority vectors
- Step 2.5: Obtaining possibility degree matrices
- Step 2.6: Calculation of criteria weights
- Step 2.7: Normalization of criteria weights

Phase 3: TOPSIS

- Step 3.1: Construction of decision matrix for alternative evaluation
- Step 3.2: Normalization of decision matrix
- Step 3.3: Calculation of weighted normalized decision matrix
- Step 3.4: Determination of ideal and anti-ideal solutions
- Step 3.5: Calculation of distances of each alternative to ideal solution
- Step 3.6: Calculation of relative closeness measure
- Step 3.7: Obtaining evaluation results

Fig. 1. Steps of the evaluation model

(C1), sanitation (C2), wastewater (C3), efficiency (C4), water stress (C5), water management(C6), and ecosystems (C7)). The SDG6 Data Portal of United Nations [10] is used to gather application data. Countries and indicators with missing values were eliminated from the analysis, that is why only eight out of 20 MENA countries [11] and seven criteria were taken into account.

Step 2.1: Pairwise comparison matrix was constructed for evaluation of criteria by using the linguistic terms in Table 1.

Pairwise comparison matrix for criteria is given in Table 2.

Steps 2.2 – 2.7: Score judgment matrix was obtained by using Eq. (9). Then, interval exponential matrix was calculated. Next, priority vectors were calculated for each criterion and possibility degree matrix was constructed. Elements of possibility degree matrix was used to obtain criteria weights. Finally, normalized criteria weight values were found by using vector normalization formula. The formulation for these calculations for interval steps of the method can be seen in Kahraman et al. [8]. Due to the page limitations, these steps was not presented in the paper. The criteria weights were presented in Table 3.

Considering the outcomes of the IVIF-AHP approach, drinking water (C1) seems to be the most important factor for evaluation of countries in views of SDG6 targets.

Step 3.1: Decision matrix of the problem constructed as it is shown in Table 4.

Steps 3.2 – 3.5: To generate a normalized decision matrix, the vector normalization formula is used to the decision matrix's component parts. Then, weight values obtained from the IVIF-AHP application were used to calculate the elements of the weighted

Table 1. Linguistic terms for pairwise comparisons.

Linguistic terms	Corresponding IVIF numbers $\left[\mu_{\tilde{I}}^{-}(x), \mu_{\tilde{I}}^{+}(x)\right], \left[v_{\tilde{I}}^{-}(x), v_{\tilde{I}}^{+}(x)\right]$
Absolutely low importance - AL	[0.10, 0.25], [0.65, 0.75]
Very low importance – VL	[0.15, 0.30], [0.60, 0.70]
Low importance – L	[0.20, 0.35], [0.55, 0.65]
Medium low importance – ML	[0.25, 0.40], [0.50, 0.60]
Equal importance – E	[0.50, 0.50], [0.50, 0.50]
Medium high importance – MH	[0.50, 0.60], [0.25, 0.40]
High importance – H	[0.55, 0.65], [0.20, 0.35]
Very high importance – VH	[0.60, 0.70], [0.15, 0.30]
Absolutely high importance – AH	[0.65, 0.75], [0.10, 0.25]

Table 2. Pairwise comparison matrix of criteria with respect to goal

	C1	C2	C3	C4	C5	C6	C7
C1	E	H	VH	MH	VH	H	AH
C2	L	E	MH	ML	MH	E	VH
C3	VL	ML	E	ML	E	MH	H
C4	ML	MH	MH	E	H	MH	VH
C5	VL	ML	E	L	E	L	MH
C6	L	E	ML	ML	H	E	VH
C7	AL	VL	L	VL	ML	VL	E

Table 3. Criteria weights

Criteria	C1	C2	C3	C4	C5	C6	C7
Weight value	0.198	0.155	0.135	0.175	0.111	0.145	0.081

normalized decision matrix. The definitions of the TOPSIS technique state that the greatest and minimum values of each column of the weighted normalized decision matrix, respectively, represent ideal and anti-ideal answers for benefit type criterion. Since the higher score values for each country are better in this analysis, all the criteria are benefit type criteria. So, ideal and anti-ideal solutions were found in this way. By using ideal and anti-ideal solutions, elements of the weighted normalized decision matrix, and Euclidean distance formula, distance of each alternative to ideal and anti-ideal solutions were calculated.

Table 4. Decision matrix

	C1	C2	C3	C4	C5	C6	C7
Algeria	72	18	76	15	138	54	11
Bahrain	99	91	96	79	134	39	100
Iraq	60	43	37	3	80	38	20
Jordan	86	82	82	33	104	64	35
Kuwait	100	100	85	107	3851	94	60
Morocco	80	39	36	9	51	71	30
Qatar	96	97	100	209	431	81	67
Tunisia	79	81	60	10	96	60	21

Steps 3.6 – 3.7: Relative closeness coefficient is the ratio of an alternative's distance to anti-ideal solution to the sum of the alternative's distance to ideal and anti-ideal solutions. Decreasing order of relative closeness coefficient provides the ranking of alternatives. In other words, the alternative with the greatest value is the most appropriate one. Distance of each alternative to the ideal and anti-ideal solutions, relative closeness coefficient values, and the rank of countries are given in Table 5, as follows:

Table 5. Decision matrix

	Distance to Ideal Solution	Distance to Anti-Ideal Solution	Relative Closeness Coefficient	Rank
Algeria	0.1937	0.0312	0.1388	7
Bahrain	0.1469	0.1029	0.4120	3
Iraq	0.2019	0.0190	0.0859	8
Jordan	0.1703	0.0673	0.2834	4
Kuwait	0.0752	0.1597	0.6798	1
Morocco	0.1928	0.0362	0.1582	6
Qatar	0.1005	0.1697	0.6281	2
Tunisia	0.1864	0.0543	0.2258	5

The rank of countries taken into account in this analysis for SDG6 targets is determined as follows:

Kuwait > Qatar > Bahrain > Jordan > Tunisia > Morocco > Algeria > Iraq.

4 Conclusion

Clean water and sanitation is an important condition for a better quality of life for people. For this reason, United Nations considers this topic among their sustainable development goals as SDG6. SDG6 contains a number of indicators, which are measured on a country basis, and evaluation of countries within this concept can be made by using MCDM approaches.

In this study, a hybrid approach based on IVIF – AHP and TOPSIS methods was proposed as a tool for evaluation of MENA countries in views of SDG6 targets. Eight countries were ranked with the proposed approach by taking seven criteria into account. Among the eight countries evaluated, Kuwait was evaluated in the first rank.

This study may be expanded in future research by taking other indicators into account. Which represents other targets related to SDG6. Also, uncertainty of countries data can be handled by fuzzy modelling, which can lead researchers to develop different hybrid methodologies based on VIKOR, WASPAS, MARCOS, etc.

References

1. Arora-Jonsson, S.: The sustainable development goals: a universalist promise for the future. Futures **146**, 103087 (2023)
2. United Nations: The 17 Goals | Sustainable Development (2023). https://sdgs.un.org/goals. Access date: 26 March 2023
3. Aydin, S., Kahraman, C., Kabak, M.: Development of harmonic aggregation operator with trapezoidal Pythagorean fuzzy numbers. Soft. Comput. **24**(15), 11791–11803 (2020). https://doi.org/10.1007/s00500-019-04638-4
4. Zadeh, L.A.: Fuzzy sets. Information and control **8**(3), 338–353 (1965)
5. Emara, N., El Said, A.: Financial inclusion and economic growth: the role of governance in selected MENA countries. Int. Rev. Econ. Financ. **75**, 34–54 (2021)
6. Masca, M.: A Comparison of Turkey's Human Capital Stock with Some Selected Mena Countries by TOPSIS Method. Int. J. Bus. Econo. Affa. **4**(1), 21–30 (2019)
7. Magableh, G.M., Mistarihi, M.Z.: Applications of MCDM approach (ANP-TOPSIS) to evaluate supply chain solutions in the context of COVID-19. Heliyon **8**(3), e09062 (2022)
8. Kahraman, C., Oztaysi, B., Cevik Onar, S.: An integrated intuitionistic fuzzy AHP and TOPSIS approach to evaluation of outsource manufacturers. J. Intell. Syst. **29**(1), 283–297 (2020)
9. Kabak, M., Saglam, F., Aktas, A.: Usability analysis of different distance measures on TOPSIS. J. Facul. Eng. Architec. Gazi Univ. **32**(1), 35–43 (2017)
10. United Nations: SDG 6 Data Portal (2023). https://www.unwater.org/our-work/sdg-6-data-portal. Access date: 10 January 2023
11. Wikipedia: MENA (2023). https://en.wikipedia.org/wiki/MENA. Access date: 10 January 2023

Application of Fuzzy AHP Method for Selection of Equipment for Concrete Works

Natasa Prascevic(✉)

Faculty of Civil Engineering, University of Belgrade, Bulevar Kralja Aleksandra, 73 11000 Belgrade, Serbia
natasa@grf.bg.ac.rs

Abstract. During the realization of a construction project, in addition to the laborious work, it is necessary to engage appropriate construction equipment. Hiring appropriate equipment is a complex and risky operation due to the existence of a number of factors that influence the right choice. These factors imply: economical aspects, operating requirements, reliability of equipment, service support, use in future projects, etc. Therefore, the problem of choosing construction equipment for the realization of a construction project should be seen as a problem of multi-criteria optimization. Since the values of factors influencing the correct choice of mechanization cannot be precisely determined, it is necessary to consider these factors as fuzzy numbers. In this paper, the application of the fuzzy AHP method with triangular fuzzy numbers for the selection of mechanization for performing concrete works is proposed. When ranking alternatives, the following quantitative and qualitative criteria were taken into account: total cost, duration of the project, availability of mechanization, environmental impact, ownership of mechanization, and conditions of maintenance of mechanization. The paper presents a case study of the selection of the best group of equipment for performing concrete works.

Keywords: Fuzzy AHP method · Selection of construction equipment · Concrete works

1 Introduction

In order to perform construction work as successfully, qualitatively, and quickly as possible, the application of construction machinery is necessary. For the correct selection of construction machinery, it is necessary to know the constructive and exploitative characteristics of the machines used for a particular job, as well as the construction technology, in order to notice the appropriate operations of the work process. In order to make the choice of machines as well as possible, there are several principles that should be adhered to: One should not choose too large or too small machines; one should procure those machines that will perform the most common type of work; one should use more standard machines than special ones because they are cheaper; one should examine whether it is more economical to buy or rent a machine; in order to avoid downtime, spare machines and spare energy sources should be provided; and care should be taken that as many machines as possible are of the same type or that as many different machines have the same type of propellant.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 319–326, 2023.
https://doi.org/10.1007/978-3-031-39777-6_39

The first step in choosing machines is a wider selection that gives us an insight into the available ones that could participate in the execution of certain processes. The second step is a shortlist, on the basis of which, out of the offered groups of machines from a wider selection, the group of machines that provide the highest economic value (the lowest price per unit of measure) is selected. In this paper, it is proposed that the shortlist of machines be made not solely on the basis of the lowest price per unit of measure but also to take into account some of the aforementioned factors. Therefore, the problem of choosing construction machinery for construction projects must be seen as a multi-criteria optimization problem.

Numerous researchers have chosen construction equipment using a variety of multi-criteria optimization techniques. Temiz and Calis [1] compared the outcomes after using the AHP and PROMETHEE methodologies to choose the best excavating machine. The AHP approach was used by Waris *et al*. [2] to address the issue of the sustainable buying of construction equipment. In order to prioritize the variables affecting the choice to purchase equipment, Petroutsatou *et al*. [3] modified the AHP model. As an efficient way to formalize the information held by qualified, seasoned practitioners, Shapira and Goldenberg [4] advocated modifying AHP to take into account the nature of equipment selection and its use. The fuzzy SWARA (Step-wise Weighted Average Regression) extensions provided by Ghorabaee *et al*. In order to choose the best handling system from a range of potential options, Yazdani-Chamzini [7] proposed an integrated fuzzy model. First, the fuzzy AHP (FAHP) method was used to determine the relative importance of the evaluation criteria. Next, the fuzzy TOPSIS (FTOPSIS) method was used to evaluate the feasible handling equipment. Ghazikalayeh *et al*. [8] established the MCDM approach for determining the best drilling, loading, and haulage equipment combinations. The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) and the fuzzy analytic network process (FANP) are merged in the suggested approach. The Neuro-Fuzzy System was suggested by Bozanic *et al*. [9] as a tool to aid in the decision-making process when choosing construction equipment. The MCDM methods, including Logarithm Methodology of Additive Weights (LMAW), VIKOR, TOPSIS, MOORA, and SAW, were used to prepare the data for Neuro-Fuzzy System modeling. Fuzzy logic was used by Hanna and Lotfallah [10] to choose the optimal crane type from among mobile, tower, and derrick cranes for a building project.

A literature review shows that various MCDM methods have been successfully used to select construction machinery. In this paper, the application of the fuzzy AHP method with triangular fuzzy numbers for the selection of mechanization for performing concrete works is proposed. Eigenvalues are determined using the expected value of fuzzy numbers and their products. The obtained eigenvalues and eigenvectors are applied in a further procedure to rank alternatives. The proposed procedure was used in a case study for the selection of machines for performing concrete works during the construction of a residential complex.

2 Methodology

2.1 AHP Method

Analytical hierarchical process (AHP) is a technique used in the field of multicriteria decision-making that was proposed by Saaty [11]. It is based on breaking down a complex problem into a hierarchy, where the goal is at the top and criteria, subcriteria, and alternatives are formed at multiple hierarchical levels. The decision-maker compares the elements in pairs at each hierarchical level with respect to the element at a higher level, using the so-called Saaty's scale. The end results of the comparison are vectors of relative importance (priority criteria) and alternatives to the goal. A central place in the evaluation of elements of hierarchy according to the AHP method are matrices of comparison—obtained transformations of semantic assessments of decision makers on the mutual significance of elements into numerical values. Thus, in relation to each element of the hierarchy, one matrix of comparison **A** is formed from a higher level by evaluating the elements at the observed level of the hierarchy. From each such matrix should be extracted a priority vector of elements that can be designated as w.

The AHP method is implemented in four steps:

Step 1: Developing a hierarchical model of decision problems by placing the top criteria and their corresponding subcriteria and alternatives at the lowest level;

Step 2: At each level of the hierarchical structure, the elements are compared to each other in pairs, whereby the decision-maker's preferences are expressed using the appropriate scale (the Saaty's scale of relative importance);

Step 3: Form a matrix of comparison. Whose element a_{ij} represents the ratio of the weight of criteria i according to the weight of criteria j, i.e.

$$a_{ij} = \frac{w_i}{w_j}, i, j = 1, 2, \ldots k \tag{1}$$

Comparison matrix **A** is reciprocal, i.e.

$$a_{ij} = \frac{1}{a_{ji}}, a_{ii} = 1, i, j = 1, 2, \ldots . k \tag{2}$$

A weight vector w is determined by solving the corresponding eigenvalue problem

$$\mathbf{A}w = kw \tag{3}$$

where $k = \lambda_{\max}$.

The vector that is obtained by solving the expression (3) is not normalized. Replacing w_i with ${}^{w_i}/[\ {}_{i=1}^{k} w_i$ can be obtained as a normalized vector to determine the relative significance of attributes. The measure of the consistency of deviations k of $\lambda_{\max}$ (consistency index) is obtained from the following expression:

$$CI = \frac{\lambda_{\max} - k}{k - 1}$$

An IC value of less than 0.10 is considered a satisfactory measure, indicating that estimates for a_{ij} are consistent and therefore a certain value for $\lambda_{\max}$ is approximately equal to the ideal value we wish to estimate.

Step 4: Determination of the composite normalized vector obtained by multiplying the weight vector of all successive levels. This composite vector is then used to find the relative priorities of all entities at the lowest (hierarchical) level.

2.2 Fuzzy AHP Method

Since some criteria are qualitative in nature, decision makers cannot simply express them. Because of this, crisp numbers are not suitable for pair-wise comparison values. To overcome this disadvantage, the fuzzy AHP method is proposed. In this paper, a triangular fuzzy number is used (Fig. 1).

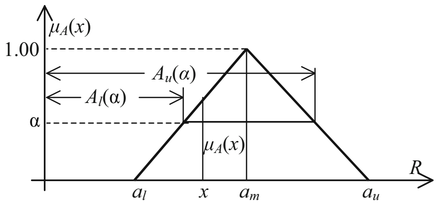


Fig. 1. Triangular fuzzy number $\tilde{A}$

A triangular fuzzy number is usually described by three characteristic values: a_l, a_m, and a_u,which are crisp numbers $\tilde{A} = (a_l, a_m, a_u)$. Reciprocal fuzzy number $\tilde{A}^{-1}$ to $\tilde{A}$ is $\tilde{A}^{-1} = (1/a_u, 1/a_m, 1/a_l)$.

The fuzzy comparison matrix $\tilde{\mathbf{A}}$ consists of elements

$$\tilde{a}_{ij} = \left]a_{ij,l}, a_{ij,m}, a_{ij,u}\right), \tilde{a}_{ii} = 1, i, j = 1, 2, \ldots, k, i \neq j,$$

is reciprocal matrix i.e.

$$\tilde{a}_{ij} = 1/\tilde{a}_{ji} = \left]1/a_{ji,u}, 1/a_{ji,m}, 1/a_{jl,l}\right) i, j = 1, 2, \ldots, k.$$

The fuzzy matrix $\tilde{\mathbf{A}}$ can be expressed by three characteristic crisp matrices $\tilde{\mathbf{A}} = (\mathbf{A}_l, \mathbf{A}_m, \mathbf{A}_u)$. Crisp matrices $\mathbf{A}_l, \mathbf{A}_m, \mathbf{A}_u$ are as follows:

$$\mathbf{A}_l = \begin{bmatrix} 1 & a_{12,l} \ldots a_{1k,l} \\ \frac{1}{a_{12,u}} & 1 \quad \ldots a_{2k,l} \\ \frac{1}{a_{1k,u}} & \frac{1}{a_{2k,u}} \ldots \quad 1 \end{bmatrix} \mathbf{A}_m = \begin{bmatrix} 1 & a_{12,m} \ldots a_{1k,m} \\ \frac{1}{a_{12,m}} & 1 \quad \ldots a_{2k,m} \\ \frac{1}{a_{1k,m}} & \frac{1}{a_{2k,m}} \ldots \quad 1 \end{bmatrix} \mathbf{A}_u = \begin{bmatrix} 1 & a_{12,u} \ldots a_{1k,u} \\ \frac{1}{a_{12,l}} & 1 \quad \ldots a_{2k,u} \\ \frac{1}{a_{1k,l}} & \frac{1}{a_{2k,l}} \ldots \quad 1 \end{bmatrix}.$$

As explained earlier, Saaty's AHP method is based on determining eigenvalues and eigenvectors at different levels of hierarchy. This means that in the case of the fuzzy AHP method, a system of homogeneous fuzzy linear equations should be solved:

$$\tilde{\mathbf{A}} \otimes \tilde{\mathbf{w}} = \tilde{\boldsymbol{\lambda}}_l \otimes \tilde{\mathbf{w}} \tag{4}$$

The calculation of the expected values of fuzzy numbers and their products provide the basis for the solution to this problem:

$$EV\left(\tilde{A}\right) = \frac{(a_l + 2a_m + a_u)}{4}.$$

$$EV(\tilde{A} \otimes \tilde{B}) = \frac{1}{12}[(2a_l + a_m)b_l + (a_l + 4a_m + a_u)b_m + (a_m + 2a_u)b_u] \tag{5}$$

It is possible to formulate a system of the fuzzy linear Eqs. (4) as follows:

$$\tilde{a}_{i1} \otimes \tilde{w}_1 \oplus \tilde{a}_{i2} \otimes \tilde{w}_2 \oplus \cdots \oplus \tilde{a}_{in} \otimes \tilde{w}_n = \tilde{\lambda} \otimes \tilde{w}_i, i = 1, 2, \ldots n$$

The fuzzy products' expected values as a result of (5) are

$$EV]\tilde{a}_{ij} \otimes \tilde{w}_{ij}) = \frac{1}{12}[]2a_{ij,l} + a_{ij,m})w_{j,l} +]a_{ij,l} + 4a_{ij,m} + a_{ij,u})w_{j,m} +]a_{ij,m} + 2a_{ij,u})w_{j,u}]$$
$$EV\left(\tilde{\lambda} \otimes \tilde{w}_i\right) = \frac{1}{12}\left[(2\lambda_l + \lambda_m)w_{i,l} + (\lambda_l + 4\lambda_m + \lambda_u)w_{i,m} + (\lambda_m + 2\lambda_u)w_{i,u}\right] i, j = 1, 2, \ldots, k$$

A system of fuzzy linear homogeneous equations is created by introducing formulas for expected values of the fuzzy products in fuzzy equations.

$$\overline{\boldsymbol{A}}_l \boldsymbol{w}_l + \overline{\boldsymbol{A}}_m \boldsymbol{w}_m + \overline{\boldsymbol{A}}_u \boldsymbol{w}_u - \overline{\lambda}_l \boldsymbol{w}_l - \overline{\lambda}_m \boldsymbol{w}_m - \overline{\lambda}_u \boldsymbol{w}_u = 0 \tag{6}$$

where are

$$\overline{\mathbf{A}}_l = 2\mathbf{A}_l + \mathbf{A}_m, \overline{\mathbf{A}}_m = \mathbf{A}_l + 4\mathbf{A}_m + \mathbf{A}_u, \overline{\mathbf{A}}_u = \mathbf{A}_m + 2\mathbf{A}_u$$

$$\overline{\lambda}_l = 2\lambda_l + \lambda_m, \overline{\lambda}_m = \lambda_l + 4\lambda_m + \lambda_u, \overline{\lambda}_u = \lambda_m + 2\lambda_u$$
$$\mathbf{w}_l = \left[w_{1,l}, w_{2,l}, \ldots, w_{n,l}\right]^T, \mathbf{w}_m = \left[w_{1,m}, w_{2,m}, \ldots, w_{n,m}\right]^T, \mathbf{w}_u = \left[w_{1,u}, w_{2,u}, \ldots, w_{n,u}\right]^T$$

Since all of the values in these equations are positive, it is possible to divide the system of Eqs. (6) into three systems, each of which represents a distinct eigenvalue problem:

$$\overline{\mathbf{A}}_l \mathbf{w}_l = \overline{\lambda}_l \mathbf{w}_l, \overline{\mathbf{A}}_m \mathbf{w}_m = \overline{\lambda}_m \mathbf{w}_m, \overline{\mathbf{A}}_u \mathbf{w}_u = \overline{\lambda}_u \mathbf{w}_u.$$

Eigenvectors $\mathbf{w}_l$, $\mathbf{w}_m$ and $\mathbf{w}_u$ and auxiliary eigenvalues $\overline{\lambda}_m$ and $\overline{\lambda}_u$ are obtained by resolving these three auxiliary eigenvalue issues. After then, the system of linear equations is solved to arrive at the requested eigenvalues, λ_l, λ_m and λ_u. The following condition $\overline{\mathbf{w}}_l \leq \overline{\mathbf{w}}_m \leq \overline{\mathbf{w}}_u$ needs to be met by normalized eigenvectors in order to satisfy the requirements for the principal eigenvalues $\lambda_I \leq \lambda_m \leq \lambda_u$.

3 Case Study

The proposed method was applied to the selection of construction machinery for the execution of concrete works on a large construction project for building a residential complex. The evaluation of alternatives and the comparison of the criteria with each other were carried out by a survey of three independent experts with many years of experience in the construction of investment projects. These experts agreed that the following criteria should be taken into account when narrowing the selection of construction machinery: total cost (C1), duration of the project (C2), availability of mechanization (C3), environmental impact (C4), and conditions of maintenance of mechanization (C5). Three alternatives (A1, A2, and A3) were considered, i.e., three different groups of machines for performing concrete work. The first two criteria are quantitative in nature, while the rest are qualitative. Qualitative criteria were reduced to quantitative grades using Saaty's scale, according to which a score of 1 is the lowest and a score of 9 is the best grade.

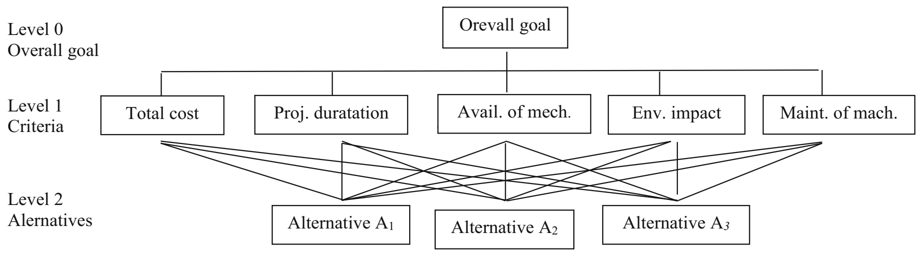


Fig. 2. Hierarchical decomposition of the model

For a defined problem, a decision hierarchy is formed (Fig. 2).

Pair-wise comparison the fuzzy matrix for criteria $\tilde{\mathbf{C}}$ and vector of criteria weights $\tilde{w}$ are

$$\tilde{\mathbf{C}} = \begin{bmatrix} (1,1,1) & (2.5,3,3.5) & (6.5,7,7.5) & (4.5,5,5.5) & (8.5,9,9.5) \\ \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (1,1,1) & (4.5,5,5.5) & (2.5,3,3.5) & (6.5,7,7.5) \\ \left(\frac{1}{7.5},\frac{1}{7},\frac{1}{6.5}\right) & \left(\frac{1}{5.5},\frac{1}{5},\frac{1}{4.5}\right) & (1,1,1) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (2.5,3,3.5) \\ \left(\frac{1}{5.5},\frac{1}{5},\frac{1}{4.5}\right) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (2.5,3,3.5) & (1,1,1) & (4.5,5,5.5) \\ \left(\frac{1}{9.5},\frac{1}{9},\frac{1}{8.5}\right) & \left(\frac{1}{7.5},\frac{1}{7},\frac{1}{6.5}\right) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & \left(\frac{1}{5.5},\frac{1}{5},\frac{1}{4.5}\right) & (1,1,1) \end{bmatrix} \tilde{w} = \begin{bmatrix} (0.467, 0.512, 0.562) \\ (0.237, 0.262, 0.289) \\ (0.058, 0.063, 0.070) \\ (0.117, 0.129, 0.143) \\ (0.031, 0.033, 0.036) \end{bmatrix}$$

Pair-wise comparison fuzzy matrices related to criteria are:

$$\tilde{\mathbf{A}}^{(1)} = \begin{bmatrix} (1,1,1) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) \\ (2.5,3,3.5) & (1,1,1) & (0.5,1,1.5) \\ (2.5,3,3.5) & \left(\frac{1}{1.5},1,\frac{1}{0.5}\right) & (1,1,1) \end{bmatrix} \tilde{\mathbf{A}}^{(2)} = \begin{bmatrix} (1,1,1) & (1,1,1) & (2.5,3,3.5) \\ \left(\frac{1}{1.5},1,\frac{1}{0.5}\right) & (1,1,1) & (2.5,3,3.5) \\ \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (1,1,1) \end{bmatrix}$$

$$\tilde{\mathbf{A}}^{(3)} = \begin{bmatrix} \left(\frac{1}{7.5},\frac{1}{7},\frac{1}{6.5}\right) & (1,1,1) & \left(\frac{1}{5.5},\frac{1}{5},\frac{1}{4.5}\right) \\ \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (4.5,5,5.5) & (1,1,1) \end{bmatrix} \tilde{\mathbf{A}}^{(4)} = \begin{bmatrix} \left(\frac{1}{5.5},\frac{1}{5},\frac{1}{4.5}\right) & (1,1,1) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) \\ \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (2.5,3,3.5) & (1,1,1) \end{bmatrix}$$

$$\tilde{\mathbf{A}}^{(5)} = \begin{bmatrix} (1,1,1) & (8.5,9,9.5) & (2.5,3,3.5) \\ \left(\frac{1}{9.5},\frac{1}{9},\frac{1}{8.5}\right) & (1,1,1) & \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) \\ \left(\frac{1}{3.5},\frac{1}{3},\frac{1}{2.5}\right) & (2.5,3,3.5) & (1,1,1) \end{bmatrix}$$

Priority fuzzy matrix $\tilde{\mathbf{P}}$ that contains alternatives' weights for each criteria is

$$P = \begin{bmatrix} (0.129, 0.143, 0.162) & (0.397, 0.438, 0.509) & (0.599, 0.649, 0.702) & (0.579, 0.636, 0.701) & (0.639, 0.692, 0.752 \\ (0.397, 0.438, 0.509) & (0.365, 0.419, 0.462) & (0.068, 0.072, 0.076) & (0.097, 0.105, 0.115) & (0.071, 0.077, 0.084) \\ (0.365, 0.419, 0.462) & (0.129, 0.143, 0.162) & (0.260, 0.280, 0.302) & (0.233, 0.289, 0.289) & (0.209, 0.231, 0.258) \end{bmatrix}$$

A vector of total weights $\tilde{g}$ was obtained by multiplying the priority fuzzy matrix $\tilde{\mathbf{P}}$ and the vector of criteria weights $\tilde{w}$:

Since the total weights are triangular fuzzy numbers their ranking can be performed according to generalized fuzzy mean i.e. expected value (7) and generalized fuzzy spread i.e. standard deviation (8)

$$g_{i,e} = \left] g_{i,l} + 2g_{i,m} + g_{i,u} \right) / 4 i = 1, 2..m \tag{7}$$

Table 1. Vector of total weights

Alt	$\tilde{g}$	Exp. Val.	Stand. Dev.	Coef. of variation
A1	(0.2763, 0.3344, 0.4145)	0.3399	0.0220	0.0647
A2	(0.2895, 0.3548, 0.4446)	0.3609	0.0247	0.0684
A3	(0.2499, 0.3108, 0.3780)	0.3124	0.0203	0.0649

$$\sigma = \left[\frac{1}{80}\left(3g_{i,l}^2 + 4g_{i,m}^2 + 3g_{i,u}^2 + 4g_{i,l}g_{i,m} \;\; 2g_{i,l}g_{i,u} \;\; 4g_{i,m}g_{i,u}\right)\right]^{1/2}, i = 1, ., m. \quad (8)$$

A fuzzy number with a higher expected value and a lower standard deviation is ranked better [12]. This means that the best ranked alternative is A2, and then the alternatives are A1 and A3. The ranking of fuzzy numbers can also be done according to the coefficient of variation [13]. A fuzzy number with a lower coefficient of variation ranks better than a number with a higher coefficient of variation. In that case, alternative A1 is the best ranked, and then there are alternatives A3 and A2.

4 Conclusions

Compared to the crisp AHP approach, the fuzzy AHP method allows for a more thorough and adaptable modeling of situations involving many criteria for decision-making. Comparisons of imprecise input factors for the selected criteria and alternatives can be integrated into the fuzzy AHP approach. It is not necessary to know the precise numerical values of the factors being examined here, in contrast to other MCDM methods. In this paper, the fuzzy AHP method is used to narrow down the list of construction equipment for concrete work. In addition to the typical evaluation of the total cost, other variables were taken into account when comparing alternatives. This paper provides a method for computing the eigenvalues and eigenvectors of the triangular fuzzy number based on the estimation of the expected values of fuzzy numbers and their products. A fuzzy eigenvalue problem with triangular fuzzy numbers is transformed into three auxiliary crisp eigenvalue problems. According to the expected value, standard deviation, and coefficient of variation, the alternative ranking is done. This method has been successfully used in the choice of equipment to complete concrete work on a substantial building project to build a residential complex.

References

1. Temiz, I., Calis, G.: Selection of construction equipment by using multi-criteria decision making methods. Procedia Engineering **196**, 286–293 (2017)
2. Waris M., et al.: An application of analytic hierarchy process (AHP) for sustainable procurement of construction equipment: Multicriteria-based decision framework for Malaysia. In: Mathematical Problems in Engineering 2019 (2019)

3. Petroutsatou, K, Ladopoulos, I., Nalmpantis, D.: Hierarchizing the criteria of construction equipment procurement decision using the AHP method. IEEE Trans. Eng. Manag., 1–12 (2021)
4. Shapira, A., Goldenberg, M.: AHP-based equipment selection model for construction projects. J. Construction Manag. Eng. **131**(12) (2005)
5. Ghorabaee, M.K., Amiri, M., Zavadskas, E.K., Antucheviciene, J.: A new hybrid fuzzy MCDM approach for evaluation of construction equipment with sustainability considerations. Arch. Civil Mecha. Eng. **18**(1), 32–49 (2017). https://doi.org/10.1016/j.acme.2017.04.011
6. Lashgari, A. et al.: Equipment selection using fuzzy multi criteria decision making model: key study of gole gohar iron min. Econ. Eng. Decisions **23**(2) (2012)
7. Yazdani-Chamzini, A.: An integrated fuzzy multi criteria group decision making model for handling equipment selection. J. Civ. Eng. Manag. **20**(5), 660–673 (2014)
8. Rahimi Ghazikalayeh, A., Amirafshari, M., Mkrchyan, H.M., Taji, M.: Application of fuzzy hybrid analytic network process in equipment selection of open-pit metal mines. Int. J. Res. Indus. Eng. **2**(3), 35–46 (2013)
9. Bozanic, D., Tešić, D., Marinković, D., Milić, A.: Modeling of neuro-fuzzy system as a support in decision-making processes. Reports Mech. Eng. **2**(1), 222–234 (2021)
10. Hanna, A.S., Lotfallah, W.B.: A fuzzy logic approach to the selection of cranes. Autom. Constr. **8**(5), 597–608 (1999)
11. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
12. Lee, E.S., Li, R.L.: Comparison of fuzzy numbers based on the probability measure of fuzzy events. Comput. Mathem. Appli. **15**, 887–896 (1988)
13. Cheng, C.-H.: A new approach for ranking fuzzy numbers by distance method. Fuzzy Sets Syst. **95**, 307–317 (1992)

Comparison of E-Commerce Sites with Pythagorean Fuzzy AHP and TOPSIS Methods

Beste Desticioglu Tasdemir[1](✉), Sena Kumcu[2], and Bahar Ozyoruk[2]

[1] Department of Operations Research, National Defence University Alparslan, Defence Sciences and National Defence Institute, Ankara, Turkey
bdesticioglu@kho.msu.edu.tr

[2] Department of Industrial Engineering, Faculty of Engineering, Gazi University, Ankara, Turkey

Abstract. Today, the rate of customers shopping on e-commerce sites is increasing day by day. Especially with the COVID-19 pandemic, which has entered our lives in recent days, a rapid trend towards e-commerce has occurred in the world and in Türkiye. For this reason, e-commerce sites need new marketing strategies and innovative practices in order to compete with each other. In order for e-commerce sites to be one step ahead of their competitors, the comments of customers who have shopped from that site before are of great importance. In this study, it is aimed to rank the e-commerce sites operating in Türkiye under the determined criteria. For this purpose, multi-criteria decision making (MCDM) methods, which are used to decide between alternatives in cases where there is more than one criterion, were used. In the study, first of all, the evaluation criteria of e-commerce sites were determined by interviewing consumers. Since the opinions of the consumers will vary from person to person, it is assumed that the information received from the consumers will be fuzzy and the weights of the criteria are determined with the Pythagorean Fuzzy Analytical Hierarchy Process (PFAHP). In the second part of the study, the criteria weights calculated with PFAHP and customer comments on e-commerce sites were calculated with the Pythagorean Fuzzy TOPSIS (PFTOPSIS) method and the ranking of e-commerce sites operating in Türkiye was made. These methods were computed by Microsoft Excel 2010. With the results obtained, the best e-commerce site was determined.

Keywords: E-commerce · FPAHP · FPTOPSIS · Multi Criteria Decision Making

1 Introduction

Online shopping platforms, the range of goods accessible, and quick internet connectivity have all increased steadily over the past ten years. Due to this evolution, both internet retail as a whole and consumer buying habits have undergone significant changes [1]. The ability to shopping at any time of the day, the easiness of comparing the quality

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 327–335, 2023.
https://doi.org/10.1007/978-3-031-39777-6_40

and prices of various brands and products, the potential to buy the same product at the online for less money than in the store and the possibility to shopping without getting tired or wasting time all contribute to an increasing focus on e-commerce in today's world where we live with an ongoing busy work routine. Therefore many researchers have examined into a number of indicators in this area that will improve the success of e-commerce sites by contributing to the growth of customer relationships. Some of these indicators in the literature can be given as follows; website design [2]; ease of use/useability [2–5]; security [2, 3, 5–7]; delivery time [2, 8]; fulfillment [2, 8, 9]; responsiveness [9]; information quality [2, 10]; customer support and services/customer relationship [3, 6, 11]; promotion activities [12, 13]; return policy [12, 14]; satisfactory shipping fees [15, 16]; contact [17]. In this study, five well-known e-commerce sites in Türkiye were assessed in terms of how well they met user expectations by utilizing both criteria drawn from the literature. Since the opinions of the consumers will vary from person to person, it is assumed that the information received from the consumers will be fuzzy and the weights of the criteria are determined with the PFAHP. In the second part of the study, the criteria weights calculated with PFAHP and customer comments on e-commerce sites were calculated with the PFTOPSIS method and the ranking of e-commerce sites operating in Türkiye has made. These methods were computed by Microsoft Excel 2010. With the results obtained, the best e-commerce site has been determined. As far as is known, no study has addressed with hybrid PFAHP-PFTOPSIS method in evaluating e-commerce websites. This study is regarded to be the first study in this field by using this method as a result. Assessing e-commerce websites evaluation can be considered as a MCDM problem because of its complex structure including qualitative and quantitative factors [18]. In this direction, studies evaluating e-commerce sites with the MCDM approach in the literature were searched in the SCOPUS and WOS database and summarized in Table 1. As a result of the literature review, the distinction between this study and the others is the use of Pythagorean fuzzy numbers (PFN), one of the expansions of fuzzy set theory, which provides a broad assessment scale, especially the subjectivity and uncertainty of human judgments.

The remainder of this paper is constructed as follows: Sect. 2 proposes the hybrid PFAHP-PFTOPSIS method and a case study of ranking e-commerce web sites in Sect. 3. Section 4 concludes the paper.

2 Methodology

In this section, the pythagorean fuzzy set (PFS) theory and proposed methodology is explained. The steps of the PFS theory have been applied [21]. Then, the steps for the proposed hybrid PFAHP-PFTOPSIS are described in Table 2.

Table 1. Literature review on evaluation of e-commerce sites with MCDM approach

Author	Study Focus	Method
[8]	Evaluating an uncertinity in multicriteria decision making (MCDM) in the framework of e-commerce website recommendation	Choquet integral-Location and uncertainty indicators
[5]	Proposing an evaluation model based on AHP, fuzzy sets and TOPSIS is proposed to tackle quality of e-commerce in e-alliance	AHP-FTOPSIS
[3]	Proposing a new methodology for evaluating three e-commerce websites, which are the most famous in Türkiye	FAHP-VIKOR
[2]	Proposing an E-commerce website evaluation model consisting hybrid MCDM techniques, specially designed for Indian market	AHP-FTOPSIS
[9]	To evaluate as well as to compare the service quality of e-commerce website	Fuzzy analytic hierarchy process (FAHP)-TOPSIS
[10]	Providing an accurate classification of E-commerce web site (EWS) performance	IFNAHP-IFNTOPSIS (IFN: Intuitionistic Fuzzy Number)
[4]	Ranking E-commerce design platform based on fuzzy theory	Fuzzy VIKOR
[19]	Ranking the top B2C e-Commerce brands of India as per the perceived risks of respondents	AHP
[6]	Assessing and prioritizing factors for designing a successful business-to-consumer (B2C) e-commerce website	FAHP-TOPSIS-Grey
[11]	Proposing an approach for evaluating six e-commerce websites	SMARTS-MULTIMOORA-IVNS
[7]	Proposing an approach to rank and assess E-commerce web sites	Fuzzy Analytic Network Process - FTOPSIS
[20]	Proposing to use fuzzy logic to allow clients to express their ratings in natural language and propose an approach for E-commerce websites ranking	FuzzyTOPSIS

Table 2. Proposed of hybrid PFAHP-PFTOPSIS Method

PFAHP	PFTOPSIS
Step 1. The matrix of binary comparisons is generated using the expert evaluations provided in accordance with the language terms scale according to [23]	**Step 1:** Alternative evaluations of a decision-making group consisting of X experts according to each criteria are made by language terms scale [24]
Step 2. Creating a pairwaise comparation matrix	**Step 2:** Construct Pythagorean fuzzy decision matrix
Step 3. Calculating the difference matrices	**Step 3:** Calculating the Pythagorean fuzzy positive ideal solution (PIS) and negative ideal solution (NIS).
Step 4. Generating the interval multiplicative matrix	**Step 4:** Computing the distance between the ideal solution and PF-PIS and PF-NIS.
Step 5. Calculating the determinacy value	**Step 5.** Calculating the revised closeness indices of each alternatives
Step 6. Obtaining the matrix of weights	**Step 6.** The algorithm is terminated by determining of the best order of alternatives
Step 7. Calculating the normalized priority weigts	
Validation Analysis: Sensitivity Analysis	

3 Case Study

The study consists of 4 stages. In the study, firstly, the studies comparing e-commerce sites in the literature were examined and the evaluation criteria were determined. These criteria can be describe as; Satisfactory shipping fees (CR1);it means that providing satisfactory shipping-free or complimentary shipping service to customers; easy exchange or return (CR2), it means that has the availability of product exchange or return policy; contact (CR3), it refers to allowing customers to speak to a live customer service representative online or by phone; reliability (CR4), it refers to enabling customers to buy from reliable vendors and make safe payments responsiveness (CR5), it refers to have the ability to respond quickly to customer requests and support the customer through the one-to-one live customer support system; website design (CR6), it means that has an easy-to-use design when customers navigate the site; Promotion activities (CR7), it refers to offering sales and discounts to online shoppers to get new customers and increase revenue.

In the second stage, experts working in e-commerce sites were interviewed and they were asked to compare the determined criteria with the pairwise comparison matrix. Afterwards, the comparison results converted to PFNs were calculated by the PFAHP method and the criteria weights were obtained. In the third part of the study, it is aimed to compare 5 e-commerce sites selling durable consumer goods in Türkiye. For privacy reasons, the names of e-commerce sites are not included in the study, and these alternative sites were shown in the study as EC1, EC2…, EC5. These 5 alternative e-commerce sites were determined among the sites with the most purchases by customers. The comparison criteria determined in the literature were compared with the pairwise comparison matrix using linguistic scale by 10 expert personnel working in e-commerce sites [22].

Table 3. Pairwise comparison matrix about e-commerce site selection criteria

	CR1	CR2	CR3	CR4	CR5	CR6	CR7
CR1	{[0.197, 0.197], [0.197, 0.197]}	{[0.33, 0.14], [0.55, 0.64]}	{[0.59, 0.42], [0.32, 0.39]}	{[0.17, 0.23], [0.74, 0.86]}	{[0.55, 0.65], [0.36, 0.46]}	{[0.56, 0.68], [0.29, 0.38]}	{[0.52, 0.35], [0.39, 0.46]}
CR2	{[0.56, 0.64], [0.33, 0.14]}	{[0.197, 0.197], [0.197, 0.197]}	{[0.59, 0.46], [0.28, 0.38]}	{[0.32, 0.39], [0.59, 0.42]}	{[0.62, 0.74], [0.27, 0.38]}	{[0.45, 0.56], [0.44, 0.56]}	{[0.46, 0.56], [0.38, 0.52]}
CR3	{[0.32, 0.39], [0.59, 0.42]}	{[0.28, 0.38], [0.59, 0.46]}	{[0.197, 0.197], [0.197, 0.197]}	{[0.28, 0.38], [0.59, 0.42]}	{[0.38, 0.46], [0.51, 0.34]}	{[0.35, 0.42], [0.55, 0.39]}	{[0.34, 0.42], [0.55, 0.40]}
CR4	{[0.74, 0.86], [0.17, 0.23]}	{[0.59, 0.42], [0.32, 0.39]}	{[0.63, 0.73], [0.28, 0.38]}	{[0.197, 0.197], [0.197, 0.197]}	{[0.60, 0.70], [0.31, 0.41]}	{[0.56, 0.63], [0.34, 0.17]}	{[0.46, 0.56], [0.41, 0.55]}
CR5	{[0.57, 0.67], [0.34, 0.44]}	{[0.27, 0.38], [0.62, 0.74]}	{[0.52, 0.36], [0.38, 0.46]}	{[0.31, 0.41], [0.60, 0.70]}	{[0.197, 0.197], [0.197, 0.197]}	{[0.55, 0.40], [0.34, 0.42]}	{[0.41, 0.55], [0.46, 0.59]}
CR6	{[0.29, 0.38], [0.56, 0.68]}	{[0.44, 0.56], [0.45, 0.56]}	{[0.55, 0.39], [0.35, 0.42]}	{[0.34, 0.17], [0.56, 0.63]}	{[0.34, 0.42], [0.55, 0.40]}	{[0.197, 0.197], [0.197, 0.197]}	{[0.34, 0.44], [0.56, 0.66]}
CR7	{[0.39, 0.46], [0.52, 0.35]}	{[0.38, 0.52], [0.46, 0.56]}	{[0.55, 0.40], [0.34, 0.42]}	{[0.41, 0.55], [0.46, 0.56]}	{[0.46, 0.59], [0.41, 0.55]}	{[0.56, 0.66], [0.34, 0.44]}	{[0.197, 0.197], [0.197, 0.197]}

The values given in Table 3 were formed by taking the averages of the scores converted into PFNs. By using these scores in the PFAHP method, the weights of the criteria used in the comparison of e-commerce sites were calculated.

As a result of the calculations made with the PFAHP method, the criteria weights were found as shown in Fig. 1. When Fig. 1 is examined, it is seen that the most important criterion in the comparison of e-commerce sites is CR4 with 0.2923. It is seen that customers who shop from e-commerce sites are the first to shop from reliable vendors. The order of importance of other criteria is seen in Fig. 1. In the third part of the study, 20 people shopping from e-commerce sites were interviewed and they were asked to rate 5 e-commerce companies according to the determined criteria. E-commerce sites were ranked with the PFTOPSIS method by using customer ratings and criterion weights found with PFAHP. In order to compare the e-commerce sites, 20 customers who shopped from these sites were interviewed and the customers were asked to rate the sites according to the determined criteria. The decision matrix created with PFNs according to linguistic scale [23] is given in Table 4. After the scores were converted to PFNs, the necessary calculations were made by applying the steps of the PFTOPSIS method and the ranking the e-commerce site given in Table 5 were obtained.

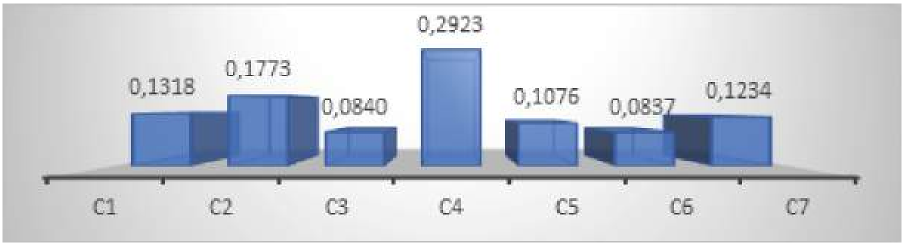


Fig. 1. e-commerce evaluation criteria weights

Table 4. Decision matrix for sorting the e-commerce site

DM	CR1		CR2		CR3		CR4		CR5		CR6		CR7	
	u	*v*	*u*	*v*	*u*	*v*	*u*	*v*	*u*	*v*	*u*	*v*	*u*	*v*
EC1	0.55	0.69	0.87	0.13	0.84	0.18	0.77	0.49	0.80	0.44	0.87	0.13	0.87	0.13
EC2	0.60	0.70	0.58	0.40	0.49	0.70	0.58	0.63	0.52	0.57	0.56	0.71	0.46	0.74
EC3	0.38	0.85	0.40	0.81	0.49	0.70	0.34	0.82	0.19	0.94	0.49	0.80	0.43	0.76
EC4	0.78	0.24	0.69	0.45	0.84	0.26	0.52	0.71	0.64	0.58	0.71	0.57	0.78	0.36
EC5	0.58	0.71	0.87	0.13	0.19	0.94	0.84	0.26	0.71	0.55	0.83	0.31	0.64	0.53

When Table 5 is examined, it is seen that EC1 is the best e-commerce site according to customers' preferences. This shows that EC1 has successful results for all 7 criteria determined in measuring the performance of e-commerce sites. The ranking of other e-commerce sites is as seen in Table 5..

In the last part of the study, sensitivity analysis was performed with the binary change of criterion weights. As the rate of shopping from e-commerce sites has increased in recent years, companies aim to make their e-commerce sites flexible to adapt to changing

Table 5. Ranking of e-commerce site alternatives

Alternatives	$\xi(X_j)$	Rank
EC1	0.0000	1
EC2	−3.5464	4
EC3	−5.7796	5
EC4	−1.6452	3
EC5	−1.0164	2

situations. Fuzzy decision models were used in this study to measure the variability of customer demands and expectations. In order to determine the effect of changing conditions on the decision model over time, a sensitivity analysis with binary changes of criterion weights is performed in this Section 21 different scenarios were created by the binary change of the criterion weights and the rankings of the e-commerce sites were determined by the calculations made. Figure 2 was formed with the rankings obtained from the calculations made with the binary changes of the criterion weights.

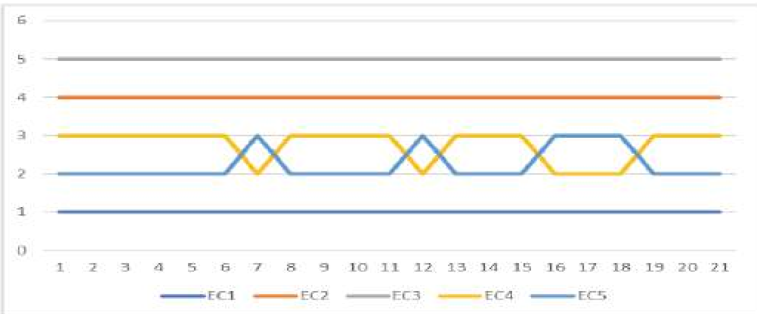


Fig. 2. Sensitivity analysis of the criteria

When Fig. 2 is examined, it is seen that EC1 is the most e-commerce site, even if the criteria weights change. This is related to the success of the EC1 site in all of the specified criteria. In addition, it is seen that the rankings of EC5 and EC4 e-commerce sites, which are in the 2nd and 3rd rank, have changed with the change of criterion weights. In addition, in the calculations made with PFTOPSIS for the scenarios, it is seen that the closeness values of the two alternatives are close to each other. Therefore, in the binary exchanges of the C4 criterion, which has the highest weight, with other criteria, the order of the alternatives has changed and the EC4 alternative has taken the 2nd place. Moreover, in order to increase the rate of preference of other e-commerce sites by customers, they must first increase their performance by starting with the criteria with the highest criterion weight.

4 Conclusion

Today, many people prefer to shop from e-commerce sites because they are faster and more economical. In this study, 5 e-commerce sites selling durable consumer goods in Türkiye were compared. The weights of the criteria determined in the comparison of e-commerce sites were calculated by the PFAHP method and the obtained criteria weights are given in Fig. 1. According to the calculations, it is seen that the criterion with the highest weight is "C4". The fact that the weight of this criterion is higher than the weights of other criteria shows that customers will first consider this criterion in determining the e-commerce sites they will shop. In addition, e-commerce sites need to improve these criteria in order of importance in order to increase the number of customers who shop. PFTOPSIS method was used in ordering the e-commerce sites. EC1 is the best e-commerce site according to customers' preferences. This shows that EC1 has successful results for all 7 criteria determined in measuring the performance of e-commerce sites. In the last part of the study, sensitivity analysis was performed with the binary change of criterion weights. In the calculations made, it was determined that e-commerce site with the highest performance. This is related to the success of the EC1 site in all of the specified criteria. In this study, e-commerce sites selling durable goods were ranked using PFAHP-PFTOPSIS methods. In future studies, e-commerce sites can be ranked by using different fuzzy sets or different MCDM methods. In addition, e-commerce sites can be compared by changing the criteria.

References

1. Le, H.T.K., Carrel, A.L., Shah, H.: Impacts of online shopping on travel demand: a systematic review. Transp Rev. **42**, 273–295 (2021)
2. Dey, S., Jana, B., Gourisaria, M.K., Mohanty, S.N., Chatterjee, R.: Evaluation of Indian B2C E-shopping websites under multi criteria decision-making using fuzzy hybrid technique. Int. J. Appl. Eng. Res. **10**, 24551–24580 (2015)
3. Aydin, S., Kahraman, C.: Evaluation of E-commerce website quality using fuzzy multi-criteria decision making approach. IAENG Int. J. Comput. Sci. **39**, 64–70 (2012)
4. Nguyen, P.A.H.: Evaluating e-commerce design platforms by a fuzzy vikor approach. In: Proceedings of 2019 International Conference on System Science and Engineering, ICSSE 2019, pp. 397–400 (2019)
5. Yu, X., Guo, S., Guo, J., Huang, X.: Rank B2C e-commerce websites in e-alliance based on AHP and fuzzy TOPSIS. Expert Syst Appl. **38**, 3550–3557 (2011)
6. Li, R., Sun, T.: Assessing factors for designing a successful B2C E-Commerce website using fuzzy AHP and TOPSIS-Grey methodology. Symmetry (Basel) **12** (2020)
7. Rekik, R.: An integrated fuzzy anp-topsis approach to rank and assess e-commerce web sites. AISC. vol. 1179, pp. 197–209 (2021). https://doi.org/10.1007/978-3-030-49336-3_20
8. Denguir-Rekik, A., Mauris, G., Montmain, J.: Propagation of uncertainty by the possibility theory in choquet integral-based decision making: Application to an E-commerce website choice support. IEEE Trans. Instrum. Meas. **55**, 721–728 (2006)
9. Wijayanti, W.R., Fajri, A., Dewi, W.R., Ulkhaq, M.M., Ardi, F., Akshinta, P.Y.: Combining the fuzzy AHP and TOPSIS to evaluate service quality of E-commerce website. In: ACM International Conference Proceeding Series, pp. 397–402 (2018)

10. Rouyendegh, B.D., Topuz, K., Dag, A., Oztekin, A.: An AHP-IFT integrated model for performance evaluation of E-commerce web sites. Inf. Syst. Front. **21**(6), 1345–1355 (2018). https://doi.org/10.1007/s10796-018-9825-z
11. Zavadskas, E.K., Bausys, R., Lescauskiene, I., Usovaite, A.: Multimoora under interval-valued neutrosophic sets as the basis for the quantitative heuristic evaluation methodology hebin. Mathematics **9**, 1–19 (2021)
12. Wang, M.: SA Huarng: An empirical study of internet store customer post-shopping satisfaction. Special Issues Inform. Syst. **3**, 632–638 (2002)
13. Agarwal, R., Venkatesh, V.: Assessing a firm's Web presence: A heuristic evaluation procedure for the measurement of usability. Inf. Syst. Res. **13**, 168–186 (2002)
14. Wang, M.: Assessment of E-service quality via E-satisfaction in E-commerce globalization. Elect. J. Inform. Syst. **11**, 1–4 (2003)
15. Lantz, B., Hjort, K.: Real e-customer behavioural responses to free delivery and free returns. Electron. Commer. Res. **13**, 183–198 (2013)
16. Lewis, M.: The effect of shipping fees on customer acquisition, customer retention, and purchase quantities. J. Retail. **82**, 13–23 (2006)
17. Wang, M.: Assessment of E-Service quality via E-satisfaction in E-commerce globalization. Electron. J. Inform. Syst. Devel. Countries. **11**, 1–4 (2003)
18. Zaim, H., Ramdani, M., Haddi, A.: Multi-criteria analysis approach based on consumer satisfaction to rank B2C E-commerce websites. In: SITA 2016 - 11th International Conference on Intelligent Systems: Theories and Applications (2016)
19. Guru, S., Nenavani, J., Patel, V., Bhatt, N.: Ranking of perceived risks in online shopping. Decision: Official J. Indian Inst. Manag. Calcutta **47**(2), 137–152 (2020). https://doi.org/10.1007/s40622-020-00241-x
20. Belouaar, H., Kazar, O., Zouai, M., Merizig, A.: A new ranking approach for E-commerce websites based on fuzzy TOPSIS algorithm. Bull. Elect. Eng. Inform.. **11**, 936–946 (2022)
21. Zhang, X., Xu, Z.: Systems of undefined: Extension of TOPSIS to multiple criteria decision making with Pythagorean fuzzy sets. Wiley Online Library **29**, 1061–1078 (2014)
22. Ilbahar, E., Karaşan, A., Cebi, S., Kahraman, C.: A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system. Saf Sci. **103**, 124–136 (2018)
23. Yucesan, M., Gul, M.: Hospital service quality evaluation: an integrated model based on Pythagorean fuzzy AHP and fuzzy TOPSIS. Soft. Comput. **24**(5), 3237–3255 (2019). https://doi.org/10.1007/s00500-019-04084-2

An Application for Evaluating Proficiency and Expert Knowledge Criteria of Information Technologies (IT) Workers with Using Fuzzy Analytic Hierarchy Process

Ali Tunç[1](✉), Sakir Taşdemir[2], Mehmet Taha Uyar[1], and Veysel Can Bingöl[1]

[1] Architecht Information Systems, Istanbul, Turkey
ali.tunc@architecht.com

[2] Computer Engineering Department, Technology Faculty, Selcuk University, Konya, Turkey

Abstract. The information technologies (IT) sector has been an area of study whose significance has grown exponentially, taking into account the effects of growing technology and the importance of technology in recent years, taking into account the effects of the pandemic. The widespread use of remote working and hybrid working models in these sectors reveals the importance of making more careful performance evaluations of employees. Human resources are essential in the upkeep, renewal, and development of existing software products, as well as the creation of new ones. Infrastructures and system algorithms created with human impact as the most important resource can be finished on time for project success and in a manner that addresses future needs. This study, it is aimed to rank and monitor the performances of software engineers and analysts working in the software development sector by measuring both their technical skills and personal competencies. The most important items accepted as evaluation criteria in the sector were weighted by experts in their fields using the Fuzzy Analytic Hierarchy Process (Fuzzy AHP) method. It was ensured that these weighted evaluation items had been scored for the people to be measured and results were obtained. In order to evaluate the employees in the roles of analyst and software developer in terms of both talent and technical knowledge, the evaluation of the criteria of the experts with Fuzzy AHP and the weighting of the criteria are shown in tables in the study. Then, newly hired people and interns to be recruited were evaluated by their mentors, and a method that could be useful for tracking their training and development experiences was provided.

Keywords: Competency · Fuzzy AHP · IT Personnel Evaluation · Technical Knowledge

1 Introduction

In today's digital age, organizations are heavily reliant on their information technology (IT) infrastructure, making IT personnel evaluation a crucial aspect of organizational success. Accurately assessing the competencies of IT personnel is essential to ensuring

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 336–343, 2023.
https://doi.org/10.1007/978-3-031-39777-6_41

high-quality IT services and maximizing organizational performance. One critical aspect of IT personnel evaluation is the assessment of both ability and technical knowledge, particularly in roles such as analysts and software developers. In this context, to achieve this, the weighting part of the evaluation process involved assessing the experts' criteria with Fuzzy AHP in this study. This ensured that the workers were assessed based on both their abilities and technical knowledge, allowing for a more comprehensive and accurate evaluation. Following this, newly hired personnel and interns were evaluated by their mentors, who mapped out a useful route for their training and growth experiences. By doing so, the organization can ensure that its IT personnel are not only competent in their roles but are also continuously improving their skills and knowledge to keep up with the rapidly evolving technological landscape. With the ongoing development of new technologies, this approach can help organizations stay ahead of the curve and ensure their IT personnel is equipped with the necessary skills to drive organizational success. However, traditional evaluation methods may not accurately capture the complexity of IT personnel competencies and may not provide an accurate representation of their abilities. To address this challenge, the study introduced the application of Fuzzy AHP as a powerful tool for evaluating IT personnel competencies. The objective was to design a decision support system that considers both ability and technical knowledge evaluations, particularly in roles such as analysts and software developers. By incorporating Fuzzy AHP, the study provided a comprehensive and accurate evaluation approach that captures the complexity of IT personnel competencies. Additionally, the study introduced the concept of mentor evaluation for newly hired personnel and interns, aiming to guide their training and growth experiences. This approach ensures that IT personnel are not only competent in their roles but also continuously improve their skills and knowledge to keep up with the rapidly evolving technological landscape. The originality of the study lies in its innovative approach to accurately assessing competencies and fostering continuous growth within organizations.

2 Evaluating Proficiency and Expert Knowledge Criteria of Information Technologies (IT)

With the developing technology in recent years, the Information Technologies (IT) sector has become a field of study that increasingly shows its importance. For this reason, personnel performance management has become the focus of today's businesses. A well-designed and implemented performance management system maximizes the efficiency of the company [1]. The Evaluation of Competency and Expertise Knowledge Criteria of Information Technologies Employees is an evaluation study with Fuzzy AHP, based on the criteria of competence and expertise in terms of both talent and technical knowledge, of analysts and software developers working in this field and intertwined with technology production. Again, newly hired people and interns to be recruited will be evaluated by their mentors with the criteria of competence and expertise, thus ensuring a healthier development and recruitment process. Evaluation of employee performance is an important element in increasing the quality of the work produced and contributes to increasing the motivation of employees to perform. The performance evaluation system helps the personnel see their targets more clearly. With this system, the personnel's self-development by recognizing their strengths and weaknesses by providing feedback comes to the forefront as a contribution to both their personal development and the efficiency of the enterprise. In addition, an organization's periodic performance evaluation of employees will also increase staff and corporate loyalty [1]. An employee's rating is based on their contribution to the success of the organization. Having an accurate and impartial evaluation system is important in terms of accurately measuring the contribution of employees to the achievement of the organization's goals. Fuzzy Logic was used in the evaluation criteria as it contains qualitative and quantitative data. During the formation of these performance determination criteria, a literature review and experts in this sector were used. The implementation and validation of the proposed model were carried out in the information technology department of a private bank operating in Turkey [1]. The pick of the most qualified employees helps to improve the company's success. On the other hand, if the selection process is flawed, the work will suffer and the cost of replacing this scenario will be expensive. To achieve the company's goals, each candidate's credentials, abilities, expertise, and overall attitudes must be evaluated [2]. Managers face numerous choices that necessitate the joint use of various kinds of data in their decision-making process. Managers must make critical decisions regarding staff success, whether as individuals or as members of a team. Performance evaluation is important for the efficient administration of an organization's human resources and staff evaluation, which helps grow people, enhance organizational performance, and feed into business planning [3]. Finally, in the area of information technology, assessing proficiency and specialist knowledge criteria is critical. (IT). Implementing a well-designed performance management system that includes tools like Fuzzy AHP and periodic staff performance assessments can considerably boost a company's productivity.

3 Literature View

Numerous studies were conducted globally on personnel evaluation with Fuzzy AHP when the literature was examined. (Wu and Fang, 2011) worked to combine the Fuzzy AHP and the Fuzzy Delphi method for developing critical competencies of electronic commerce professional managers. Their study proposes a novel approach that combines qualitative, quantitative, and fuzzy set theory to develop a framework of critical managerial competencies for electronic commerce (EC) professional managers [4]. (Faisal et al., 2022) worked on a novel framework for leadership competencies development in megaprojects. Their study proposes a framework to categorize and prioritize leadership competencies for project managers in megaprojects [5]. (Dink et al., 2019) worked on using the Fuzzy AHP method to empirically study the key competency and capabilities affecting the selection of middle managers for global shipping logistics service providers (GSLSPs) [6]. (Akyol and Güler, 2017) worked on the role of competencies in the employee selection function. Their study discusses how the competency approach has become an essential point in HRM functions, particularly in employee selection. Their study presents a two-hierarchical model to help decision-makers choose the best project manager alternative based on their competence, with Achievement and Action and Impact and Influence being the essential criteria with higher priority weights [7]. (Samanlioglu et al., 2018) discusses the personnel selection process in a Turkish dairy company's IT department as a group multicriteria decision-making problem and presents an approach that integrates the Fuzzy AHP with Chang's extent analysis and fuzzy The Technique for Order of Preference by Similarity to Ideal Solution (Fuzzy TOPSIS). Their approach takes into account decision makers' verbal evaluations using intuitionistic fuzzy numbers and hierarchical level weights and ranks five IT personnel alternatives based on the importance weights of thirty subcriteria determined by Fuzzy AHP [8]. (Lee et al., 2008) proposes an approach based on the FAHP and BSC for evaluating the performance of the IT department in the manufacturing industry in Taiwan. The authors propose a Fuzzy AHP and balanced scorecard (BSC) approach as an effective evaluation method to measure the IT department's performance [9].

4 Method

In this study, it was tried to reach the results by using two main methods. The fuzzy logic algorithm [10] and the Fuzzy AHP [11] which has a large usage area in the literature, are among these approaches. Fuzzy logic is a mathematical framework that allows for the representation and manipulation of uncertain and vague information. Fuzzy AHP is a decision-making technique that employs fuzzy logic to address uncertainty and imprecision in decision-making procedures. These methods are widely used in decision-making processes and have been applied to many fields, including logistics and transportation.

4.1 Fuzzy Logic

The concept of Fuzzy Logic was first introduced by Lotfi A. Zadeh in an article titled "The Theory of Fuzzy Logic and Fuzzy Sets" published in 1965. However, it was used as a more frequently used term in the 1970s. It became much more popular in the 1980s when the Japanese used fuzzy logic principles in their production. Fuzzy Logic applications also expressed as fuzzy ravioli, consist of computer-assisted artificial intelligence-oriented applications that reason in a way that imitates human behavior and the functioning of nature. It enables the grading of imprecise sentences produced by humans and used in everyday language. For example, ambiguous sentences such as "somewhat cold", "almost wrong", and "too slow" cannot be expressed numerically, although they help solve problems. In such cases, the fuzzy logic algorithm imitates the human mind and produces ambiguous analyzes and numerical models.

4.2 Fuzzy AHP

The Fuzzy AHP is a decision-making technique that employs fuzzy logic to address uncertainty and imprecision in decision-making procedures. It is an AHP technique extension that provides for more flexible and precise decision-making, especially when the available data is partial or ambiguous. The input data in Fuzzy AHP are stated as linguistic words rather than numeric values, making it more appropriate for assessing complicated and subjective criteria. The technique entails breaking down complicated choices into a hierarchical structure of criteria and sub-criteria, followed by pairwise comparisons to determine the relative significance of each criterion in relation to others at the same level. The imprecision and ambiguity in the data are then handled using fuzzy logic, allowing for more precise and flexible decision-making [12].

5 Study and Results

This study aimed to evaluate the competency and expertise knowledge criteria of employees in Information Technologies, particularly those in roles of analysts and software developers, using Fuzzy AHP. The assessment considered technical knowledge, such as programming languages and software testing, as well as talent evaluation criteria like problem-solving and communication skills. Additionally, the study assessed new hires and interns with the same criteria to enhance the recruitment and development process. The results provided insights to improve employee performance, leading to a more productive and competitive workforce. To determine the weights of the criteria mentioned, various weighting methods were reviewed from the literature. Fuzzy AHP was found to be the most commonly used method, so it was utilized to weigh the criteria in this study. The five most relevant criteria for the calendar feature were selected, and seven expert analysts and software developers with at least five to ten years of experience provided their opinions using pairwise comparisons with linguistic variables. Buckley's [13] geometric mean approach was used to integrate the opinions of the evaluators. This method was applied to weigh the outcomes of group evaluations based on the criteria. Ozkan's study [14] definitions and procedures were used during the cascade phase of the process. The fuzzy comparison was conducted using a 9-point scale, which is shown in Table 1.

Table 1. Fuzzy comparison scale.

Order	Linguistic Scale	Triangular Fuzzy-Scale	Triangular Fuzzy Reciprocal Scale
1	Equally Important	(1,1,1)	(1,1,1)
2	Very Little Important	(1,2,3)	(1/3,1/2,1)
3	Less Important	(2,3,4)	(1/4,1/3,1/2)
4	Important	(3,4,5)	(1/5,1/4,1/3)
5	High Significant	(4,5,6)	(1/6,1/5,1/4)
6	Very important	(5,6,7)	(1/7,1/6,1/5)
7	Very High Important	(6,7,8)	(1/8,1/7,1/6)
8	Strong Important	(7,8,9)	(1/9,1/8,1/7)
9	Very Strong Important	(8,9,10)	(1/10,1/9,1/8)

The assessment criteria for analysts are chosen to ensure that they have the necessary skills and knowledge to successfully perform their job functions. SQL knowledge is important for analysts who work with databases, as they need to be able to write queries and understand how to optimize database performance. Analysis definition information is necessary for analysts to understand the business requirements and specifications for the project they are working on. Test case creation information is important for analysts to ensure that the software functions as intended and meets the required standards. The performance impact analysis helps analysts understand how changes to the software may affect its performance. By assessing analysts based on these criteria, organizations can ensure that they are hiring individuals with the necessary skills and knowledge to perform their job functions effectively, leading to successful software development projects. Table 2 displays the assessment and weighting results for the analysts. According to the results, “Analysis Definition Information” had the highest weight value, and “Perform Impact Analysis” had the lowest weight value.

Table 2. Criteria assessment weights for analysts.

Criteria	Weights
BOA Base Definition Information	20.3%
SQL Knowledge	19.9%
Analysis Definition Information	21.8%
Test Case Creation Information	19.4%
Perform Impact Analysis	18.5%

Code quality, performance impact analysis, and follow-up development packages are important assessment criteria for software developers. Code quality ensures efficient, reliable, and maintainable code, while performance impact analysis helps developers understand how changes may affect system performance. Follow-up development packages ensure developers have the necessary skills to use tools effectively. These criteria help organizations hire skilled developers for successful software development. Table 3 displays the assessment and weighting results for the software developers. According to the results, “React, WPF, C# Knowledge” had the highest weight value, and “Follow-up of Development Packages.” had the lowest weight value.

Table 3. Criteria assessment weights for software developers.

Criteria	Weights
Code Quality	17.1%
Perform Impact Analysis	16.7%
SQL Knowledge	17.0%
Follow-up of Development Packages	15.3%
React, WPF, C# Knowledge	17.3%
Mastery of BOA Architecture	16.6%

The assessment criteria for competencies are chosen to ensure that individuals possess the necessary skills and traits to be successful in their roles. Problem-solving is important for identifying and resolving issues that may arise in software development projects. Communication and negotiation skills are crucial for effective collaboration with team members, stakeholders, and clients. Business and time management skills are necessary for meeting project deadlines and ensuring that the project remains within budget. Adaptability to teamwork is essential for contributing effectively in team environments. Self-motivation is important for driving individuals to achieve their goals and meet project objectives. Requirements management and comprehension awareness ability ensure that individuals can understand project requirements and specifications, and build solutions that meet these requirements. Finally, the ability to build algorithms is essential for developing efficient and effective software solutions. By assessing individuals based on these competencies, organizations can ensure that they hire individuals with the necessary skills and traits to be successful in their roles. Table 4 displays the assessment and weighting results for the competencies. According to the results, "Problem-Solving" had the highest weight value, while "Adaptability to Teamwork" and "Self-Motivation" had the lowest weight values. Fuzzy AHP was utilized to determine the weight values.

Table 4. Criteria assessment weights for competencies.

Criteria	Weights
Problem Solving	13.0%
Communication/Negotiation	12.4%
Business and Time Management	12.5%
Adapt to Teamwork	12.0%
Self-Motivation	12.0%
Requirements Management	12.7%
Comprehension Awareness Ability	12.8%
Ability to Build Algorithms	12.6%

6 Conclusion and Future Works

In conclusion, as the IT sector continues to grow and evolve, the importance of careful performance evaluations of employees becomes increasingly crucial. Human resources are vital in the development, upkeep, and renewal of software products, and it is essential to have skilled and knowledgeable individuals in analyst and software developer roles. The use of Fuzzy AHP in this study provided a valuable tool for assessing employees' abilities and technical knowledge, and the assessment of newly hired employees and interns provided a useful roadmap for their training and growth experiences. In the future, further research could explore the application of these assessment criteria to other roles within the IT sector and the impact of ongoing training and development on employee performance. Additionally, this study was tested on a sample of 22 trainee students and 13 new employees, with promising results in terms of identifying areas for improvement and mapping out a useful training and growth plan. Moving forward, this assessment criteria will be applied to a larger pool of individuals, with plans to use it for assessing 150 future hires. This will provide a more comprehensive understanding of the effectiveness of this approach and its potential for optimizing employee performance and development within the IT sector. In the future, the study will not only be limited to IT, but will also be applied in different business units and business areas.

References

1. Kurt, N., Ayvaz, B.: BİLGİ TEKNOLOJİLERİ PERSONELİ PERFORMANS DEĞERLENDİRME KARAR DESTEK SİSTEMİ TASARIMI. İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi **20**(40), 178–194 (2021)
2. Kalem, R.N., Kpinar, M.E.: Personnel performance assessment using entropy based MABAC method: an application in the food sector. Equinox J. Econ. Bus. Polit. Stud. **9**(1), 89–106 (2022)
3. Ahmed, I., et al.: Employee performance evaluation: a fuzzy approach. Int. J. Prod. Perform. Manag. **62**, 718–734 (2013)
4. Wu, C.-H., Fang, W.-C.: Combining the fuzzy analytic hierarchy process and the fuzzy Delphi method for developing critical competences of electronic commerce professional managers. Qual. Quant. **45**, 751–768 (2011)
5. Faisal, M.N., et al.: PMBOK, IPMA and fuzzy-AHP based novel framework for leadership competencies development in megaprojects. Benchmarking Int. J. Ahead-of-Print (2022)
6. Ding, J.-F., Kuo, J.-F., Tai, W.-H.: Using fuzzy AHP method to evaluate key competency and capabilities of selecting middle managers for global shipping logistics service providers. Pomorstvo **33**(1), 3–10 (2019)
7. Güler, M.E., Akyol, E.M.: Role of competencies in employee selection function: a fuzzy analytical hierarchy process approach. Ege Acad. Rev. **17**(2), 201–214 (2017). https://dergipark.org.tr/en/pub/eab/issue/39977/475143
8. Samanlioglu, F., et al.: A fuzzy AHP–TOPSIS-based group decision-making approach to IT personnel selection. Int. J. Fuzzy Syst. **20**, 1576–1591 (2018)
9. Lee, A.H.I., Chen, W.-C., Chang, C.-J.: A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan. Exp. Syst. Appl. **34**(1), 96–107 (2008)
10. Zadeh, L.: Fuzzy sets. Inform Control **8**, 338–353 (1965)
11. Wind, Y., Saaty, T.L.: Marketing applications of the analytic hierarchy process. Manage. Sci. **26**(7), 641–658 (1980)
12. Tavana, M., Khalili-Damghani, K., Rahmatian, R.: A hybrid fuzzy MCDM method for measuring the performance of publicly held pharmaceutical companies. Ann. Oper. Res. **226**(1), 589–621 (2014). https://doi.org/10.1007/s10479-014-1738-8
13. Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**(3), 233–247 (1985)
14. Ulutaş, A., Özkan, A.M., Tağraf, H.: Personnel selection by using fuzzy analytic hierarchy process and fuzzy grey relational analysis. Electron. J. Soc. Sci. **17**(65), 223–232 (2018)

Classification of Fuzzy MCDM Literature Applied to Humanitarian Logistics Problems

İbrahim Miraç Eligüzel and Eren Özceylan(✉)

Industrial Engineering Department, Gaziantep University, Gaziantep 27100, Turkey
erenozceylan@gmail.com

Abstract. Humanitarian logistics refers to the processes and systems involved in mobilizing people, resources, skills and knowledge to help vulnerable people affected by natural disasters and complex emergencies. There are many sub-problems (such as shelter location decision, service provider selection) in humanitarian logistics where more than one criterion is effective. The fuzzy multi-criteria decision making (MCDM) methods are effective tools in order to cope with uncertainties in human opinions under emergency situations. This paper aims to present a classification review and bring together existing literature and the most recent advances to lead researchers about the methodologies and applications of fuzzy MCDM in the field of humanitarian logistics. Using the relevant keywords (humanitarian logistics, disaster management, relief operations, fuzzy MCDM and etc.) in Scopus database, more than 35 papers are collected, analyzed and illustrated. Text mining approach Latent Dirichlet allocation (LDA) and fuzzy c-means clustering (FCM) analysis are provided using tabular and graphical illustrations. In addition to the descriptive analysis, the results are classified in terms of considered problems, applied methodologies, and handled uncertain issues. As a result, 35 papers are categorized into four clusters and the relationships between them are discussed.

Keywords: Literature Review · Fuzzy MCDM · Humanitarian Logistics · Disaster Management · Classification

1 Introduction

The field of humanitarian logistics is consistently emphasized and prioritized in literature. The efficacy of humanitarian logistics is contingent upon the specific circumstances or nature of the catastrophe at hand. Hence, the process of decision making takes into account multiple factors. The adoption of a multidisciplinary approach has emerged as a pivotal measure in the effective management of humanitarian logistics, owing to its ability to offer diverse perspectives. The integration of MCDM principles and fuzzy set theory has the potential to enhance the sensitivity, concreteness, and practicality of results. Fuzzy set theory is a theoretical framework that accounts for the uncertainties inherent in human judgments. Consequently, an in-depth understanding is necessary to establish a foundation for subsequent enhancements. For that reason, a comprehensive

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 344–352, 2023.
https://doi.org/10.1007/978-3-031-39777-6_42

survey for fuzzy MCDM approaches for humanitarian logistics is conducted in the proposed study. In order to achieve this objective, 35 abstracts are lemmatized and cleaned from stop words, and then a text mining approach called LDA is utilized. The generative probabilistic model known as LDA is employed to ascertain the feature vector of each text within a given corpus. LDA is a statistical model that is employed to identify latent topics within a corpus of documents and estimate the likelihood of individual terms within those topics. Therefore, each abstract is represented by numeric values indicated by topic probabilities. After the LDA application, clustering approach is utilized in order to gather a group of abstracts that are close to each other with respect to the concept. Therefore, a systematic approach to analyze literature for fuzzy MCDM applications on humanitarian logistics is provided. The structure of the proposed study consists of literature, methodology and analysis, and conclusion.

2 Literature

A methodical approach to reading and analyzing the literature is necessary for a thorough comprehension of the particular topics. In order to analyze literature with a large number of studies, text mining techniques are an essential tool to use. Several studies are conducted in this aspect. One of the studies proposed by Luling et al. [1] that applies LDA, Berts, and K-means modeling method to obtain topic and extracting the information about COVID-19 thorough related literature. From the same perspective, Sirajzade et al. [2] conducted a study that compares the classical text mining algorithms with modern ones for the COVID-19 literature. Another study is proposed to analyze the literature about smart and sustainable city research trends [3]. Also, study proposes a deep auto encoders-based fuzzy C-means, a novel topic-detecting algorithm based on deep learning and clustering techniques, is presented for examining the trend in research topics. Compared to subjects produced by LDA, non-negative matrix factorization, and Eigen space-based fuzzy C-means, those produced by this suggested technique have relatively higher coherence values. For the different and recent focuses, the aforementioned systematic approaches with text mining consideration can be applied [4, 5] to have efficient and effective way to analyze trends and gaps. The main concern in the proposed study is to analyzing fuzzy MCDM applications on humanitarian logistics. In the concept of humanitarian logistics, decision making occupies an important role. However, it is difficult for decision makers to make decision with certain borders. Therefore, fuzzy approaches became a favorable tool to utilize the process of decision making for humanitarian logistics applications. In that aspect several studies are proposed in the literature with a wide range of application. Qin et al. [6] proposed research that has objective to introduce a novel approach for even swaps utilizing the prospect theory framework in conjunction with hesitant fuzzy linguistic term sets and COVID-19 pandemic outbreak is utilized as a case. Ayyildiz [7] proposed a study that considers the green supply chain resilience in the post COVID-19 era. This study has extended the Supply Chain Operations Reference model by incorporating new performance attributes to assess the resilience of green supply chains. The multi-criteria decision-making problem is utilized to handle the model and ascertain the significance of the performance attributes. The Best Worst Method an integrated Interval Valued Intuitionistic Fuzzy Analytic Hierarchy Process is employed to assess the significance of performance attributes. From the logistics perspective, Sheu [8] has proposed a model for managing relief demand in emergency logistics operations during natural disasters of significant magnitude, where information is incomplete. The model is characterized by its dynamic nature. The methodology that has been put forth comprises of a tripartite process consisting of three distinct steps. The proposed methodology involves three key techniques: (1) the utilization of data fusion to predict the demand for relief across various regions, (2) the application of fuzzy clustering to categorize affected areas into distinct groups, and (3) the implementation of MCDM to determine the prioritization order of these groups. The study proposed by Habib and Sarkar [9] presents a two-stage model aimed at ensuring sustainable debris management in the aftermath of disasters during the response phase. The present study introduces a methodology for the selection of Test Data Management and Delivery Systems in the initial phase. This methodology comprises of two components, namely the Analytical Network Process and a fuzzy Technique for Order Preference by Similarity to the ideal solution. Besides the aforementioned focus of studies MCDM with integration of fuzzy approach is applied widely on humanitarian logistics processes. In that aspect, studies [10 11] that covers type-2 fuzzy MCDM approach are presented in the literature. All in fuzzy approach with MCDM has a crucial place in the literature due its capability of handling decision with uncertainties.

3 Methodology and Analysis

At the first stage, LDA is applied to gather the numerical presentation of each text and the coherence score is calculated to find the number of topic. Then, a fuzzy c-means clustering approach is used to obtain a collection of the text.

3.1 Methodology

In this study, two main approaches are utilized. The first one is the text mining approach called as LDA. The feature vector of each text is determined using LDA, a generative probabilistic model of a collection. In order to determine the probability of terms contained inside given subjects, LDA is used to identify the underlying topics in the provided documents. The "fcmeans", "sklearn" and "nltk" Python libraries are used in conjunction with a computer that has the following specifications: Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz 8 GB RAM. The formula (Eq. 1) is given as follows [12]:

$$p(D|\alpha, \beta) = \prod_{d=1}^{M} \int P(\theta_d|\alpha)\left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn}|\theta_d)p(w_{dn}|z_{dn}, \beta)\right)d\theta_d \quad (1)$$

The concentration parameter of each document topic distribution is denoted by α in the Dirichlet prior. The topic assignment for $w(d, n)$ is referred to as $z(d, n)$, where w (d, n) represents the nth word in the d_{th} document. The number of words in the document is represented by N, while M represents the number of documents to be analyzed. The corpus of collection M documents is represented by D. After the LDA application, each document presented with the feature vectors and fuzzy c-means clustering is applied by utilizing these feature vectors. The formation of fuzzy clusters is achieved through the partitioning of training samples based on the membership functions matrix U, which is denoted as [μ_{ki}]. The degree of membership of x_k in cluster i is denoted by v_i. Here, x_k represents the k_{th} measured data point in a d-dimensional space. The standard Fuzzy C-Means algorithm employs the Euclidean distance as a cost function that is to be minimized, and can be mathematically represented by the following Eq. (2):

$$J_{FCM}(U,V) = \sum_{k=1}^{c} \sum_{i=1}^{n} \mu_{ki}^{m} |x_k - v_i|^2 \quad (2)$$

3.2 Analysis and Results

In this part of the study, LDA and fuzzy c-means clustering are applied. At the first stage, the number of topics is decided in accordance with the coherence score. Coherence score close the one indicates that the number of topics leads to better interpretation. In this application, the best coherence score found as -1.1955 with four topics. After that, fuzzy c-means is applied with the extracted features consists of topic probabilities. Using silhouette analysis, one can examine the distance between the resulting clusters. These values, which are referred to as silhouette coefficients, are near to 1 imply that the sample is distant from the neighboring clusters. A value of 0 indicates that the sample is on or very near to the decision boundary between two neighboring clusters, whereas negative values indicate that the sample may have been misclassified. In that aspect, the closest score to 0 is found as 0.81 with four clusters. In Table 1, label for each documents and their reference.

Table 1. Labeling and references for 35 abstracts

Label	Ref	Label	Ref	Label	Ref
D1	Zang et al. [13]	D13	Rezaei et al. [41]	D25	Trivedi and Singh [40]
D2	Kumar et al. [16]	D14	Mzougui et al. [14]	D26	Habin and Sakar [9]
D3	Le et al. [19]	D15	Geng et al. [17]	D27	Yu [15]
D4	Ayyildiz and Taskin [11]	D16	Yariyan et al. [20]	D28	Ugo [18]
D5	Qin et al. [6]	D17	Jayant [22]	D29	Kristianto et al. [21]
D6	Hossain and Thakur [26]	D18	Sari et al. [24]	D30	Rabenasolo and Zeng [23]
D7	Alkahtani et al. [29]	D19	Onar et al. [27]	D31	Tuzkaya et al. [25]
D8	Alkahtani et al. [31]	D20	Ozdemir et al. [30]	D32	Tirado et al. [28]
D9	Ayyildiz [7]	D21	Venkatesh et al. [32]	D33	Sheu [8]
D10	Göçer [36]	D22	Yadar and Barve [34]	D34	Jing and Lindu [33]
D11	Kaur and Singh [38]	D23	Drakaki et al. [37]	D35	Han et al. [35]
D12	Yılmaz and Kabak [10]	D24	Drakaki et al. [39]		

After the fuzzy c-means clustering application, documents are allocated to four clusters. The details for clusters are given in Table 2. Table 2 illustrates the summary of each cluster. From Table 2, it can be said that the largest number of documents belongs to cluster 2, while lowest number of document allocated to cluster 4. When it comes to topic probabilities, topic 2, topic 3, topic 1, and topic 4 have the highest probabilities in clusters 1-2-3-4 respectively. Therefore, it can be expressed that each cluster has its unique characteristics considering the focuses of the documents in the clusters. Topics are given in Table 3 with their probabilities of words.

Table 2. Summary of clusters

Clusters	Central feature	Central paper	List of papers
#1	[1.61625615e-03; **9.54643005e-01;** 1.98975911e-03; 3.26608165e-02]	A multi-criteria dynamic flow model for relief operations	D34, D33, D31, D28, D26, D20, D13, D6
#2	[1.14590639e-02; 2.08684653e-02; **9.39926502e-01;** 2.14759890e-02]	Humanitarian relief supply chain performance evaluation by a SCOR based Trapezoidal type-2 fuzzy multi-criteria decision making methodology: An application to Turkey	D1, D29, D21, D18, D16, D15, D14, D9, D7, D5, D3, D2
#3	[**9.38116727e-01;** 1.83319599e-02; 2.49170053e-02; 1.19957166e-02]	Prioritizing distribution centers in humanitarian logistics using type-2 fuzzy MCDM approach	D0, D10, D32, D30, D27, D24, D23, D17, D11
#4	[2.29921892e-02; 3.11694124e-04; 4.32570971e-02; **9.28628720e-01**]	Interval valued intuitionistic fuzzy analytic hierarchy process-based green supply chain resilience evaluation methodology in post COVID-19 era	D4, D8, D12, D25, D22, D19

From Table 3, it can be expressed that topic 1 is related to the decision on location from the supply chain perspective with a fuzzy approach. Topic 2 has the same perspective with topic 1 with the difference of focusing healthcare supply chain. In topic 3, supply chain perspective is underlined more than other topics and hybrid approaches are focused. For topic 4, supply chain processes are expressed. A basic illustration of each cluster with documents is given the Fig. 1.

Table 3. Topic word probabilities

Topic Label	Word Probabilities
Topic#1	(‘0.015 * “criterion” + 0.011 * “decision” + 0.011 * “location” + 0.010 * “fuzzy” + 0.008 * “disaster” + 0.007 * “approach” + 0.007 * “site” + 0.007 * “supply” + 0.007 * “use” + 0.006 * “chain”’)
Topic#2	(‘0.011 * “criterion” + 0.011 * “decision” + 0.011 * “fuzzy” + 0.008 * “supply” + 0.008 * “hcsc” + 0.008 * “disaster” + 0.006 * “multicriteria” + 0.006 * “risk” + 0.006 * “use” + 0.006 * “logistics”’),
Topic#3	(‘0.017 * “chain” + 0.016 * “supply” + 0.009 * “disaster” + 0.008 * “decision” + 0.008 * “criterion” + 0.007 * “fuzzy” + 0.006 * “process” + 0.006 * “hybrid” + 0.005 * “use” + 0.005 * “multicriteria”’),
Topic#4	(‘0.015 * “supply” + 0.011 * “chain” + 0.010 * “decision” + 0.010 * “criterion” + 0.009 * “fuzzy” + 0.007 * “result” + 0.006 * “multicriteria” + 0.006 * “use” + 0.006 * “location” + 0.005 * “disaster”’)

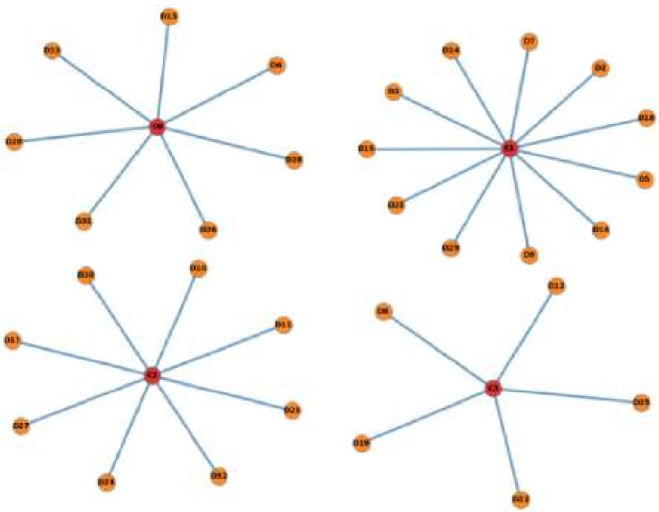


Fig. 1. Representation of clusters

4 Conclusion

All in all, a two stage approach is utilized in the proposed study to analyses the literature about MCDM in humanitarian logistics with fuzzy methods. In the first stage of the study, LDA is applied in order to have a numeric presentation of each document. The topic probabilities for documents are gathered and the number of topics decided by the calculation of coherence score, which is four. After that, fuzzy c-means clustering is applied in order to provide a systematic approach for analyzing the considered literature. The number of clusters is decided in accordance with the silhouette analysis, which equals to four. Then, four clusters with attended documents are gathered. The findings of the proposed study are given as follows with respect to clusters. The largest number of documents attended to cluster 2. The emphasis in this cluster is on the supply chain perspective, with particular attention given to hybrid approaches. From the analyses, it can be said that studies in cluster 1 have the area related to MCDM application on healthcare supply chain during the disaster. For cluster 2, it can be expressed that studies are mostly related to hybrid MCDM approaches for supply chain during disaster. Cluster 3, the MCDM application on deciding location in supply chain processes for disasters is the main focus. Lastly, Cluster 4 has a focus of application of MCDM on supply chain for disasters. The limitation of our study is that the number of studies is sufficient but large number of documents can provide a more precise interpretation. For the future study, number of documents can be increased, full text can be gathered and different text mining and clustering approaches can be applied in order to make comparison and obtain better results. For future studies, different text mining approaches can be applied and comparison can be made to gather more meaningful clusters. In addition, the full text of studies can be considered in order to accomplish a deeper analysis of the objective of the proposed study.

References s

1. Luling, X., Zixi, W., Yeqiu, C., Yichao, W., Di, Z.: COVID-19 literature mining and analysis research. In: 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms, EEBDA 2022. IEEE, pp. 1045–1052 (2022)
2. Sirajzade(B), J., Bouvry, P., Schommer, C.: Deep mining COVID-19 literature. In: Florez, H., Gomez, H. (eds.) Applied Informatics (ICAI 2022), vol. 1643, pp. 121–133. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19647-8_9
3. Parlina, A., Ramli, K., Murfi, H.: Exposing emerging trends in smart sustainable city research using deep autoencoders-based fuzzy c-means. Sustain. **13**, 1–28 (2021)
4. Van Nguyen, T., Cong Pham, H., Nhat Nguyen, M., Zhou, L., Akbari, M.: Data-driven review of blockchain applications in supply chain management: key research themes and future directions. Int. J. Prod. Res. (2023)
5. Sindhusuta, S., Chi, S.W., Derrible, S.: A text-mining-based approach for conducting literature review of selected Meshfree methods. Comput. Assist. Methods Eng. Sci. **28**, 265–290 (2021)
6. Qin, R., Liao, H., Jiang, L.: An enhanced even swaps method based on prospect theory with hesitant fuzzy linguistic information and its application to the selection of emergency logistics plans under the COVID-19 pandemic outbreak. J. Oper. Res. Soc. **73**, 1227–1239 (2022)
7. Ayyildiz, E.: Interval valued intuitionistic fuzzy analytic hierarchy process-based green supply chain resilience evaluation methodology in post COVID-19 era. Environ. Sci. Pollut. Res. **30**, 42476–42494 (2021)
8. Sheu, J.B.: Dynamic relief-demand management for emergency logistics operations under large-scale disasters. Transp. Res. Part E Logist. Transp. Rev. **46**, 1–17 (2010)
9. Habib, M.S., Sarkar, B.: An integrated location-allocation model for temporary disaster debris management under an uncertain environment. Sustain. **9**, 716 (2017)
10. Yılmaz, H., Kabak, Ö.: Prioritizing distribution centers in humanitarian logistics using type-2 fuzzy MCDM approach. J. Enterp. Inf. Manag. **33**, 1199–1232 (2020)
11. Ayyildiz, E., Taskin, A.: Humanitarian relief supply chain performance evaluation by a SCOR based Trapezoidal type-2 fuzzy multi-criteria decision making methodology: an application to Turkey. Sci. Iran. **29**, 2069–2083 (2022)

12. Eligüzel, N., Çetinkaya, C., Dereli, T.: Comparative analysis with topic modeling and word embedding methods after the Aegean Sea earthquake on Twitter. Evol. Syst. **14**, 245–261 (2022)
13. Zhang, J.Z., Srivastava, P.R., Eachempati, P.: Evaluating the effectiveness of drones in emergency situations: a hybrid multi-criteria approach. Ind. Manag. Data Syst. **123**, 302–323 (2023)
14. Mzougui, I., Carpitella, S., Certa, A., El Felsoufi, Z., Izquierdo, J.: Assessing supply chain risks in the automotive industry through a modified MCDM-Based FMECA. Processes **8**, 579 (2020)
15. Yu, D.: Softmax function based intuitionistic fuzzy multi-criteria decision making and applications. Oper. Res. Int. J. **16**(2), 327–348 (2015). https://doi.org/10.1007/s12351-015-0196-7
16. Kumar, D., Soni, G., Joshi, R., Jain, V., Sohal, A.: Modelling supply chain viability during COVID-19 disruption: a case of an Indian automobile manufacturing supply chain. Oper. Manag. Res. **15**, 1224–1240 (2022)
17. Geng, S., Hou, H., Zhang, S.: Multi-criteria location model of emergency shelters in humanitarian logistics. Sustain. **12**, 1759 (2020)
18. Ugo, P.D.: A multi-criteria decision making for location selection in the niger delta using fuzzy TOPSIS approach. Int. J. Manag. Bus. Res. **5**, 215–224 (2015)
19. Le, M.T., Nhieu, N.L., Pham, T.D.T.: Direct-use geothermal energy location multi-criteria planning for on-site energy security in emergencies: a case study of Malaysia. Sustain. **14**, 15132 (2022)
20. Yariyan, P., Avand, M., Soltani, F., Ghorbanzadeh, O., Blaschke, T.: Earthquake vulnerability mapping using different hybrid models. Symmetry (Basel) **12**, 405 (2020)
21. Kristianto, Y., Gunasekaran, A., Helo, P., Hao, Y.: A model of resilient supply chain network design: a two-stage programming with fuzzy shortest path. Expert Syst. Appl. **41**, 39–49 (2014)
22. Jayant, A.: Decision support framework for smart implementation of green supply chain management practices. In: Patnaik, S. (ed.) New Paradigm of Industry 4.0: Internet of Things, Big Data & Cyber Physical Systems, vol. 64, pp. 49–98. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-25778-1_4
23. Rabenasolo, B., Zeng, X.: Handbook on Decision Making (2012)
24. Sari, D.P., Wicaksono, P.A., Handayani, N.U., Siahaan, Y.E.: Integrating AHP, cluster analysis, and fuzzy TOPSIS to evaluate emergency warehouse locations of Mount Merapi eruption victims. Int. J. Adv. Sci. Eng. Inf. Technol. **10**, 1839–1845 (2020)
25. Tuzkaya, U.R., Yilmazer, K.B., Tuzkaya, G.: Location selection of emergency logistics centers using an integrated dematel-ANP approach. In: Uncertainty Modeling in Knowledge Engineering and Decision Making, pp. 143–148 (2012)
26. Maqsoom, A., et al.: Efficiency of multiple hybrid techniques for the earthquake physical susceptibility mapping: the case of Abbottabad District. Pakistan. Environ. Earth Sci. **80**, 1–20 (2021)
27. Onar, S.C., Kahraman, C., Oztaysi, B.: Multi-criteria spherical fuzzy regret based evaluation of healthcare equipment stocks. J. Intell. Fuzzy Syst. **39**, 5987–5997 (2020)
28. Tirado, G., Vitoriano, B., Ortuño, M.T.: A multicriteria dynamic flow model for relief operations. In: World Scientific Proceedings Series on Computer Engineering and Information Science 7. Uncertainty Modeling in Knowledge Engineering and Decision Making - Proceedings 10th International FLINS Conference, vol. 7, pp. 9–14 (2012)
29. Hossain, M.K., Thakur, V.: Benchmarking health-care supply chain by implementing industry 4.0: a fuzzy-AHP-DEMATEL approach. Benchmarking **28**, 556–581 (2021)
30. Ozdemir, A.I., et al.: The role of blockchain in reducing the impact of barriers to humanitarian supply chain management. Int. J. Logist. Manag. **32**, 454–478 (2020)

31. Alkahtani, M., et al.: A COVID-19 supply chain management strategy based on variable production under uncertain environment conditions. Int. J. Environ. Res. Public Health **18**, 1–23 (2021)
32. Venkatesh, V.G., Zhang, A., Deakins, E., Luthra, S., Mangla, S.: A fuzzy AHP-TOPSIS approach to supply partner selection in continuous aid humanitarian supply chains. Ann. Oper. Res. **283**(1–2), 1517–1550 (2018). https://doi.org/10.1007/s10479-018-2981-1
33. Hou, J., Zhao, L.: Logistics service mode selection based on decision-tree model and grey multi-criteria risk decision-making method. In: Proceedings 2007 IEEE International Conference on Grey Systems and Intelligent Services, GSIS 2007, pp. 820–825 (2007)
34. Yadav, D.K., Barve, A.: Prioritization of cyclone preparedness activities in humanitarian supply chains using fuzzy analytical network process. Nat. Hazards **97**, 683–726 (2019)
35. Han, R., Cao, G., Lu, B.: Performance evaluation for emergency logistic outsourcing vendor based on fussy analytical hierarchical process. J. Southeast Univ. **37**, 348–354 (2007)
36. Gocer, F.: A novel interval value extension of picture fuzzy sets into group decision making: an approach to support supply chain sustainability in catastrophic disruptions. IEEE Access **9**, 117080–117096 (2021)
37. Drakaki, M., Gören, H.G., Tzionas, P.: An intelligent multi-agent based decision support system for refugee settlement siting. Int. J. Disaster Risk Reduct. **31**, 576–588 (2018)
38. Kaur, H., Prakash Singh, S.: Multi-stage hybrid model for supplier selection and order allocation considering disruption risks and disruptive technologies. Int. J. Prod. Econ. **231**, 107830 (2021)
39. Drakaki, M., Gören, H. G., Tzionas, P.: An intelligent multi-agent system using fuzzy analytic hierarchy process and axiomatic design as a decision support method for refugee settlement siting. In: Dargam, F., Delias, P., Linden, I., Mareschal, B. (eds.) ICDSST 2018. LNBIP, vol. 313, pp. 15–25. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-90315-6_2
40. Trivedi, A., Singh, A.: Prioritizing emergency shelter areas using hybrid multi-criteria decision approach: a case study. J. Multi-Criteria Decis. Anal. **24**, 133–145 (2017)
41. Rezaei, M., Khalilpour, K.R., Jahangiri, M.: Multi-criteria location identification for wind/solar based hydrogen generation: the case of capital cities of a developing country. Int. J. Hydrogen Energy **45**, 33151–33168 (2020)

Circular Intuitionistic Fuzzy PROMETHEE Methodology: A Case of Smart Cities Evaluation

Esra Çakır(✉) and Emre Demircioğlu

Department of Industrial Engineering, Galatasaray University, 34349 Ortakoy/Istanbul, Turkey
{ecakir,edemircioglu}@gsu.edu.tr

Abstract. Due to economic, environmental, social and cultural reasons, city administrations comply with the trends suitable for the needs of the high population with a human-oriented perspective. In addition to the central government, metropolitan municipalities are keeping up with this transformation by starting to create smart city units within their own structure. Adoption of smart city strategies in the management of crowded cities is prioritized to increase the level of welfare. In order to evaluate how smart cities are adapting to this strategies, it would be appropriate to periodically inspect themselves according to expert opinion and certain criteria. In this study, the suitability of seven cities for smart cities was investigated according to six smart city indicators. Circular intuitionistic fuzzy set (C-IFS), which have recently been included in the literature and are the extension of intuitionistic fuzzy set (IF), including the uncertainty of IF group decisions, have been applied for the first time with the PROMETHEE I and PROMETHEE II methods. The proposed method is intended to support the integration of C-IFS into MCDM models.

Keywords: Circular intuitionistic fuzzy set · PROMETHEE I · PROMETHEE II · Smart Cities · Smart City Indicators

1 Introduction

Smart cities can be defined as the integration of information and communication technologies of economy, transportation and traffic management in order to increase the quality of life of people living in the city [1]. A smart city supports sustainable and livable cities [10] and its main mission is to make life easier and efficient activities [2]. In this adaptation; solutions such as sensors, software that optimizes traffic flow, ride sharing, scheduling, monitoring public transport, air conditioning systems are favorable [3]. Coordinated and periodically checked approaches should be adopted by organizations and governments to create smart cities [4]. Cities in the developed countries are trying to maintain their leadership in smart city applications in order to increase their preferability [5].

In the literature, it is a current issue to evaluate smart cities using many MCDM methods according to various indicators. Anand et al. evaluate smart cities in India according to sustainability indicators via fuzzy AHP methodology [6]. Adali et al. assest European

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 353–361, 2023.
https://doi.org/10.1007/978-3-031-39777-6_43

smart cities by proposing Level Based Weight Assessment (LBWA) and Evaluation Based on Distance from Average Solution (EDAS) methodology in grey environment [7]. Şeker evaluates IoT based sustainable smart waste management system in Istanbul with entropy method including interval-valued q-rung orthopair fuzzy sets [8]. Hajduk and Jelonek are applied TOPSIS methodology for ranking smart cities in the context of urban energy [9]. Ogrodnik employed PROMETHEE method to analyze the largest cities in terms of the smart city indicators in Poland [10].

The PROMOTHEE method is a multi-criteria decision-making method consisting of 2 main stages, PROMOTHEE 1 (partial ranking) and PROMOTHEE 2 (complete ranking), first developed by Brans in 1982 [11]. The method determines the order of decision points with the main phases of PROMETHEE 1 and PROMETHEE 2. The PROMETHEE method is based on pairwise comparisons of decision points according to evaluation factors. The main difference that distinguishes the PROMOTHEE method from other multivariate decision-making methods is that it takes into account the importance weights that show the relations of the evaluation factors with each other and the internal relationship of each evaluation factor. The interrelationship of the evaluation factors is revealed by the distribution of the data set and 6 different distributions are foreseen for this purpose in the method [12]. The PROMETHEE method has been extended many times in the literature with fuzzy sets in order to integrate the uncertain statements of decision makers [13–17].

With Zadeh's [18] fuzzy set theory, which brings a new perspective to set theory, a great deal of space has been opened to the literature in expressing the uncertainty in human nature. Zadeh attributed the belonging of an element to the set with imprecise values ranging from 0–1. With the intutionistic numbers put forward by Atanassov [19], the situation of not belonging to the cluster began to be expressed by taking the uncertainty one step further. In recent years, Atanassov [20] has expanded his IF set and introduced C-IFS numbers by creating a convex structure that can also express the uncertainty created by numbers in IF grup decision making. As in other fuzzy sets, it is clear that fuzzy numbers will increase the efficiency of MCDM methods by using C-IFS in many MCDM methods. In order to develop this last fuzzy set, some studies have taken place in the literature [21–26].

This study aims to extend MCDM literature by integrating C-IFS, which is novel extension of IFS. In addition to the C-IF AHP [27, 28], C-IF VIKOR [28, 29], C-IF TOPSIS [30–32] methods suggested in the literature, the study contributes to the literature by using the C-IF PROMETHEE I and PROMETHEE II approaches for the first time. It is possible to obtain a meaningful perspective by using C-IFS in MCDM methods, since it also includes group decision uncertainties in views expressed in IF numbers.

This paper is organized as follows. Section 2 gives the preliminaries of circular intuitionistic fuzzy set. Section 3 introduces the novel C-IF PROMTHEE I and PROMETHEE II methodology. Section 4 applies the proposed approach on evaluation of smart cities according to city indicators. Section 5 concludes the study and proposes future directions.

2 Circular Intuitionistic Fuzzy Set

Circular intuitionistic fuzzy set (C-IFS) was developed by Atanassov as an extension of IF numbers with a convex structure that expresses the uncertainty in the IF numbers together with the group uncertainty [20, 21]. The preliminaries of circular intuitionistic fuzzy numbers are given as follows:

Definition 1: [22] Let E be a fixed universe, and a generic element a C-IFS C_r in E is denoted by x; $C_r = \{<\mathrm{x} : u_C(x), v_C(x);\ \mathrm{r}>,\ \mathrm{x} \in E\}$ is the form of an object that is the circular intuitionistic fuzzy set (C-IFS), where the functions "u, v: E $\rightarrow$ [0, 1]" define respectively the membership function and the non-membership function of the element $x \in E$ to the set C-IFS with condition:

$$0 \leq u_C(x) + v_C(x) \leq 1 \text{ and r } \in \left[0, \sqrt{2}\right] \tag{1}$$

where r is the radius of the circle around each element $x \in E$ [31]. The indeterminacy function is defined as $\alpha_C(x) = 1 - u_C(x) - v_C(x)$. A C-IFS is reduced to a standard IFS when r = 0.

Definition 2: [33] Let $\tilde{a}_i = \left(u_{\tilde{a}_i}, v_{\tilde{a}_i}\right)$ (i = 1,2,…,n) be an IFS, then their aggregated value by using the IF weighted averaging (IFWA) operator is also an IF value [26]:

$$\mathrm{IFWA}(\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_n) = \left(1 - \prod\nolimits_{j=1}^{n}\left(1 - u_{\tilde{a}_j}\right)^{w_j}, \prod\nolimits_{j=1}^{n} v_{\tilde{a}_j}^{w_j}\right) \tag{2}$$

where $\mathrm{w} = \{\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n\}$ is the weighting vector of $\tilde{a}_i$ (i = 1,2,…,n) with $\mathrm{w}_j \in [0, 1]$ and $\sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{w_j} = 1$.

Definition 3: [20] Let $\left\{<m_{\mathrm{i},1}, n_{\mathrm{i},1}>, <m_{\mathrm{i},2}, n_{\mathrm{i},2}>, \ldots\right\}$ is a set of intuitionistic fuzzy (IF) pairs. The C-IFS C_i is calculated from pairs where i is the number of the IF set and k_i is the number of IF pairs in each set. The arithmetic average of the set is as follows:

$$C_\mathrm{i} = <u_C(C_\mathrm{i}), v_C(C_\mathrm{i})> = <\frac{\sum_{j=1}^{k_i} m_{\mathrm{i,j}}}{k_i}, \frac{\sum_{j=1}^{k_i} n_{\mathrm{i,j}}}{k_i}> \tag{3}$$

The radius r_i of the C_i is obtained by the maximum of the Euclidean distances as follows:

$$r_\mathrm{i} = \max_{1 \leq j \leq k_i} \sqrt{\left(u_C(C_\mathrm{i}) - m_{\mathrm{i,j}}\right)^2 + \left(v_C(xC_\mathrm{i}) - n_{\mathrm{i,j}}\right)^2} \tag{4}$$

Definition 4: [20] Let $L^* = \{<a, b> \mid a, b \in [0, 1]\ \&\ a + b \leq 1\}$. Therefore C_r can be rewritten in the form $C_r^* = \{<\mathrm{x} : O_r(u_C(x), v_C(x)))>,\ \mathrm{x} \in E\}$ where "$O_r(u_C(x), v_C(x)) = \{<a, b> \mid a, b \in [0, 1], \sqrt{(u_C(x) - a)^2 + (v_C(x) - b)^2} \leq r, a + b \leq 1\}$" is a function of circle representation. The C-IFSs are geometrical represented in Fig. 1.

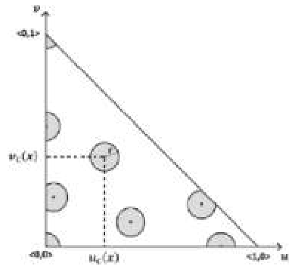


Fig. 1. Illustration of circular intuitionistic fuzzy numbers.

Definition 5: [21] New score and accuracy functions for C-IFSs are presented. Let $c = (u_c, v_c; r)$ be an C-IFV, a score function S_{C-IFS} and an accuracy function H_{C-IFS} of the C-IFV c is defined as follows with respect to the decision maker's (or manager's) preference information $\pi \in [0, 1]$:

$$S_{C-IFS}(c) = \frac{u_c - v_c + \sqrt{2}r(2\pi - 1)}{3} \text{ where } S_{C-IFS}(c) \in [-1, 1] \tag{5}$$

$$H_{C-IFS}(c) = u_c + v_c \quad \text{where } H_{C-IFS}(c) \in [0, 1] \tag{6}$$

Here, π reflects the perspective of the manager to the model. If π is equal to zero, it shows the most pessimistic scenario, and if π is equal to one, it shows the most optimistic scenario for the problem. Generally, $\pi \in [0, 0.5)$ represents a pessimistic scenario, and $\pi \in (0.5, 1]$ represents an optimistic scenario. $\pi = 0.5$ represents the unbiased scenario of manager [21].

3 Circular Intuitionistic Fuzzy PROMETHEE

Recently, the use of C-IFS numbers in the literature has increased as it highlights the uncertainty in IF numbers. In addition to its theoretical development [20, 21], its integration in MCDM methods and applications in industrial subjects have been realized [23–33]. In this study, C-IFS is integrated for the first time in the PROMETHEE method. The PROMETHEE method was introduced to the literature by Brans and was developed by Brans and Vincke [12]. With the PROMETHEE method, it is possible to perform both partial ranking (PROMETHEE I) and full ranking (PROMETHEE II) on a finite number of alternatives [34]. The MCDM method enables pairwise comparisons of alternatives according to six types of preference functions through determined criteria [12].

Based on the IF-PROMETHEE methods [16–35] in the literature, the implementation steps of the C-IFS PROMETHEE I and PROMETHEE II approach are proposing below:

Step 1: Identify the set of decision makers $DM = \{DM_1, DM_2, \ldots, DM_l\}$, alternatives $X = \{x_1, x_2, \ldots, x_n\}$, criteria $C = \{c_1, c_2, \ldots, c_m\}$, and criterion weights $W = \{w_1, w_2, \ldots, w_m\}$.

Step 2: Gather IF decision matrix from DMs and aggregate the IF matrices to construct C-IFS decision matrix by using Eq. (2), Eq. (3), Eq. (4).

Step 3: Calculate the deviations of C-IF evaluations for each pair of alternatives over different criteria c_j (j = 1, 2,…, m) via following equation:

$$d_j(x, y) = S_{C-IFS}(c_j(x)) - S_{C-IFS}(c_j(y)) \tag{7}$$

where $d_j(x, y)$ denotes the difference between the assessments of the alternatives x and x on the criterion c_j. The score function for C-IFS is given in Eq. (5) and the decision maker's perspective (π) should be determined.

Step 4: Set preference functions as follows:

$$P_j(x, y) = f_j\left[d_j(x, y)\right] \forall x, y \in X \tag{8}$$

where f_j is a preference function, which translates the difference between the assessments of the alternatives x and y on the criterion c_j into a preference degree ranging from 0 to 1, i.e., $0 \leq P_j(x, y) \leq 1$, starting at 0 if $c_j(x) = c_j(y)$ and increasing with $d_j(x, y)$ to reach 1 when the deviation is big enough.

Establish the case specific parameters q as an indifference threshold and p as a strict preference threshold. Then compute the preferences μ_{ik}^j for the alternative x_i against alternative x_k with respect to the criterion c_j by using the V-shape (see in Fig. 2) with indifference criterion as in Eq. (9). After that, the preference matrix U^j (j = 1, 2,…,m) can be established.

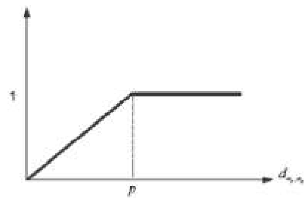


Fig. 2. Type V Function: Criterion with linear preference and indifference area.

$$P\left(d_j\left(x_i x_j\right)\right) = \mu_{ij}^j = \begin{cases} 0, & d_j\left(x_i x_j\right) \leq q \\ \frac{s(c_j(x_i)) - s(c_j(x_j)) - q}{p - q} & q < d_j\left(x_i x_j\right) \leq p \\ 1, & d_j\left(x_i x_j\right) > p \end{cases} \tag{9}$$

Step 5: Construct the intuitionistic fuzzy preference relation $R^j = \left(r_{ij}^j\right)_{nxn}$ over the criteria c_j (j = 1, 2,…,m) By using the equations $v_{ki} = \mu_{ik}$ and $v_{ik} = \mu_{ki}$. Then, aggregate the IF matrices to construct overall circular intuitionistic fuzzy preference relation by using Eq. (2).

Step 6: Calculate the intuitionistic fuzzy positive outranking flow $\phi^+(x_i)$, the intuitionistic fuzzy negative outranking flow $\phi^-(x_i)$ for the alternative xi by using following equations:

$$\phi^+(x_i) = \frac{1}{n-1} \bigoplus_{k=1, k \neq i}^{n} r(x_i, x_k) = \frac{1}{n-1} \bigoplus_{k=1, k \neq i}^{n} r_{ik} \tag{10}$$

$$\phi^{-}(x_i) = \frac{1}{n-1} \overset{n}{\underset{k=1,k\neq i}{\oplus}} r(x_k,\ x_i) = \frac{1}{n-1} \overset{n}{\underset{k=1,k\neq i}{\oplus}} r_{ki} \tag{11}$$

Since each alternative is compared with (n − 1) other alternatives in X, the intuitionistic fuzzy positive outranking flow and the intuitionistic fuzzy negative outranking flow can be derived analogously to the classical PROMETHEE.

Step 7. The partial ranking is established by comparing $\beta^{+}(x_i)$ and $\beta^{-}(x_i)$ of the alternatives.

- Case I- Partial ranking: x_i outranks x_k if $\beta^{+}(x_i) \geq \beta^{+}(x_k)$ and $\beta^{-}(x_i) \leq \beta^{-}(x_k)$.
- Case II- Equality: $\beta^{+}(x_i) = \beta^{+}(x_k)$ and $\beta^{-}(x_i) = \beta^{-}(x_k)$ hold simultaneously implies indifference between two alternatives.
- Case III- Incomparability: this takes place if $\beta^{+}(x_i) > \beta^{+}(x_k)$ and $\beta^{-}(x_i) > \beta^{-}(x_k)$ or $\beta^{+}(x_i) < \beta^{+}(x_k)$ and $\beta^{-}(x_i) < \beta^{-}(x_k)$.

Step 8: Since the score function of an IFS is in Eq. (12), the complete ranking is derived according to the deviation between the score values of the intuitionistic fuzzy positive outranking flow and that of the intuitionistic fuzzy negative outranking flow from PROMETHEE II as in Eq. (13).

$$S_{IFS}(c) = u_c - v_c \text{ where } S_{C-IFS}(c) \in [-1, 1] \tag{12}$$

$$S_{IFS}(\beta(x_i)) = S_{IFS}\left(\beta^{+}(x_i)\right) - S_{IFS}\left(\beta^{-}(x_i)\right) \tag{13}$$

4 Application

The proposed C-IF PROMETHEE model in the previous section is applied on seven big cities in Turkey according to the six smart city indicators with five decision makers. The π parameter is set 0.75 to observe optimistic view. Since there is a page limit in the paper, some information about the application is included. The hierarchical structure of the smart city evaluation case is illustrated in Fig. 3.

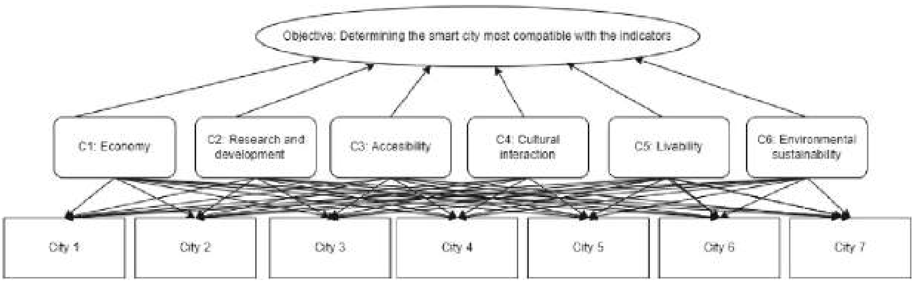


Fig. 3. Hierarchical structure of the smart city evaluation case.

The aggregated C-IFS decision matrix in Step 2 is calculated as in Table 1.
The final ranking results are given in Table 2.

Table 1. The aggregated C-IFS decision matrix.

	C1 $w_{C1} = 0.09$	C2 $w_{C2} = 0.24$	C3 $w_{C3} = 0.12$	C4 $w_{C4} = 0.25$	C5 $w_{C5} = 0.17$	C6 $w_{C6} = 0.13$
City 1	<0.23, 0.55; 0.05>	<0.51, 0.39; 0.08>	<0.41, 0.35; 0.21>	<0.56, 0.21; 0.05>	<0.15, 0.54; 0.06>	<0.71, 0.27; 0.02>
City 2	<0.25, 0.44; 0.08>	<0.34, 0.55; 0.07>	<0.35, 0.40; 0.12>	<0.35, 0.40; 0.14>	<0.24, 0.41; 0.09>	<0.25, 0.60; 0.06>
City 3	<0.45, 0.41; 0.15>	<0.26, 0.30; 0.10>	<0.29, 0.55; 0.14>	<0.21, 0.55; 0.23>	<0.36, 0.38; 0.21>	<0.24, 0.41; 0.24>
City 4	<0.32, 0.43; 0.09>	<0.51, 0.37; 0.07>	<0.37, 0.27; 0.22>	<0.24, 0.61; 0.12>	<0.28, 0.17; 0.25>	<0.16, 0.50; 0.18>
City 5	<0.46, 0.36; 0.07>	<0.22, 0.43; 0.06>	<0.15, 0.31; 0.08>	<0.36, 0.48; 0.04>	<0.61, 0.28; 0.15>	<0.19, 0.56; 0.10>
City 6	<0.12, 0.57; 0.11>	<0.45, 0.10; 0.09>	<0.25, 0.52; 0.09>	<0.37, 0.18; 0.15>	<0.34, 0.24; 0.17>	<0.53, 0.28; 0.09>
City 7	<0.47, 0.24; 0.12>	<0.22, 0.64; 0.12>	<0.31, 0.24; 0.15>	<0.28, 0.37; 0.32>	<0.31, 0.55; 0.20>	<0.26, 0.46; 0.08>
p value	<0.09, 0.8; 0.03>	<0.07, 0.14; 0.02>	<0.07, 0.08; 0.02>	<0.09, 0.11; 0.07>	<0.12, 0.1; 0.05>	<0.14, 0.08; 0.06>
q value	<0.26, 0.25; 0.8>	<0.22, 0.41; 0.05>	<0.20, 0.23; 0.59>	<0.26, 0.32; 0.21>	<0.35, 0.29; 0.14>	<0.41, 0.25; 0.17>

Table 2. Final ranking results of C-IF PROMETHEE I and PROMETHEE II.

	C-IF PROMETHEE I						C-IF PROMETHEE II	
	$\beta^{+}(x_i)$	$S_{IFS}(\beta^{+}(x_i))$	Ranking	$\beta^{-}(x_i)$	$S_{IFS}(\beta^{-}(x_i))$	Ranking	$S_{IFS}(\beta(x_i))$	Ranking
City 1	<0.15, 0.28>	−0.13	1	<0.02, 0.74>	−0.72	7	0.59	1
City 2	<0.00, 1.00>	−1.00	7	<0.04, 0.64>	−0.59	2	−0.41	7
City 3	<0.01, 0.95>	−0.94	6	<0.04, 0.64>	−0.60	3	−0.34	6
City 4	<0.3, 0.79><	−0.76	5	<0.03, 0.74>	−0.71	5	−0.06	4
City 5	<0.3, 0.65>	−0.63	4	<0.04, 0.66>	−0.62	4	−0.01	3
City 6	<0.6, 0.47>	−0.41	2	<0.01, 0.72>	−0.71	6	0.30	2
City 7	<0.2, 0.60>	−0.58	3	<0.11, 0.36>	−0.26	1	−0.33	5

5 Conclusion

In this study, C-IF PROMETHEE method is applied for the first time for the evaluation of alternative cities thanks to the indicators used in the city indexes to evaluate the smart infrastructures of big cities. The C-IF numbers, which Atanassov added to the literature, use IF numbers to create a structure that includes the uncertainty in the group decision-making process. As a result of the application, 7 alternative cities are evaluated according to 6 indicators and rankings are given. City 1 is more adapted the smart city indicators than other alternative cities. This study expanded the use of C-IFS in MCDM methods and created a comparable field with the use of PROMETHEE I and PROMETHEE II methods. In future studies, it is useful to compare the results of methods such as C-IFS VIKOR, C-IFS TOPSIS, C-IFS AHP, which have already been introduced in the literature, to analyze their effectiveness.

References

1. Ismagilova, E., Hughes, L., Dwivedi, Y.K., Raman, K.R.: Smart cities: advances in research—an information systems perspective. Int. J. Inf. Manage. **47**, 88–100 (2019)

2. Asimakopoulou, E., Bessis, N.: Buildings and crowds: forming smart cities for more effective disaster management. In: 2011 Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, pp. 229–234 (2011)
3. Ang, K.L., Seng, J.K.P., Ngharamike, E., Ijemaru, G.K.: Emerging technologies for smart cities' transportation: geoinformation, data analytics and machine learning approaches. ISPRS Int. J. Geo Inf. **11**(2), 85 (2022)
4. Nevens, F., Frantzeskaki, N., Gorissen, L., Loorbach, D.: Urban Transition Labs: cocreating transformative action for sustainable cities. J. Clean. Prod. **50**, 111–122 (2013)
5. Pellicer, S., Santa, G., Bleda, A.L., Maestre, R., Jara, A.J., Skarmeta, A.G.: A global perspective of smart cities: a survey. In: 2013 Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, pp. 439–444 (2013)
6. Anand, A., Rufuss, D.D.W., Rajkumar, V., Suganthi, L.: Evaluation of sustainability indicators in smart cities for India using MCDM approach. Energy Procedia **141**, 211–215 (2017)
7. Adali, E.A., Öztaş, G.Z., Öztaş, T., Tuş, A.: Assessment of European cities from a smartness perspective: an integrated grey MCDM approach. Sustain. Cities Soc. **84**, 104021 (2022)
8. Şeker, Ş: IoT based sustainable smart waste management system evaluation using MCDM model under interval-valued q-rung orthopair fuzzy environment. Technol. Soc. **71**, 102100 (2022)
9. Hajduk, S.: Multi-criteria analysis of smart cities on the example of the Polish cities. Resources **10**(5), 44 (2021)
10. Ogrodnik, K.: The application of the PROMETHEE method in evaluation of sustainable development of the selected cities in Poland. Ekonomia i Środowisko **62**(3), 18 (2017)
11. Brans, J.P., Mareshal, B.: Multiple Criteria Decision Analysis, Chapter 5, Promethee Methods (1982)
12. Brans, J.P., Vincke, P.: A preference ranking organisation method. Manage. Sci. **31**(6), 647–656 (1985)
13. Akram, M., Shumaiza, S.: Multi-criteria decision making based on q-rung orthopair fuzzy promethee approach. Iran. J. Fuzzy Syst. **18**(5), 107–127 (2021)
14. Tong, L., Pu, Z., Chen, K., Yi, Z.: Sustainable maintenance supplier performance evaluation based on an extend fuzzy PROMETHEE II approach in petrochemical industry. J. Clean. Prod. **273**, 122771 (2020)
15. Wan, S.-P., Zou, W.-C., Zhong, L.-G., Dong, J.-Y.: Some new information measures for hesitant fuzzy PROMETHEE method and application to green supplier selection. Soft. Comput. **24**(12), 9179–9203 (2019). https://doi.org/10.1007/s00500-019-04446-w
16. Demircioğlu, M.E., Ulukan, H.Z.: A novel hybrid approach based on intuitionistic fuzzy multi criteria group-decision making for environmental pollution problem. J. Intell. Fuzzy Syst. **38**(1), 1013–1025 (2020)
17. Akram, M., Shumaiza, Alcantud, J.C.R.: An m-polar fuzzy PROMETHEE approach for AHP-assisted group decision-making. Math. Comput. Appl. **25**(2), 26 (2020)
18. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
19. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets Syst. **20**(1), 87–96 (1986)
20. Atanassov, K.: Circular intuitionistic fuzzy sets. J. Intell. Fuzzy Syst. **39**(5), 5981–5986 (2020)
21. Çakır, E., Taş, M.A.: Circular intuitionistic fuzzy decision making and its application. Expert Syst. Appl. **225**, 120076 (2023). https://doi.org/10.1016/j.eswa.2023.120076
22. Atanassov, K., Marinov, E.: Four distances for circular intuitionistic fuzzy sets. Mathematics **9**(10), 1121 (2021)
23. Khan, M.J., Kumam, W., Alreshidi, N.A.: Divergence measures for circular intuitionistic fuzzy sets and their applications. Eng. Appl. Artif. Intell. **116**, 105455 (2022)
24. Çakır, E., Taş, M.A., Ulukan, Z.: A new circular intuitionistic fuzzy MCDM: a case of Covid-19 medical waste landfill site evaluation. In: 2021 IEEE 21st International Symposium on Computational Intelligence and Informatics (CINTI), pp. 143–148 (2021)

25. Çakır, E., Taş, M.: Circular intuitionistic fuzzy multi-criteria decision making methodology. Avrupa Bilim ve Teknoloji Dergisi **28**, 900–905 (2021)
26. Çakır, E., Taş, M.A., Ulukan, Z.: Circular intuitionistic fuzzy sets in multi criteria decision making. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F.M. (eds.) ICSCCW 2021. LNNS, vol. 362, pp. 34–42. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-92127-9_9
27. Çakır, E., Taş, M.A.: Circular intuitionistic fuzzy analytic hierarchy process for remote working assessment in Covid-19. In: Kahraman, C., Tolga, A.C., Cevik Onar, S., Cebi, S., Oztaysi, B., Sari, I.U. (eds.) Intelligent and Fuzzy Systems: Digital Acceleration and the New Normal-Proceedings of the INFUS 2022 Conference, vol. 1. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09173-5_68
28. Otay, İ, Kahraman, C.: A novel circular intuitionistic fuzzy AHP&VIKOR methodology: an application to a multi-expert supplier evaluation problem. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi **28**(1), 194–207 (2022)
29. Kahraman, C., Otay, İ.: Extension of VIKOR method using circular intuitionistic fuzzy sets. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds.) Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation, Held 24–26 August 2021, Volume 2, INFUS 2021. LNNS, vol. 308, pp. 48–57. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-85577-2_6
30. Alkan, N., Kahraman, C.: Circular intuitionistic fuzzy TOPSIS method with vague membership functions: supplier selection application context. Notes Intuitionistic Fuzzy Set **27**(1), 24–52 (2021)
31. Alkan, N., Kahraman, C.: Circular intuitionistic fuzzy TOPSIS method: pandemic hospital location selection. J. Intell. Fuzzy Syst. **42**(1), 295–316 (2022)
32. Chen, T.Y.: Evolved distance measures for circular intuitionistic fuzzy sets and their exploitation in the technique for order preference by similarity to ideal solutions. Artif. Intell. Rev. 1–55 (2022)
33. Xu, Z., Liao, H.: Intuitionistic fuzzy analytic hierarchy process. IEEE Trans. Fuzzy Syst. **22**(4), 749–761 (2013)
34. Brans, J.P., Vincke, P., Mareshal, B.: How to select and how to rank projects: the PROMETHEE method. Eur. J. Oper. Res. **24**, 228–238 (1986)
35. Liao, H., Zeshui, X.: Multi-criteria decision making with intuitionistic fuzzy PROMETHEE. J. Intell. Fuzzy Syst. **27**(4), 1703–1717 (2014)

An International Hub Airport Selection Problem Using Fuzzy Analytic Hierarchy Process (F-AHP): Real Case Study in Turkey

Melis Tan Taçoğlu[1(✉)], Mert Paldrak[2], Mustafa Arslan Örnek[2], and Caner Taçoğlu[3]

[1] Logistics Management Department, Yaşar University, Bornova, İzmir, Turkey
melis.tacoglu@yasar.edu.tr
[2] Industrial Engineering Department, Yaşar University, Bornova, İzmir, Turkey
mert.paldrak@yasar.edu.tr
[3] Industrial Engineering Department, Izmir University of Economics, Balçova, İzmir, Turkey

Abstract. In recent years, air transportation volume has been significantly growing with the help of recent technological development related to the airline industry. Airline companies utilize their scarce resources such as aircraft, crew, and slot time based on airport regulation and capacity constraints efficiently to meet passenger demand; however, the limited capacity of the airport's immobile resources leads airlines to search for new hub alternatives. Hub selection is one of the most crucial decisions for airline companies' strategy and also the selected hub's future planning and operational strategy. This research aims to propose a solution method for selecting the best potential hub alternative among many potential hubs in a single hub system considering six main criteria: the potential hub's city population, the distance between the existing hub (İstanbul airport), the capacity, the intensity, the distance between city location, and the accessibility. F-AHP is used as a solution methodology to evaluate each criterion and the proposed hubs. The computational experiment is conducted based on the potential twelve international airports in Turkey for the selection of an international hub airport considering the six aforementioned criteria. The results indicate that accessibility of the airport is the most crucial criterion among the six criteria, and Izmir Adnan Menderes Airport is the best eligible alternative to be an international hub.

Keywords: Hub selection · F-AHP · Airport · Multi-Criteria Decision Modeling

1 Introduction

Air transportation has become an essential part of global travel, trade, and commerce. One of the key decisions for airlines is the selection of an airport hub, which serves as a central point for connecting flights and facilitating the transportation of passengers and cargo. The HS is a popular strategy that airlines use to maximize the efficiency of their operations. This network involves connecting multiple destinations through a central hub, which allows airlines to consolidate their operations and reduce costs

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 362–370, 2023.
https://doi.org/10.1007/978-3-031-39777-6_44

while providing greater connectivity to passengers. The HS is especially effective for international passengers as it enables them to travel to their final destination with fewer connections, reducing travel time and increasing convenience.

Airport hub selection is a critical decision for airlines as it determines their ability to efficiently transport passengers and cargo across multiple destinations. The selection process involves various factors such as airport infrastructure, location, capacity, operational efficiency, and access to key markets. Airlines must carefully evaluate these factors to select the optimal airport hub that can support their business goals and enhance their competitiveness. International passengers represent a significant market segment for airlines and airports, with a growing demand for air travel. The selection of an optimal airport hub and the implementation of an HS can provide airlines with a competitive advantage in attracting international passengers by offering more efficient and convenient travel options. In this study, hub selection is conducted solely considering international passengers.

In this study, 12 potential international airports are evaluated considering five main criteria to select an eligible hub, and these criteria are given in the following: the potential hub's city population, the distance between the existing hub (İstanbul airport), capacity, intensity, the distance between city location, operational cost, and accessibility. The Fuzzy Analytic Hierarchy Process (F-AHP) method is used to evaluate the hub airport selection problem. This study contributes to the hub selection literature in two ways: different criterion considerations compared to former articles, and Turkey-based real case study implementation. The computational results indicate that accessibility is the most essential criterion among all criteria. Moreover, Izmir Adnan Menderes International Airport is chosen as the best alternative to be a hub airport to serve international passengers. This paper is organized as a literature review, fuzzy analytical hierarchy process, comparative case study, computational results, and conclusion, respectively.

2 Literature Review

Since the regulation of the domestic air transport industry in the United States changed in 1978, hub-and-spoke (HS) networks have become an essential component of airline networks globally. According to Gillen et al. [1], the most effective utilization of the HS network occurs when airlines have a broad air transportation network and numerous destinations in their network. Some of the well-known global companies such as Delta, Lufthansa, Turkish Airlines, and Emirates use HS as a network business model [2].

Passenger demand in the aviation industry is growing exponentially, and airlines deal with complex problems such as network planning, fleet scheduling, and flight schedules. Airline companies must make significant decisions, and the best decision is selected considering various alternatives and criteria [3]. Multi-Criteria Decision Making (MCDM) is selecting the best choice among many alternatives, including many criteria and decision-makers [4]. According to [3], most airline planners frequently use MCDM methodologies for the decision-making process while evaluating different alternatives or criteria. Indeed, MCDM is used in the literature in the aviation sector to decide on different types of problems, such as efficiency assessment in airlines [5], international new route selection [4], service quality [6], and financial performance [7]. [3] reviewed 166 papers that were published between 2000 to 2018, and decisions on the problem must be made among many criteria and alternatives in air transportation. He claimed that among the articles related to the airline, 32% of them related to service quality, 20% of them related to safety, 13% of them related to a partner, 11% of them related to a fleet, 11% of them related with competition, 6% of them affiliated with financial performance and remaining 6% articles are related with other issues such as route [8], landing path figuration for a given air station [9].

The hub location problem was first introduced by O'Kelly [10]. In the hub location selection literature, various types of this problem were studied such as taking into consideration of various allocation versions, capacity limitations, network features, etc. The uncapacitated multiple allocations p-hub median problems (UMAp-HM), and is one of the most well-established problems in this literature [11]. This problem aims to minimize the total transportation cost while locating p numbers of hubs and distributing airports to hubs in the airline flight network [11]. Although many heuristic and meta-heuristic approaches have been developed to solve one of the most popular types of the hub location selection problem which is UMAp-HMP, hub location problems with more than one objective are relatively scant in the literature [11, 12]. The Journal of Air Transport Management underlines the importance of MCDM and Multiple Objective Optimization (MOO) applications in the aviation industry, and this journal is the first to publish in this field [3].

3 Fuzzy Analytical Hierarchy Process

Fuzzy Analytical Hierarchy Process (F-AHP) employs the fuzz theory with basic Analytical Hierarchy Process (AHP) which was first introduced by Saaty [13]. AHP is a vastly used decision-making technique to solve various multi-criteria decision-making problems ranging from supplier selection to worker recruitment problems. It considers pair-wise comparisons of different alternatives concerning some criteria and provides a decision-support tool for multi-criteria decision-making problems.

Basic AHP does not involve vagueness due to personal judgments, it has been ameliorated with the help of the fuzzy logic approach. In F-AHP, the pair-wise comparisons of both criteria and alternatives are performed through linguistic variables. Most F-AHP applications consider the triangular fuzzy membership functions for pair-wise comparisons. The steps of the F-AHP employed in this study are recapitulated as follows:

Step 1: A decision-maker compares the criteria or alternatives using linguistic terms shown in Table 1.

Based on the corresponding triangular fuzzy members of these linguistic terms, for example, if the decision maker claims that "Criterion 1 (C1) is Strongly Important than Criterion 2 (C2)" then it takes the fuzzy triangular scale as (6, 7, 8). On the contrary, the pair-wise comparison matrix of the criteria of Criterion 2 to Criterion 1, takes the fuzzy triangular scale as (1/8, 1/7, 1/6).

Table 1. Linguistic Variables of This Study

Saaty Scale	Definition	Fuzzy Triangular Scale
1	Equally Important	(1, 1, 1)
3	Weakly Important	(2, 3, 4)
5	Fairly Important	(4, 5, 6)
7	Strongly Important	(6, 7, 8)
9	Absolutely Important	(8, 9, 9)

The pair-wise contribution matrix is shown in Eq. 1, where (a_ij^k)˜ indicates the preference of k^th decision member for i^th criterion over j^th criterion, via fuzzy triangular numbers.

$$\widetilde{A_k} = \begin{bmatrix} \widetilde{a_{11}^k} & \widetilde{a_{12}^k} & \dots & \widetilde{a_{1n}^k} \\ a_{21}^k & \dots & \dots & a_{2n}^k \\ \dots & \dots & \dots & \dots \\ \widetilde{a_{n1}^k} & \widetilde{a_{n1}^k} & \dots & a_{nn}^k \end{bmatrix} \tag{1}$$

Step 2: In case there is more than one decision maker, the preferences of each decision maker are averaged $\left(\tilde{a}_{ij}\right)$ and calculated by the Eq. 2.

$$\tilde{a}_{ij} = \frac{\sum_{k=1}^{K} \widetilde{a_{ij}^k}}{K} \tag{2}$$

Step 3: According to these preferences, pair-wise contribution matrix is updated as in Eq. 3 below. According to these preferences, the pair-wise contribution matrix is updated as in Eq. 3 below:

$$\tilde{A} = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \dots & \dots & \tilde{a}_{2n} \\ \dots & \dots & \dots & \dots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \dots & \tilde{a}_{nn} \end{bmatrix} \tag{3}$$

$$\tilde{a}_{ij} = \left(l_{ij}, m_{ij}, u_{ij}\right)$$

$$\tilde{a}_{ji} = (\tilde{a}_{ij})^{-1} = \left(u_{ij}^{-1}, m_{ij}^{-1}, l_{ij}^{-1}\right) \tag{4}$$

$$l_{ij} = \left(\prod_{t=1}^{k} a_{l_{ij}}^{t}\right)^{1/k} \approx t = 1, 2, \dots, k \tag{5}$$

$$m_{ij} = \left(\prod\nolimits_{t=1}^{k} a_{m_{ij}}^{t}\right)^{1/k} \approx t = 1, 2, \ldots, k \tag{6}$$

$$u_{ij} = \left(\prod\nolimits_{t=1}^{k} a_{u_{ij}}^{t}\right)^{1/k} \approx t = 1, 2, \ldots, k \tag{7}$$

$A^k = \left(a_{l_{ij}}^k, a_{m_{ij}}^k, a_{u_{ij}}^k\right)$ is the fuzzy comparison matrix from decision maker k by comparing criteria i and j.

Step 4: All alternatives are compared based on each criterion using linguistic variables in Table….

Step 5: For each alternative based on a criterion, the value of fuzzy synthetic extent can be calculated as follows:

$$\begin{aligned} S_{ij} &= \sum\nolimits_{j=1}^{m} a_{ij} \int \left[\sum\nolimits_{l=1}^{m} \sum\nolimits_{j=1}^{n} a_{lj}\right]^{-1} \\ &= \left(\frac{\sum_{j=1}^{m} l_{ij}}{\sum_{l=1}^{m} \sum_{j=1}^{n} u_{lj}}, \frac{\sum_{j=1}^{m} m_{ij}}{\sum_{l=1}^{m} \sum_{j=1}^{n} m_{lj}}, \frac{\sum_{j=1}^{m} u_{ij}}{\sum_{k=1}^{m} \sum_{j=1}^{n} l_{lj}}\right) \approx i \end{aligned} \tag{8}$$

Step 6: Since the weights of the criteria are still fuzzy, they need to bed defuzzified and normalized. After normalization of the weight vector, the importance weights of criteria are as follows:

$$W = (W_1, W_2, \ldots, W_n)^T \tag{9}$$

Step 7: Since S_{ij} are still fuzy triangular members, they need to be defuzzified using the formula below:

$$S_{ij} = \left(\frac{l_{ij} + 2\nabla m_{ij} + u_{ij}}{3}\right) \quad i = 1, 2, .., n, j = 1, 2.., m \tag{10}$$

Step 8: Calculate the overall score of each alternative using Eq. (11)

$$V_i = \sum\nolimits_{i=1}^{n} S_{ij} \nabla W_i \quad i = 1, 2..m \tag{11}$$

Step 9: The alternative with the highest score is selected.

4 Comparative Case Study

In this paper, a decision-making process is handled in Turkey where the aim is to select an international hub location to meet incoming and outgoing passengers' demands and increase convenience to them. In this part, the outlines of hub selection and the extent analysis with fuzzy AHP are given for this case study.

Turkey is a large country where there are two existing hub airports, namely Istanbul International Airport and Esenboga International Airport to serve domestic and international passengers throughout the world. However, due to the increase in air traffic volume, Turkey needs one more hub-and-spoke airport by which passengers can transit to their final destination. This study aims to choose another hub airport among all possible alternatives while considering various criteria. These criteria are determined after holding a meeting with experts of this study and they are named: airport intensity, the distance between a candidate hub airport and existing hub airports, the population of the city where the candidate airport is situated, airport capacity, airport location, and airport accessibility.

There are numerous domestic and international airports in Turkey, yet the most international passenger traffic observed in the top 12 international airports between March 2022 and March 2023 are selected as candidate hubs in this study. The candidate airports are given as follows: Izmir Adnan Menderes International Airport, Antalya International Airport, Dalaman International Airport, Kayseri International Airport, Adana International Airport, Trabzon International Airport, Gaziantep International Airport, Alanya International Airport, Diyarbakır International Airport, Hatay International Airport, Konya International Airport, and Samsun International Airport.

The first criterion is airport intensity and it is associated with the international passenger traffic of the airport. To measure airport intensity, the total number of international passenger arrival and departure between March 2022–2023 for each candidate airport is considered and provided by the General Directorate of State Airports Authority (DHMI). The second criterion, namely the distance between the existing hub airports and a candidate hub is measured by averaging the distance between each candidate airport and operating hubs through Google Maps. The third criterion is city population which might be proportional to the international market size. The pertinent data is obtained through the Turkish Statistical Institution (TUIK). The next criterion is the existing capacities of the candidate airports. Since this study focuses on finding the most eligible international hub airport in Turkey among all alternatives, airport capacities for serving international passengers are taken into consideration. The data related to airport capacity for each candidate is provided by DHMI. The fifth criterion is airport location where the distance between the city center and the candidate airport is considered. The relevant distance values are collected using Google Maps. The last criterion is the accessibility of candidate hubs which is associated with the number of public transports between the city to the airport. The necessary data is found online through DHMI.

The AHP method is a useful technique to analyze complex decision-making problems by dividing them into three layers: object layer, criteria layer, and alternatives layer. When we assess our hub selection problem, the first layer is the object layer which is the aim of this study. The second layer involves all six criteria, namely intensity, distance, population, capacity, location, and accessibility. The third layer consists of all possible alternatives among which the best candidate will be selected. Based on the information and analysis above, a hierarchical structure for the hub selection problem is established in Fig. 1.

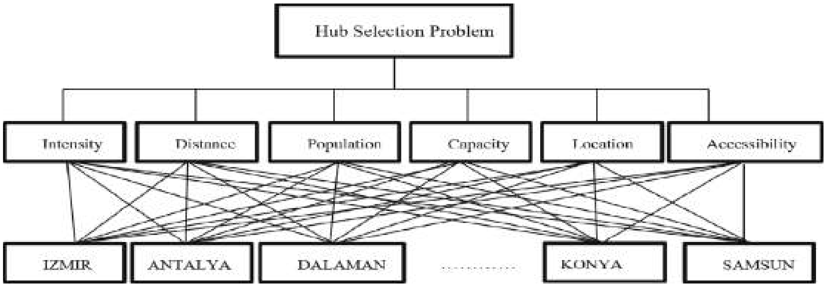


Fig. 1. Hierarchical Structure of Hub Selection Problem

5 Computational Results

In this study, there are three decision-makers whose evaluations are taken into consideration. Firstly, each decision maker has performed pair-wise comparisons of all criteria, and their evaluations are aggregated to obtain the resulting weight of each criterion. Based on the F-AHP technique, the weight of each criterion is calculated and given in the table below Table 2. According to evaluations of decision-makers, the most essential criterion is accessibility which highly influences the selection of a hub airport. In addition to accessibility, airport capacity, and intensity are important factors as well.

Table 2. Criteria Weights

Criteria	Criteria Weights
Intensity (I)	0,20
Distance (D)	0,05
Population (P)	0,07
Capacity (C)	0,27
Location (L)	0,09
Accessibility (A)	0,34

After pair-wise comparisons of each alternative based on each criterion, the resulting table is attained in which the overall scores of each alternative are given. To calculate the final column of this table, each alternative-criterion score is multiplied by the corresponding criterion weight and summed up. In Table 3, the resulting table is provided.

Result shows that the most suitable airport to be used as a hub is chosen as Izmir Adnan Menderes Airport. Moreover, Antalya International Airport is very close to the best alternative, this airport can also be thought of as an alternative hub.

Table 3. Alternative-Criterion Pairwise Comparison Results

Airports	I	D	P	C	L	A	V_i
İzmir	0,16	0,05	0,18	0,18	0,09	0,18	0,16
Antalya	0,22	0,06	0,10	0,22	0,09	0,10	0,15
Dalaman	0,05	0,06	0,05	0,05	0,03	0,18	0,09
Kayseri	0,10	0,05	0,05	0,08	0,09	0,06	0,08
Adana	0,10	0,07	0,10	0,12	0,09	0,06	0,09
Trabzon	0,07	0,12	0,07	0,08	0,09	0,07	0,08
Gaziantep	0,07	0,12	0,08	0,06	0,09	0,06	0,07
Alanya	0,05	0,06	0,08	0,04	0,09	0,06	0,06
Diyarbakır	0,05	0,18	0,06	0,05	0,09	0,06	0,06
Hatay	0,05	0,11	0,06	0,04	0,09	0,06	0,06

6 Conclusion

This study aims to support airline decision-makers in their new hub investment strategies by considering various criteria. The problem is formulated as a multi-criteria decision-making problem where there exist twelve alternatives and six main criteria to evaluate them. To solve this problem, F-AHP is employed where pair-wise comparisons are performed using linguistic variables. Based on the results obtained through F-AHP, Izmir Adnan Menderes Airport is chosen as the best alternative to be a hub and the accessibility is the most important criterion. As a future direction, the problem can be solved using other techniques such as Fuzzy TOPSIS, Fuzzy PROMETEE, etc.

References

1. Gillen, D., Morrison, W.G.: Regulation, competition and network evolution in aviation. J. Air Transp. Manag. **11**(3), 161–174 (2005)
2. Çiftçi, M.E., Özkır, V.: Optimising flight connection times in airline bank structure through Simulated Annealing and Tabu Search algorithms. J. Air Transp. Manag. **87**, 101858 (2020)
3. Dožić, S.: Multi-criteria decision making methods: application in the aviation industry. J. Air Transp. Manag. **79**, 101683 (2019)
4. Deveci, M., Demirel, N.C., Ahmetoğlu, E.: Airline new route selection based on interval type-2 fuzzy MCDM: a case study of new route between Turkey-North American region destinations. J. Air Transp. Manag. **59**, 83–99 (2017)
5. Wanke, P., Barros, C.P., Chen, Z.: An analysis of Asian airlines efficiency with two-stage TOPSIS and MCMC generalized linear mixed models. Int. J. Prod. Econ. **169**, 110–126 (2015)
6. Tsafarakis, S., Kokotas, T., Pantouvakis, A.: A multiple criteria approach for airline passenger satisfaction measurement and service quality improvement. J. Air Transp. Manag. **68**, 61–75 (2018)
7. Pineda, P.J.G., et al.: An integrated MCDM model for improving airline operational and financial performance. J. Air Transp. Manag. **68**, 103–117 (2018)

8. Yilmaz, A.K., et al.: Modeling regional routes with Greek airlines for flight operations to AOE airport. Aircr. Eng. Aerosp. Technol. **90**(8), 1227–1237 (2018)
9. Rezaei, J., Hemmes, A., Tavasszy, L.: Multi-criteria decision-making for complex bundling configurations in surface transportation of air freight. J. Air Transp. Manag. **61**, 95–105 (2017)
10. O'kelly, M.E.: The location of interacting hub facilities. Transp. Sci. **20**(2), 92–106 (1986)
11. Soylu, B., Katip, H.: A multiobjective hub-airport location problem for an airline network design. Eur. J. Oper. Res. **277**(2), 412–425 (2019)
12. Farahani, R.Z., et al.: Hub location problems: a review of models, classification, solution techniques, and applications. Comput. Ind. Eng. **64**(4), 1096–1109 (2013)
13. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**(3), 234–281 (1977)

A Novel Continuous Intuitionistic Fuzzy CODAS Method and Its Application to IOT Service Provider Selection for Sustainable Supply Chain

Nurşah Alkan(✉) and Cengiz Kahraman

Department of Industrial Engineering, Istanbul Technical University, Macka, Besiktas, 34367 Istanbul, Turkey
nalkan@itu.edu.tr

Abstract. Intuitionistic fuzzy sets (IFSs), which are the most widely used fuzzy set extension in the literature, are represented by membership and non-membership degrees. Since IFSs and their direct extensions are usually expressed with crisp membership and non-membership degrees, it is known that the IFSs existing in the literature are discrete IFSs or linear continuous IFSs. However, although IFSs have been used to tackle different real-life problems, it is often difficult to estimate crisp values for membership and non-membership degrees. Therefore, in this study, it is aimed to develop a multi-criteria decision-making model based on novel nonlinear continuous intuitionistic fuzzy sets. CINFUSs, consisting of membership and non-membership degrees represented by second-order nonlinear functions, have been used to develop the CODAS method in a fuzzy environment. The CINFUS-based CODAS method developed in the study has been applied to the Internet of things (IOT) service provider selection for sustainable supplier selection.

Keywords: Continuous Intuitionistic Fuzzy Sets · CINFUS · CODAS · IOT · sustainable supply chain

1 Introduction

Since the information about today's real-life problems is generally not crisp and cannot be explained with deterministic models, getting the right decisions in the problems considered can cause errors. For this reason, it is necessary to use and develop decision making methods based on fuzzy sets in order to take into account the uncertainty and precision in the data. After the fuzzy sets were developed by Zadeh [1], many extensions have been put forward by various researchers, but the biggest innovation in fuzzy set extensions has been the development of intuitionistic fuzzy sets (IFSs) defined with membership and non-membership degrees by Atanassov [2]. IFSs are represented by two parameters, membership and non-membership degrees satisfying the condition that the sum of the membership and non-membership degrees of each element in a fixed set of the universe is at most one [3]. Although IFSs have been considered for different problems in many studies, it is often difficult to assign single membership and non-membership

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 371–381, 2023.
https://doi.org/10.1007/978-3-031-39777-6_45

degrees [4]. In addition, intuitionistic fuzzy sets based on nonlinear membership and non-membership functions have not been developed in the literature. Moreover, these degrees must be of a continuous structure in order to take into account the uncertainty at a continuous interval.

We proposed continuous intuitionistic fuzzy sets (CINFUSs) and used them in developing a new fuzzy extension of CODAS method which is an effective method among MCDM methods. Thanks to their flexible structure, CINFUSs have the ability to model the hesitancy of decision makers better than IFSs, and they can add uncertainty to fuzzy membership and non-membership functions, as in Gaussian fuzzy sets. With CINFU-CODAS method, the alternatives are ranked in a more flexible structure according to their distance from the negative ideal solutions to find the most suitable solution. In the study, the developed method is implemented to IoT service provider selection for sustainable supply chain.

The remaining of this paper is organized as follows. Section 2 presents continuous intuitionistic fuzzy sets and CINFU-CODAS method. The developed CINFU-CODAS method is applied to IoT based service provider selection for sustainable supply chain and comparative analysis is presented in Sect. 3. Finally, the study ends in Sect. 4 with conclusions and recommendations for further research.

2 Continuous Intuitionistic Fuzzy CODAS

2.1 Preliminaries of Continuous Intuitionistic Fuzzy Sets

This section includes the basic definitions related to CINFUSs. The basic operations of CINFUSs are mentioned as follows.

Definition 1. Let X be a fixed set. A CINFUS C in X is an object having the form:

$$C = \{\langle x, F(\mu_C(x), \tau), F(\vartheta_C(x), \tau)\rangle | x \in X\} \tag{1}$$

where the function $F(\mu_C(x), \tau) : X \to [0, 1]$ and $F(\vartheta_C(x), \tau) : X \to [0, 1]$ denote continuous membership and non-membership degrees of x to the set C, respectively. Also, where $F(\mu_C(x), \tau) = \alpha\tau - \alpha\tau^2$ and $F(\vartheta_C(x), \tau) = (4-\alpha)\tau - (4-\alpha)\tau^2$. It satisfies the condition

$$\begin{gathered} 0 \leq F(\mu_C(x), \tau) + F(\vartheta_C(x), \tau) \leq 1 = 0 \leq \alpha\tau - \alpha\tau^2 \\ + (4-\alpha)\tau - (4-\alpha)\tau^2 \leq 1 = 0 \leq 4\tau - 4\tau^2 \leq 1, \ \mathit{for}\ \forall\ x \in X \end{gathered} \tag{2}$$

The degree of uncertainty in judgement is defined as $F(\pi, \tau) = 1 - \alpha\tau + \alpha\tau^2 - (4-\alpha)\tau + (4-\alpha)\tau^2 = 1 - 4\left(\tau - \tau^2\right)$.

Definition 2. Let $CINFUS_A = \left\langle\left(\alpha_A\tau - \alpha_A\tau^2\right), \left((4-\alpha_A)\tau - (4-\alpha_A)\tau^2\right)\right\rangle$ and $CINFUS_B = \left\langle\left(\alpha_B\tau - \alpha_B\tau^2\right), \left((4-\alpha_B)\tau - (4-\alpha_B)\tau^2\right)\right\rangle$ be two CINFUSs, then the mathematical operations on these two continuous intuitionistic fuzzy numbers (CINFUNs) are defined as follows:

$$\begin{aligned} & CINFUS_A \oplus CINFUS_B = \\ & \Big\langle\Big((\alpha_A+\alpha_B)\tau - (\alpha_A+\alpha_B)\tau^2 - \left(\alpha_A\tau - \alpha_A\tau^2\right)\left(\alpha_B\tau - \alpha_B\tau^2\right)\Big), \\ & \left((4-\alpha_A)\tau - (4-\alpha_A)\tau^2 \left((4-\alpha_B)\tau - (4-\alpha_B)\tau^2\right)\right)\Big\rangle \end{aligned} \tag{3}$$

$$\begin{aligned}&CINFUS_A \otimes CINFUS_B = \left(\alpha_A\tau - \alpha_A\tau^2\right)\left(\alpha_B\tau - \alpha_B\tau^2\right), (4-\alpha_A)\tau - (4-\alpha_A)\tau^2 \\ &+(4-\alpha_B)\tau - (4-\alpha_B)\tau^2 - \left((4-\alpha_A)\tau - (4-\alpha_A)\tau^2\right)\left((4-\alpha_B)\tau - (4-\alpha_B)\tau^2\right) \\ &= \left\langle\left((\alpha_A+\alpha_B)\left(\tau-\tau^2\right)\right), \left((8-\alpha_A-\alpha_B)\left(\tau-\tau^2\right) - (4-\alpha_A)(4-\alpha_B)\left(\tau-\tau^2\right)^2\right)\right\rangle\end{aligned} \tag{4}$$

$$\lambda \times CINFUS_A = \left\langle\left(1-\left(1-\alpha_A\tau+\alpha_A\tau^2\right)^{\lambda}\right), \left((4-\alpha_A)^{\lambda}\left(\tau-\tau^2\right)^{\lambda}\right)\right\rangle \tag{5}$$

$$CINFUS_A^{\lambda} = \left\langle\left(\alpha_A\tau-\alpha_A\tau^2\right)^{\lambda}, \left(1-\left(1-(4-\alpha_A)\left(\tau-\tau^2\right)\right)^{\lambda}\right)\right\rangle \tag{6}$$

Definition 3. Let $\tilde{A}_i = \left\langle\left(\alpha\tau - \alpha\tau^2\right)_i, \left((4-\alpha)\tau - (4-\alpha)\tau^2\right)_i\right\rangle, (i = 1, 2, \ldots\ldots, n)$ be a set of CINFUSs and $w = (w_1, w_2, \ldots., w_n)^T$ be weight vector of $\tilde{A}_i$ with $\sum_{i=1}^{n} w_i = 1$, then a continuous intuitionistic fuzzy weighted geometric (*CINFUWG*) the operator is a mapping (CINFUWG): $\tilde{A}^n \rightarrow \tilde{A}$, where.

$$\begin{aligned}&CINFUWG\left(\tilde{A}_1, \tilde{A}_2, \ldots., \tilde{A}_n\right) \\ &= \left(\prod_{i=1}^{n}\left(\alpha\tau-\alpha\tau^2\right)_i^{w_i}, 1-\prod_{i=1}^{n}\left(1-\left((4-\alpha)\tau-(4-\alpha)\tau^2\right)_i\right)^{w_i}\right)\end{aligned} \tag{7}$$

2.2 Proposed Method: CINFU-CODAS

Step 1. Construct the decision matrices from each expert E_k where $k = 1, 2, \ldots K$. The set of alternatives is given by $A_i = \{A_1, A_2, \ldots\ldots, A_m\}$. Each alternative is evaluated based on the set of criteria denoted by $C_j = \{C_1, C_2, \ldots\ldots, C_n\}$ where their weights are indicated by $w_j = (w_1, w_2, \ldots., w_n)$ where $w_j > 0$ and $\sum_{j=1}^{n} w_j = 1$ and also, $i = 1, 2, \ldots., m$, $j = 1, 2, \ldots., n$.

Step 2. Collect the importance assessments for criteria and the decision matrices from each expert. The importance assessments and decision matrices are completed by using the linguistic terms given in Table 1. Then, the linguistic variables expressed for both criteria and alternatives are converted to CINFU values, CINFU criteria weight matrix $\tilde{\Phi}_k = \left(\tilde{0}_{jk}\right)_{n\times 1}$ and CINFU decision matrix $\tilde{\Psi}_k = \left(\tilde{\psi}_{ijk}\right)_{n\times m}$ based on $k^{th}E$ is constructed.

Step 3. Obtain the aggregated weights for criteria by using CINFUWG operator given in Eq. (7). The individual criteria weight matrices of each expert are aggregated in a single criteria weight matrix. Aggregated criteria weight matrix $\tilde{\Omega} = \left(\tilde{\omega}_j\right)_{n\times 1}$ is constructed for each τ value, whose ranges from zero to one are determined by the expert. Here, $\tilde{\Omega} = \left(\tilde{\omega}_j\right)_{n\times 1}$ in which $\tilde{\omega}_j = \left(\mu_{j,\tau}, \vartheta_{j,\tau}\right)$ show the aggregated CINFU numbers for criteria according to j^{th} criterion.

Step 4. Obtain the aggregated decision matrix $\tilde{X} = \left(\tilde{\chi}_{ij}\right)_{n\times m}$ by using CINFUWG operator given in Eq. (7). The individual decision matrices of experts are integrated to aggregate in a single decision matrix according to each τ value. Here, $\tilde{X} = \left(\tilde{\chi}_{ij}\right)_{n\times m}$

Table 1. Linguistic scale for criteria weighting and decision matrix

Linguistic terms for criteria	Linguistic terms for alternatives	CINFUNs	
		μ	ϑ
Absolutely less important than – (ALI)	Absolutely Low Value – (ALV)	$0\tau - 0\tau^2$	$4\tau - 4\tau^2$
Very strongly less important than – (VSLI)	Very Low Value – (VLV)	$0.5\tau - 0.5\tau^2$	$3.5\tau - 3.5\tau^2$
Strongly Less important than – (StLI)	Low Value – (LV)	$\tau - \tau^2$	$3\tau - 3\tau^2$
Slightly less important than – (SlLI)	Slightly Low Value – (MLV)	$1.5\tau - 1.5\tau^2$	$2.5\tau - 2.5\tau^2$
Exactly equal importance – (EEI)	Medium Value – (AV)	$2\tau - 2\tau^2$	$2\tau - 2\tau^2$
Slightly more important than – (SlMI)	Slightly High Value – (MHV)	$2.5\tau - 2.5\tau^2$	$1.5\tau - 1.5\tau^2$
Strongly more important than – (StMI)	High Value – (HV)	$3\tau - 3\tau^2$	$1\tau - 1\tau^2$
Very strongly more important than – (VSMI)	Very High Value – (VHV)	$3.5\tau - 3.5\tau^2$	$0.5\tau - 0.5\tau^2$
Absolutely more important than – (AMI)	Absolutely High Value – (AHV)	$4\tau - 4\tau^2$	$0\tau - 0\tau^2$

in which $\tilde{\chi}_{ij} = \left(\mu_{ij,\tau}, \vartheta_{ij,\tau}\right)$ indicate the aggregated CINFU numbers of i^{th} alternative according to j^{th} criterion.

Step 5. Obtain the normalized aggregated decision matrix $\tilde{X}^N = \left(\tilde{\chi}_{ij}^N\right)_{n\times m}$ by replacing the membership and non-membership degrees for cost criteria. It is not necessary to make this replacement for benefit criteria. The normalized values for aggregated decision matrix are computed as presented in Eq. (8).

$$\tilde{\Psi}^N = \left(\tilde{\psi}_{ij}^N\right) = \begin{cases} \left(\mu_{ij,\tau}, \vartheta_{ij,\tau}\right), \textit{for benefit criteria} \\ \left(\vartheta_{ij,\tau}, \mu_{ij,\tau}\right), \textit{for cost criteria} \end{cases} \tag{8}$$

Step 6. Calculate the weighted normalized aggregated decision matrix $\tilde{X}_w^N = \left(\tilde{\chi}_{w,ij}^N\right)_{n\times m}$ by Eq. (9).

$$\tilde{X}_w^N = \left(\tilde{\chi}_{w,ij}^N\right) = \tilde{\omega}_j \otimes \tilde{\chi}_{w,ij}^N \tag{9}$$

where $\tilde{\chi}_{w,ij}^N = \left(\mu_{ij,\tau}^w, \vartheta_{ij,\tau}^w\right)$ shows the weighted normalized CINFU decision matrix of *ith* alternative with respect to *jth* criterion.

Step 7. Determine the CINFU negative ideal solution ($CINFU - NIS$) based on the weighted normalized aggregated decision matrix by Eq. (10).

$$(CINFU - NIS) = \tilde{\zeta}_j = \min_i \left(\tilde{\chi}^N_{w,ij} \right) \tag{10}$$

where $\tilde{\zeta}_j^- = \left(\mu^{w^-}_{ij,\tau}, \vartheta^{w^-}_{ij,\tau} \right)$ is the minimum weighted normalized aggregated CINFU numbers with the lowest score value among alternatives for j^{th} criterion.

Step 8. Compute the Euclidean and taxicab distances of the alternatives to the CINFU-NIS by using Eqs. (11) and (12).

$$\gamma_i^E = \sqrt{\frac{1}{2n} \sum\nolimits_{j=1}^{n} \left(\begin{array}{l} \left(\left(\mu^{w}_{ij,\ \tau} \right)^2 - \left(\mu^{w^-}_{ij,\ \tau} \right)^2 \right)^2 + \left(\left(\vartheta^{w}_{ij,\ \tau} \right)^2 - \left(\vartheta^{w^-}_{ij,\ \tau} \right)^2 \right)^2 \\ + \left(\left(\pi^{w}_{ij,\ \tau} \right)^2 - \left(\pi^{w^-}_{ij,\ \tau} \right)^2 \right)^2 \end{array} \right)} \tag{11}$$

$$\gamma_i^T = \sqrt{\frac{1}{2n} \sum_{j=1}^{n} \left(\left| \left(\mu^{w}_{ij,\ \tau} \right)^2 - \left(\mu^{w^-}_{ij,\ \tau} \right)^2 \right| + \left| \left(\vartheta^{w}_{ij,\ \tau} \right)^2 - \left(\vartheta^{w^-}_{ij,\ \tau} \right)^2 \right| + \left| \left(\pi^{w}_{ij,\ \tau} \right)^2 - \left(\pi^{w^-}_{ij,\ \tau} \right)^2 \right| \right)} \tag{12}$$

Step 9. Determine the relative assessment matrix $\Phi = (\xi_{ik})_{mxm}$ by using Eqs. (13).

$$\xi_{ik} = \left(\gamma_i^E - \gamma_k^E \right) + \left(\xi \left(\gamma_i^E - \gamma_k^E \right) \times \left(\gamma_i^T - \gamma_k^T \right) \right) \tag{13}$$

where $k = \{1, 2, \ldots, n\}$. ξ is a threshold function that is given in Eq. (14) and θ is a threshold parameter of this function which can be set by the expert.

$$\xi(x) = \begin{cases} 1, \ if\ |x| \geq \theta \\ 0, \ if\ |x| < \theta \end{cases} \tag{14}$$

Step 10. Calculate the assessment scores of each alternative by using Eq. (15).

$$\Omega_i = \sum\nolimits_{k=1}^{m} \xi_{ik} \tag{15}$$

Step 11. Rank the alternatives according to the decreasing values of assessment scores (Ω_i). The alternative with the highest Ω_i is the best alternative.

3 Application

3.1 Problem Definition

A firm is to choose an IOT service provider among five alternatives based on seven criteria. Three experts whose experiences are different evaluate the alternatives and criteria. Expert weights have been determined as (0.5, 0.3, 0.2) based on the experiences. Three experts assess the criteria by using the linguistic terms in Table 1 as given in Table 2. Seven criteria have been determined based on expert opinions and a comprehensive literature review to evaluate five IoT service providers (A1, A2, A3, A4, A5) as presented in the hierarchy in Fig. 1. The decision matrices collected from three experts are given in Table 3.

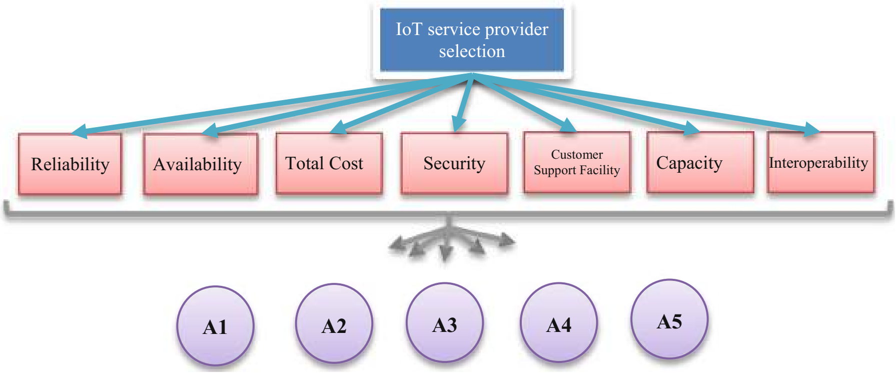


Fig. 1. Hierarchical structure of the problem

Table 2. Assessments of criteria by experts

	E1	E2	E3
C1	MI	SHI	MI
C2	MI	SHI	HI
C3	SLI	LI	LI
C4	HI	HI	VHI
C5	SHI	HI	VHI
C6	HI	MI	SLI
C7	VHI	AHI	VHI

Table 3. Decision matrices

	E1					E2					E3				
	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5
C1	LV	SLV	HV	VLV	VHV	SLV	LV	VHV	LV	AHV	SLV	LV	VHV	VLV	HV
C2	HV	LV	VHV	HV	SHV	MV	MV	SHV	SHV	HV	HV	SLV	HV	VHV	HV
C3	SHV	MV	LV	MV	LV	HV	SHV	MV	SHV	VLV	VHV	HV	SLV	HV	SLV
C4	VLV	HV	SLV	VLV	MV	SLV	VHV	LV	ALV	MV	SLV	HV	LV	VLV	SHV
C5	MV	VHV	AHV	ALV	ALV	MV	HV	AHV	LV	VLV	HV	SHV	VHV	LV	MV
C6	VHV	HV	HV	VHV	SHV	AHV	MV	VHV	AHV	HV	VHV	SHV	SHV	HV	MV
C7	MV	LV	MV	HV	MV	LV	VLV	SLV	SHV	HV	MV	SLV	MV	VHV	SHV

3.2 Problem Solution

Step 1. The linguistic assessments in Table 2 and Table 3 are aggregated by Eq. (7). The aggregated results are given in Tables 4 and 5, respectively.
Step 2. The normalized decision matrix is presented in Table 6.

Table 4. Aggregated criteria weights matrix

	C1		C7	
τ	μ	ϑ	μ	ϑ
0.1	0.192	0.165	0.328	0
0.2	0.342	0.294	0.583	0
0.3	0.449	0.385	0.765	0
0.4	0.513	0.440	0.874	0
0.5	0.535	0.459	0.911	0
0.6	0.513	0.440	0.874	0
0.7	0.449	0.385	0.765	0
0.8	0.342	0.294	0.583	0
0.9	0.192	0.165	0.328	0

Table 5. Aggregated decision matrix

		A1		A2		A3		A4		A5	
	τ	μ	ϑ	μ	ϑ	μ	ϑ	μ	ϑ	μ	ϑ
C1	0.1	0.110	0.246	0.110	0.246	0.292	0.064	0.055	0.301	0.318	0
	0.2	0.196	0.438	0.196	0.438	0.518	0.113	0.098	0.535	0.565	0
	0.3	0.257	0.575	0.257	0.575	0.680	0.148	0.129	0.702	0.742	0
	0.4	0.294	0.657	0.294	0.657	0.778	0.170	0.148	0.802	0.848	0
	0.5	0.306	0.685	0.306	0.685	0.810	0.177	0.154	0.835	0.883	0
	0.6	0.294	0.657	0.294	0.657	0.778	0.170	0.148	0.802	0.848	0
	0.7	0.257	0.575	0.257	0.575	0.680	0.148	0.129	0.702	0.742	0
	0.8	0.196	0.438	0.196	0.438	0.518	0.113	0.098	0.535	0.565	0
	0.9	0.110	0.246	0.110	0.246	0.292	0.064	0.055	0.301	0.318	0
C7	0.1	0.146	0.157	0.079	0.273	0.165	0.192	0.264	0.088	0.213	0.138
	0.2	0.260	0.279	0.141	0.485	0.294	0.342	0.469	0.157	0.378	0.245
	0.3	0.341	0.366	0.185	0.636	0.385	0.449	0.615	0.206	0.496	0.322
	0.4	0.390	0.418	0.211	0.727	0.440	0.513	0.703	0.236	0.567	0.368
	0.5	0.406	0.435	0.220	0.757	0.459	0.535	0.732	0.246	0.590	0.383
	0.6	0.390	0.418	0.211	0.727	0.440	0.513	0.703	0.236	0.567	0.368
	0.7	0.341	0.366	0.185	0.636	0.385	0.449	0.615	0.206	0.496	0.322
	0.8	0.260	0.279	0.141	0.485	0.294	0.342	0.469	0.157	0.378	0.245
	0.9	0.146	0.157	0.079	0.273	0.165	0.192	0.264	0.088	0.213	0.138

Table 6. Normalized decision matrix

		A1		A2		A3		A4		A5	
	τ	μ	ϑ	μ	ϑ	μ	ϑ	μ	ϑ	μ	ϑ
C1	0.1	0.110	0.246	0.110	0.246	0.292	0.064	0.055	0.301	0.318	0
	0.2	0.196	0.438	0.196	0.438	0.518	0.113	0.098	0.535	0.565	0
	0.3	0.257	0.575	0.257	0.575	0.680	0.148	0.129	0.702	0.742	0
	0.4	0.294	0.657	0.294	0.657	0.778	0.170	0.148	0.802	0.848	0
	0.5	0.306	0.685	0.306	0.685	0.810	0.177	0.154	0.835	0.883	0
	0.6	0.294	0.657	0.294	0.657	0.778	0.170	0.148	0.802	0.848	0
	0.7	0.257	0.575	0.257	0.575	0.680	0.148	0.129	0.702	0.742	0
	0.8	0.196	0.438	0.196	0.438	0.518	0.113	0.098	0.535	0.565	0
	0.9	0.110	0.246	0.110	0.246	0.292	0.064	0.055	0.301	0.318	0
C7	0.1	0.146	0.157	0.079	0.273	0.165	0.192	0.264	0.088	0.213	0.138
	0.2	0.260	0.279	0.141	0.485	0.294	0.342	0.469	0.157	0.378	0.245
	0.3	0.341	0.366	0.185	0.636	0.385	0.449	0.615	0.206	0.496	0.322
	0.4	0.390	0.418	0.211	0.727	0.440	0.513	0.703	0.236	0.567	0.368
	0.5	0.406	0.435	0.220	0.757	0.459	0.535	0.732	0.246	0.590	0.383
	0.6	0.390	0.418	0.211	0.727	0.440	0.513	0.703	0.236	0.567	0.368
	0.7	0.341	0.366	0.185	0.636	0.385	0.449	0.615	0.206	0.496	0.322
	0.8	0.260	0.279	0.141	0.485	0.294	0.342	0.469	0.157	0.378	0.245
	0.9	0.146	0.157	0.079	0.273	0.165	0.192	0.264	0.088	0.213	0.138

Step 3. The weighted normalized decision matrix is given in Table 7.
Step 4. The negative ideal solution from Table 7 with different values of τ is obtained as given in Table 8.
Step 5. The Euclidean and taxicab distances of the alternatives to the negative ideal solution are calculated as follows (Table 9).
Step 6. The relative assessment matrix is calculated by using $\theta = 0.2$. This matrix is not given here since the space limitation.
Step 7. The assessment score of each alternative is computed and the ranking of alternatives are obtained as given in Table 10.

Table 7. Weighted normalized decision matrix

		A1		A2		A3		A4		A5	
	τ	μ	ϑ	μ	ϑ	μ	ϑ	μ	ϑ	μ	ϑ
C1	0.1	0.021	0.371	0.021	0.371	0.056	0.218	0.011	0.416	0.061	0.165
	0.2	0.067	0.603	0.067	0.603	0.177	0.373	0.034	0.671	0.193	0.294
	0.3	0.116	0.739	0.116	0.739	0.306	0.477	0.058	0.817	0.333	0.385
	0.4	0.151	0.808	0.151	0.808	0.399	0.535	0.076	0.889	0.435	0.440
	0.5	0.164	0.829	0.164	0.829	0.433	0.554	0.082	0.911	0.472	0.459
	0.6	0.151	0.808	0.151	0.808	0.399	0.535	0.076	0.889	0.435	0.440
	0.7	0.116	0.739	0.116	0.739	0.306	0.477	0.058	0.817	0.333	0.385
	0.8	0.067	0.603	0.067	0.603	0.177	0.373	0.034	0.671	0.193	0.294
	0.9	0.021	0.371	0.021	0.371	0.056	0.218	0.011	0.416	0.061	0.165
C7	0.1	0.048	0.157	0.026	0.273	0.054	0.192	0.086	0.088	0.070	0.138
	0.2	0.152	0.279	0.082	0.485	0.171	0.342	0.273	0.157	0.220	0.245
	0.3	0.261	0.366	0.142	0.636	0.295	0.449	0.471	0.206	0.379	0.322
	0.4	0.341	0.418	0.185	0.727	0.385	0.513	0.615	0.236	0.496	0.368
	0.5	0.370	0.435	0.201	0.757	0.418	0.535	0.667	0.246	0.538	0.383
	0.6	0.341	0.418	0.185	0.727	0.385	0.513	0.615	0.236	0.496	0.368
	0.7	0.261	0.366	0.142	0.636	0.295	0.449	0.471	0.206	0.379	0.322
	0.8	0.152	0.279	0.082	0.485	0.171	0.342	0.273	0.157	0.220	0.245
	0.9	0.048	0.157	0.026	0.273	0.054	0.192	0.086	0.088	0.070	0.138

Table 8. Negative ideal solutions

τ	C1		C7	
	μ	ϑ	μ	ϑ
0.1	0.011	0.416	0.026	0.273
0.2	0.034	0.671	0.082	0.485
0.3	0.058	0.817	0.142	0.636
0.4	0.076	0.889	0.185	0.727
0.5	0.082	0.911	0.201	0.757
0.6	0.076	0.889	0.185	0.727
0.7	0.058	0.817	0.142	0.636
0.8	0.034	0.671	0.082	0.485
0.9	0.011	0.416	0.026	0.273

Table 9. Euclidean and taxicab distances

τ	A1		A5	
	Euclidean	Taxicab	Euclidean	Taxicab
0.1	0.062	0.209	0.100	0.266
0.2	0.113	0.296	0.169	0.364
0.3	0.156	0.353	0.222	0.424
0.4	0.186	0.386	0.257	0.458
0.5	0.197	0.397	0.270	0.469
0.6	0.186	0.386	0.257	0.458
0.7	0.156	0.353	0.222	0.424
0.8	0.113	0.296	0.169	0.364
0.9	0.062	0.209	0.100	0.266

Table 10. Assessment scores

τ	A1		A2		A3		A4		A5	
	Score	Ranking	Score	Ranking	Score	Ranking	Score	Ranking	Score	Ranking
0,1	−0,129	5	0,022	2	0,124	1	−0,077	4	0,060	3
0,2	−0,213	5	0,070	2	0,203	1	−0,125	4	0,065	3
0,3	−0,274	5	0,124	2	0,264	1	−0,165	4	0,052	3
0,4	−0,316	5	0,163	2	0,305	1	−0,194	4	0,041	3
0,5	−0,330	5	0,178	2	0,320	1	−0,205	4	0,037	3
0,6	−0,316	5	0,163	2	0,305	1	−0,194	4	0,041	3
0,7	−0,274	5	0,124	2	0,264	1	−0,165	4	0,052	3
0,8	−0,213	5	0,070	2	0,203	1	−0,125	4	0,065	3
0,9	−0,129	5	0,022	2	0,124	1	−0,077	4	0,060	3

3.3 Comparative Analysis

The comparative analysis is conducted by IF-CODAS [5] method. The obtained ranking by IF-CODAS is presented in Table 11.

The ranking of the alternatives from the best to the worst is obtained as A2 > A3 > A5 > A1 > A4. The ranking by the proposed CINFU-CODAS method was A3 > A2 > A5 > A4 > A1. This difference in ranking is due to the discrete vs. continuous approaches of the used methods.

Table 11. Relative assessment matrix and ranking of alternatives

	A1	A2	A3	A4	A5	Score	Ranking
A1	0	−0.082	0.113	−0.017	−0.066	−0.052	4
A2	0.082	0	−0.031	0.066	0.016	0.133	1
A3	−0.113	0.031	0	0.097	0.047	0.062	2
A4	0.017	−0.066	−0.097	0	−0.050	−0.196	5
A5	0.066	−0.016	−0.047	0.050	0	0.054	3

4 Conclusion

Nonlinear continuous intuitionistic fuzzy sets were used for the first time in the development of a new fuzzy extension of the CODAS method in this paper. The usage of continuous and nonlinear membership functions has produced more comprehensive and more accurate results than using discrete functions. Although there were not great differences in the ranking results, it was seen that the first alternative in the ranking changed. We recommend that the proposed CINFUS sets also be used in the development of fuzzy extensions of other multi-criteria decision-making methods, for instance CINFU TOPSIS, CINFU AHP or CINFU VIKOR.

References

1. Zadeh, L.: Fuzzy set. Inf. Control **8**(3), 338–353 (1965)
2. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets. Syst. **20**(1), 87–96 (1986)
3. Alkan, N., Kahraman, C.: An intuitionistic fuzzy multi-distance based evaluation for aggregated dynamic decision analysis (IF-DEVADA): ıts application to waste disposal location selection. Eng. Appl. Artif. Intell. **111** (2022)
4. Alkan, N., Kahraman, C.: Risk analysis for digitalization oriented sustainable supply chain using interval-valued Pythagorean fuzzy AHP. Intelligent and Fuzzy Techniques: Smart and Innovative Solutions: Proceedings of the INFUS 2020 Conference, Istanbul, Turkey (2021)
5. Büyüközkan, G. Göçer, F.: Prioritizing the strategies to enhance smart city logistics by ıntuitionistic fuzzy CODAS. Atlantis Stud. Uncertainty Model. **1**, 805–811 (2020)

An Initial Research on Comparison of Fuzzy AHP and Classical AHP Methods

Basar Oztaysi(✉), Sezi Cevik Onar, and Cengiz Kahraman

Industrial Engineering Department, Istanbul Technical University, Istanbul, Turkey
oztaysib@itu.edu.tr

Abstract. Analytic Hierarchy Process (AHP) is one of the most popular method in multi-criteria decision making and it has been extended by various extensions of fuzzy sets. The method uses linguistic terms for expert judgements and these judgements are represented by crisp or fuzzy numbers in different methods. Although there are various extension of AHP, such as Type-2 Fuzzy AHP, Hesitant Fuzzy AHP, Intuitionistic fuzzy AHP, Spherical fuzzy AHP, there are no studies on the comparison of the methods. In this study, we try to compare five classical AHP and one Fuzzy AHP (Bukley's AHP) for 3×3 and 4×4 matrices. To this end, all possible matrices are created and solved. In order to make a comparison the ranking result and weight results are used. The results show that the difference between the methods increase as the matrix size increase. It is also observed that, the methods provide similar results for consistent matrices.

Keywords: Analytic Hierarch Process · Fuzzy Sets · Fuzzy AHP

1 Introduction

Analytic Hierarchy Processes (Saaty 1980) is one of the most popular multi criteria decision making method which is used to solve complex decision making problems by breaking down the problem into a hierarchical structure of criteria and sub-criteria. The method uses decision-maker's judgments in pairwise comparisons which are then converted to numerical values to calculate relative weights. Saaty (1996) later proposed Analytic Network Process (ANP) method which can be used to solve decision problems where dependencies exist within the decision model.

Both AHP and ANP use linguistic scale to collect expert judgements and then use them in the solution steps by converting them into numeric values. One branch of studies in the literature focus on the numeric representation of linguistic scale and propose using fuzzy sets instead of crisp values. These studies propose fuzzy AHP and ANP methods to use fuzzy sets in the solution phase. The initial studies on fuzzy AHP are Laarhoven and Pedrycz (1983), Buckley (1985), Habenicht (1989). Later, AHP has been extended by using extensions of fuzzy sets such as Type-2 fuzzy sets, intuitionistic fuzzy sets.

In this study, we present an experimental study on the results of crisp AHP and fuzzy AHP in order to show the effects of methods on the results. To this end a software is

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 382–388, 2023.
https://doi.org/10.1007/978-3-031-39777-6_46

developed and different AHP methods are used to solve all possible 3×3 and 4×4 pairwise comparison matrices.

The rest of the paper is as follows; in Sect. 2 a brief literature review is provided, in Sect. 3 the methods which are involved into the study are explained. In Sect. 4, the application is explained and the results are discussed. In Sect. 5, the conclusions and directions for future studies are given.

2 Literature Review

AHP, proposed by Saaty (1980), is based on dividing complex problems into small problems and use pairwise comparisons which compares the relative importance of each criterion over other criteria. In the classical AHP the pairwise comparison matrix is constructed using expert evaluations, then these evaluations are used to derive the weights of each criterion. In classical AHP, there are five methods (Topcu 2023):

- Approximate Method: The row averages are calculated and then normalized to find the weights of the criteria.
- Inverse of column Max: Sum of values in each column is calculated then the values are inverted and finally the inverse values are normalized to find the weights of the criteria.
- Mean of Normalized values: Each column is normalized, then row averages are calculated to find the weights.
- Geometric Mean: Geometric mean of each row is calculated, then geometric mean values are normalized to find the weights.
- Power Method: The pairwise comparison matrix is raised to a sufficiently large power, then the row averages of the new matrix are calculated. The process ends when the weight values converge.

The most commonly adopted AHP technique is Power method but due to calculation simplicity other methods can be used.

Another branch of AHP studies focus on utilizing the advantages of fuzzy sets on representing vagueness of information. The initial studies on Fuzzy AHP are as in the following:

- Laarhoven and Pedrycz (1983): The authors use triangular fuzzy sets and use Lootsma's logarithmic least square method to derive fuzzy weights and fuzzy performance scores.
- Buckley (1985): The author propose a Fuzzy AHP method using trapezoidal fuzzy numbers based on geometric mean approach.
- Chang (1996): The author propose a new fuzzy AHP, with the use of triangular fuzzy numbers for pairwise comparison scale, and the use of the extent analysis method for the synthetic extent value of the pairwise comparison

Later, novel AHP methods are proposed by using extensions of fuzzy sets such as: Intuitionistic fuzzy AHP (Silavi et al. 2006) interval valued Fuzzy AHP (Chen and Lin 2010), Interval type-2 fuzzy AHP (Kahraman et al. 2014), Pythagorean Fuzzy AHP (Ilbahar et al. 2018), Hesitant Fuzzy Ahp (Oztaysi et al. 2015), q-Rung orthopair fuzzy AHP (Garg et al. 2020).

A literature survey on Classical AHP&ANP and FuzzyAHP&ANP shows papers on classical AHP&ANP is significantly more than Fuzzy AHP (Fig. 1).

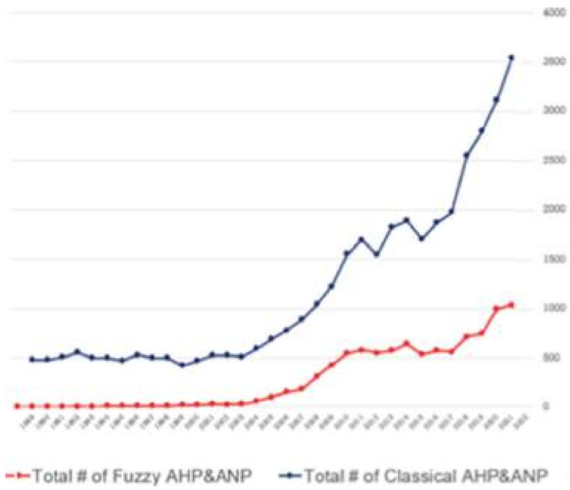


Fig. 1. Comparison of Fuzzy and Classical AHP&ANP

A deeper look at the journals and application fields, it can be seen that Classical AHP&ANP has been used in various fields and has been published in different domain journals, however Fuzzy AHP&ANP papers has been published in fuzzy oriented journals and conferences. One of the possible reasons for this result can be the existence of AHP Specific software such as ExperChoice and SuperDecisions which enable researchers to easily utilize the technique in their studies.

3 Methodology

In this study, we focus on comparing different AHP methods. AHP method enables complex problems to be divided into small problems and work on these problems in terms of pairwise comparison matrices. Although, the hierarchy levels in the decision model differs based on the complexity of the problem, decision makers deal with pairwise comparison matrices. So we prefer to focus on unit pairwise comparison matrices instead of decision problems.

A pairwise comparison matrix is formed based on the number of criteria/alternatives. In a comparison matrix, the diagonal is known since a criterion/alternative is equal to itself. Knowing the values in the upper side of the diagonal is enough to construct a pairwise comparison matrix since $a_{ij} = 1/a_{ji}$. As a result, when the upper diagonal values are known, whole pairwise comparison matrix can be derived. The number of cells in the upper diagonal of the matrix (N(UDV)) can be calculated as in the following.

$$\mathrm{N(UDV)} = \sum_{n=1}^{m-1} n$$

The number of all possible matrices (N(APPM(m)) can be calculated based on N(UDV) and scale used for expert evaluations. In classical AHP, the scale includes numbers from 1 to 9 and their inversed values. In other terms, there are 17 different possibilities for an expert to use for judgements. The number of all possible matrices with dimension m can be calculated as in the following.

$$\mathrm{N(APPM(m))} = 17^{m}$$

A list of N(UDV) and N(APPM(m)) for different dimensions are given in Table 1.

Table 1. N(UDV) and N(APPM(m)) values for different pairwise comparison matrices.

Number of Elements (m)	N(UDV)	N(APPM(m))
3	3	4913
4	6	24137569
5	10	2.01599×10^{12}
6	15	2.86242×10^{18}
7	21	6.90919×10^{25}
8	28	2.83511×10^{34}
9	36	1.9777×10^{44}
10	45	2.34532×10^{55}

In order to compare the results of different AHP techniques we develop a software to generate all possible pairwise comparison matrices and calculate the results using different AHP techniques. In this study we only focus on m = 3 and m = 4, and we include the AHP Methods; M1: Approximate Method, M2: Inverse of column Max, M3: Mean of Normalized values, M4: Geometric Mean,M5: Power Method, M6: Buckley's Fuzzy AHP. The steps of the study are as follows:

Step 1: Set the dimension of the matrix (In this study we only focus on m = 3 and m = 4)
Step 2: Generate all possible matrices using linguistic terms
Step 3: Calculate the results

Step 3.1: Convert each matrix to numeric pairwise comparison matrix for classical and fuzzy AHP techniques.
Step 3.2: Apply the calculation steps of the selected AHP Technique
Step 3.3: Defuzzify the weights for Fuzzy AHP method.
Step 3.4: Save the weight values.

Step 4: Analyze the results.

4 Application and Results

Saaty (1980) identify the term consistency for pairwise comparison matrices. Consistency ratio for each matrix can be calculated and the matrix is entitled consistent if the Consistency Ratio is below 0.1. As a part of the study, we identify the consistent and inconsistent matrices. Table 2 shows that 22.12% of 3 × 3 matrices are consistent and 3.47% of 4 × 4 matrices are found to be consistent.

Table 2. Number of consistent matrices.

	All matrices	Consistent matrices	Ratio
3 × 3	4.913	1.087	22.12%
4 × 4	24.137.569	838.223	3.47%

Comparing of the results of different methods is hard to measure since there is no objective true result or "Ground Truth" in this type of problems. The power method is most commonly adopted AHP technique, so we choose Power Method as a benchmark and compare the results of other methods with its results. We make two types of comparisons, the weights and the ranks. For weight comparison we use root mean square error of the weight values. For ranking, we use the term ranking error which means there is a difference between the rankings of the methods and calculate the ratios.

When all matrices are considered, for 3 × 3 matrices best method seems to be Geometric Mean method (Table 3). However, the method is the worst method for 4 × 4 in terms of ranking errors. For 4 × 4 matrices the best method is mean of normalized values method (Table 4).

Table 3. Ranking Error values of the methods for all matrices

	3 × 3 Ranking Error	4 ×4 Ranking Error
M1: Approximate Method	18.07%	34.76%
M2: Invers of Column Max	17.95%	35.01%
M3: Mean of Norm. Values	1.83%	33.12%
M4: Geometric Mean	0.00%	51.15%
M6: Buckley's Fuzzy AHP	1.95%	47.68%

When the consistent matrices are considered only different results are observed. When ranking error is considered mean of normalized values, geometric mean and Buckley's Fuzzy AHP gives the same results. When 4 × 4 matrices are considered Mean of Normalized values presents the most similar results with Power method (Table 5).

In terms of root mean square, Geometric Mean method presents the best technique for 3 × 3 and 4 × 4 consistent matrices. The method is followed by Mean of Normalized Values and Buckley's fuzzy AHP (Table 6).

Table 4. Root Mean Square Error of the methods for all matrices

	3 × 3 Weight RMSE	4 × 4 Weight RMSE
M1: Approximate Method	7.53%	5.69%
M2: Inverse of Column Max	5.35%	7.63%
M3: Mean of Norm. Values	2.92%	3.09%
M4: Geometric Mean	0.00%	4.52%
M6: Buckley's Fuzzy AHP	0.57%	4.44%

Table 5. Ranking Error values of the methods for consistent matrices

	3 × 3 Ranking Error	4 × 4 Ranking Error
Method 1: Approximate Method	0.24%	15.16%
Method 2: Invers of Column Max	0.24%	15.00%
Method 3: Mean of Norm. Values	0.00%	1.56%
Method 4: Geometric Mean	0.00%	3.07%
Method 6: Buckley's Fuzzy AHP	0.00%	4.08%

Table 6. Root Mean Square Error of the methods for consistent matrices

	3 × 3 Weight RMSE	4 × 4 Weight RMSE
Method 1: Approximate Method	3.19%	3.91%
Method 2: Invers of Column Max	2.08%	3.05%
Method 3: Mean of Norm. Values	0.36%	0.74%
Method 4: Geometric Mean	0.00%	0.46%
Method 6: Buckley's Fuzzy AHP	0.73%	0.77%

5 Conclusions

The literature provides various crisp and fuzzy AHP methods. In this paper, we try to benchmark the results of most commonly adopted AHP Methods. As a comparison method we compare the results of the methods with the Power Method to see the differences in results. The results show that the results show higher difference in 4 × 4 matrices when compared to 3 × 3 matrices. It can also be observed that the results of all matrices show a higher error when compared to consistent matrices.

For future studies, other AHP extensions such as Type-2 fuzzy AHP, Intuitionistic fuzzy AHP and hesitant fuzzy AHP can be added to the comparison matrix. Another future direction can be to provide other comparison metrics. The current comparison method is based on the results of Power Method and it provides comparative results.

Providing an absolute metric for comparison may give better insight about the results of different AHP Methods.

References

Saaty, T.L.: The Analytic Hierarchy Process. McGraw Hill, New York (1980). Reprinted by RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA, 15213, 2000a

Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process. RWS Publications, Pittsburgh (1996)

Habenicht, W.: Multiple criteria decision analysis in regional planning, concepts, methods and applications. Eur. J. Oper. Res. **42**(1) (1989)

Topcu, I.: (2023). https://web.itu.edu.tr/topcuil/ya/AHP.htm

van Laarhoven, P.J.M., Pedrycz, W.: A fuzzy extension of Saaty's priority theory. Fuzzy Sets Syst. **11**(1–3), 229–241 (1983)

Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**(3), 233–247 (1985)

Chang, D.Y.: Applications of the extent analysis method on fuzzy AHP. Eur. J. Oper. Res. **95**(3), 649–655 (1996)

Oztaysi, B., Onar, S.C., Boltürk, E., Kahraman, C.: Hesitant fuzzy analytic hierarchy process. In: 2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Istanbul, Turkey, pp. 1–7 (2015). https://doi.org/10.1109/FUZZ-IEEE.2015.7337948

Silavi, T., Malek, M.R., Delavar, M.R.: Multicriteria map overlay in geospatial information system via intuitionistic fuzzy AHP method. In: Applied Artificial Intelligence - Proceedings of the 7th International FLINS Conference, FLINS 2006, pp. 401–408 (2006)

Chen, C.T., Lin, K.H.: A decision-making method based on interval-valued fuzzy sets for cloud service evaluation. In: 4th International Conference on New Trends in Information Science and Service Science, Gyeongju, Korea (South), pp. 559–564 (2010)

Kahraman, C., Oztaysi, B., Ucal sari, I., Turanoglu, E.: Fuzzy analytic hierarchy process with interval type-2 fuzzy sets. Knowl. Based Syst. **59**, 48–57 (2014)

Ilbahar, E., Karaşan, A., Cebi, S., Kahraman, C.: A novel approach to risk assess-ment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system. Saf. Sci. **103**, 124–136 (2018)

Garg, H., Ali, Z., Mahmood, T.: Algorithms for complex interval-valued q-rung ortho-pair fuzzy sets in decision making based on aggregation operators. AHP TOPSIS Expert Syst. **38**(1), e12609 (2020)

Sustainable Future Technology Evaluation with Digital Transformation Perspective in Air Cargo Industry Using IF ANP Method

Celal Alpay Havle[1(✉)] and Gülçin Büyüközkan[2]

[1] Professional Flight Department, Özyeğin University, 34794 Istanbul, Turkey
celal.alpay.havle@gmail.com

[2] Industrial Engineering Department, Galatasaray University, 34349 Istanbul, Turkey
gbuyukozkan@gsu.edu.tr

Abstract. Sustainability is one of the most critical issues of recent years. Many companies that contribute to the global economy focus on sustainability studies more than ever. On the other hand, technological developments lead companies to digital transformation. This situation necessitated the adoption of a sustainability approach that includes digital transformation and digital technologies. Therefore, it is critical for companies embarking on a transformation journey to invest in technologies that support sustainability. Advances in technology and digital transformation have contributed to the growth of e-commerce volume globally. However, this has led to the growth and complexity of supply chain networks. One of the biggest players in this network structure is undoubtedly the air cargo industry. The air cargo industry plays a critical role in the sustainability of the global economy, employment, social development, and the environment. Therefore, this study proposes a new model for the evaluation of sustainable future technologies in the air cargo industry from the perspective of digital transformation. Multi-criteria decision-making (MCDM) approach was adopted in the study. The intuitionistic fuzzy analytical network process (IF ANP) is used to validate the proposed model in a real case conducted in the Turkish air cargo industry.

Keywords: Air Cargo Industry · Digital Transformation · IF ANP · Sustainability · Sustainable Future Technology Evaluation

1 Introduction

The sudden acceleration in technological developments is the beginning of a new era. Technology has become an indispensable part of both people and the business world, and digital transformation has become the most popular topic today. Digital technologies used for products and services have radically changed the classical business processes [1–3]. Although digitalization will change both supply and demand, the course of technological progress is now harder to predict than disruptions in the past [4]. While all changes and transformations affect the world, sustainability, which is always important,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 389–397, 2023.
https://doi.org/10.1007/978-3-031-39777-6_47

is on the agenda of companies that contribute to the global economy. Digital technologies that bring along digital transformation have begun to be seen as a strategy for sustainability [5]. Digital technologies have started to be used in many processes such as more efficient use of resources, providing environmental, economic, and social benefits, reducing carbon footprint, reducing time losses without added value in processes. The use of digital technology has become popular in almost every industry. One of them is undoubtedly the air cargo industry, which is one of the most important players in the complex supply chain networks created by the increasing e-commerce volume. Airlines serving in the air cargo industry are using digital transformation as a tool to get closer to carriers and avoid using carriers as intermediaries in some parts of the business [4]. Air cargo is a key contributor to sustainable development [5]. The air cargo industry plays a vital role in supporting social development, biodiversity conservation, health systems and global peace by driving inclusive growth, job creation and the world economy [6]. The air cargo industry has a responsibility to minimize or compensate for its negative impacts by reducing its environmental footprint through decarbonization and waste reduction programs. To achieve internationally accepted goals such as the UN Sustainable Development Goal and the Paris Agreement, it is critical that all companies in the air cargo industry initiate or accelerate their sustainable transformation [6]. At this point, it is a fact that digital transformation is inevitable. Sustainability has become more important than ever before. Therefore, the use of technologies that support sustainability is a critical issue. From this point of view, a new model has been proposed for the evaluation of sustainable technologies through the literature survey and the opinions of the decision makers (DMs) who are industrial experts. The proposed model has been applied to a real case in the air cargo industry in Turkey, considering the importance of sustainability for the air cargo industry. Sustainable technology evaluation for the air cargo industry is a decision-making problem that consists of multiple criteria. For this reason, a multi-criteria decision-making (MCDM) approach has been adopted. The analytic network process (ANP) method [7] validates the proposed model. The opinions of the DMs regarding the structure of the problem used a basis in the selection of the technique used. Modeling and solving real-life problems are complex. Moreover, decision-making processes involve human perception. Therefore, uncertainty, intuition and hesitation are likely to occur. To eliminate this, the classical ANP method has been extended to the intuitionistic fuzzy (IF) [8] environment. The structure of the study is organized as follows: Sect. 2 presents the relationship between sustainability, digital transformation, and air cargo. Details regarding the proposed method are provided in Sect. 3. The conducted case study is explained in Sect. 4 before the study is concluded in Sect. 5.

2 Sustainability, Digital Transformation and Air Cargo

Technology evolves rapidly during the last few years. Digital platforms such as social media exist for over a decade. The use of such digital channels results in the generation of massive amounts of data. The adoption of big data is significant to achieve digital transformation. This develops sustainable societies. The sustainable development of society is accelerated by the relationship between big data analytics and business transformation for value creation [9]. The society's need for accessible, fundable, interoperable, and reusable infrastructures and data management standards that provide greater access to information is increasing day by day [10]. The digitalization of services and the transformation of contemporary business models are necessary to accelerate the creation of sustainable societies. New digital business models have the potential to deliver accuracy and efficiency and are expected to get ahead of economic needs. Moreover, these new business models are expected to address collective value-producing societal challenges affecting organizations, companies, consumers, and the public [11]. Today, cities optimize their services and performances through digital technologies that foster the knowledge economy, data-driven innovation and achieve a resource-efficient society [12]. Digital infrastructure integration is seen in many areas such as health, finance, and transportation along with digital transformation. Investing in these infrastructures enables the innovation and digitalization of city services, while enabling the deployment of a wide variety of technology ecosystems (Pappas et al. 2018). Adopting digital transformation technologies and benefiting from their advantages have positive effects on sustainability [13]. Sustainability is significant for air cargo companies. However, there is no common definition of sustainability across the industry. Sustainability is more in line with the priorities, geographical locations, roles, and sizes of each company [5]. Especially with the COVID-19 crisis, sustainability has become more important than ever for almost every industry, including the air cargo industry. The air cargo industry both contributes to building a better world and can continually evolve to reduce its environmental footprint and become a more sustainable transportation system [5]. With the development of mindsets, corporate leaders have come to realize that a holistic approach is more important than short-term profitability for the survival of businesses. But a safe, secure, and digital air cargo industry is profitable and sustainable [5]. The air cargo community is focused on initiatives driving operational excellence, quality improvements and digital transformation. Furthermore, the industry is committed to reduce its environmental footprint [5]. The International Air Cargo Association (TIACA) emphasizes that the air cargo industry has a very important role in social development and aims to reach a more sustainable world, but this is largely unknown to people who are not involved in the industry [5]. Furthermore, TIACA stated that sustainability and digital transformation is opportunity for the business to reboot better. TIACA stated in its sustainability report published in 2021 that digital transformation is now a necessity for the survival of companies in the air cargo industry [5]. It is emphasized that a sustainable air cargo industry must be fully digital to be effective and profitable. Moreover, 61% of industry experts state that they have a digital transformation plan. It is stated that the rest should increase its speed [5].

3 Proposed Method

3.1 Intuitionistic Fuzzy Sets

This section briefly describes intuitionistic fuzzy (IF) sets and arithmetic operations with intuitionistic fuzzy numbers (IFN) in these sets. The IF sets that were developed by Atanassov in 1986 [8] are based on fuzzy set theory developed by Zadeh in 1965 [14] and expanded from it. A membership degree is constituted to each element by IFSs. A crisp (fixed) set $\boldsymbol{E}$ is presented by Atanassov where $\boldsymbol{A} \subset \boldsymbol{E}$ and $\boldsymbol{A}^*$ denotes an IFS. Then $\boldsymbol{A}^*$ can be defined in $\boldsymbol{E}$ as follows:

$$A^* = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in E\} \tag{1}$$

The functions $\mu_A : E \to [0, 1]$ and $\nu_A : E \to [0, 1]$ define the degree of membership. Additionally, the non-membership degree of an object is defined as $0 \le \mu_A(x), \nu_A(x) \le 1$, $\{\langle x, \mu_A(x), 1 - \mu_A(x)\rangle | x \in E\}$, where $x \in E$ for $\forall x \in E$ in the set A.

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \tag{2}$$

Then $\pi_A(x)$ indicates the uncertainty degree of the membership of element x where $x \in E$. $\pi_A(x) = 0$ for $\forall x \in E$ for classical fuzzy sets. An IFS can be defined as $A = \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in E\}$ in $X = \{x\}$ where $\mu_A(x) \to [0, 1]$, $\nu_A(x) \to [0, 1]$, $\pi_A(x) \to [0, 1]$.

3.2 Computational Steps of the IF ANP Method

The computational steps of the IF ANP method are given as follows [15–17]:

Step 1 & 2. Determination of the DMs, evaluation criteria and the alternatives.
Step 3. Calculation of the weights of DMs: The weights of the DMs are computed using Eq. (3) that is adapted from Boran et al. [18].

$$\lambda_k = \frac{\left[\mu_k + \pi_k\left[\frac{\mu_k}{1-\pi_k}\right]\right]}{\sum_{k=1}^{K}\left[\mu_k + \pi_k\left[\frac{\mu_k}{1-\pi_k}\right]\right]} \tag{3}$$

where $\sum_{k=1}^{K} \lambda_k = 1$. For this purpose, the linguistic terms in Table 1 are used.

Table 1. Linguistic scale for the influence weights of DMs [19]

Linguistic Terms	IF Values (μ, υ, π)
Extremely Unimportant (EU)	(0.00, 0.95, 0.05)
Very Unimportant (VU)	(0.05, 0.90, 0.05)
Unimportant (U)	(0.25, 0.70, 0.05)
Somewhat Unimportant (SU)	(0.40, 0.55, 0.05)
Medium Importance (MI)	(0.50, 0.45, 0.05)
Somewhat Important (SI)	(0.60, 0.35, 0.05)
Important (I)	(0.75, 0.20, 0.05)
Very Important (VI)	(0.90, 0.05, 0.05)
Extremely Important (EI)	(0.95, 0.00, 0.05)

Step 4. Construction of network-based hierarchy: A network-like structure for the evaluating the criteria is constructed based on the opinions of the DMs.
Step 5. Evaluation of the criteria: A pairwise comparison process is conducted by each DM using the linguistic terms given in Table 2.
Step 6. Aggregation: Individual evaluations of the DMs are aggregated into group IF values using IF weighted averaging (IFWA) operator [20] given in Eq. (4)

$$R_{ij} = IFWA_{\lambda}\left(R_{ij}^{(1)}, R_{ij}^{(2)}, R_{ij}^{(3)}, \ldots, R_{ij}^{(k)}\right) = \lambda_1 R_{ij}^{(1)} \oplus \lambda_2 R_{ij}^{(2)} \oplus \lambda_3 R_{ij}^{(3)} \oplus \cdots \oplus$$

$$\lambda_k R_{ij}^{(k)} = \begin{pmatrix} 1 - \prod_{k=1}^{K}\left(1 - \mu_{ij}^{(k)}\right)^{\lambda_k}, \\ \prod_{k=1}^{K}\left(\upsilon_{ij}^{(k)}\right)^{\lambda_k}, \prod_{k=1}^{K}\left(1 - \mu_{ij}^{(k)}\right)^{\lambda_k} - \prod_{k=1}^{K}\left(\upsilon_{ij}^{(k)}\right)^{\lambda_k} \end{pmatrix} \tag{4}$$

Table 2. Linguistic scale for criteria evaluation [19]

Linguistic Terms	IF Values (μ, υ, π)	Reciprocal
Equally Important (EI)	(0.02, 0.18, 0.80)	(0.18, 0.02, 0.80)
Intermediate Value (IV)	(0.06, 0.23, 0.70)	(0.23, 0.06, 0.70)
Weakly More Important (WMI)	(0.13, 0.27, 0.60)	(0.27, 0.13, 0.60)
Intermediate Value (IV2)	(0.22, 0.28, 0.50)	(0.28, 0.22, 0.50)
Strongly More Important (SMI)	(0.33, 0.27, 0.40)	(0.27, 0.33, 0.40)
Intermediate Value (IV3)	(0.47, 0.23, 0.30)	(0.23, 0.47, 0.30)
Very Strong More Important (VSMI)	(0.62, 0.18, 0.20)	(0.18, 0.62, 0.20)
Intermediate Value (IV4)	(0.80, 0.10, 0.10)	(0.10, 0.80, 0.10)
Absolutely More Important (AMI)	(1.00, 0.00, 0.00)	(0.00, 1.00, 0.00)

Table 3. Randomness indices (RI) [23]

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

Step 7. Checking the consistency: Consistency ratio ($C.R.$) is calculated using Eq. (5) that is adopted from Abdullah and Najib [21]. Here, n represent the number of elements in the matrix and π_{ij} indicates the hesitation value. Here, if $CR \leq 0.1$, then the consistency is satisfied [22]. Otherwise, evaluations should be revised (Table 3).

$$CR = \frac{RI - \frac{\sum \pi_{ij}}{n}}{n - 1} \tag{5}$$

Step 8. Computation of the criteria weights: In this step, IF entropy weights are calculated using Eq. (7) by employing Eq. (6).

$$\overline{\overline{w_i}} = -\frac{1}{nln2}[\mu_i ln\mu_i + \upsilon_i ln\upsilon_i - (1-\pi_i)ln(1-\pi_i) - \pi_i ln2] \tag{6}$$

$$w_i = \frac{1-\overline{\overline{w_i}}}{n-\sum_{j=1}^{n}\overline{\overline{w_i}}} \tag{7}$$

where $\sum_{j=1}^{n} w_i = 1$.

Step 9. Construction of the general sub-matrix: Here, a general sub-matrix is constructed based on the identified relations in the network-based hierarchy to form a supermatrix.

Step 10. Computation of the weighted supermatrix: The initial supermatrix based on the weights reached in Step 8 is normalized by reaching unity for the sum of each matrix column.

Steps 11 & 12. Calculation of limit supermatrix and prioritization of the alternatives: The limit supermatrix is obtained through the weighted supermatrix that is raised to its limiting powers until the weights converge and remain constant. Limit supermatrix shows the priority weights of the alternatives.

4 Case Study

In this study, a new model is proposed for sustainable future technology evaluation with a digital transformation perspective. It was applied to Turkish air cargo industry to analyze its validity by applying computational steps of the IF ANP method given above. For this purpose, the main structure of the proposed model is organized based on the sustainability model introduced by Büyüközkan et al. [24]. The criteria [5, 6, 25] of the model shown in Fig. 1. Three DMs were selected from the industry. All of them are experts. Their academic background and experiences justify the reason of their selection as DMs. Further information regarding the DMs cannot be provided due to confidentiality, privacy concerns and page limitation.

Importance degrees of DMs were assigned as VI, SI, and I, respectively using linguistic terms in Table 1. The weights of the DMs were calculated as 0.40, 0.27 and 0.33, respectively. For this, Eq. (3) was utilized. Pairwise comparisons were conducted using the linguistic terms in Table 2 by the DMs and the computational steps of the IF ANP method given above were applied. General sub-matrix notation for IF ANP steps is shown in Table 4. The convergence is reached when the supermatrix is raised to its 51st power. Results that were obtained from the limit supermatrix (M^{51}) are given in Table 5. Importance degrees of the alternatives are $A_1 \succ A_2 \succ A_3$. The most appropriate sustainable future technology for air cargo industry is *Augmented Reality & Artificial Intelligence* based on the obtained results with respect to the goal.

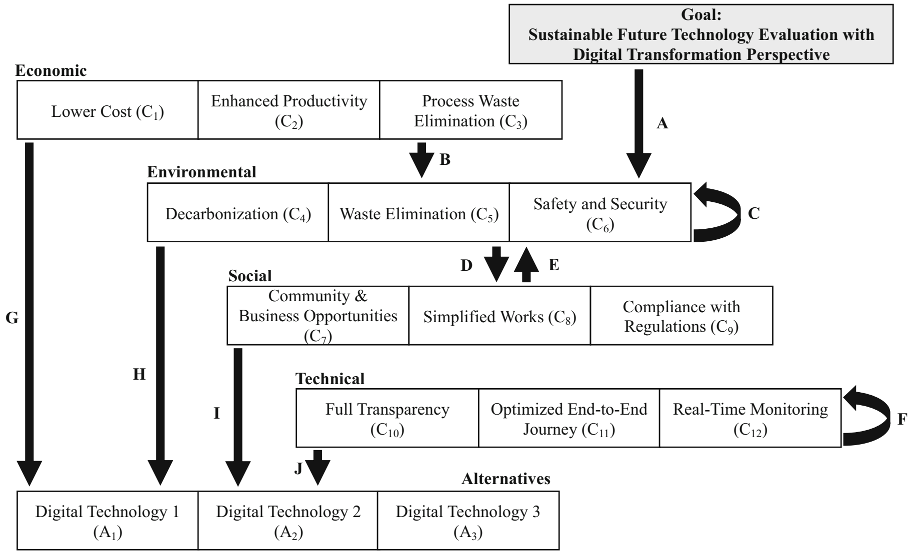


Fig. 1. ANP evaluation network of the proposed model by DMs [5, 6, 24, 25]

Table 4. General sub-matrix notation

	Goal	Economic	Environmental	Social	Technical	Alternatives
Goal						
Economic						
Environmental	A	B	C	E		
Social			D			
Technical					F	
Alternatives		G	H	I	J	*I*

Table 5. Obtained results from the limit supermatrix (M^{51})

With respect to	A_1	A_2	A_3
Goal	**0.3344**	**0.3332**	**0.3324**
C_1	0.3338	0.3322	0.3340
C_2	0.3346	0.3328	0.3325
C_3	0.3341	0.3329	0.3330
C_4	0.3353	0.3336	0.3312
C_5	0.3335	0.3339	0.3325
C_6	0.3344	0.3321	0.3335
C_7	0.3381	0.3295	0.3324

(*continued*)

Table 5. (*continued*)

With respect to	A_1	A_2	A_3
C_8	0.3339	0.3328	0.3333
C_9	0.3335	0.3330	0.3335
C_{10}	0.3351	0.3312	0.3337
C_{11}	0.3360	0.3320	0.3321
C_{12}	0.3348	0.3323	0.3329

5 Conclusion and Perspective

There is a strong relationship between sustainability and digital transformation. It is seen that digital transformation is a strategy and opportunity for sustainability. Hence, digital technologies should be employed for sustainability. Air cargo is one of the industries where the sustainability is focused. However, it is seen that there is no specific sustainability definition for air cargo when the academic literature is examined. Furthermore, digital transformation perspective in sustainability is limited. In this study, a new sustainable future technology evaluation model has been proposed with digital transformation perspective. The proposed model is applied to Turkish air cargo industry to analyze its validity. The classical ANP method has been extended to IF environment to deal with uncertainty, intuition, and hesitation in the decision-making process. The applied method revealed that the *Augmented Reality & Artificial Intelligence* may be employed as a future technology for sustainable air cargo industry. For future studies, the proposed model can be improved. Different analytical methods can be utilized, and the results can be compared. Scenario and sensitivity analysis can be conducted. Calculations can be performed in more complex fuzzy environments.

Acknowledgments. This research has been financially supported by Galatasaray University Research Fund (FOA-2021-1059 and FOA-2023-1181).

References

1. Büyüközkan, G., Havle, C.A., Feyzioğlu, O.: A new digital service quality model and its strategic analysis in aviation industry using interval-valued intuitionistic fuzzy AHP. J. Air Transp. Manag. **86**, 101817 (2020)
2. Büyüközkan, G., Havle, C.A., Feyzioğlu, O.: Digital competency evaluation of low-cost airlines using an integrated IVIF AHP and IVIF VIKOR methodology. J. Air Transp. Manag. **91**, 101998 (2021)
3. Büyüközkan, G., Havle, C.A., Feyzioğlu, O.: An integrated SWOT based fuzzy AHP and fuzzy MARCOS methodology for digital transformation strategy analysis in airline industry. J. Air Transp. Manag. **97**, 102142 (2021)
4. Bäuml, M., Hausmann, L.: Air-freight forwarders move forward into a digital future. McKinsey & Company (2018) https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/air-freight-forwarders-move-forward-into-a-digital-future/. Accessed 23 Apr 2023

5. Hourcade, C., Messeiller, J.: Air Cargo Sustainability Report. Accelerating sustainable transformation of air Cargo. Change Horizon & TIACA (2021). https://tiaca.org/wp-content/uploads/2021/04/TIACASustainabilityReport-ChangeHorizon-April2021.pdf. Accessed 23 Apr 2023
6. Hourcade, C., Messeiller, J., Hughes, G.: The Air Cargo Sustainability Roadmap, Accelerating the transformation of air cargo, The International Air Cargo Association (TIACA). Change Horizon & TIACA (2021). https://tiaca.org/wp-content/uploads/2021/11/TIACA-Air-Cargo-Sustainability-Roadmap-FINAL04November2021.pdf. Accessed 23 Apr 2023
7. Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process, vol. 4922, no. 2. RWS Publications, Pittsburgh (1996)
8. Atanassov, K.: Intuitionistic fuzzy sets. In: Fuzzy Sets and Systems, 20th edn. (1986)
9. Pappas, I.O., Mikalef, P., Giannakos, M.N., Krogstie, J., Lekakos, G.: Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. Inf. Syst. e-Bus. Manag. **16**, 479–491 (2018)
10. Wilkinson, M.D., et al.: The FAIR Guiding principles for scientific data management and stewardship. Sci. Data **3**(1), 1–9 (2016)
11. Porter, M.E., Kramer, M.R., Lenssen, G., Smith, N.: Creating shared value. Manag. Sustain. Bus. (2019)
12. van Stijn, A., Gruis, V.: Towards a circular built environment: an integral design tool for circular building components. Smart Sustain. Built Environ. **9**(4), 635–653 (2020)
13. Gomez-Trujillo, A.M., Gonzalez-Perez, M.A.: Digital transformation as a strategy to reach sustainability. Smart Sustain. Built Environ. **11**(4), 1137–1162 (2022)
14. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
15. Büyüközkan, G., Güleryüz, S., Karpak, B.: A new combined IF-DEMATEL and IF-ANP approach for CRM partner evaluation. Int. J. Prod. Econ. **191**, 194–206 (2017)
16. Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process, pp. 83–135. RWS Publications, Pittsburgh (2001)
17. Kayakutlu, G., Büyüközkan, G.: Effective supply value chain based on competence success. Supply Chain Manag. Int. J. **15**(2), 129–138 (2010)
18. Boran, F.E., Genç, S., Kurt, M., Akay, D.: A multi-criteria intuitionistic fuzzy group decision making for supplier selection with TOPSIS method. Expert Syst. Appl. **36**(8), 11363–11368 (2009)
19. Büyüközkan, G., Göçer, F., Karabulut, Y.: A new group decision making approach with IF AHP and IF VIKOR for selecting hazardous waste carriers. Measurement **134**, 66–82 (2019)
20. Xu, Z.: Intuitionistic fuzzy aggregation operators. IEEE Trans. Fuzzy Syst. **15**(6), 1179–1187 (2007)
21. Abdullah, L., Najib, L.: A new preference scale MCDM method based on interval-valued intuitionistic fuzzy sets and the analytic hierarchy process. Soft. Comput. **20**(2), 511–523 (2014). https://doi.org/10.1007/s00500-014-1519-y
22. Saaty, T.L.: How to make a decision: the analytic hierarchy process. Eur. J. Oper. Res. **48**(1), 9–26 (1990)
23. Saaty, T.L.: A scaling method for priorities in hierarchical structures. J. Math. Psychol. **15**(3), 234–281 (1977)
24. Büyüközkan, G., Karabulut, Y., Mukul, E.: A novel renewable energy selection model for United Nations' sustainable development goals. Energy **165**, 290–302 (2018)
25. Future Air Cargo Executives (FACE) IATA: Vision for Air Cargo in 2030. How will the world of air cargo look like in 2030? https://sway.office.com/s1YaCALy1w8dIQBC. Accessed 23 Apr 2023

Check for updates

Multi-criteria Performance Evaluation Framework for Humanitarian Relief Organizations Using Spherical Fuzzy-TOPSIS

Sezin Güleryüz(✉)

Bartın University, 74100 Bartın, Turkey
sguleryuz@bartin.edu.tr

Abstract. In recent decades, the world has seen a rise in global disasters' variety, frequency, and intensity. As a result, humanitarian relief organizations (HROs) around the world are working hard to save lives and help survivors in disaster-prone areas with inadequate infrastructure. In these processes, the performance of HROs is crucial. Hence, well-planned and well-executed preparedness and response efforts of HROs can greatly reduce the negative consequences of disasters on society, the economy, and the environment. In this study, the performance evaluation of HROs is discussed by using Technique for Order of Preference by Similarity to the Ideal Solution (TOPSIS) with spherical fuzzy sets. The originality of the study stems from the fact that the Spherical fuzzy TOPSIS (SF-TOPSIS) technique, which is new to the literature, is applied for the first time in a case study to the problem of evaluating the performance HROs.

Keywords: Performance Evaluation · Humanitarian Relief Organizations · Spherical fuzzy TOPSIS

1 Introduction

Humanitarian relief organizations (HROs)' efficient and effective preparedness and response capabilities are crucial to achieving effective relief operations and their performance needs to be improved. After disasters, essential relief supplies such as food or medicine are vital items to helping and saving people. These items could, however, be ruined by the catastrophe or quickly run out, creating a huge demand for them. In these situations, humanitarian relief logistics (HRL) can have a major impact on community preparedness and disaster mitigation [1].

Performance evaluation of HROs in the scope of HRL supports decision-making by practitioners and managers and helps improve aid operations' effectiveness and efficiency. As a result, the transparency and accountability of HRL in disaster response increases. Given that logistics is crucial to relief operations and also a costly aspect of any relief operation, it has become critical for all organizations engaged in disaster management to evaluate the performance of their relief chains [2]. Due to limited resources, all actors involved in humanitarian relief operations must enhance their performance to

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 398–406, 2023.
https://doi.org/10.1007/978-3-031-39777-6_48

respond efficiently to disasters [3]. Therefore, this paper assesses HROs' performance focus on a Multi-Criteria Decision Making (MCDM) framework to identify their shortcomings and characteristics that need improvement. The goal is for the findings to be useful for both Turkey and HROs to assess their effectiveness in providing humanitarian aid.

MCDM techniques are appropriate for handling complex problems that have high levels of uncertainty, conflicting objectives, and multi-interests [4]. The appropriate solution to this complexity causes the need for experts who are familiar with the subject. However, since the Decision Makers' (DMs) evaluations and the information provided are uncertain, it is highly difficult to select the best options. In this study, Technique for Order Preference by Similarity to Ideal Solutions (TOPSIS) under spherical fuzzy sets (SFS) which is Spherical fuzzy TOPSIS (SF-TOPSIS) proposed by Kutlu Gündoğdu and Kahraman [5] is utilized to overcome the uncertainty of experts' judgments and expressions in their evaluations.

SFS is a relatively newly developed technique that combines Pythagorean and Neutrosophic Fuzzy Sets address the membership function on a spherical surface [5]. Different from traditional fuzzy sets, SFS can model more complex uncertainty and subjectivity, as they allow for the representation of multiple conflicting opinions and views. This makes SFS particularly useful in decision-making situations where the information is uncertain or subjective.

Based on the information presented above, the following are the study's main scientific contributions:

- The paper's novelty stems from its capacity to manage high uncertainty in a performance evaluation problem and provide a real case study with SF-TOPSIS in the performance evaluation of HROs for the first time.
- This study aims to assess the performance of HROs based on a proposed model. The proposed model assists HROs in recognizing the aspects that require enhancements and implementing necessary modifications. This can result in more efficient and effective relief operations and can help to increase transparency by providing stakeholders with information about the performance of HROs.

This paper is structured as follows: The literature survey on HROs performance evaluation and the proposed evaluation model are introduced in the 2nd section. The 3rd section deals with a real case study. Finally, the conclusion and future recommendations are presented in the fourth section.

2 Literature Review

Logistics and humanitarian supply chain are critical for responding to humanitarian crises. HROs must deliver aid promptly, efficiently, and reasonably priced. The accuracy of items and quantities provided, as well as the speed and cost of the response, are all critical factors in humanitarian aid delivery. To balance these considerations, it is essential to have a method for measuring performance [6]. In the literature, the initial efforts to measure performance in the humanitarian sector were mainly centered on evaluating the outcome characteristics of relief efforts. The Balanced Scorecard framework has been adopted in the humanitarian field due to its ability to balance various performance aspects. Davidson [7] suggested a framework that includes four performance metrics. In a similar vein, Beamon and Balcik [2] proposed a performance evaluation model comprising resource performance, output performance, and flexibility. Santarelli et al. [8] introduced a performance measurement system. Lu et al. [1] developed a model to analyze the logistics performance of organizations. This model has five attributes namely: reliability, responsiveness, agility, cost, and asset management. Roh et al. [9] presented indicators for humanitarian relief logistics. These are; reliability, responsiveness, cost, agility, and dialog with authorities. Ayyıldız and Taşkın [10] utilized to evaluate NGOs' performances by a new SCOR-based trapezoidal Type-2 Fuzzy MCDM. In their study reliability, responsiveness, and flexibility criteria are utilized. Based on what has been discussed above, the number of studies about performance evaluation and its extensions based on supply chain and logistics increasing. But none of these studies measure the overall performance of the humanitarian organization. They focus more specifically on humanitarian logistics or the humanitarian chain. This paper is therefore initiated to address the above issues by measuring the performances of HROs using the MCDM technique with the help of a proposed model.

2.1 Proposed Model

The main criteria are categorized into technical, economic, and social aspects as summarized in Table 1.

In this study proposed model is first discussed with researchers. The model's efficiency is improved by eliminating some criteria based on the literature review and DMs' opinions. The proposed model is strengthened by consulting experts, online journals, project reports, and conference papers.

2.2 MCDM Approaches Based on Spherical Fuzzy Sets

To show the applicability of different MCDM techniques with SFS, a literature review was conducted. It is seen that MCDM techniques with SFS present a powerful method to cope with uncertainty in various application areas. Many of the studies are based on case studies where SF-TOPSIS is used for selecting the alternatives. There are many applications integrated with different extensions of fuzzy sets such as; intuitionistic interval-valued fuzzy sets used for the comparison of techniques. Consequently, it is seen that no study so far focuses on the performance evaluation of HROs in SFS. This study contributes to the literature by presenting a novel model with SF-TOPSIS in HROs performance evaluation for the first time.

Table 1. Evaluation criteria of the proposed model.

Criteria	Definition
Technical aspects	
Reliability (C_{11})	The ability to deliver supplies at the correct time, quantity, and quality is a fundamental requirement for HROs [1, 2, 9]
Responsiveness (C_{12})	Responsiveness relates to the speed with which humanitarian supply chain management can respond to crises [2, 9, 10]
Technical Capability (C_{13})	The availability of infrastructure and equipment to provide satisfactory humanitarian relief operations [3, 11]
Economic aspects	
Cost (C_{21})	Costs associated with humanitarian aid arise from managing the supply chain, procurement of resources, inventory and warehouse management, and material delivery [2, 8, 9, 11]
Flexibility (C_{22})	HROs encounter multiple uncertainties and must swiftly adapt their responses to meet local requirements. This encompasses delivery time and flexibility in the upstream supply chain [1, 8–11]
Financial Capability (C_{23})	The funding mechanism coordinates and controls all activities related to relief operations [11]
Social aspects	
Dialog with Authorities (C_{31})	Establishing a good dialogue and coordination with authorities can help build a relationship that facilitates the allocation and management of supplies and the protection of individuals in disaster areas [8, 9]
Impact on Society (C_{32})	Prior experience of HROs in providing community-related projects creates a good vision for the public [9]
Skills & Experience of Personnel (C_{33})	This criterion encompasses technical competence, training, and the provision of competent personnel for humanitarian assistance [11]

3 Methodology

This section outlines the computational procedures of SF-TOPSIS adapted from Kutlu Gündoğdu and Kahraman [5]. Detailed steps are given in Fig. 1, as follows:

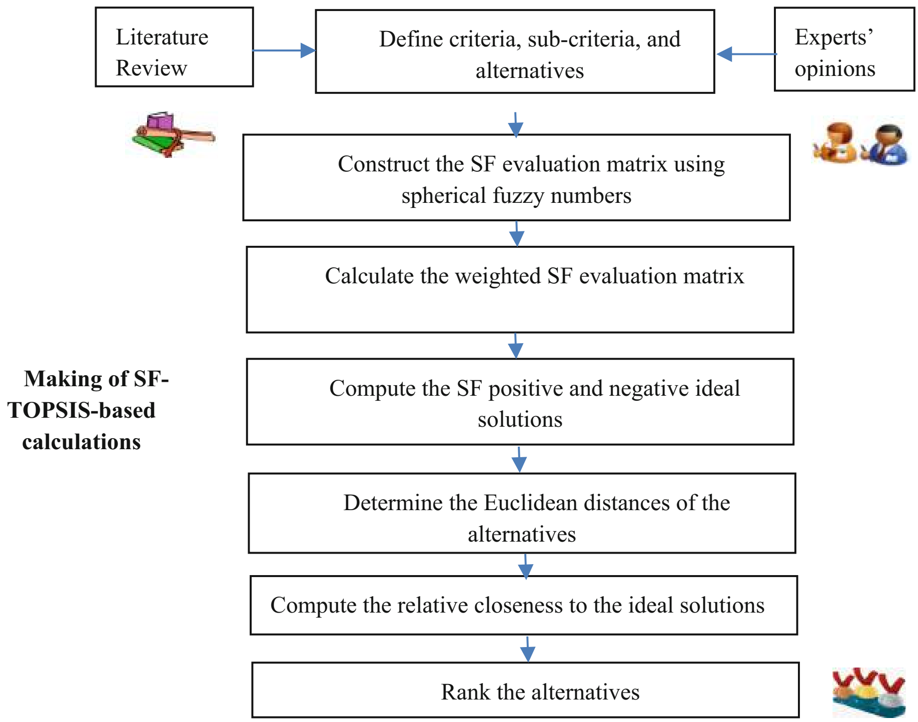


Fig. 1. Computational steps of SF-TOPSIS.

3.1 Application

The aim of the study is to develop a framework for performance measurement in HROs and to recommend key elements using an appropriate model. This study presents a real case study to evaluate the performances of HROs in Turkey. Data was collected from different independent HROs in Turkey and then analyzed with SF-TOPSIS. The study draws on the experience of three experts in the field, who are DMs (Logistician, Academician, and Logistics responsible). Interviews were conducted to get their ideas via phone or e-mail. Considering the rules of confidentiality and privacy, the evaluations were scaled with real data, and the HRO name was kept confidential. Three experts were named in the decision process (DM1, DM2, and DM3) and their evaluations of 4 alternative HROs (A1, A2, A3, A4) were analyzed.

Step 1: The application of SF-TOPSIS is illustrated by a real case. The selection criteria for performance evaluation were determined through literature review and expert opinions and given in Table 1. The scale used in the study is given in Table 2.

The weights of DMs are 0.3, 0.2, and 0.5 respectively. DMs' evaluations of the alternatives are shown in Tables 3, 4, and 5.

Step 2: DMs' spherical fuzzy assessments are combined using SWAM operator, taking into account the weights of the DMs.

Table 2. Definition and fuzzy scales of linguistic variables.

Linguistic variables	Abbrv	μ	v	π
Absolutely more important	AMI	0,9	0,1	0,1
Very high important	VHI	0,8	0,2	0,2
High important	HI	0,7	0,3	0,3
Slightly more important	SMI	0,6	0,4	0,4
Equally important	EI	0,5	0,5	0,5
Slightly low important	SLI	0,4	0,6	0,4
Low important	LI	0,3	0,7	0,3
Very low important	VLI	0,2	0,8	0,2
Absolutely low important	ALI	0,1	0,9	0,1

Table 3. Linguistic data used by DM1 For SF-TOPSIS.

DM1	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
A1	VHI	VHI	VHI	VHI	VHI	HI	HI	VHI	VHI
A2	VHI	VHI	HI	VHI	VHI	SMI	VHI	HI	VHI
A3	HI	VHI	HI	VHI	HI	VHI	HI	HI	SMI
A4	VHI	VHI	HI	VHI	HI	VHI	VHI	VHI	HI

Step 3: The evaluations of the criteria are presented in Table 6. Then, according to these assessments, the combined criteria weights of each criterion were created by using SWAM operator.

Step 4: This step involves the defuzzification of criteria weights followed by the calculation of score functions as shown in Table 7.

Table 4. Linguistic data used by DM2 For SF-TOPSIS.

DM2	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
A1	VHI	VHI	VHI	VHI	HI	HI	HI	VHI	VHI
A2	HI	VHI	HI	VHI	HI	HI	AMI	AMI	AMI
A3	HI	HI	HI	VHI	VHI	SMI	HI	HI	SMI
A4	VHI	HI	VHI	VHI	VHI	VHI	VHI	VHI	SMI

Step 5: Positive and negative Euclidean distance outcomes for each criterion are presented in Table 8.

Table 5. Linguistic data used by DM3 For SF-TOPSIS.

DM3	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
A1	VHI	VHI	VHI	AMI	HI	HI	HI	VHI	AMI
A2	HI	VHI	VHI	AMI	HI	HI	HI	VHI	AMI
A3	VHI	HI	HI	VHI	VHI	SMI	SMI	AMI	SMI
A4	HI	HI	HI	VHI	VHI	VHI	AMI	VHI	SMI

Table 6. The importance weights of criteria.

Crt.	DM1	DM2	DM3
C_{11}	VHI	AMI	AMI
C_{12}	VHI	VHI	AMI
C_{13}	HI	HI	HI
C_{21}	VHI	VHI	VHI
C_{22}	HI	HI	HI
C_{23}	SMI	HI	SMI
C_{31}	AMI	VHI	VHI
C_{32}	HI	HI	HI
C_{33}	VHI	VHI	AMI

Table 7. Score functions.

	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
A1	0,22	0,20	0,05	0,20	0,02	0,00	0,06	0,05	0,29
A2	0,12	0,20	0,03	0,20	0,02	0,00	0,15	0,05	0,32
A3	0,15	0,11	0,01	0,13	0,03	0,00	0,03	0,06	0,01
A4	0,15	0,11	0,01	0,13	0,03	0,01	0,25	0,05	0,02

Step 6–7: For each alternative, the final results are obtained by calculating the CR values as seen in Table 9.

Taking into account all criteria and their evaluations using SF-TOPSIS, the final ranking shows that A2 > A1 > A4 > A3. Therefore, it is recommended to select Alternative 2 as the best option.

Table 8. Distances to the ideal solution based on SWAM

Alt.	D PIS	D NIG
A1	0,52	0,89
A2	0,42	0,97
A3	1,12	0,20
A4	0,80	0,67

Table 9. Closeness ratio of alternatives and ranking

Alt	Closeness r.	Rank
A1	0,31	2
A2	0,00	1
A3	2,43	4
A4	1,21	3

4 Conclusion

Evaluating the performance of HROs can aid decision-making processes for practitioners and managers and enhance the efficiency and effectiveness of aid operations. This can increase transparency and accountability in HRL during disaster responses.

In, some of the proposed frameworks and models may not capture all the important aspects of humanitarian performance, such as the impact on local communities. Therefore, there is a need to develop a comprehensive framework that captures all the relevant aspects of humanitarian performance. This study developed a new evaluation model, based on a detailed literature review. Additionally, the opinions of DMs were incorporated into the model. This study utilized SF-TOPSIS to address the uncertainty in the decision process and hesitations in DM assessments to obtain the most appropriate alternative. This technique provides a comprehensive and flexible framework for evaluating multiple criteria in performance evaluations.

During the research process, some limitations were encountered. For example, the evaluations of DMs and possible deviations which could have affected the results were not taken into account. To mitigate the impact of this limitation and for future studies, it is necessary to either benefit from the opinions of more experts or to perform a comparative analysis with the SF-VIKOR method, which is another MCDM method. Additionally, another perspective for future studies could be to consider the criteria. A comprehensive analysis can be conducted by researching in detail which criterion affects the alternative results more by using the AHP technique.

References

1. Lu, Q., Goh, M., De Souza, R.: A SCOR framework to measure logistics performance of humanitarian organizations. J. Humanit. Logist. Supply Chain Manag. **6**(2), 222–239 (2016)
2. Beamon, B.M., Balcik, B.: Performance measurement in humanitarian relief chains. Int. J. Public Sect. Manag. **21**(1), 4–25 (2008)
3. Celik, E., Taskin Gumus, A.: An assessment approach for non-governmental organizations in humanitarian relief logistics and an application in Turkey. Technol. Econ. Dev. Econ. **24**(1), 1–26 (2018)
4. Büyüközkan, G., Güleryüz, S.: A new GDM-based AHP framework with linguistic interval fuzzy preference relations for renewable energy planning. J. Intell. Fuzzy Syst. **27**(6) (2014)
5. Kutlu Gündoğdu, F., Kahraman, C.: Spherical fuzzy sets and spherical fuzzy TOPSIS method. J. Intell. Fuzzy Syst. **36**(1), 337–352 (2019)
6. Sahebjamnia, N., Torabi, S.A., Mansouri, S.A.: A hybrid decision support system for managing humanitarian relief chains. Decis. Support Syst. **95**, 12–26 (2017)
7. Davidson, A.L.: Key performance indicators in humanitarian logistics: master thesis of engineering in logistics, Massachusetts Institute of Technology, Cambridge (2006)
8. Santarelli, G., Abidi, H., Klumpp, M., Regattieri, A.: Humanitarian supply chains and performance measurement schemes in practice. Int. J. Product. Perform. Manag. **64**(6), 784–810 (2015)
9. Roh, S., Lin, H.H., Jang, H.: Performance indicators for humanitarian relief logistics in Taiwan. Asian J. Shipp. Logist. **38**(3), 173–180 (2022)
10. Ayyildiz, E., Taskin, A.: Humanitarian relief supply chain performance evaluation by a SCOR-based Trapezoidal type-2 fuzzy multi-criteria decision-making methodology: an application to Turkey. Scientia Iranica **29**(4), 2069–2083 (2022)
11. Kim, S., Ramkumar, M., Subramanian, N.: Logistics service provider selection for disaster preparation: a socio-technical systems perspective. Ann. Oper. Res. **283**(1–2), 1259–1282 (2018). https://doi.org/10.1007/s10479-018-03129-3

Signal Recognition Using Weighted Additive Convolution of Evaluation Criteria

Ramin Rzayev(✉) and Azer Kerimov

Institute of Control Systems, Vahabzadeh Street 68, AZ1141 Baku, Azerbaijan
raminrza@yahoo.com

Abstract. Existing signal recognition methods have both their advantages and disadvantages, which are found when recognizing signals from classes defined by different characteristic standards. Therefore, for signals from different classes, the indicators of recognition quality by one method or another can differ significantly. There is a need to create a more balanced method capable of providing the necessary stability relative to the accuracy and reliability of the final results in the process of recognizing signals from various classes. As such signal recognition method, the article proposes to use an approach based on the use of weighted additive convolution of signal proximity criteria using the Euclidean metric. Euclidean distances between reference points are used as evaluation criteria, which are used in the context of applying the four most well-known recognition methods: the amplitude method (the trivial Euclid method), the DDTW method using the values of the first derivatives, and methods based on the Wavelet transform and the Fourier transform.

Keywords: Signal Recognition · Euclidean Metric · Additive Convolution

1 Introduction

In [1, 2], we considered some well-known signal recognition methods, the accuracy of which was compared on the basis of artificially generated families of sequentially shifted signals. It was noted that each of the analyzed recognition methods, due to its characteristic feature, answers the question of the proximity of signals in two main positions: by the amplitude characteristics of the curves reflecting the signals, or by their orientation in space, determined by the corresponding values of the derivatives of the 1^{st} and 2^{nd} orders. So, for example, when recognizing using the Euclidean distance, only the amplitude characteristic of the signals is used. At the same time, the DDTW (Derivative Dynamic Time Warping) method [3] uses only the characteristic of signals associated with the spatial orientation of signal curves (values of 1^{st} order derivatives). In particular, the 1^{st} order derivative of the signal is used in solving many recognition problems [4–6]. In the review article [4], using the 1^{st} order derivative, methods for recognizing signals received from spectrometers are analyzed in the context of solving biological problems. In [5], 1^{st} order derivatives are used to correct the baseline of the signal, as well as to remove jumps in signals received from spectrometers. In [6],

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 407–416, 2023.
https://doi.org/10.1007/978-3-031-39777-6_49

using the 1st order derivative, issues related to the appearance of various noises after the sampling of the analog signal are considered. Based on the existing developments in the subject area, the importance and relevance of further research the signal recognition methods become obvious.

2 Criteria for Assessing the Proximity of Signals

At the preliminary stage of signal recognition, as a rule, the main features of recognition are identified and, on their basis, the corresponding distance norm is selected. Next, the recognition procedure is carried out by comparing the recognizable signals with the standard by calculating the pairwise distances between them based on the selected metric. The choice of recognition features depends on the nature of the problem being solved (the family of recognizable signals) and the applied approach. However, in all cases, the Euclidean metric is used as the basic norms of the distance between reference points of corresponding signals. So, to form the weighted additive convolution, four distance norms were chosen as evaluation criteria, using which the following known methods of signal recognition are used: the amplitude method, the method using 1st order derivative (DDTW), Wavelet and Fourier transforms. At the same time, for each of the listed methods, the corresponding features of recognition are determined.

Amplitude Recognition Method (ARM). The values of reference points are chosen as recognition features. In particular, if for two arbitrary signals x and y the points a_i and b_i ($i = 0, 1, \ldots, N$) are the reference points, respectively, then the Euclidean metric is chosen as the norm of the distance between them in the form

$$D_1(x, y) = \sqrt{\sum\nolimits_{i=1}^{N} (a_i - b_i)^2}. \tag{1}$$

DDTW Recognition Method. As recognition features the values of the 1st derivatives at the reference points are chosen. In the discrete case, the expression $\dot{a}(i) = [a(i) - a(i-1)]/T$ is taken as the 1st order derivative, where $a(i) = a(iT)$, $i = 0, 1, \ldots, N$; T is the sampling period of the analog signal a [7]. In particular, if for two arbitrary signals x and y the reference points are respectively the values of the 1st derivatives p_i and q_i ($i = 0, 1, \ldots, N$), then the Euclidean metric is chosen as the norm of the distance between them as follows

$$D_2(x, y) = \sqrt{\sum\nolimits_{i=1}^{N} (p_i - q_i)^2}. \tag{2}$$

Wavelet Transform (WT). According to this recognition method, each signal is decomposed into high-frequency and low-frequency components. Moreover, each component is characterized by the values of the so-called detailing and approximating coefficients. For example, for the signal shown in Fig. 1(a), which includes 256 reference points, the WT at four levels looks like that shown in Fig. 1(b).

Here, average values of characteristics (coefficients) in each filtering band are chosen as recognition features. In the case under consideration (see Fig. 1(b)), where four

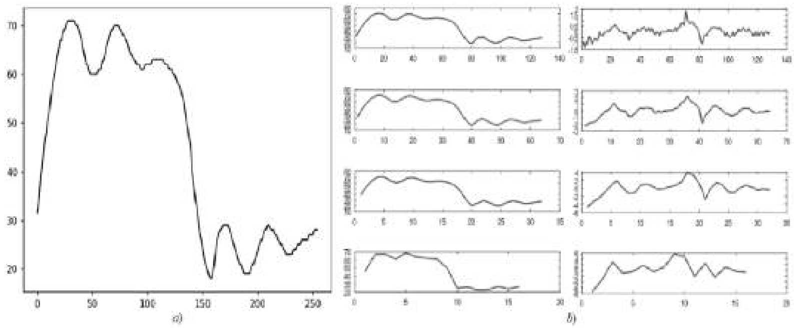


Fig. 1. Signal including 256 reference points (a) and its WT at 4 levels (b).

decomposition levels are chosen, there are eight values of recognizable features [8, 9]. Assuming for two arbitrary signals x and y the average values and standard deviations of the coefficients in the high-frequency and low-frequency bands, respectively, in terms of H_{1i}, L_{1i}, H_{2i}, L_{2i}, ($i = 0, 1, \ldots, N$), where N is the number of decomposition levels, the following Euclidean metric is chosen as distance norm

$$D_3(x, y) = \sqrt{\sum\nolimits_{i=1}^{N} (H_{1i} - H_{2i})^2 + \sum\nolimits_{i=1}^{N} (L_{1i} - L_{2i})^2}. \tag{3}$$

Fourier Transform (FP). The use of FT implies the creation of a spectral image for a recognizable signal. In particular, for the signal that includes 256 reference points (see Fig. 1(a)), the FT generates the corresponding amplitude spectral image (see Fig. 2) that includes 128 reference points. When applying the FT, the variables of the amplitude spectrum are considered as recognizable features [10], which in common form for two signals are denoted as f_{1i} and f_{2i} ($i = 0, 1, \ldots, N$), where N is the number of variables. In this case, the following Euclidean metric is chosen as the norm of the distance between two arbitrary signals x and y

$$D_4(x, y) = \sqrt{\sum\nolimits_{i=1}^{N} (f_{1i} - f_{2i})^2}. \tag{4}$$

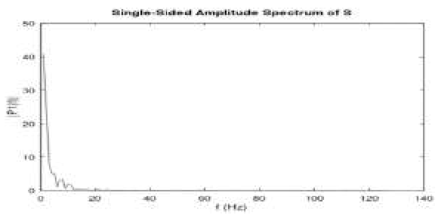


Fig. 2. Amplitude spectrum of the signal obtained using the FT.

Thus, in view of the foregoing, to recognize the signals the weighted additive convolution (WAC) of criteria D_k (x, y) $(k = 1 \div 4)$ is proposed in the following form

$$D = \sum\nolimits_{k=1}^{4} w_k D_k(x, y), \tag{5}$$

where w_k are the weights of the evaluation criteria, reflecting the contribution of each of the above metrics (1)–(4) in solving the problem.

To identify the weights w_k, which a priori must satisfy the conditions: $\sum_{k=1}^{4} w_k = 1$, $0 \int w_k \int 1$, the application of the mentioned signal recognition methods is considered on the individual basis using the example of a single class of curves. As such a class, it is chosen the artificial family of signals formed by the uniform displacement of the curves horizontally. A general analysis of recognition results is carried out on the basis of 2 criteria for assessing the adequacy of recognition method.

Criterion 1 (*sensitivity*): for a particular recognized signal, the Euclidean distances from the left standing and from the right standing signals should be approximately equal, that is, their ratio should be approximately equal to one. If the standing signals on the left and on the right are symmetrical with respect to this signal, then these distances will be absolutely equal.

Criterion 2 (*stability*): increasing the step of signal shifts cannot improve the satisfaction of recognition methods, that is, the accuracy of the recognition method must remain the same or deteriorate.

Obviously, the weights of the evaluation criteria correlate with the recognition results, that is, they are in a certain proportion with the results obtained using the specified recognition methods separately. Therefore, we have chosen the following statement as the main paradigm: how many times the recognition results will differ using the particular method, the weight coefficients corresponding to it in the additive convolution (5) will differ so many times. Based on this paradigm, the identification of weights w_k is carried out on the basis of Criteria 1 and 2.

3 Analysis of Recognition Results

To form the family of recognizable signals, the signal s_0 was chosen as the base (standard) signal. Recognizable signals are built relative to s_0 by successive uniform horizontal displacement [1]. Assuming a shift of 10 units as the step *h,* the artificial family signals $S_{10} = \{s_1, s_2, \ldots, s_6\}$ is obtained as shown in Fig. 3.

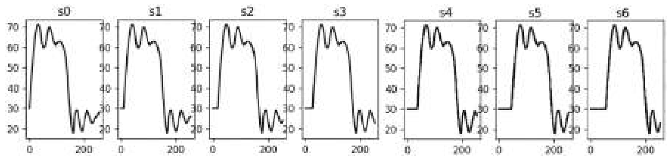


Fig. 3. Standard s_0 and family of recognizable signals $S_{10} = \{s_1, s_2, \ldots, s_6\}$.

For each $k = 1 \div 4$, let us introduce the following notation: $D_k^h(x, y)$ is the distance between signals; $R_{ki}^h = \frac{D_k^h(s_i, s_{i+1})}{D_k^h(s_{i-1}, s_i)}$ $(i = 1 \div 5)$ is the ratio between adjacent distances (that is, between the distances from the right standing $(i + 1)$-th and from the left standing $(i$-$1)$-th signals to the i-th signal). Then, in these notations, the satisfaction of the methods in terms of compliance with Criteria 1 and 2 can be formulated as follows:

- the adequacy of the recognition method for compliance with Criterion 2 is evaluated based on the value of the maximum deviation, defined as $G_k^h = \max\limits_{i=1\div 5}\{1 - R_{ki}^h\}$;
- the adequacy of the recognition method for compliance with Criterion 4 is evaluated based on the fulfillment of the condition $R_{ki}^{h_1} < R_{ki}^{h_2}$ $(i = 1 \div 5; k = 1 \div 4)$, if $h_1 < h_2$ is satisfied, where h_1 and h_2 are the steps of curve displacements (for example, horizontally) in two different families of recognizable signals.

Results of pairwise comparison of signals from the family $S_{10} = \{s_1, s_2, \ldots, s_6\}$ using metrics (1)–(4) sufficiently satisfy Criterion 2. This is confirmed by the comparative estimates summarized in Tables 1, 2, 3 and 4, as well as subsequent calculations.

Table 1. Pairwise comparisons of signals using the metric (1).

	$s0$	s_1	s_2	s_3	s_4	s_5	s_6
s_0	0	113.10	197.99	252.89	291.48	328.66	367.74
s_1	113.10	0	112.96	197.67	252.69	291.45	328.64
s_2	197.99	112.96	0	112.80	197.58	252.59	291.41
s_3	252.89	197.67	112.80	0	112.66	197.16	252.39
s_4	291.48	252.69	197.58	112.66	0	112.24	196.89
s_5	328.66	291.45	252.59	197.16	112.24	0	112
s_6	367.74	328.64	291.41	252.39	196.89	112	0

Table 2. Pairwise comparisons of signals using the metric (2).

	$s0$	s_1	s_2	s_3	s_4	s_5	s_6
s_0	0	15.572	21.303	21.575	19.527	18.680	19.658
s_1	15.572	0	15.490	21.243	21.461	19.405	18.609
s_2	21.303	15.490	0	15.441	21.136	21.401	19.326
s_3	21.575	21.243	15.441	0	15.287	21.046	21.341
s_4	19.527	21.461	21.136	15.287	0	15.199	20.800
s_5	18.680	19.405	21.401	21.046	15.199	0	14.925
s_6	19.658	18.609	19.326	21.341	20.800	14.925	0

Tables 5 and 6 present the values of the ratios of adjacent distances R_{ki}^{10} and R_{ki}^{20}, as well as the values of the maximum deviations G_k^{10} and G_k^{20} for two basic families of curves S_{10} and S_{20}, constructed, respectively, by uniformly shifting the curves horizontally by step $h_1 = 10$ and step $h_2 = 20$.

Table 3. Pairwise comparisons of signals using the metric (3).

	$s0$	s_1	s_2	s_3	s_4	s_5	s_6
s_0	0	47.200	91.698	134.15	178.20	223.49	267.04
s_1	47.200	0	45.804	90.048	135.27	181.64	226.26
s_2	91.698	45.804	0	46.112	93.499	141.03	185.51
s_3	134.15	90.048	46.112	0	49.772	98.313	141.90
s_4	178.20	135.27	93.499	49.772	0	49.025	93.822
s_5	223.49	181.64	141.03	98.313	49.025	0	47.586
s_6	267.04	226.26	185.51	141.90	93.822	47.586	0

Table 4. Pairwise comparisons of signals using the metric (4).

	$s0$	s_1	s_2	s_3	s_4	s_5	s_6
s_0	0	0.51072	0.92713	1.38810	1.60620	1.58330	1.65850
s_1	0.51072	0	0.75580	1.28820	1.59980	1.56370	1.61820
s_2	0.92713	0.75580	0	1.04720	1.46540	1.47470	1.52160
s_3	1.38810	1.28820	1.04720	0	0.91938	1.02530	1.11600
s_4	1.60620	1.59980	1.46540	0.91938	0	0.46027	0.79208
s_5	1.58330	1.56370	1.47470	1.02530	0.46027	0	0.71235
s_6	1.65850	1.61820	1.52160	1.11600	0.79208	0.71235	0

Table 5. Indicators of fulfillment of Criteria 2 and 4 based on the family S_{10}.

Method	k	R_{k1}^{10}	R_{k2}^{10}	R_{k3}^{10}	R_{k4}^{10}	R_{k5}^{10}	G_k^{10}
ARM	1	0.99883	0.99851	0.99878	0.99625	0.99789	0.0037495
DDTW	2	0.99611	0.99743	0.99162	0.99650	0.98405	0.0159500
WT	3	0.97043	1.00670	1.07940	0.98498	0.97066	0.0793870
FT	4	1.47990	1.38550	0.87795	0.50063	1.54770	0.5476600

Further, based on the G_k^{10} and G_k^{20} (see Tables 5, 6), to reflect the so-called "deteriorations" from the application of distance norms D_k in the process of recognizing signals from S_{10} and S_{20} the corresponding coefficients u_k are calculated as follows:

$$u_1 = \frac{G_1^{20}}{G_1^{10}} = 4.8508;\ u_2 = \frac{G_2^{20}}{G_2^{10}} = 9.1154;\ u_3 = \frac{G_3^{20}}{G_3^{10}} = 0.9800;\ u_4 = \frac{G_4^{20}}{G_4^{10}} = 1.8698.$$

Table 6. Indicators of fulfillment of Criteria 2 and 4 based on the family S_{20}.

Method	k	R_{k1}^{20}	R_{k2}^{20}	R_{k3}^{20}	R_{k4}^{20}	R_{k5}^{20}	G_k^{20}
ARM	1	0.99794	0.99640	0.98876	0.98694	0.98181	0.018188
DDTW	2	0.99385	0.98399	0.97312	0.94591	0.85461	0.145390
WT	3	1.03360	0.98548	1.06750	0.93302	1.07780	0.077800
FT	4	1.49200	0.81074	1.95100	0.38305	2.02400	1.024000

The factors u_k are used to identify the weights of the evaluation criteria by follows:

$$\begin{cases} \frac{w_k}{w_l} = \frac{u_l}{u_k}, \text{if } k, l = 2, 3, 4; \\ \sum_{k=1}^{4} w_k = 1. \end{cases} \tag{6}$$

As a result of solving the system of Eqs. (6), the following values were obtained: $w_1 = 0.11018$, $w_2 = 0.058632$, $w_3 = 0.54535$, $w_4 = 0.28584$. Then, the weighted additive convolution formula (5) can be rewritten as follows

$$D = 0.11018 \cdot D_1(x, y) + 0.058632 \cdot D_2(x, y) + 0.54535 \cdot D_3(x, y) + 0.28584 \cdot D_4(x, y) \tag{7}$$

Further, the convolution (7) was tested on the families $S_h = \{s_0, s_1, s_2, \ldots, s_6\}$, where in each case s_0 is the standard, relative to which these families are formed by steps $h = 5, 10, 15, 20, 30$. For each family the maximum deviations from unity of adjacent distance ratios were obtained and summarized in Table 7.

Table 7. Values of maximum deviations from unity of ratios of adjacent distances between reference points of curves from families S_h.

Method	k	G_k^5	G_k^{10}	G_k^{15}	G_k^{20}	G_k^{30}
ARM	1	0.00069881	0.0037495	0.013154	0.018188	0.25759
DDTW	2	0.00390210	0.0159500	0.020242	0.145390	0.14935
WT	3	0.01730900	0.0793870	0.132200	0.077800	0.46902
FT	4	0.37622000	0.5476600	1.043600	1.024000	2.86460
WAC (7)		0.01148900	0.0500170	0.098663	0.056458	0.39412

As can be seen from Table 7 and Fig. 4, for all families of curves S_h ($h = 5$, 10, 15, 20, 30), WAC (7) demonstrates smaller deviations from unity than the WT criterion with the highest weight (0.54535).

The proposed approach was tested on the example of other artificial families of curves constructed in a similar way on the basis of the standard c_0 (see Fig. 5), which, as it is easy to see, is fundamentally different from s_0. As such families of curves, C_5

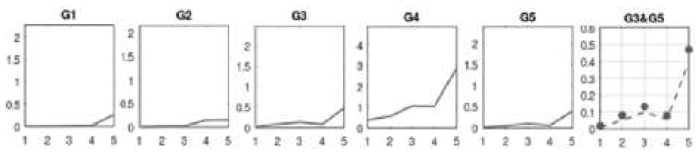


Fig. 4. Demonstration of deviations from unity of ratios of adjacent distances between reference points of curves from families S_h.

$= \{c_1, c_2, \ldots, c_6\}$ and $C_{10} = \{c_1, c_2, \ldots, c_6\}$ are chosen, which are constructed by successive uniform displacement of signals by steps $h = 5$ and $h = 10$, respectively. Based on calculations similar to the previous case, the maximum deviations from unity of the ratios of adjacent distances between curves from the C_5 and C_{10} families were determined and summarized in Table 8.

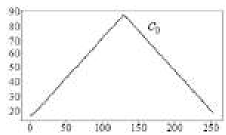


Fig. 5. Base signal c_0 for constructing artificial families of curves.

Table 8. Values of maximum deviations from unity of ratios of adjacent distances between reference points of curves from families C_5 and C_{10}.

Method	k	G_k^5	G_k^{10}
ARM	1	0.00069881	0.0037495
DDTW	2	0.00390210	0.0159500
WT	3	0.01730900	0.0793870
FT	4	0.37622000	0.5476600
WAC		0.01148900	0.0500170

As in the previous case, the corresponding "deterioration" coefficients were obtained, which made it possible for the given class of recognized signals to identify the weights of the proximity assessment criteria in the additive convolution (5) as following numbers: $w_1 = 0.19766$, $w_2 = 0.24375$, $w_3 = 0.19795$, $w_4 = 0.36064$. As a result, the weighted additive convolution formula is rewritten as:

$$D = 0.19766 \cdot D_1(x, y) + 0.24375 \cdot D_2(x, y) + 0.19795 \cdot D_3(x, y) + 0.36064 \cdot D_4(x, y) \tag{8}$$

WAC (8) was tested on the families $C_h = \{c_1, c_2, \ldots, c_6\}$, where in each case c_0 is the standard, relative to which these families are formed by sequential shifting the curves by steps $h = 5, 10, 15, 20, 30$. In particular, for each family of curves using all recognition methods, the values of maximum deviations from unity of the ratios of adjacent distances between the corresponding reference points of the curves were obtained and summarized in Table 9. As can be seen from Table 9 and Fig. 6, for all families of curves C_h ($h = 5$, 10, 15, 20, 30) the convolution (8) demonstrates smaller deviations from unity than the FT-method with the highest weight (0.36064) in (8).

Table 9. Values of maximum deviations from unity of ratios of adjacent distances between reference points of curves from families C_h.

Method	k	G_k^5	G_k^{10}	G_k^{15}	G_k^{20}	G_k^{30}
ARM	1	0.010082	0.021531	0.039536	0.065291	0.11418
DDTW	2	0.013796	0.023892	0.018411	0.027104	0.28476
WT	3	0.051320	0.109440	0.271830	0.212770	0.38641
FT	4	0.688910	0.806350	0.838810	0.829050	0.84601
WAC (8)		0.024082	0.052365	0.112500	0.109440	0.18649

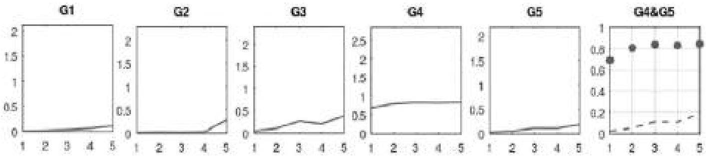


Fig. 6. Demonstration of deviations from unity of ratios of adjacent distances between reference points of curves from families C_h.

4 Conclusion

Theoretical prerequisites for creating a method using weighted additive convolution of proximity criteria on specific examples have proved their consistency. Numerical experiments have shown that the weighted additive convolution of evaluation criteria, adapting to a specific class of signals, demonstrates better results than traditional recognition methods, even those that have the greatest weight in the WAC.

References

1. Kerimov, A.B.: Comparison of some signal recognition methods for their adequacy. In: Proceedings on 8th International Conference on Control and Optimization with Industrial Applications, Baku, Azerbaijan, vol. I, 273–275 (2022)
2. Kerimov, A.B.: Accuracy comparison of signal recognition methods on the example of a family of successively horizontally displaced curves. Inf. Control Probl. **42**(2), 80–91 (2022)
3. Keogh, E.J., Pazzani, M.J.: Derivative Dynamic Time Warping. https://epubs.siam.org/doi/epdf/10.1137/1.9781611972719.1

4. Santos, M., Morais, C., Nascimento, Y., Araujo, J., Lima, K.: Spectroscopy with computational analysis in biological studies: a decade (2006–2016). TrAC Trends Anal. Chem. **97**, 244–256 (2017)
5. Liu, Y., Lin, J.: A general-purpose signal processing algorithm for biological profiles using only first-order derivative information. BMC Bioinf. **20**, 611 (2019). https://doi.org/10.1186/s12859-019-3188-4
6. Cedro, L., Janecki, D.: Determining of signal derivatives in identification problems, FIR differential filters. Acta Montanistica Slovaca Rocník **16**(1), 47–54 (2011)
7. Novozhilov, B.: Calculation of the derivative of an analog signal in a programmable logic controller. Aerosp. Sci. J. Moscow State Techn. Univ. **4**, 1–12 (2016). (in Russian)
8. Saraswat, S., Srivastava, G., Shukla, S.: Wavelet transform based feature extraction and classification of atrial fibrillation arrhythmia. Biomed. Pharmacol. J. **10**(4), 1715–1725 (2017)
9. Zhao, M., Chai, Q., Zhang, S.: A method of image feature extraction using wavelet transforms. In: Huang, DS., Jo, KH., Lee, HH., Kang, HJ., Bevilacqua, V. (eds.) ICIC 2009. LNCS, vol. 5754, pp. 187–192. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04070-2_21
10. Hindarto, Anshory, I., Efiyanti, A.: Feature extraction of heart signals using fast Fourier transform. In: Proceeding on 1st IBSC: Towards the Extended Use of Basic Science for Enhancing, Indonesia, pp. 165–167. The University of Jember (2016)

An Application of the Fuzzy TODIM Method for Evaluating Closed-Loop Supply Chain's Reproduction Process in E-Waste

Yiğit Kazançoğlu and Melis Tan Taçoğlu(✉)

Logistics Management Department, Yaşar University, Bornova, İzmir, Turkey
{yigit.kazancoglu,melis.tacoglu}@yasar.edu.tr

Abstract. A closed-loop supply chain (CLSC) is a circular system that keeps resources and materials in use for as long as possible, reducing waste and creating economic and environmental benefits with the collaboration of many stakeholders. Designing CLSC is especially crucial in electronic waste or e-waste. This study aims to use the Fuzzy TODIM technique to examine the uncertainty parameters used in modeling the CLSC and choose the best alternative to CLSC's reproduction process for Waste Electrical and Electronic Equipment (WEEE) with regard to uncertain criteria. The usage of the Fuzzy TODIM application in CLSC's reproduction process is contributed to the literature on a new implementation area of this approach. This study emphasizes six uncertain criteria (cost, demand, type of returned items, returned quality, returned quantity, and time delay) while comparing three distinct forms of the WEEE reproduction process (recycling, remanufacturing, reuse/repair). This study demonstrated that reuse/repair reproduction is the best alternative, followed by remanufacturing and recycling. Cost and return quality are the essential criteria, among others.

Keywords: Closed Loop Supply Chain · Fuzzy TODIM · E-Waste

1 Introduction

A closed-loop supply chain is a sustainable approach to designing the entire product lifecycle, from its production to its disposal, using multiple times as long as possible through reproduction processes. One crucial area of the closed-loop supply chain is e-waste, which refers to electronic products that have reached the end of their life and are discarded, also called WEEE. Waste Electrical and Electronic Equipment (WEEE) is a growing concern in today's society due to the rapid pace of technological improvements and the resulting increase in the production and disposal of electronic devices. WEEE refers to any electronic product that has reached the end of its life cycle and is no longer useful or safe. Many uncertainties in designing and implementing reproduction processes have emerged from the environment, stakeholders' disagreement, etc. Modeling the complex system with the involvement of many uncertainties is achieved by defining a vast number of uncertain parameters. The main goal of this research is to analyze

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 417–425, 2023.
https://doi.org/10.1007/978-3-031-39777-6_50

the uncertainty parameters or variables in CLSC and select the best alternative CLSC's reproduction process for WEEE among alternatives concerning uncertain criteria using the Fuzzy TODIM approach. It has contributed to the literature using the Fuzzy TODIM application in CLSC's reproduction process. In this paper, three different types of reproduction processes (re-cycling, remanufacturing, reuse/repair) for WEEE consider six criteria: cost, demand, type of returned goods, returned quality, return quantity, and time delay. This study indicates that reuse/repair reproduction is the best CLSC design, and cost and collected product quality are the most crucial elements in this design.

This paper is organized as a literature review, methodology, application, and conclusion, respectively.

2 Literature Review

Regarding supply chain operations, management, and stakeholders, a Reverse Supply Chain (RSC) differs significantly from a Forward Supply Chain (FSC). FSC and RSC operations impact and assist one another in a well-established CLSC. Various stakeholders and actors are involved in each stage of CLSC [1]. To examine the uncertainty of the CLSC, sources of the uncertainty must be clearly defined and classified based on the uncertainties.

The studies of [2] and [3] are of paramount importance in the analysis of CLSC/RL uncertainty literature due to evaluating all areas of the RL. These two studies were publicized in 2009; hence CLSC had attracted academicians in 2009, and they had only begun working on CLSC that year. Then, an extensive literature review for CLSC for environmental, constitutional, social, and economic aspects was conducted [4]. The authors analyzed 382 published articles between 2007 to 2013. According to the research, demand and return amounts were mainly used uncertain parameters in the articles with the following rates, delivery time, lead time, transportation time, waste generation, environmental problems, risk elements, and weights.64 CLSC/RL articles published in 2017 were analyzed for dynamic complexity and deep uncertainty in the CLSC [5]. They used "Deep Uncertainty" to evaluate many possibilities regardless of leveling factors in the environment. Many stakeholders and actors are involved in examining sources of uncertainty in the decision-support literature based on the Walker et al. notion [6]. They classified sources of uncertainty in CLSC into four groups: the context, the structure and system behavior, the parameters, the system model outcome, and the uncertainty related to the CLSC actors [5]. In the forthcoming years, 302 published CLSC studies from 2004 to 2018 analyzed inevitable uncertainties from different perspectives in the CLSC: design, modeling and implementation point of view [7]. This paper categorizes uncertainties into four groups: parameter, context, CLSC model structure, and CLSC model result based [8].

Some researchers used different sources of uncertainty classification in CLSC. [9] classify uncertainty into three groups: internal organization (production operation and product properties), internal supply chain (need, supply chain structure, and supply chain), and external (environmental, catastrophic events, and global economic issues). [10] categorized uncertainties based on modeling approaches that can be either stochastic or deterministic by analyzing CLSC literature between 2001 to 2019. He claimed that a vast number of uncertain parameters manage uncertainty in the CLSC structure: demand, return qualities, cost, capacity, return quality, price, lead times, risk, disposal rate, supply, collection rate, manufacturing rate, carbon emissions, material flow, transportation distance within the facility, transportation location, flexibility, facility location, and supplier selection. It concludes that there is neither a certain class identification nor a specific division for the uncertain factors for CLSCs [7].

3 Methodology

The TODIM approach (an acronym in Portuguese of Interactive and Multi-criteria Decision Making) is a discrete multi-criteria approach entrenched in Prospect Theory. This approach compared each alternative pair based on each criterion to discard infrequent divergence from these relationships [11]. Fuzziness is used to mitigate the impact of decision-makers' prejudicial opinions and prejudices in ordering alternatives. This study examines the Fuzzy TODIM method proposed by Tosun et al. [12]. The suggested method's algorithm is summarized as follows:

1. Two sets is defined for alternatives and criteria: $\mathrm{M} = \{1, 2, .., \mathrm{m}\}$ and $\mathrm{N} = \{1, 2, .., \mathrm{n}\}$. Assume $\mathrm{A} = \{\mathrm{A}_1, \mathrm{A}_2, \ldots, \mathrm{A}_\mathrm{m}\}$ represents an alternative set, where A_i represents the alternative i and $\mathrm{C} = \{\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_\mathrm{n}\}$is a finite criterion set, where C_j indicates the criteria j. Let $\mathrm{w} = \{\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_\mathrm{n}\}^\mathrm{T}$ is a weight vector, where w_n is the weight of criterion n.
2. Linguistic variables are expressed. Let $\mathrm{S} = \{\mathrm{s}_\mathrm{f} | \mathrm{f} = 0, 1, .., \mathrm{T}\}$ represents linguistic variable set. Each member in this set is showed by a triangular fuzzy number $\tilde{A} = (\mathrm{l}, \mathrm{m}, \mathrm{u})$. Values are created by using equation given in the following:

$$\tilde{A} = (l, m, u) = \left(\max\left(\frac{f-1}{T}, 0\right), \frac{f}{T}, min\left(\frac{f+1}{T}, 1\right)\right) \quad (1)$$

3. Let d represents the number of decision-makers. Each criterion and alternative is calculated based on the number of decision-makers by using the equation given in the following. cw_j (fuzzy criteria weights) indicates the average weight of criterion j with respect to d number of decision makers $\tilde{x}_{ij}$ is the integrated linguistic evaluation of alternative i as to criterion j, where $\tilde{x}_{ij} \int S, i\nu M, j\nu N$.

$$\tilde{x}_{ij} = \frac{1}{d}\left[\sum\nolimits_{e=1}^{d} \tilde{x}_{ij}^{e}\right] \quad i = 1, 2, 3, \ldots, m \quad (2)$$

$$cw_j = \frac{1}{d}\left[\sum\nolimits_{e=1}^{d} cw_j^{e}\right] \quad j = 1, 2, 3, \ldots, n \quad (3)$$

4. Defuzzification of cw_j is done by optimism index. In this approach, optimism index ($\pi\nu[0, 1]$) is implemented to the equation to assess decision-makers attitudes towards risk. If $\alpha = 0.5$, it is claimed that the decision maker is neutral. In this study, $\alpha = 0.5$. If the decision maker is pessimistic and risk-averse then α is close to zero. Let $\tilde{A}_j$ is the triangular fuzzy number of alternative j and each alternative value after

defuzzification is represented $V\left(\tilde{A}_j\right)$ and equation is shown in the following where $S = \bigcup_{j=1}^{n} S_j$; $S_j = (l_j, m_j, u_j)$, $x_{min} = \inf S$, $x_{max} = \sup S$:

$$V\left(\tilde{A}_j\right) = m_j\left\{\pi\left[\frac{u_j - x_{min}}{x_{max} - x_{min} + u_j - m_j}\right] + (1-\pi)\left[1 - \frac{x_{max} - l_j}{x_{max} - x_{min} + m_j - l_j}\right]\right\} \tag{4}$$

5. Each criterion weight is normalized with the formula given below:

$$w_j = \frac{V\left(\tilde{A}_j\right)}{\sum_{j=1}^{n} V\left(\tilde{A}_j\right)} \tag{5}$$

6. Maximum w_j value where j νN is selected as a reference point and regarding this value each criterion relative weight (w_{jr}) is calculated by given formula below:

$$w_{jr} = w_j / w_r \quad \forall j \int N, \max\{w_j | j \int N\} \tag{6}$$

7. Gain and loss matrices are constructed by $\tilde{x}_{ij}$ values. Suppose s_b and s_d, where b, d = 0, 1, …, T, are the linguistic evaluation of alternative A_i and A_g where i, g $\int M$, corresponding criterion j, where j $\int N$ respectively. Fuzzy values relation is shown below:
   1. If $s_b > s_d$, then $\tilde{x}_{ij} > \tilde{x}_{gj}$
   2. If $s_b < s_d$, then $\tilde{x}_{ij} < \tilde{x}_{gj}$
   3. If $s_b = s_d$, then $\tilde{x}_{ij} = \tilde{x}_{gj}$

The gain and loss of alternative A_i with respect to A_g for criterion j (C_j) is calculated to generate gain matrix ($[G_{ig}^{j}]_{mm}$) and loss matrix ($[L_{ig}^{j}]_{mm}$) by following equations where $d\left(\hat{x}_{ij}, \hat{x}_{gj}\right)$ is the distance between two fuzzy variables that represented by a triangular number:

$$d\left(\hat{x}_{ij}, \hat{x}_{gj}\right) = \sqrt{1/3\left[\left(x_{ij}^{l} - x_{gj}^{l}\right)^2 + \left(x_{ij}^{m} - x_{gj}^{m}\right)^2 + \left(x_{ij}^{u} - x_{gj}^{u}\right)^2\right]} \tag{7}$$

For benefit criterion:

$$G_{ig}^{j} = \begin{cases} d\left(\dot{x}_{ij}, \dot{x}_{gj}\right), & if\, \tilde{x}_{ij} \geq \tilde{x}_{gj} \\ 0, & otherwise \end{cases} \tag{8}$$

$$L_{ig}^{j} = \begin{cases} 0, & if\, \tilde{x}_{ij} \geq \tilde{x}_{gj} \\ -d\left(\dot{x}_{ij}, \dot{x}_{gj}\right), & otherwise \end{cases} \tag{9}$$

For cost criterion

$$G_{ig}^{j} = \begin{cases} 0, & if\, \tilde{x}_{ij} \geq \tilde{x}_{gj} \\ -d\left(\dot{x}_{ij}, \dot{x}_{gj}\right), & otherwise \end{cases} \tag{10}$$

$$L_{ig}^{j} = \begin{cases} -d\left(\dot{x}_{ij}, \dot{x}_{gj}\right), & if\, \tilde{x}_{ij} \geq \tilde{x}_{gj} \\ 0, & otherwise \end{cases} \tag{11}$$

8. Dominance degree of gain ($\tau_{ig}^{j(+)}$) and dominance degree of loss ($\tau_{ig}^{j(-)}$) is examined based on gain and loss matrixes and the dominance degree matrix ($[\tau_{ig}^{j}]_{mm}$) is constructed by their summation shown as follows: (ω is the attenuation factor and it is accepted 1 in this study)

$$\tau_{ig}^{j(+)} = \sqrt{G_{ig}^{j} w_{jr} / \sum\nolimits_{j=1}^{n} w_{jr}} \tag{12}$$

$$\tau_{ig}^{j(-)} = -1/\psi \sqrt{-L_{ig}^{j} (\sum\nolimits_{j=1}^{n} w_{jr}) / w_{jr}} \tag{13}$$

$$\tau_{ig}^{j} = \tau_{ig}^{j(-)} + \tau_{ig}^{j(+)} \tag{14}$$

9. General dominance matrix calculated summation over all criterion j for alternative A_i relative A_g by following equation:

$$\delta_{ig} = \sum\nolimits_{j v N} \tau_{ig}^{j} \tag{15}$$

10. Each of the A_i's aggregated value is calculated and ranked:

$$\xi(A_i) = \frac{\sum_{g \int M} \delta_{ig} - {}_{i \int M}{}^{min}\{\sum_{g \int M} \delta_{ig}\}}{{}_{i \int M}{}^{max}\{\sum_{g \int M} \delta_{ig}\} - {}_{i \int M}{}^{min}\{\sum_{g \int M} \delta_{ig}\}} \tag{16}$$

4 Application

This section explains criteria, alternatives and computational results, respectively.

4.1 Criteria Definition

The uncertain factors mentioned in at least three literature cited above review articles are accepted as a criterion for the Fuzzy TODIM analysis and shown in Table 1. These literature review articles cover CLSC literature between 2001–2018 years. That's why considering these articles solely to define criteria assumed to be adequate to validate CLSC interpretation.

Each criterion is explained briefly, respectively. The first criterion is cost. The cost might occur in different stages of the CLSC, such as transporting goods to the customers, CLSC planning costs, etc. The second criterion is market demand. Lack of information for the newly emerged market or incapability of collecting accurate customer demand are some reasons that make customer demand uncertain factors in CLSC. The third criterion is the type of returned goods. Companies are not sure which kind of product to collect from the customer. The fourth and fifth criteria are return quality and quantity. Vagueness about customers' behavior of using the good and the lifecycle of the good brings uncertainty in the amount and quality of the collected product. The last criterion is the time delay. Time delay uncertainties are occurred due to delivery lead time uncertainty, throughput time uncertainty, transportation time uncertainty, etc.

Table 1. Criterion regarding CLSC literature review articles

Single Uncertain Factors	Criteria	[4]	[5]	[7]	[10]
Cost	C1		X	X	X
Demand	C_2	X	X	X	X
Type of returned goods	C_3		X	X	X
Return quality	C_4		X	X	X
Return quantity	C_5	X	X	X	X
Time delay	C_6	X	X	X	X

4.2 Reproduction Processes in WEEE

A closed-loop supply chain is a system that aims to create a sustainable and circular economy by reducing waste and reusing materials. Recovering these materials from waste is known as the reproduction process(R-imperatives). There are four R-imperatives in ten different WEEE: repair, remanufacturing, reuse, and recycling [13]. This study determines three alternatives based on R-imperatives in WEEE: A_1-Recycling, A_2-Remanufacturing, A_3-Repair/Reuse shown in Fig. 1. Collected goods are transported to the supplier for the recycling process, whereas a remanufactured process is conducted in the manufacturing plant. Moreover, if repair or reuse processes are involved in reproduction, goods are directly transferred to distributors or retailers.

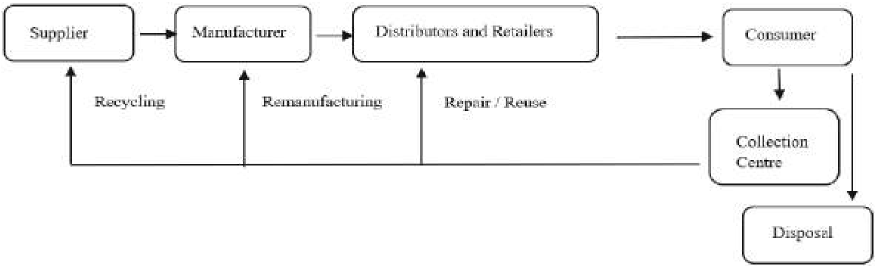


Fig. 1. R-imperatives in WEEE

4.3 Computational Results

Three decision maker who are specialized in e-waste is selected for this study. Evaluation document is sent for experts to fill criterion weight matrix and alternative matrix based on five linguistic variables shown in Table 2. Triangular fuzzy numbers ($\tilde{A}$) are generated by Eq. (1).

The integrated linguistic evaluation of alternative i as to criterion j ($\tilde{x}_{ij}$) is computed by equation by Eq. (2) for three decision makers and shown in Table 3.

Table 2. Linguistic variables with fuzzy numbers

Linguistic Variables	l	m	u
Unimportant	0	0	0,25
Weakly Important	0	0,25	0,5
Fairly Important	0,25	0,5	0,75
Important	0,5	0,75	1
Absolutely Important	0,75	1	1

Table 3. Integrated fuzzy assessment values ($\tilde{x}_{ij}$) in the fuzzy alternative matrix

Criteria	R_1			R_2			R3		
	l	m	u	l	m	u	l	m	u
C1	0,67	0,92	1,00	0,58	0,83	1,00	0,67	0,92	1,00
C2	0,25	0,50	0,75	0,25	0,50	0,75	0,58	0,83	0,92
C3	0,50	0,75	0,92	0,58	0,83	1,00	0,75	1,00	1,00
C4	0,33	0,58	0,83	0,42	0,67	0,92	0,58	0,83	1,00
C5	0,17	0,42	0,67	0,33	0,58	0,83	0,25	0,50	0,67
C6	0,08	0,33	0,58	0,25	0,50	0,75	0,25	0,50	0,67

Table 4. Average fuzzy weights of each criterion and relative weights (w_{jr})

Criteria	Fuzzy weights (cw_j)			wj	wjr
	l	m	u		
C1	0,67	0,92	1	0,76	1
C2	0,33	0,5833	0,83	0,34	0,44
C3	0,33	0,58	0,83	0,33	0,44
C4	0,67	0,92	1	0,76	1
C5	0	0,25	0,5	0,08	0,1
C6	0,17	0,42	0,67	0,18	0,24

Fuzzy weights (c_{wj}), criteria weights (w_j) and relative weights (w_{jr}) are calculated using Eqs. (3)–(6) represented in Table 4. C_1 or C_4 accepted as reference.

Gain matrix ($[G_{ig}^{j}]_{mm}$) and loss matrix ($[L_{ig}^{j}]_{mm}$) are constructed in the comparison of two alternative based on same criterion j (j ∫ 1, 2, 3, 4, 5, 6) by Eqs. (7)–(11). C_1 and C_6 are accepted as cost criteria, while the rest are accepted as benefit. Then dominance degree of gain ($\tau_{ig}^{j(+)}$) and dominance degree of loss ($\tau_{ig}^{j(-)}$) are constructed by using

Eqs. (12)–(14). After general dominance matrix (δ_{ig}) is calculated summation over all 6 criteria for each alternative pair by Eq. 15 and shown in below.

$$\delta_{ig} = \begin{bmatrix} 0 & -1,01 & -0,71 \\ -4,44 & 0 & -1,81 \\ -4,93 & -2,74 & 0 \end{bmatrix} \tag{17}$$

Each of the alternatives aggregated value is obtained by using Eq. 16, shown in the following: $\xi(A_1) = 0, \xi(A_2) = 0,82, \xi(A_3) = 1$.

5 Conclusion

A closed-loop supply chain is a circular system that maximizes resource and material use while minimizing waste production with the help of several stakeholders. This study focused on WEEE because the design of WEEE has a vital role in a sustainable future. Six criteria with three different reproduction processes are evaluated by using the Fuzzy TODIM method. This study contributes the literature by Fuzzy TODIM method implementation in a new field: evaluation of the CLSC's reproduction processes. It is shown that the minimum backward chain reuse/repair reproduction is the best CLSC considering six criteria. Cost and return quality are the most crucial parameters in the design of CLSC. Future research could investigate CLSC's reproduction processes with more than six criteria that defined in this study.

References

1. Schenkel, M., et al.: Creating integral value for stakeholders in closed loop supply chains. J. Purch. Supply Manag. **21**(3), 155–166 (2015)
2. Pokharel, S., Mutha, A.: Perspectives in reverse logistics: a review. Resour. Conserv. Recycl. **53**(4), 175–182 (2009)
3. Sasikumar, P., Kannan, G.: Issues in reverse supply chain, Part III: classification and simple analysis. Int. J. Sustain. Eng. **2**(1), 2–27 (2009)
4. Govindan, K., Soleimani, H., Kannan, D.: Reverse logistics and closed-loop supply chain: a comprehensive review to explore the future. Eur. J. Oper. Res. **240**(3), 603–626 (2015)
5. Coenen, J., van Der Heijden, R.E.C.M., van Riel, A.C.R.: Understanding approaches to complexity and uncertainty in closed-loop supply chain management: past findings and future directions. J. Clean. Prod. **201**, 1–13 (2018)
6. Walker, W.E., et al.: Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. Integr. Assess. **4**(1), 5–17 (2003)
7. Peng, H., et al.: Uncertainty factors, methods, and solutions of closed-loop supply chain—a review for current situation and future prospects. J. Clean. Prod. **254**, 120032 (2020)
8. Liao, H., et al.: A non-probabilistic model of carbon footprints in remanufacture under multiple uncertainties. J. Clean. Prod. **211**, 1127–1140 (2019)
9. Simangunsong, E., Hendry, L.C., Stevenson, M.: Supply-chain uncertainty: a review and theoretical foundation for future research. Int. J. Prod. Res. **50**(16), 4493–4523 (2012)
10. MahmoumGonbadi, A., Genovese, A., Sgalambro, A.: Closed-loop supply chain design for the transition towards a circular economy: a systematic literature review of methods, applications and current gaps. J. Clean. Prod. **323**, 129101 (2021)

11. Autran Monteiro Gomes, L.F., Duncan Rangel, L.A.: An application of the TODIM method to the multicriteria rental evaluation of residential properties. Eur. J. Oper. Res. **193**(1), 204–211 (2009)
12. Tosun, Ö., Akyüz, G.: A fuzzy TODIM approach for the supplier selection problem. Int. J. Comput. Intell. Syst. **8**(2), 317–329 (2015)
13. Islam, Md.T., Huda, N.: Reverse logistics and closed-loop supply chain of waste electrical and electronic equipment (WEEE)/E-waste: a comprehensive literature review. Resour. Conserv. Recycl. **137**, 48–75 (2018)

A Proposed Method for Aircraft Selection Using Inter-valued Spherical Fuzzy AHP

Murat Levent Demircan[1](✉) and Tuba Kaplan[2]

[1] Faculty of Engineering and Technology, Galatasaray University, Ortaköy, İstanbul, Turkey
ldemircan@gsu.edu.tr
[2] Turkish Airlines-Cargo Strategic Planning and Data Analysis Department, İstanbul, Turkey
t.kaplan@thy.com

Abstract. Air transport is a growing industry today, and passengers frequently prefer it because it is faster than other modes. As a result, there is an increasing demand for air transportation, not only in terms of passengers but also in terms of logistics, creating a competitive environment in the industry. Airline companies are willing to enhance their fleet to meet this thriving demand. Therefore, fleet planning and aircraft selection are crucial issues to be managed by airline firms. Fleet planning experts should analyze aircraft types under vagueness based on criteria to select the most appropriate one according to the company's needs. This paper evaluates aircraft alternatives considering three main financial, technical, and environmental categories under fuzzy and vague conditions. This study helps high-level executives make appropriate decisions among alternatives by considering necessary conditions. The multi-criteria decision-making method, Analytic Hierarchical Process (AHP), uses inter-valued spherical fuzzy sets. The method is explained in detail, and the results are analyzed.

Keywords: Aircraft Selection · Multi-Criteria Decision Making · Spherical Fuzzy AHP · Fleet Panning

1 Introduction

In recent years, the airline industry has had a growing trend both in passenger and cargo transportation. Even though Covid-19 has damaged this industry profoundly, there is a recovery in passenger and an increase in air cargo transportation. In 2021, there has been a 49% decline in the total number of passengers [1], while air cargo has increased by 6.9% [2]. According to Airbus' Global Market Forecast [3], passenger and freight traffic will rise 3.6% and 3.2%, respectively. This increasing demand has created a competitive market in the industry. As a result, airline firms should make their decisions professionally to succeed in this competitive environment.

For airlines to meet demand and acquire a competitive edge, fleet planning is essential. An airline firm must figure out how to choose the greatest aircraft because there are numerous factors to consider when selecting an aircraft from among the available aircraft alternatives. Multi-Criteria Decision Making (MCDM) assists decision-makers in

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 426–437, 2023.
https://doi.org/10.1007/978-3-031-39777-6_51

selecting the best course of action from various potential solutions to a problem. Fuzzy set theory is useful in evaluating the ambiguity and unpredictability of the decision maker's nominative judgment [4]. By addressing this uncertainty, the suggested MCDM methodologies will assist decision-makers in evaluating alternatives. MCDM techniques were created in the 1950s to help airline planners choose the right aircraft. The goal of planners is to choose the best option from a variety of planes while taking numerous factors into account [5]. Numerous techniques are utilized in the evaluation process, and MCDM methods have ensured convincing solutions in choosing aircraft [6]. There are several MCDM techniques. The optimal aircraft alternative will be chosen in this study using the AHP approach and spherical fuzzy sets. In the Saaty-developed AHP [7], an issue is organized hierarchically, starting with a purpose and progressing through criteria, sub-criteria, and successive levels. The hierarchy guarantees that the decision-makers have a comprehensive understanding of the intricate relationships that are part of the issue. They can use it to help them decide whether components of the same grade are equivalent. Elements are presented with nine-level scales on which to place their weights. Pairwise comparison, however, introduces uncertainty because expert opinions are required [8]. In many real-world situations, the decision-makers might be unable to give their choices accurate numerical numbers [9]. As a result, each pairwise comparison decision in AHP is represented by fuzzy numbers [10].

This study aims to find the appropriate aircraft for an airline company in Turkey. The motivations behind this study are as follows. First, increasing passenger and air cargo demand creates the necessity to expand the fleet. Since there will be much vagueness during the decision process, this study will help firms make their decisions professionally and scientifically. Secondly, this paper will consider the criteria of belly capacity, which has not been mentioned in other studies. Belly capacity has been crucial recently because there has been an enormous increase in air cargo during the pandemic. The majority of cargo is transported in the aircraft's belly. Therefore, when purchasing an airplane, belly capacity should be considered. The rest of this essay will proceed as follows; A overview of the literature on choosing aircraft using MCDM methods is included in Sect. 2. The specifics of the suggested methodology are explained in Sect. 3. Applying the MCDM technique to a case study is explained in Sect. 4 with numerical examples. A summary of the outcomes is provided in Sect. 5. The conclusion is provided in Sect. 6.

2 Literature Review

Numerous studies in the literature use various multi-criteria decision-making techniques to choose the best airplane from the choices while considering certain factors.

Yeh and Chang put up a new fuzzy MCDM algorithm [11] to address the aircraft selection issue. They established eleven sub-criteria in addition to three core criteria: economic efficiency, social responsibility, and technological advancement. Using the suggested algorithm, B757-200, A321, B767-200, MD-82, and A310-300 are chosen as alternatives and assessed. According to the system, B757-200 was ranked #1. Analytic Network Process (ANP) was utilized by Ozdemir et al. [12] to assess the Turkish Airlines aircraft options. After the interactions between the criteria, sub-criteria, and alternatives are described in ANP, the weighted supermatrix and limit super matrix are built.

Three alternatives—the A319, A320, and B737—are examined in this study based on three primary criteria: cost, time, and physical characteristics. Applying the strategy made it seem more attainable to buy the B737. Three MCDM techniques—Simple Additive Weighting (SAW), Elimination and Choice Expressing Reality (ELECTRE), and TOPSIS—were used by Sun et al. [13] to tackle an airline's hypothetical aircraft selection problem. As a further criterion for judgment, the author added robustness. To deal with economic ambiguity, the study employs the Taguchi loss function. Six criteria are used to evaluate three different types of aircraft. The most preferred aircraft is the B747-400. Gomes et al.'s [14] solution to the Brazilian investment in regional charter flights utilized NAIDE (Novel Approach to Imprecise Assessment and Decision Environments) techniques. The authors classified the critical criteria into financial, logistical, and quality. Each category has multiple sub-criteria, including operational cost, range, and safety. It is decided that LET-410 is the best substitute. Dozic and Kalic [15] use the AHP approach to choose passenger aircraft. From the fuzzy pairwise comparison, crisp priorities are derived using logarithmic fuzzy preference programming (LFPP). Three criteria and ten sub-criteria are used to evaluate seven possibilities. Among the alternatives, the CRJ 700 is chosen as the top choice.

AHP and fuzzy set theory were combined by Bruno et al. [16] to choose aircraft to support airline strategic decisions. The primary factors include the impact on the environment, technical performance, interior quality of the airplane, and economic performance. For a case study of Air Italy, the suggested methodologies are used to analyze three aircraft. Each criterion is weighted according to importance by four Air Italy experts. According to the survey, the Sukhoi SSJ 1000 had the highest aircraft evaluation score. Dožić and Kalić [17] comparison of MCDM and AHP methodologies. Fuzzy ANP and the generalized Choquet integral approach were used by Ozdemir and Başlgil [18] to select the best airplane. Fuzzy AHP is used to compare the outcomes. A broad averaging operator that shows how different criteria interact is the chocquet integral. The trapezoidal fuzzy numbers are employed in the Choquet integral approach to quantify language words and contrast with experts. Three options are compared based on three primary criteria: cost, physical characteristics, and time. The outcome was found using both the fuzzy ANP and Choquet integral approach. Bakır and Kırac [5] employed the TOPSIS approach to identify the appropriate aircraft. The TOPSIS method places the chosen alternative closest to the positive ideal solution and farthest away from the negative ideal solution, where the positive ideal solution is made up of all the best values of the criteria, and the negative ideal solution is made up of the worst values of the criteria. The criteria listed were range, cost, speed, passenger capacity, and fuel consumption. The evaluation of four options revealed that the B737-800 is the most experimental aircraft. Sener and Yilanli [19] employed the Decision Making Trial and Evaluation Laboratory (DEMATEL) technique to determine the efficacy of the criteria. In pairs, the fourteen criteria are evaluated and contrasted. They then determined their effects, said the decision-makers. Four options were assessed. Ilgin [20] suggested TOPSIS and Linear Physical Programming (LPP) approaches for choosing aircraft. In TOPSIS, decision-makers subjective weightings of the criteria provide the basis for ranking. LPP does not account for the ambiguity of DMs' preferences and does away with subjective weighting. The author compared six different aircraft options based on five factors: cost, fuel consumption, range, seating capacity, and luggage volume. The A321 (neo) ranked #1 in both the TOPSIS and LPP rankings, which were identical. Kartal [21] proposed the EDAS and COPRAS-based Stepwise Weight Assessment Ratio Analysis (SWARA) approach to assess the preferred narrow-body and wide-body passenger aircraft. The purchase price, fuel capacity, maximum seat capacity, range, maximum take-off weight, and cargo capacity are the six criteria that have been established. Alternatives with seven narrow-body seats and nine wide-body seats were taken into consideration. The study administered a SWARA questionnaire to 10 specialists, all academics with prior experience in the aviation sector. The experts were asked to rank the criteria in order of relevance before calculating the relative importance values of the criteria based on the results. The criteria weights were calculated using the SWARA approach to analyze the data gathered for each expert. The final weights for each criterion were then calculated by averaging the criteria weights estimated for ten experts. Rankings were obtained using the EDAS and COPRAS procedures for narrow-body and wide-body passenger aircraft after establishing the criterion weights. According to the study, the A321neo and the A350-1000 airplane were the best alternatives in EDAS and COPRAS techniques for passenger aircraft with narrow and wide bodies, respectively. To analyze the various types of short-to-medium range aircraft, Ardil [22] employed the Process Using Preference Analysis for Reference Ideal Solution (PARIS) approach and the TOPSIS method before settling on the A312 (neo). In this study, six alternatives are assessed using the decision-making criteria of the airplane's cost, fuel efficiency, seat count, and luggage volume. Type-2 fuzzy sets (IT2F) were employed by AHP and TOPSIS in interval Kiraci and Akan [23] to choose commercial aircraft. Four aircraft with short and medium ranges were assessed using eight sub-criteria and three core criteria: economic performance, technical performance, and environmental effect. The weight of the criterion for the IT2FTOPSIS approach, which ranks the aircraft alternatives, was determined using IT2FAHP. According to the results, the A321neo is the best alternative aircraft. For Canada's regional operations, Ahmed et al.'s integrated Fuzzy AHP and Efficacy Method framework was put forth [24]. FAHP eliminates all ambiguity in defining the weight of the chosen criteria. The total efficacy coefficient is then calculated using the weight of the criteria. The criteria include purchase price, seating capacity, aircraft range, payload, and carbon monoxide emissions. The ARJ21-700 is chosen as a suitable aircraft. Akyurt and Kabadayi [25] employed fuzzy AHP and fuzzy grey relational techniques to choose the best cargo airplane. The advantage of the fuzzy GIA method is that it can produce successful results for solving problems even when relatively little data is available or there is a large amount of variability between factors. The fuzzy AHP method was used to calculate the weight values that determine the importance of the criteria. According to the criteria determined by the fuzzy AHP method, the most suitable aircraft was selected by the fuzzy GIA method. The three basic criteria—cost, operational compatibility, and time—are further broken down into 15 sub-criteria. As alternatives, B777F, A330-200F, B747-400F, and A310-300F have been chosen. Following the application of the approach, the B777F was determined to be the aircraft that was most appropriate based on its highest grey relational degree. Deveci et al. [26] employed the Entropy-based Weighted Aggregated Sum Product Assessment (WASPAS) technique to identify appropriate aircraft for a certain route between Kuwait and

Istanbul. The approach is put out to determine the best strategy to generate the weight vector and combine the preferences of the DM. Ten sub-criteria are used to choose between the five essential criteria: income, capacity, customer expectation, cost, and competitiveness. The analysis revealed that Airbus32C ought to be chosen for this route. To choose the best regional airplane, Bakir et al. [27] suggested integrating fuzzy Pivot Pairwise Relative Criteria Importance Assessment (F-PIPRECIA) and fuzzy Measurement Alternatives and Ranking according to the Compromise Solution (F-MARCOS). Six regional aircraft possibilities were compared using fourteen criteria in the study. For group decisions with numerous decision-makers and criteria, the PIPRECIA is a useful strategy. The criteria weights are determined for F-PIPRECIA, and the alternatives were evaluated based on the criteria weights for F-MARCOS. The best suitable aircraft was determined to be the CRJ1000.

3 Methodology

In this part, the mathematical operations of spherical fuzzy sets (SFSs) and the proposed spherical fuzzy AHP method are presented.

3.1 Inter-valued Spherical Fuzzy Sets (IVSFSs)

Spherical Fuzzy Sets (SFSs) have been developed by Kutlu Gundogdu and Kahraman [28, 29], allowing decision-makers to justify their hesitancy in their linguistic selections of an option. To determine membership degrees for criteria and sub-criteria, SFSs can be employed to obtain a broader range. Decision-makers in SFS assign degrees of membership, non-membership, and reluctance. SFSs allow squared sums to be 1.0, so DM may provide clarification in an SFSs field.

SFSs has been developed as an inference of Pythagorean Fuzzy Sets Gündoğdu and Kahraman [28, 29]. The definition of SFSs is described as follows:

Definition 1. Let X be a universal set, let $\tilde{A}s$ be the single-valued SFS which is described as,

$$\tilde{A}s = \left\{ ʃx, \mu_{\tilde{A}s}(x), v_{\tilde{A}s}(x), \lambda_{\tilde{A}s}(x) \forall | x\sigma X \right\} \tag{1}$$

Where $\mu_{\tilde{A}s}: X \geq\ [0,1\ ]$, $v_{\tilde{A}s}: X \geq\ [0,1\ ]$, $\lambda_{\tilde{A}s}: X \geq\ [0, 1]$ and

$$0 \leq \mu^2{}_{\tilde{A}s}(x) + v^2{}_{\tilde{A}s}(x) + \lambda^2{}_{\tilde{A}s}(x)\ \leq 1, \forall \mathrm{x} \in X.$$

Here, $\mu_{\tilde{A}s}(x)$ is the membership degree, $v_{\tilde{A}s}(x)$ is non-membership degree and $\lambda_{\tilde{A}s}(x)$ is the degree of hesitancy of x to $\tilde{A}s$. Gündoğdu and Kahraman [28] described the arithmetic calculation of IVSFSs. The formulas to defuzzify and aggregate IVSFSs are also presented.

Definition 2. Equation (2) defines and IVSFS $\tilde{A}s$ of the universal set X.

$$\tilde{A}s = \left\{ \left(x, \begin{matrix} \left[\mu^L_{\tilde{A}s}(x), \mu^U_{\tilde{A}s}(x)\right], \\ \left[v^L_{\tilde{A}s}(x), v^U_{\tilde{A}s}(x)\right], \\ [\lambda^L_{\tilde{A}s}(x), \lambda^U_{\tilde{A}s}(x)] \end{matrix} \right) | x \in X \right\} \tag{2}$$

where; $0 \leq \mu^L{}_{\tilde{A}s}(x) \leq \mu^U{}_{\tilde{A}s}(x) \leq 1$, $0 \leq v^L{}_{\tilde{A}s}(x) \leq v^U{}_{\tilde{A}s}(x) \leq 1$ and $0 \leq \left(\mu^U{}_{\tilde{A}s}(x)\right)^2 + \left(v^U{}_{\tilde{A}s}(x)\right)^2 + \left(\lambda^U{}_{\tilde{A}s}(x)\right)^2 \leq 1$. For each $x\sigma X$, $\mu^U{}_{\tilde{A}s}(x)$ is the upper degrees of membership, $v^U{}_{\tilde{A}s}(x)$ is the upper degrees of non-membership and $\lambda^U{}_{\tilde{A}s}(x)$ is the hesitancy of x to $\tilde{A}s$. An IVSFS is defined as;

$$\Big\langle [\mu^L_{\tilde{A}s}(x), \mu^U_{\tilde{A}s}(x)], [v^L_{\tilde{A}s}(x), v^U_{\tilde{A}s}(x)], [\lambda^L_{\tilde{A}s}(x), \lambda^U_{\tilde{A}s}(x)]$$

is indicated by $\tilde{\alpha} = \int [a, b], [c, d], [e.f] \forall$where [a, b] ⊂ [0, 1], [c, d] ⊂ [0, 1] ⊂ [e, f] ⊂ [0, 1], $b^2 + d^2 + f^2 \leq 1$.

3.2 Spherical Fuzzy AHP (SFAHP)

The steps of SFAHP are as follows;

Step 1. The first step starts with the conventional AHP which is forming a 4-level structure. The first level is the definition of the problem. The second and third levels are the division of main and sub-criteria, respectively. The last level represents selection of alternatives (see Fig. 1).

Step 2. The IVSFSs based measurement reference based on linguistic preferences of DMs is developed. This measurement reference would be used to set pairwise comparisons.

Step 3. The consistency ratio of the pairwise comparison matrix should be controlled by following the steps of conventional AHP and it must be ensured that CRs are below 10%. If any CR is above 10%, then the linguistic preferences should be revised until reaching the expected CR.

Step 4. The DMs' preferences are considered to grip the numeric importance of IVSFSs. IVSWAM formula (Eq. 6) is used to calculate the criteria and sub-criteria weights.

$$\text{IVSWAM}_{\omega}(\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_k) = \omega_1\tilde{a}_1 \ \omega_2\tilde{a}_2 \ldots \omega_k\tilde{a}_k \text{ where } \omega = 1/n$$

Step 5. The global weights are calculated. The preference importance of IVSFSs is aggregated for each level of the hierarchy to acquire the score ranks. At that point, there are two approaches which are using partial IVSFAHP and using complete IVSFAHP. In this study the first approach is used. The Eq. (3) is used to defuzzify the criteria weightings: Score $\omega_j^S = \text{S}(\tilde{\omega}_j^S)$.

$$S(\tilde{\omega}_j^S) = \frac{a^2 + b^2 - c^2 - d^2 - (e/2)^2 - (f/2)^2}{2} + 1 \tag{3}$$

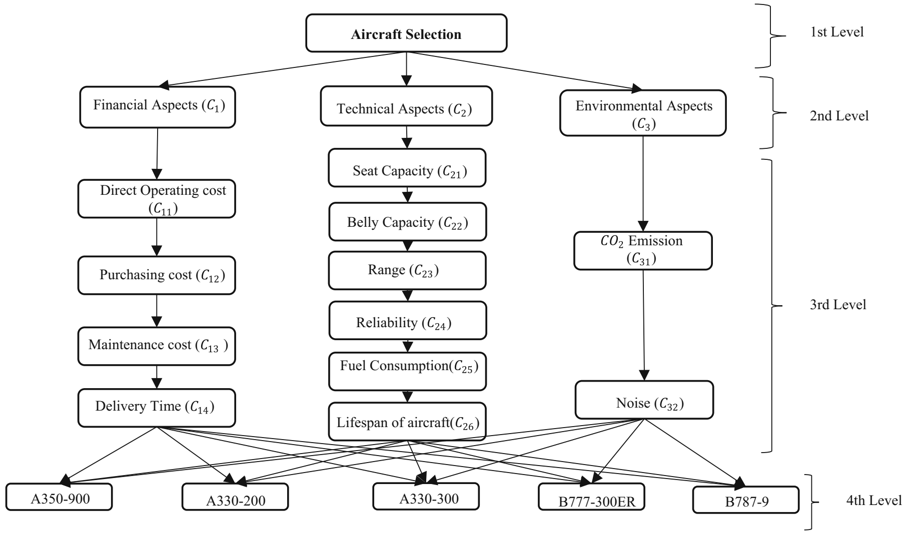


Fig. 1. Criteria and Sub-Criteria Framework.

For normalizing the criteria weights:

$$\overline{\omega}_j^S = \frac{\mathrm{S}(\tilde{\omega}_j^S)}{\sum_{j=1}^{n} \mathrm{S}(\tilde{\omega}_j^S)} \tag{4}$$

For weighting the decision matrix were $\tilde{a}_{S_{ij}} = \overline{\omega}_j^S . \tilde{a}_{S_i}$,

$$\tilde{a}_{S_{ij}} = \left\{ \begin{array}{c} \left[\left(1 - \left(1 - \alpha_{S_i}^2\right)^{\overline{\omega}}\right)^{\frac{1}{2}}, \left(1 - \left(1 - b_{S_i}^2\right)^{\overline{\omega}_j^S}\right)^{\frac{1}{2}} \right], \\ \left[c_{S_i}^{\overline{\omega}_j^S}, d_{S_i}^{\overline{\omega}_j^S} \right], \\ \left[\left(\left(1 - \alpha_{S_i}^2\right)^{\overline{\omega}_j^S} - \left(1 - \alpha_{S_i}^2 - e_{S_i}^2\right)^{\overline{\omega}_j^S} \right)^{\frac{1}{2}}, \right. \\ \left. \left(\left(1 - b_{S_i}^2\right)^{\overline{\omega}_j^S} - \left(1 - b_{S_i}^2 - f_{S_i}^2\right)^{\overline{\omega}_j^S} \right)^{\frac{1}{2}} \right] \end{array} \right\} \tag{5}$$

$$IVSWAM_{\omega} = \left\{ \begin{array}{c} \left[(1 - \prod_{j=1}^{k} (1 - \alpha_j^2)^{\omega_j})^{1/2}, (1 - \prod_{j=1}^{k} (1 - b_j^2)^{\omega_j})^{1/2} \right], \left[\prod_{j=1}^{k} c_j^{\omega_j}, \ \prod_{j=1}^{k} d_j^{\omega_j}, \right] \\ \left[\left(\prod_{j=1}^{k} (1 - \alpha_j^2)^{\omega_j} - \prod_{j=1}^{k} (1 - \alpha_j^2 - e_j^2)^{\omega_j} \right)^{1/2}, \left(\prod_{j=1}^{k} (1 - b_j^2)^{\omega_j} - \prod_{j=1}^{k} (1 - b_j^2 - f_j^2)^{\omega_j} \right)^{1/2} \right] \end{array} \right\} \tag{6}$$

Step 6. The calculation and defuzzification of each alternative are done.
Step 7. The best decision is found.

4 Numerical Analysis

After describing the theoretical foundation of IVSF-AHP, we will use the numerical application of choosing an aircraft from a private firm to illustrate the process. Like the first step of the technique, three decision-makers are used to define the significant criterion and sub-criteria. There are three key factors to consider: economic factors, technological characteristics, and environmental factors. The following are the criteria summaries:

Financial Aspects (C_1): Financial criteria is a critical factor that should be considered before purchasing an aircraft since it is an essential investment for airline companies. An aircraft has several financial criteria such as operating cost, purchasing expense, maintenance cost and delivery time. There is also indirect operating cost, which is ignored in this study because it depends on airline operations and every airline has different operations. The costs which are considered in this study are related to aircraft type. *Direct Operating Cost* (C_{11}): This criterion is the cost of running the aircraft. Direct operating cost includes fuel expense, ownership cost, handling expenses, and crew fees. Fuel cost is based on the aircraft's fuel price and fuel consumption [30]. Ownership cost includes depreciation and insurance cost of the aircraft. Handling cost is applied by ground handling agencies and airport authorities in order to provide services like check-in, boarding, baggage/cargo loading, cleaning of aircraft and so on [30]. Handling cost varies on aircraft type. It increases when the unit volume of aircraft is high. Crew fee is the total cost of cockpit and cabin crew which differs from one airline company to another. *Purchasing Cost* (C_{12})*:* The aircraft's market price assigned by the manufacturer. In general, manufacturers apply discounts as airlines request. However, in this study, the price is the list price of aircrafts, and the discount is ignored. *Maintenance Cost* (C_{13})*:* Aircraft maintenance includes repair, overhaul on engines and components, modification, inspection, and condition determination. Aircraft maintenance ensures a fully serviceable aircraft at minimum cost when required by an airline. Aircraft maintenance is inevitable in the airline industry to meet safety necessities. Also, punctuality performance is very crucial for airlines. It is required to execute the maintenance to avoid delay in a flight schedule. Therefore, investing heavily in spare parts is necessary for efficient punctuality performance [30]. *Delivery Time* (C_{14})*:* The time that the aerospace companies can deliver the aircraft. Each aircraft type has different delivery time. Airline firms want to receive the aircraft as soon as possible to operate the flight to get more profit.

Technical Aspects (C_2): Technical aspects are the operational capability of an aircraft including seat capacity, belly capacity, range, reliability, fuel consumption and the life of aircraft. This criterion is essential to make long-term plans and strategic decisions. Airlines can decide the frequency and the route (direct or with stops) for the destination points based on technical aspects of aircrafts. *Seat Capacity* (C_{21})*:* Aircraft cabin is classified as narrow-body if only one aisle provides the total passenger capacity up to 250 and wide-body if two aisles provide the total passenger capacity of 200 to 850 in fuselage. This criterion is one of the crucial factors that affects airlines' profitability. DMs must select the aircraft with the right seats to match demand with supply. *Belly Capacity* (C_{22})*:* The capacity reserved for luggage and freight in a passenger aircraft. Air cargo transportation is a growing trend and a critical revenue parameter. Although airlines have cargo aircrafts, belly of aircrafts are also used to carry cargoes. Even though Covid-19 seriously damaged the airline industry, air cargo was important to transport goods, especially pharmaceutical products, during the pandemic. *Range* (C_{23})*:* The maximum distance an aircraft can fly from departure to arrival point. The larger the aircraft gets, the higher the range becomes. The range of an aircraft may vary depending on the number of passengers, cargo load and fuel amount carried by the aircraft. Range is also critical factor for airlines if they want to fly to new destinations. *Reliability* (C_{24})*:* Reliability is the property of a component, item, system, or aircraft stated by the probability that it will perform a necessary function in a given environment for a certain period. The reliability of engines and systems of an aircraft is important to actualize operations safely during the flight. New aircrafts have priority of being chosen since they have higher reliability as the technological developments increase the performance of aircraft equipment. *Fuel Consumption* (C_{25})*:* It measures the aircraft's burned fuel (kg/hr) during the flight. Fuel consumption is crucial for selecting an aircraft to decrease environmental impact and cost. Therefore, aircrafts with low fuel consumption are preferable for airlines. *The Life of Aircraft* (C_{26})*:* The number of years that an aircraft can be operated before its retirement.

Environmental Aspects (C_3): CO_2 emission and noise level are the two fundamental environmental factors that are considered in aviation industry. With serious impacts, airlines and aircraft manufacturers try to diminish each factor's level to protect the environment. *CO_2 Emission* ($\boldsymbol{C_{31}}$)*:* The carbon emission is a serious issue considered by airlines, and it is tried to be reduced to decrease the level of damage caused by fuel of the aircrafts. CO_2 is the most common gas released into the atmosphere. Noise ($\boldsymbol{C_{32}}$): Since the population has grown in the cities and the residential areas have spread across airports, the number of people affecting by aircraft noise has increased in recent years. The noise obtained by aircrafts have damage on both human and animals. There are two main reasons why aircrafts produce noise: Aeronautical noise which results from the air friction throughout the aircraft and the noise of engine. Some airport authorities in some countries have sound limit for aircraft's landing and taking off operations [11]. Therefore, the noise level of aircraft is important to use it in certain countries.

We selected four different alternatives which are A350-900, A330-200, A330-300, B777-300ER and B787-9. These alternatives are all wide body. The features of the aircrafts are given in Table 1.

Table 1. Aircrafts' Features

	A350-900	A330-200	A330-300	B777-300ER	B787-9
Maximum Range	15,372 km	15,094 km	11,750 km	13,649 km	14,010 km
Seat Capacity	348	293	335	396	330
Belly Capacity	17 tons	14 tons	16 tons	25.2 tons	11 tons
Fuel Consumption	5800 kg/hr	5590 kg/hr	5700 kg/hr	6630 kg/hr	5600 kg/hr
Price	$308.1M	$231.5M	$256.4M	$375.5M	$292.5M

5 Result

After establishing the criteria and sub-criteria, three decision-makers with experience in aircraft procurement process management in the private airline industry carefully examined each. Using relevant linguistic indices, we constructed the consistency ratio (CR) values for the pairwise comparison matrix of the IVSFS. The pairwise comparison and IVSFSs weights and their corresponding CR values are calculated. The final rating results for comparable competing alternative aircrafts are shown in Table 2 and 3.

Table 2. Final spherical fuzzy global priority weights

	C_{11}	C_{12}	C_{13}	C_{14}	C_{21}	C_{22}	C_{23}	C_{24}	C_{25}	C_{26}	C_{31}	C_{32}
A	([0.40, 0.47], [0.60, 0.66], [0.20, 0.27])	([0.15, 0.18], [0.90, 0.92], [0.10, 0.14])	([0.33, 0.40], [0.71, 0.75], [0.14, 0.19])	([0.27, 0.32], [0.81, 0.83], [0.11, 0.16])	([0.31, 0.38], [0.73, 0.77], [0.14, 0.19])	([0.23, 0.27], [0.84, 0.87], [0.12, 0.17])	([0.23, 0.29], [0.88, 0.90], [0.07, 0.12])	([0.40, 0.49], [0.66, 0.71], [0.12, 0.20])	([0.20, 0.24], [0.83, 0.87], [0.13, 0.18])	([0.40, 0.49], [0.66, 0.71], [0.12, 0.20])	([0.56, 0.67], [0.43, 0.49], [0.15, 0.23])	([0.53, 0.63], [0.48, 0.54], [0.17, 0.25])
B	([0.25, 0.28], [0.77, 0.83], [0.18, 0.24])	([0.23, 0.27], [0.85, 0.87], [0.12, 0.16])	([0.27, 0.32], [0.78, 0.81], [0.15, 0.20])	([0.21, 0.25], [0.84, 0.87], [0.12, 0.16])	([0.19, 0.23], [0.84, 0.88], [0.13, 0.17])	([0.18, 0.22], [0.88, 0.90], [0.11, 0.14])	([0.16, 0.19], [0.92, 0.94], [0.08, 0.11])	([0.14, 0.17], [0.90, 0.94], [0.11, 0.15])	([0.40, 0.49], [0.66, 0.71], [0.12, 0.20])	([0.20, 0.24], [0.83, 0.86], [0.14, 0.18])	([0.31, 0.36], [0.68, 0.74], [0.20, 0.26])	([0.20, 0.24], [0.80, 0.86], [0.16, 0.21])
C	([0.25, 0.28], [0.77, 0.83], [0.18, 0.24])	([0.22, 0.25], [0.85, 0.88], [0.12, 0.16])	([0.20, 0.23], [0.83, 0.87], [0.14, 0.18])	([0.18, 0.22], [0.86, 0.89], [0.12, 0.15])	([0.14, 0.17], [0.89, 0.93], [0.11, 0.15])	([0.24, 0.28], [0.83, 0.86], [0.12, 0.17])	([0.08, 0.09], [0.97, 0.98], [0.06, 0.09])	([0.20, 0.24], [0.83, 0.87], [0.13, 0.18])	([0.30, 0.36], [0.73, 0.77], [0.14, 0.19])	([0.14, 0.17], [0.90, 0.93], [0.11, 0.15])	([0.21, 0.25], [0.77, 0.84], [0.17, 0.23])	([0.43, 0.50], [0.57, 0.63], [0.19, 0.26])
D	([0.40, 0.47], [0.60, 0.66], [0.20, 0.27])	([0.12, 0.14], [0.93, 0.95], [0.09, 0.13])	([0.27, 0.32], [0.78, 0.81], [0.15, 0.20])	([0.12, 0.14], [0.92, 0.94], [0.10, 0.14])	([0.39, 0.48], [0.67, 0.72], [0.12, 0.19])	([0.30, 0.38], [0.78, 0.81], [0.11, 0.17])	([0.11, 0.13], [0.95, 0.96], [0.07, 0.10])	([0.28, 0.34], [0.77, 0.81], [0.14, 0.18])	([0.15, 0.18], [0.88, 0.92], [0.12, 0.16])	([0.30, 0.36], [0.76, 0.80], [0.14, 0.20])	([0.31, 0.36], [0.68, 0.74], [0.20, 0.26])	([0.34, 0.39], [0.67, 0.73], [0.19, 0.26])
E	([0.40, 0.47], [0.60, 0.66], [0.20, 0.27])	([0.30, 0.38], [0.79, 0.82], [0.10, 0.16])	([0.15, 0.17], [0.88, 0.92], [0.12, 0.16])	([0.16, 0.19], [0.89, 0.91], [0.11, 0.14])	([0.25, 0.30], [0.79, 0.82], [0.14, 0.18])	([0.12, 0.14], [0.93, 0.95], [0.09, 0.13])	([0.18, 0.22], [0.90, 0.92], [0.08, 0.11])	([0.35, 0.44], [0.71, 0.75], [0.13, 0.20])	([0.26, 0.31], [0.78, 0.82], [0.14, 0.19])	([0.30, 0.36], [0.75, 0.79], [0.15, 0.20])	([0.38, 0.46], [0.57, 0.63], [0.19, 0.25])	([0.31, 0.36], [0.70, 0.76], [0.19, 0.26])

Table 3. Score values and rankings of alternatives

Aircraft Options	Scores	Ranks
A. A350-900	0.2834	1
B. A330-200	0.2336	4
C. A330-300	0.2297	5
D. B777-300ER	0.2533	3
E. B787-9	0.2542	2

6 Conclusion

As the air traffic grows, airline companies need to widen their fleet. Therefore, fleet planning and aircraft selection become a crucial issue for airlines. Selecting an appropriate aircraft helps airlines enhance their profit and reduce costs. Since aircraft selection is a complex decision-making process with multiple criteria, MCDM methods is one way to handle this complexity. In this study, we have proposed the inter- valued spherical fuzzy analytic hierarchy process that includes uncertainty and ambiguity into the decision-making process to evaluate aircrafts. This method allows decision makers to state their opinions and choices flexibly. The model consists of three main criteria which are financial aspects, technical aspects, and environmental aspects and twelve sub-criteria (direct operating cost, purchasing cost, maintenance cost, delivery time, seat capacity, belly capacity, range, reliability, fuel consumption, lifespan of aircraft, CO2 emission and noise). We focused on wide-body aircraft alternatives and evaluated them under the criteria that we selected. A350-900, A330-200, A330-300, B777-300ER and B787-9 are selected as alternatives. Three decision makers evaluated alternatives and A350-900 is selected the most appropriate one. For future studies, other MCDM methods under spherical fuzzy sets can be used to solve aircraft selection problems. Also, the effectiveness of spherical fuzzy AHP can be compared with other decision-making methods such as traditional AHP, TOPSIS, and ELECTRE in future studies. The comparison can be made regarding accuracy, robustness, and reliability.

References

1. ICAO: Effects of Novel Coronavirus (COVID-19) on Civil Aviation: Economic Impact Analysis (2022)
2. Air IATA: Air Cargo Market Analysis (2021)
3. Airbus: Global Market Forecast 2022 (2022)
4. Zadeh, L.: Fuzzy sets. Inf. Control, 338–353 (1965)
5. Bakır, M. Kiracı, K.: Using the multi criteria decision making methods in aircraft selection problems and an application. J. Transp. Logist., 14–24 (2018)
6. Dožić, S., Lutovac, T., Kalić, M.: Fuzzy AHP approach to passenger aircraft. J. Air Transp. Manag. **68**, 165–175 (2018)
7. Saaty, T.L.: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. McGraw-Hill International Book Company (1980)

8. Liu, Y., Eckert, C., Earl, C.: A review of fuzzy AHP methods for decision-making. Expert Syst. Appl. **161** (2020)
9. Chan, F.T.S., Kumar, N.: Global supplier development considering risk factors using fuzzy extended AHP-based approach. Omega, 417–431 (2007)
10. Xu, Z., Liao, H.: Intuitionistic fuzzy analytic hierarchy process. IEEE Trans. Fuzzy, 749–761 (2014)
11. Yeh, C.H., Chang, Y.H.: Modeling subjective evaluation for fuzzy group multicriteria decision making. Eur. J. Oper. Res. **194**(2), 464–473 (2009)
12. Özdemir, Y., Basligil, H., Karaca, M.: Aircraft selection using analytic network process: a case for Turkish airlines. In: World Congress on Engineering (2011)
13. Sun, X., Gollnick, V., Stumpf, E.: Robustness consideration in multi-criteria decision making to an aircraft selection problem. J. Multi-Criteria Decis. Anal. **18**, 55–64 (2011)
14. Gomes, L.F.A.M., Mattos Fernandes, J.E., Mello, J.C.C.: A fuzzy stochastic approach to the multicriteria selection of an aircraft for regional chartering. J. Adv. Transp, 223–237 (2014)
15. Dožić, S., Kalić, M.: An AHP approach to aircraft selection process. Transp. Res. Procedia **3**, 165–174 (2014)
16. Bruno, G., Esposito, E., Genovese, A.: A model for aircraft evaluation to support strategic decisions. Expert Syst. Appl., 5580–5590 (2015)
17. Dožić, S., Kalić, M.: Comparison of two MCDM methodologies in aircraft type selection problem. Transp. Res. Procedia 910–919 (2015)
18. Özdemir, Y.: Aircraft selection using fuzzy ANP and the generalized Choquet integral. J. Intell. Fuzzy Syst., 589–600 (2016)
19. Şener, Z., Yılanlı, M.: Aircraft selection in airline fleet management (2020)
20. İlgin, M.A.: Aircraft selection using linear physical programming. J. Aeronaut. Space Technol. **12**, 121–128 (2019)
21. Kartal, M.: Havayolu İşletmeleri İçin Uçak Seçim Stratejisinde Swara Tabanli Edas Ve Copras Yöntemlerinin Kullanimi, Yöktez (2022)
22. Ardil, C.: Aircraft selection process using preference analysis for reference ideal solution (PARIS). Int. J. Aerosp. Mech. Eng. **3**, 80–90 (2020)
23. Akan, E., Kiracı, K.: Aircraft selection by applying AHP and TOPSIS in interval type-2 fuzzy sets. J. Air Transp. Manag. **89** (2020)
24. Ahmed, S.K, Sivakumar, G., Kabir, G., Ali, S.M.: Regional aircraft selection integrating FAHP and efficacy method. J. Prod. Syst. Manuf. Sci., 63–86 (2020)
25. Akyurt, İ.Z., Kabadayı, N.: Bulanık AHP ve Bulanık Gri İlişkiler Analizi Yöntemleri ile Kargo Uçak Tipi Seçimi: Bir Türk Havayolu Firmasında Uygulama **15**(57), 38–55 (2020)
26. Deveci, M., Öner, S.C., Çiftçi, M.E., Özcan, E.: Interval type-2 hesitant fuzzy Entropy-based WASPAS approach for aircraft type selection. Appl. Soft Comput. (2022)
27. Mahmut, B., Şahap, A., Emircan, Ö.: Regional aircraft selection with fuzzy Piprecia and fuzzy Marcos: a case study of the Turkish airline industry. Facta Universitatis Ser. Mech. Eng. **19**, 423–445 (2021)
28. Kutlu Gündoğdu, F., Kahraman, C.: A novel spherical fuzzy analytic hierarchy process and its renewable energy application. Soft. Comput. **24**(6), 4607–4621 (2019). https://doi.org/10.1007/s00500-019-04222-w
29. Gündoğdu, F.K.: Principals of spherical fuzzy sets. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A., Sari, I. (eds.) INFUS 2019. AISC, vol. 1029, pp. 15–23. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-23756-1_4
30. Camilleri, M.A.: Aircraft operating costs and profitability. In: Travel Marketing, Tourism Economics and the Airline Product, Chapter 12, pp. 191–204 (2018)

An Application with Fuzzy AHP for Making the Assignment Aimed at the Holistic Benefit Over the Person Evaluation as Required by the Requirements and Evaluations of Individuals

İbrahim Yel[1](✉), Hilal Şahin[1], Mehmet Emin Baysal[2], and Ahmet Sarucan[2]

[1] Architecht Information Systems, İstanbul 34906, Turkey
ibrahimyel@gmail.com
[2] Konya Technical University, Konya 42250, Turkey

Abstract. It is assumed that the success rates of information technology (IT) projects will increase with the management of development packages called Most Valuable Product (MVP). This is considered to be a real requirement. Based on this assumption, this study identified different criteria based on the requirements or MVPs to be developed in an organization. The criteria are differentiated for analyst and software developer positions. The level of each criterion was defined by the experts for the development of MVPs. The same criteria were used to evaluate both analyst and software developer positions. After evaluating the MVPs and determining the levels of the individuals, the weights of the criteria were determined using fuzzy AHP to determine the desired level and the levels of the individuals. The minimization problem of which person should be assigned to which MVP package over the weights was solved using the Excel solver. The sums were performed. The result is a balanced distribution of multiple MVP packages within a unit; with the goal of assigning MVP packages to the positions that can best improve themselves and increase overall benefit. Evaluation of employee competencies and needs definitions of projects with fuzzy MCDM method constitutes an important contribution of this study. Under the guidance of the evaluations, erroneous assignments that would lead the project to failure were avoided and more competent people were appointed for more critical requirements. Thus, the potentials of competent persons were utilized at the highest level.

Keywords: MVP · Assignment of Requirements · Personnel Evaluation

1 Introduction

Today's organizations must deliver more effective products or services while meeting increasing customer and service expectations and competing more intensely for market share. To this end, companies should have the most efficient use of their resources, the highest output at the lowest cost and the highest level of customer satisfaction. This can

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 438–444, 2023.
https://doi.org/10.1007/978-3-031-39777-6_52

only be achieved through the successful management and completion of projects in the field of information technologies. Software systems are now fundamental components of many businesses. As a matter of fact, in some sectors, such as banking and telecommunications, it is not possible for the organization to survive without the necessary software. Therefore, the need for advanced software systems is increasing day by day as the capabilities of software organizations grow. Without processes to support software development, these projects are unlikely to be successful [1]. CHAOS report, prepared by Standish Group in 2015, shares that the success rate of software development projects is 36%. Even though we believe that this rate has increased today, the success rates are still not at the desired level [2]. It can be assumed that the correct identification of requirements and the appropriate assignment of tasks according to these requirements have a significant impact on the success of software development projects. This study aims at assigning the requirements to the right people. First, requirements were evaluated by 5 experts from Architect Information Systems according to the specified criteria. Then, the employees were evaluated according to the competence criteria that correspond to the criteria of the requirements. Fuzzy AHP method was used for the weighting of the evaluation criteria among each other. Based on the differences between the criteria-based scores of the requirements and the competency criteria of the employees, an assignment model was created.

The remaining paper is organized as follows. The relevant literature is presented in the second section. The methodology is also part of this section. The third section provides a case study. Section four presents the concluding remarks.

2 Literature Review

When reviewing the literature, it is seen that AHP and Multi-Criteria Decision Making methods are often used to solve problems such as determining roles and responsibilities in projects, selecting, assigning and scheduling personnel. Danışman, states that team assignment problems are most commonly encountered and solutions are sought in the areas of health care, manufacturing, and project management, which we are also interested in [3]. Samanlioglu and Taskaya, weighted the criteria with Fuzzy AHP and used Fuzzy TOPSIS method to evaluate the personnel in the information technology department of a Turkish dairy company [4]. Aksakal and Dağdeviren, used criteria evaluation with Fuzzy AHP to plan jobs according to the needs and desires of the personnel in a manufacturing company. In the study, which aimed at personnel satisfaction, the assignments were made by maintaining workload balance and fairness [5]. Kozanoğlu and Özok, considered the effect of job requirements on personnel selection as important. Since the evaluation of job requirements levels and personnel qualifications is subjective and uncertain, they used Fuzzy Analytic Hierarchy Process (FAHP) and Fuzzy TOPSIS (FTOPSIS) and Fuzzy VIKOR (FVIKOR) methods to mathematically evaluate the most suitable candidate for personnel selection [6]. Bozat, states that if the assignments are not made according to the needs of the human resources department, the work will be delayed and the company will suffer losses. Otherwise, they have seen that redundant assignments cause unnecessary resource utilization. They found the assignment results with Centralized Data Envelopment Analysis [7]. Ceylan and Gürsev, analyzed the success criteria of 4 different application development projects with different project management methods of a company in the banking sector. By using AHP and TOPSIS methods from multicriteria decision making methods, it was found which methods had higher success rates and it was aimed to increase the success rate of future projects [8]. Varlı and Tamer divided the supervisors working in a factory into seniority levels using the Analytic Hierarchy Process (AHP) method. The aim of the study was to distribute the workforce in a balanced and fair way and an assignment model was created using Goal Programming. As a result of the study, employee satisfaction was achieved and it was seen that a balanced and experienced workforce was provided to the shifts [9].

2.1 Fuzzy AHP

As a multi-criteria decision making method, the fuzzy AHP method was chosen to weight the criteria. The geometric mean method, which consists in having more than one evaluator evaluating the criteria, was selected as the integrating method [10]. The calculations and formulas that are used in the fuzzy AHP method have been adapted from the study in the literature [11].

3 Case Study

With the data collected according to the literature, four criteria were identified to evaluate the requirements of software projects using the knowledge of senior analysts, software developers working at Architect Information Systems. These criteria are described in Table 1.

Table 1. Requirements criteria

No	Criteria
1	Level of communication requirements
2	Level of technical requirement
3	Difficulty of the requirement
4	Size of project

After the requirement criteria were identified, employee evaluation criteria were determined to evaluate analysts and software developers, each corresponding to a requirement criterion. The first three criteria were scored on a five-point Likert scale by five experts. This scoring was based on the information contained in the requirement definition.

Based on the requirements, the number of development areas such as Definition Screen, Listing Screen, Flow, Report Definition, Parameter Definition, etc. that are actively used in the effected information technologies is determined. This number is multiplied by the scale score assigned to the difficulty of the requirement, and the depth score is obtained. Thus, the depth of the requirement definition is also included in the model. For each requirement, the distribution of man-days between tasks was made.

In the Software Development Life Cycle (SDLC) process applied in Architect Information Systems, there are four main steps that affect the development part. Four tasks, namely analysis, software, test and bug fix, are considered in this study. In each step of the SDLC, man-days were allocated based on fixed coefficients given by the organization. After the man-days were determined, the multipliers of the requirements were determined.

The criteria-based score to be used in the allocation was obtained by normalizing the values of these criteria based on the requirements and the values of these values based on the project formed by the combination of the requirements. Since different positions are involved in each of the steps of the SDLC processes, a study was conducted on the steps belonging to analysts and software developers. Since the man-day values assigned to the projects of software developers and analysts are different, the man-days and requirements evaluation criteria were distributed based on the position.

After the requirements man-days were calculated, the employee ratings were scored separately for analysts and software developers. The scoring was done by the employees' managers. The scoring results are shown in Table 2.

Table 2. The scoring results

Name	Role	Coping with Difficulties	Communication Skills	Technical Ability	Experience
A	Analyst	4	3	3	3
B	Analyst	3	2	3	1
C	Software Developer	3	3	4	3
D	Software Developer	5	3	3	3
E	Software Developer	3	3	3	3
F	Software Developer	3	2	3	1
G	Analyst	3	3	3	3
H	Software Developer	3	2	3	1

3.1 Application of Fuzzy AHP

The criteria, determined by interviewing experienced employees in the field of software development, were compared with each other using the pairwise comparison matrix method and evaluated in such a way that they could receive {1,3,5,7,9} points. Each employee's score was recorded, and the final results were determined by arithmetic average. The weighted comparison matrix values of the criteria are given in Table 3 (evaluated according to the nine-point triangular scale).

The weights of the criteria were determined by comparing the competencies defined by five different experts. Table 3 shows the pairwise comparison results of two experts.

Table 3. The pairwise comparison results of two experts

Expert	Req. Criteria \Competencies	Coping with Diff	Com. Skill	Tech. Ability	Experience
E1	Level of Diff	Equal Importance	High Important	Less Important	High Important
E1	Level of Comm	Very Little Important	Equal Importance	Equal Importance	Less Important
E1	Tech. Ability	Less Important	Important	Equal Importance	Less Important
E1	Size of project	Equal Importance	Equal Importance	Very Little Important	Equal Importance
E2	Level of Diff	Equal Importance	High Important	Less Important	High Important
E2	Req. of Comm	Equal Importance	Equal Importance	Very Little Important	Less Important
E2	Tech. Ability	High Important	High Important	Equal Importance	High Important
E2	Size of project	Equal Importance	Equal Importance	Very Little Important	Equal Importance

The scores were averaged using the geometric mean method, fuzzified and normalized. Weights were obtained over the scores obtained by defuzzification.

3.2 Assignments

Excel Solver was used to compare the results of the requirements and staff assessments of the projects. The project-employee assignments were made in such a way as to minimize the positive and negative difference between the values.

After obtaining the values of the criteria Man-Days, Demand Size, Difficulty Level, Communication Requirement, Technical Requirement of the project requirements, the values of the employee criteria that match these criteria one-to-one; Coping with Difficulty, Communication Skills, Technical Skills, Experience were also obtained.

The obtained requirement and employee criteria values were matched using Excel Solver in such a way that the positive and negative differences between the requirement criteria values and the employee competencies values were minimized. The constraints of one analyst per project and a minimum of one and a maximum of two software developers per project were also added. The assignment results are shown in Table 4 and Table 5.

Table 4. Assignment of Analysts

Analysts Name			
Project No	A	B	G
1	1	0	0
2	1	0	0
3	0	0	1
4	0	1	0
5	0	1	0
6	0	0	1

Table 5. Assignment of Software developers

Software developers Name					
Project No	C	D	E	F	H
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	0	0	1
4	0	0	0	1	0
5	0	0	1	0	0
6	0	0	1	0	0

4 Results

Besides the research on tasking, the evaluation of employee competencies and project requirements is one of the important criteria of this study with regard to its contribution to the literature.

The objectives such as "project success rate", "employee satisfaction" and "utilization of employee potential" have drawn attention to the need for task assignment in project management.

Through the application of this systematic to the results obtained, project requirement criteria and employee evaluation criteria can be determined, which can be agreed upon by all organs within the organization; after this study, task assignment in subsequent projects can be carried out in a healthier and more efficient way. The first goal is to ensure that people who do not have sufficient skills to meet the project requirements are not assigned to the project and are not the cause of the project's failure.

The second goal is to utilize the employee's potential at the maximum level by not assigning the person who has very high competencies according to the project requirements to this project.

The third goal is to maintain the employee's motivation by assigning according to the employee's competencies. The assignment models created as a result of the studies will be used for project assignments in the company Architect. They will contribute to the company's efficiency and employee satisfaction.

In advanced studies that will be built on this work in the future, data analysis can be made to increase the success of the project obtained after the implementation of the current system. Afterwards, according to this analysis, estimations for the secondary benefit of the completion times of the projects over the position levels of the experience as well as the competencies of the employees or the outputs with simulation can be discussed.

References

1. Nalbant, S.: Improving the efficiency of the software development process: an information systems proposal, 1.ULUSAL YAZILIM MÜHENDİSLİĞİ SEMPOZYUMU, pp. 1–7, EMO, İzmir (2003)
2. Standish Group Homepage. https://www.standishgroup.com. Accessed 2 April 2023
3. Danışman, O.: Investigation of team aassignment problem with multi criteria decision making methods, Master thesis, Başkent University Institute of Science and Engineering Department of Engineering and Technology Management, Ankara (2020)
4. Samanlioglu, F., Taşkaya, Y.E., Çokcan O.: A fuzzy AHP-TOPSIS based group decision-making approach to personnel selection, vol. 20, pp. 1576–1591 (2018)
5. Aksakal, E., Dağdeviren, M.: Talent management based personnel assignment model and solution propasal. J. Facul. Eng. Architect. Gazi Univ. **30**(2), 249–262 (2015)
6. Kozanoğlu, O., Özok, A.F.: A fuzzy human resource allocation model in quality function deployment, Master thesis, ITU Institute of Science and Technology, Industrial Engineering Program, İstanbul (2011)
7. Bozat, S.B.: A Data envelopment analysis approach for efficient human resource allocation in project management, Master thesis, Hacettepe University Institute of Science and Technology Department of Industrial Engineering (2022)
8. Ceylan, Z., Gürsev, S.: Comparison of scrum-kanban-waterfall applications in information technology projects using AHP and TOPSIS methods. J. Inform. Technol. **13**(3), 329–339 (2020)
9. Varlı, E., Eren, T.: Shift scheduling problems and a case study. J. Inform. Technol. **10**(2), 185–197 (2017)
10. Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**(3), 233–247 (1985)
11. Ulutaş, A, Özkan, Ahmet M., Tağraf, H.: Personnel selection by using fuzzy analytic hierarchy process and fuzzy grey relational analysis. Electron. J. Soc. Sci. (17)65 (2018)

An Application for Selecting and Evaluating Organization Based KPIs with Fuzzy MCDM in a Large-Scale Organization Operating in the IT Sector

İbrahim Yel[1(✉)], Mehmet Emin Baysal[2], and Ahmet Sarucan[2]

[1] Architecht Information Systems, İstanbul 34906, Turkey
ibrahimyel@gmail.com
[2] Konya Technical University, Konya 42250, Turkey

Abstract. In today's world, where many people use the outputs of the Information Technologies (IT) industry, it has become a problem for large-scale organizations to compare their IT departments objectively. The use of fuzzy MCDM methods for such a multidimensional problem is evaluated within the scope of this study. As a first step, twenty KPIs were determined by literature review and expert opinion to evaluate organizational units. Subsequently, KPI selection criteria were defined based on the opinions of ten expert personnel working in Architecht Information Systems for the selection of the most appropriate KPIs. These criteria are weighted by the fuzzy AHP method. Eight of the KPIs over the Neutrosophic Z Number (NZN) sets were selected for comparison by the same experts. Four different organizations were ranked by scoring over eight KPIs. It has been confirmed that the results of objective evaluations, instead of subjective evaluations, are more embraced by employees and managers, according to the outputs of the common sense and the infrastructural suitability of the institution.

Keywords: Fuzzy MCDM · IT Organizations · KPI Selection

1 Introduction

The increase in the effectiveness of information technologies (IT) in recent years brings with it new searches in the performance measurement of organizational units consisting of employees who play an active role in the successful completion of pro-jects. In large-scale software development organizations where agile transformations are implemented, determining the measurement criteria on the basis of products and organizational units suitable for this transformation has become a more specific problem.

This study investigates the solution of the problem of determining performance indicators on the basis of organizational units. In short, as solution steps, 20 alternatives that could be performance indicators were suggested by 10 experts in their fields working at Architecht Information Systems. Comparison criteria are weighted with fuzzy AHP. Afterwards, based on the evaluations, selection of eight KPIs were made with neutrosophic Z number sets. Comparisons were made by the experts over eight selected KPI's.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 445–453, 2023.
https://doi.org/10.1007/978-3-031-39777-6_53

Four different organizational units were compared KPI points and their weights and than ranked.

After a brief introduction, the second section of the study is dedicated to related studies in the literature. Methodological short identifications of Fuzzy AHP and Neutrosophic Z numbers are also included in this section. The third section is dedicated to a case study. The results and discussion are the subject of the fourth section.

2 Literature Review

When the literature is examined, there are not so many studies on KPIs in the IT field. It is also possible to find certain KPIs, which are defined by the opinions of valuable experts on various websites anonymously. In addition, recent studies in the field of fuzzy AHP and Neutrosophic Z Number Sets (NZN), which mostly appeal to the field of IT, are included in this section.

In a study figured out in the literature, writers stated that KPIs should be directly related to the matches of the organization. However, it was stated that it should be measurable and quantifiable [1]. In another study, KPIs investigated were divided into groups. In this grouping, some definitions were made such as the risk of the application, ability to work, business continuity and whether there is a coordinated strategy with the corporate architecture [2]. Another study, the authors were developed a simulation-based approach for IT service management. In this study, more customer-oriented and service-oriented criteria were examined on the basis of external factors such as competition, market pressure and customer requirements [3]. On the other hand, were expressed, the necessity of understanding the KPIs and SLAs of the employees, focusing on the role of the organization they were affiliated with in the organization as a whole. Subsequently, the author identified the five principles that should be used in KPI definitions [4]. In another study in the field of KPI evaluation on IT sector, the writers made a comparative compilation of the changes in the ITIL performance indicators on the basis of before and after the ITIL adaptation, over a case study [5].

Another literature study handled by authors is project selection problem in the field of information technologies, claiming that they have defined an applicable framework through single valued neutrosophic numbers and AHP methods [6]. Similarly, in [7], while selection of the project through TOPSIS and DEMATEL methods in the neutrosophic environment, were stated that choosing the project as the most important step of the project life cycle. From another perspective, the authors were used neutrosophic statistics in the strategic planning of information systems. In this context, they pointed out that information systems simplify business processes, reduce costs, and are more secure for transactions [8]. In a study about IT services, the writers emphasized the importance of optimizing the working time of software developers and/or technicians by right prioritization in IT services. In their studies, they used Saaty's AHP method from a neutrosophic perspective for weighting [9].

Another example of using AHP in a neutrosophic environment about the selection of the best candidate to reach a decision over the differences of decision makers are made in [10].

Finally, another study in the field of IT using AHP and neutrosophic Z number sets involved determining the competencies of talented personnel working in the IT sector and to evaluate the employees in the role of system analysts based on these competencies [11].

2.1 Fuzzy AHP

Emphasized in [12] that the fuzzy AHP method is the most preferred method in the criterion weighting phase. The geometric mean used for group decision-making techniques was used to combine the evaluations [13]. The steps regarding this method were applied by being inspired by the study involved in [14].

2.2 Neutrosophic Z Number Sets

The calculation steps are obtained from [15]. The scale used in the calculations in NZN steps is the same as the study of [16].

3 Case Study

With the knowledge gained in the scope of literature reviews and opinions of three analysts and seven software development experts who have at least five years of experience in Architecht Information Systems, five criteria identified and used in the evaluation of KPIs are shared in Table 1.

Table 1. KPI evaluation criteria.

Criteria Code	Criteria	Weights
1	Data Availability	%24,7
2	Organizational Segregation	%17,8
3	Intelligibility	%18,0
4	Compliance with Corporate Culture	%19,3
5	Representation Ability	%20,2

During the determination of the KPI evaluation criteria, the experience of the expert personnel was used with the brainstorming, which was made with evaluations based on the criteria included in the studies in the literature. These criteria in Table 1 were weighted with fuzzy AHP by the same expert group. A nine-level scale was used for weighting. The results of the weights are again in the rightmost column of Table 1.

After the criterion weights were determined, evaluations were made on the Neutrosophic Z number sets(NZN) method during the evaluation of the alternative KPI's. In these evaluations, the concepts of Truth, Falsity and Indeterminacy regarding the mathematical equivalents of the scale expressed with linguistic variables were conveyed to the experts in the evaluator and it was also ensured that their knowledge about the evaluation systematic was refreshed.

Table 2. Detailed information about alternatives.

KPI Code & Identification	KPI Description & Direction
1. Auto Finding's Density	The ratio of the Number of Findings with the "Critical" or "High Significance" rating in a Project to the number of lines of code in the Project(Decrease)
2. Infrastructure Projects Planning Effectiveness	The ratio of the total number of person days belonging to the infrastructure projects that will affect the continuity of the system to the difference between the start and end dates of the relevant project(Increase)
3. Customer satisfaction	The weighted average of the results of the satisfaction survey sent to the stakeholders affected by the developed project over the project man-day cost(Increase)
4. Requirement Definition Effectiveness	Correlation value between the number of person days incurred for a project and the number of requirements defined for the project(Increase)
5. Test Scenario Generation Effectiveness	Correlation value between the number of person days incurred for a project and the number of test cases defined for the project(Increase)
6. Project Progress Consistency	The standard deviation between the weekly progress rate between the start and end days of the projects and the rate of progress over the person day (Decrease)
7. Development Effectiveness	The sum of the results of the change effect made within the scope of the project with the score determined on the basis of the module, on the basis of all projects(Increase)
8. Defect Generation Rate	The average of the ratio of the number of defects defined as a result of the developments made on the basis of the project to the man-day value of the project.(Decrease)
9. Defect Resolution Speed	According to the level of the defect defined (Low - Medium - High), lower or higher resolution speed than the average resolution time in the institution(Increase)
10. Financial Impact of Requirements	The ratio of the personnel labor cost to be incurred in the realization of the requirements to the financial benefit to be obtained by the realization of the requirement(Decrease)

(*continued*)

Table 2. (*continued*)

KPI Code & Identification	KPI Description & Direction
11. Error Solution Cost	The ratio of the resolution time of errors with error records opened by end users to the total working time in the relevant working period(Decrease)
12. Agile Maturity Level	The standard deviation of the differences between the story points of burndown chart, and the realized story points, in the sprints of agile projects(Decrease)
13. Average Waiting Time of Projects	The average of subtracting the working days between the assignment of the projects and the transition to the next phase from the working days planned (Decrease)
14. Module Size Ratio	The ratio of the total number of lines of code in the modules under the responsibility of the teams to the number of classes in the relevant modules(Increase)
15. ISF Generation Capacity	The ratio of the number of findings to the total number of lines of code(Decrease)
16. Creating Added Value	The ratio of the difference between the working hours that add value to the work and the activities that do not add value to the working time(Increase)
17. Competency Development	The slope of the increase or decrease in the scores obtained by the employees as a result of the competency evaluations(Increase)
18. Project Deadline Compliance	Standard deviation of the difference between the time determined as the project deadline and the completion time of the relevant project(Decrease)
19. System Availability	The ratio of the total working time and the time difference of possible system interruptions to the total working time(Increase)
20. Education Importance	The standard deviation of the difference between the planned training periods and the trainings received in certain periods(Decrease)

Details of the definitions determined as alternative KPIs are given in Table 2. In order to better understand the definitions in the "KPI Description" column in Table 2, the following definitions should be known. **Focus:** Refers to the error or review requests encountered in the live environment, opened by end users, especially for IT personnel to review. **Defect:** It refers to the records of structural or business errors detected during the development phase of the application, opened by both analysts and test engineers. **Findings:** It refers to automatic findings opened by internal or external applications that scan code development environments at certain periods. **Module:** It is the name given to the grouping that expresses the code groups that IT personnel are responsible for. **SDLC:** Abbreviation of the term "Software Development Life Cycle". It expresses the life cycle of software development. **Information Security Findings (ISF):** It is the record of the points that need to be corrected as a result of standard or transitional scans by information security teams.

The 11-level linguistic scale used in the assessments is given in Table 3. The evaluations of all evaluators according to the scales in the Table 3, were combined with the geometric mean.

Table 3. NZN scale set [16].

No	Linguistic Terms	Neutrosophic Set					
		TL	TU	IL	IU	FL	FU
1	Equal importance	0,5	0,5	0,5	0,5	0,5	0,5
2	Weakly more importance	0,5	0,6	0,35	0,45	0,4	0,5
3	Moderate importance	0,55	0,65	0,3	0,4	0,35	0,45
4	Moderately more importance	0,6	0,7	0,25	0,35	0,3	0,4
5	Strong importance	0,65	0,75	0,2	0,3	0,25	0,35
6	Strongly more importance	0,7	0,8	0,15	0,25	0,2	0,3
7	Very strong importance	0,75	0,85	0,1	0,2	0,15	0,25
8	Very strongly more importance	0,8	0,9	0,05	0,1	0,1	0,2
9	Extreme importance	0,9	0,95	0	0,05	0,05	0,15
10	Extremely high importance	0,95	1	0	0	0	0,1
11	Absolutely more importance	1	1	0	0	0	0

As a result of the evaluations, the result points on the far right of Table 4 and the ranking results of these alternatives were obtained. The ranking results were calculated as a result of the evaluation of all experts' opinions.

Since the evaluation results of ten evaluators cannot be shared due to page limitations, only the results of the third evaluator are included in Table 4. Table 4 also includes the details of the evaluation results of the third evaluator over the five criteria of twenty KPI definitions. These details are coded over the numbers of the scales shared in Table 3 and the KPI numbers shared in Table 2.

Eight of the alternatives with the highest weight were selected as previously decided by the experts. With this selection, KPI definitions with high intelligibility and acceptability, which were actually evaluated by the expert staff within the institution, were reached.

Table 4. Evaluation information of alternatives by third evaluator.

KPI Code	First Criterion	Second Criterion	Third Criterion	Fourth Criterion	Fifth Criterion	Order	NZN Results
1	6	2	9	4	9	2	0.974
2	7	11	7	2	11	1	0.982
3	2	9	7	10	10	3	0.970
4	4	3	4	4	3	14	0.699
5	3	3	5	4	4	15	0.698
6	2	3	3	1	3	18	0.597
7	1	1	1	1	1	20	0.537
8	3	1	3	1	1	19	0.596
9	3	3	4	4	4	13	0.720
10	5	2	5	2	3	17	0.680
11	6	4	5	6	6	5	0.785
12	7	2	7	7	7	7	0.765
13	6	6	5	7	2	8	0.761
14	2	2	7	7	7	12	0.748
15	7	2	7	2	2	11	0.751
16	2	7	7	2	7	6	0.781
17	4	4	6	5	6	9	0.761
18	4	5	3	3	2	16	0.683
19	6	4	6	7	7	4	0.789
20	2	7	2	7	2	10	0.758

As a result, the results of the eight selected KPI definitions belonging to four different organizational units that actively carry out software development activities in the relevant institution were obtained. While reaching the results, the scores and values on the basis of KPI were normalized according to the increasing or decreasing direction identified in Table 2. The weights of the eight selected KPIs were calculated over the values in the "NZN Results" column in Table 4 used eight chosen KPI. The relevant results are given in Table 5. For instance, the calculation steps of Over All row is simply multiplying the value of "Org2" and "Weights" then sum the results of KPI based multiplied points.

Table 5. Organizations' evaluation results over selected KPI's.

KPI Code	Org1	Org2	Org3	Org4	Weights
2	0,76	0,82	0,65	0,85	%14,4
1	1	0,8	0,6	0	%14,3
3	0,85	0,88	0,82	0,77	%14,2
19	0,99	1	0,98	0,96	%11,6
11	0,99	1	0,2	0	%11,5
16	0,72	0,85	0,65	0,68	%11,5
12	0	0,01	1	0,7	%11,2
13	0	1	0,89	0,72	%11,2
Over All	0,69	0,8	0,72	0,58	

4 Results and Discussion

It is one of the outputs of this study that it has been applied in an institution with MCDM methods as an additional contribution to the studies on KPI definition, especially for the evaluation of organizations in the IT sector. In addition, in large-scale organizations that tend to compartmentalize their organizational units on a product basis with agile transformation, this study has an exemplary feature in order to compare organizational units.

The results in Table 5 are also a reference for continuous improvement within the company by comparing organizational units that have different areas of responsibility. Over identified and selected eight KPIs, gives the opportunity of scoring and comparing these units separately. t would be appropriate to compare the second organization with the highest score and the fourth organization with the lowest score. The fourth organization's score(0.85) is higher than the second organization's score(0.82) over KPI number 2 with the highest weight(14.4). In terms of KPI number 12, the fact that the second organization's score(0,01) was very low compared to the fourth organization's score(0,7) did not prevent the second organization from taking the first place by integrating all KPI evaluations over weights.

For future research, based on the results of this study, KPIs that all organizations will agree on can be determined within the institution. So, the comparisons can be made on the data that will be drawn from these KPIs. Additionally, the infrastructures for the production of corporate balanced score cards can be created. By going into more detail over the detailed explanations in the KPI definitions the model can be more inclusive and acceptable.

References

1. Mazlum, A.R.: Determining of key performance indicators in business. İstanbul Ticaret Üniversitesi Girişimcilik Dergisi **1**(2), 1–14 (2017)

2. Sofiyabadi, J., Kolahi, B., Valmohammadi, C.: Key performance indicators measurement in service business: a fuzzy VIKOR approach. Total Qual. Manag. Bus. Excell. **27**(9–10), 1028–1042 (2016)
3. Cater-Steel, A., et al.: Decision support systems for IT service management. Int. J. Inform. Decision Sci. **8**(3), 284–304 (2016)
4. Brooks, P.: Metrics for IT service management Van Haren (2006)
5. Spremic, M., Zmirak, Z., Kraljevic, K.: IT and business process performance management: case study of ITIL implementation in finance service industry. In: ITI 2008–30th International Conference on Information Technology Interfaces. IEEE (2008)
6. Alava, M.V., et al.: Single valued neutrosophic numbers and analytic hierarchy process for project selection. Infinite Study (2018)
7. Abdel-Basset, M., Atef, A., Smarandache, F.: A hybrid Neutrosophic multiple criteria group decision making approach for project selection. Cogn. Syst. Res. **57**, 216–227 (2019)
8. Ponce Ruiz, D.V., et al.: Neutrosophic statistics in the strategic planning of information systems. Neutrosophic Sets Syst. **44**(1), 44 (2021)
9. Molina Chalacan, L.J., et al.: Selection of criteria for the prioritization of information technology services in a neutrosophic environment. Neutrosophic Sets Syst. **44**(1), 41 (2021)
10. Abdel-Basset, M., et al.: Multi-criteria group decision making based on neutrosophic analytic hierarchy process. J. Intell. Fuzzy Syst. **33**(6), 4055–4066 (2017)
11. Yel, İ., Sarucan, A., Baysal, M.E.: Determination of competencies with fuzzy multi-criteria decision making methods for determining the development program for analyst position in a participation bank. In: Kahraman, C., Tolga, A.C., Cevik Onar, S., Cebi, S., Oztaysi, B., Sari, I.U. (eds.) Intelligent and Fuzzy Systems. INFUS 2022. LNNS, vol. 504. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09173-5_51
12. Heidary Dahooie, J., et al.: Competency-based IT personnel selection using a hybrid SWARA and ARAS-G methodology. Hum. Factors Ergon. Manufac. Serv. Indust. **28**(1), 5–16 (2018)
13. Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**(3), 233–247 (1985)
14. Ulutaş, A., Özkan, A.M., Tağraf, H.: Personnel selection by using fuzzy analytic hierarchy process and fuzzy grey relational analysis. Electron. J. Soc. Sci. **17**(65) (2018)
15. Ye, J.: Similarity measures based on the generalized distance of neutrosophic Z-number sets and their multi-attribute decision making method. Soft. Comput. **25**(22), 13975–13985 (2021). https://doi.org/10.1007/s00500-021-06199-x
16. Bolturk, E., Kahraman, C.: A novel interval-valued neutrosophic AHP with cosine similarity measure. Soft. Comput. **22**(15), 4941–4958 (2018). https://doi.org/10.1007/s00500-018-3140-y

Fuzzy Cognitive Maps

Analysis of the Factors Behind the Value of an Agency for an Airline

Irem Ucal Sari(✉) and M. Fatih Özer

Industrial Engineering Department, Istanbul Technical University, Macka, Istanbul, Turkey
ucal@itu.edu.tr

Abstract. Generally, discounted fares are offered for agencies requesting group reservations compared to individual fares. This is because group bookings are seen as a more stable source of revenue for airlines. Especially in the Covid-19 period, the agencies whose business was seriously damaged requested the lowest prices they can get, to be able to find passengers and make a profit. The airlines were also in a financial crisis and must have balanced the prices to make a profit while attracting the agencies. From this perspective, the need for customer segmentation and defining the value of agencies for airlines arises. The purpose of this study is to define "which dynamics, that the agency creates while making business with airlines, will be more important to form its value" by using Fuzzy Cognitive Mapping (FCM), which is one of the decision-making techniques. This study examines which variables might affect the value of agencies to an airline through the weight of these variables in order to create insights into the pricing strategies of airlines to travel agencies.

Keywords: Fuzzy Cognitive Mapping · Agency Valuation · Civil Aviation

1 Introduction

The civil aviation industry has been one of the fastest-growing industries of the 21st century, both in terms of companies' capacities and passenger demand [7]. Commercial aviation has become a part of people's lifestyles in the last century, as an indicator of economic prosperity, as a sector that stimulates trade and encourages the development of tourism and has become relatively more accessible in recent years [8]. Although the sector is growing rapidly in terms of capacity and demand, companies have always had to struggle with low-profit margins. The main reasons for this situation are the fragility of the aviation industry to external factors, high and fluctuating fuel prices, which are one of the most important cost items for airlines, currency fluctuations, interest rates, and high competition in the market. The civil aviation sector, where competition and external factors are high, is an area where profitability and stable income are very difficult, together with costs such as labor, safety, and operation costs, as well as fuel costs.

An example of the effect of external events on the airline industry is the terrorist attack of September 11, which depressed the entire industry, caused many airline companies to go bankrupt, and those who survived this effect were caught in the 2002 oil crisis

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 457–464, 2023.
https://doi.org/10.1007/978-3-031-39777-6_54

and had a very difficult time [12]. The factors that play an important role in the further increase of high competition can be shown as the entry of low-cost carriers (LCC) into the market in 2006 and the development of pricing strategies [2]. On the other hand, price competition has also contributed significantly to the success of airline companies. In addition, the availability of flight tickets and online reservations played an important role in this success [9].

The recent outbreak of Covid19 has caused great damage to the global economy, socially and financially. One of the industries most affected by this epidemic has been the global airline industry [1]. In the report "Covid-19's Impact on European Aviation on 2020 and Outlook on 2021" published by Eurocontrol [4], it is stated that 2020 is the worst year the industry has experienced and that only the European aviation industry has suffered a financial loss of 56.2 billion Euros. The report also states that more than 6 million flights were canceled and 190,000 aviation workers were laid off. It was stated in the report that six million and one hundred thousand flights were canceled along with the flights to or from Europe, which decreased by 55% last year due to Covid19. Fuel prices, sharp fluctuations and devaluations in the currencies of countries, and the demand shock, which decreased due to the effect of lockdowns, caused 3 airline companies to enter a difficult period again which were observed to be vulnerable to the deterioration in the cyclical economy.

For all these reasons, in the aviation industry with low-profit margins and frequent crises, institutions are looking for ways to increase their revenue as well as reduce their costs. At this point, the importance of revenue management emerges once again. Revenue management is an area where it is aimed to maximize revenue by selling the right product to the right customer at the right price at the right time. Revenue management started in 1978 with the aviation industry, where liberalization began in the USA, and continued its development. The seat inventory on the aircraft is managed and the seat fees are differentiated according to this principle, considering the distribution of seats, the passengers who would like to pay low fares and would not otherwise choose the flight, and the passengers who are willing to pay high fares to participate in the flight. In this way, it is tried to maximize the occupancy rate on the plane and the total income obtained from the passengers boarding the plane.

Most of the studies in the literature are mainly made for individual customers. When the studies in the literature on inventory control are examined, it is seen that the optimization of group reservations is generally not taken into account in revenue management models [10]. However, considering the "Load Factor", if the strategy of taking only individual passengers on the plane at higher fares is applied, there may be a loss of income due to the empty seats. For this reason, group requests are important for airlines and priced appropriately to provide a win-win situation for both agents and their passengers as well as the airlines. Airlines aim to increase their revenue by using some analytical techniques to best control the inventory of the seats they offer to the market in various fare classes.

Generally, discounted fares are offered for agencies requesting group reservations compared to individual fares. This is because group bookings are seen as a more stable source of revenue for airlines. Especially in the Covid-19 period, the agencies whose business was seriously damaged requested the lowest prices they can get, to be able to find passengers and make a profit. The airlines were also in a financial crisis and must have balanced the prices to make a profit while attracting the agencies. From this perspective, the need for customer segmentation and defining the value of agencies for airlines arises. The purpose of this study is to define "which dynamics, that the agency creates while making business with airlines, will be more important to form its value" by using Fuzzy Cognitive Mapping (FCM), which is one of the decision-making techniques. This study examines which variables might affect the value of agencies to an airline through the weight of these variables in order to create insights into the pricing strategies of airlines to travel agencies. A fuzzy cognitive map is implemented to explore the factors affecting the value of agencies. According to the results of this study, the sensitivity of agencies to the fare offers of airlines, and the annual reservation volume of agencies are particularly important for the value of the agency to the airline. On the other hand, the applied two scenarios show that there is not only one criterion for the value of an agency to the airline, the chosen concepts are related and have effects on each other.

The paper is structured as follows: Sect. 2 provides a brief overview of fuzzy cognitive mapping. Section 3 then presents a detailed explanation of the application of FCM to the factors influencing the value of an agency for an airline. Finally, the paper concludes with suggestions for future research.

2 Fuzzy Cognitive Mapping

The term "Fuzzy Cognitive Map" was created by Kosko [6] and it is a modification of a cognitive map by using fuzzy causal relationships with real numbers [−1, 1]. FCM is being used to model several different decision-making problems in various fields. According to Chandana et al. [3], FCM represents domain knowledge with a graphical representation formed of nodes and arcs that are joining various nodes. The nodes are defined as concepts or variables that are connected with directed arcs. The relationship of these variables is shown by arcs show with a weight in the range of [−1, 1]. These weights explain the relative strength between variables or so-called concepts. When there is a positive relation between two concepts, it shows that if one concept increase, the connected concept will rise according to the one increased. On the other hand, when there is a negative relation, an increase in one concept will lead to a decrease in the connected concept. A weight near 1 or −1 means a strong relationship regardless of the sign. The concepts in the cognitive map can be factors, actors, or entities which have an effect on the system. Both intuitive and quantifiable factors can be modeled in Fuzzy Cognitive Maps [11]. With FCM, the changes in the system over time can be simulated after the initial state is represented with a matrix. This method simulates the behavior of the system in the long run with an appropriate number of iterations until the system reaches a steady state. Equation (1) represents the simulation process:

$$A_i^{(k+1)} = f\left[A_i^{(k)} + \sum_{\substack{j=1 \\ j \geqq i}}^{n} A_j^{(k)} w_{ji} \right] \tag{1}$$

where $A_i^{(k+1)}$ is value of variable i at $(k+1)^{th}$ iteration, $A_i^{(k)}$ is value of variable i at k^{th} iteration. W represents the weight between two variables. f is a threshold function that can be a bivalent, trivalent, hyperbolic tangent, and sigmoid function and reduces the concept values in a comparable range [5].

3 Application

In this study, the concepts are determined as variables that define the value of an agency for an airline to express the crucial concepts that are most effective on the value of an agency. The variables included in the model are; annual reservation volume (C1), the peak/off-peak distribution of reservations (C2), materialization rate of an agency (C3), reliability of an agency on reservation payments (ease of collection) (C4), the history of an agency with an airline (duration of business) (C5), diversity of passenger portfolio (leisure, ethnic, business, economy) (C6), origin and destination diversity (volume of requests for more revenue-generating lines) (C7), penalty collection rate / ease of penalty collection (C8), the ratio of the airline to the total business volume of the agency (C9), the agency's sensitivity to the fare offers/being very negotiable (C10).

These concepts are all variables that form the value of an agency to the airline. As the first concept, annual reservation volume is determined. The volume of reservations means the number of passengers an agency brings to the airline, which is one of the most important parts of the relationship between the airline and the agency. The second concept is the peak/off-peak distribution of agency reservations. On peak dates, the airline already has enough demand from several agencies and individual passengers. However, during the off-peak dates, the demand is very low and airline needs this bulk of passenger groups very much. The third concept is the materialization rate of the agency. During the protocol agreements with agencies, airlines offer a price for a certain period of time and flights in certain destinations and expect a percentage of materialization from an agency for the offered price. If the agency agrees to the offer and can't realize the agreed rate of materialization, the airline applies a penalty for every passenger that has not been fulfilled. The fourth concept is the reliability of an agency on reservation payments. The payment history of an agency can be seen easily from the records. The payment of reservations at the exact agreed dates is proof of reliability for an agency. The fifth concept is the history of an agency with an airline. It can be explained as the duration of time that an airline is working with a specific agency. The sixth concept is the diversity of the agency's passenger portfolio. There are various types of passengers such as leisure, ethnic, seaman, student, business, etc. These passenger groups act together. Having a market share from these passenger groups is valuable for airlines. The seventh concept is origin and destination diversity. Some routes create more revenue for airlines. The agency that brings passengers for more revenue generated ONDs is creating more revenue for airlines. The eighth concept is the penalty collection rate and ease of penalty collection. When the agency can't realize the agreed materialization rate, a penalty is applied. The ninth concept is the ratio of the airline to the total business volume of the agency. This is an important variable because if the airline has the biggest volume in the business of an agency, it creates a more reliable and profitable relationship for both sides. On the other hand, if the volume is low, the agency generally makes requests for breaking the price with other airlines or only makes requests when there is no availability for the flights of other airlines. The final concept is the agency's sensitivity to fares. When an agency is sensitive to the prices, the process of the agreement takes more time and the agency becomes very bargainer. These agencies are less profitable and require longer work time for analysts. This study analyzes these variables to see which one of them will be more crucial to the airline while working on discount rates and fare rules for agencies.

In the application of FCM, five experts that are working in the revenue management department of the airline company determine the relationships. The causal relationships among these variables are assessed. Causal map of the aggregated judgments of the experts is shown in Fig. 1.

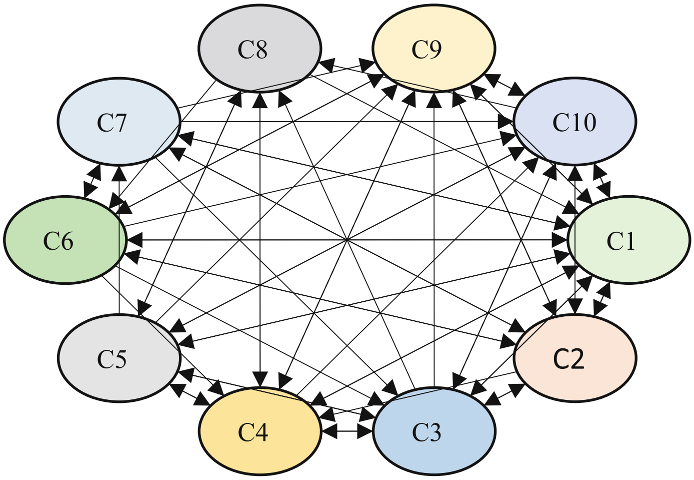


Fig. 1. Aggregated map of experts

The strength of the existing relationships is indicated by the causality as linguistic variables in the interval [−1,1], defined as "very very high", "very high", "high", "medium", "low", very low", and "very very low". The membership functions, μ, of the fuzzy sets characterizing these linguistic terms are given in Fig. 2 [5]. In the defuzzification of fuzzy variables graded mean integration approach is used.

The aggregated adjacency matrix that is constructed using the majority rule is given in Table 1.

In order to trigger FCM simulations, an initial vector, $A^{(0)} = [1,0,1,0,0,0,1,0,0,1]$, where Annual reservation volume (C1), materialization rate of an agency (C3), origin and destination diversity (C7), the agency's sensitivity to the fare offers/being very negotiable (C10) are activated, is formed. The sigmoid function is used as a threshold function to reduce the concept values into a normalized range through iterations in Eq. 1. The steady state values are calculated as $A^{(5)} = [0.999, 0.993, 0.995, 0.995, 0.992, 0.968, 0.953, 0.968, 0.997, 0.997]$.

The behavior of the system during time is also represented in Fig. 3.

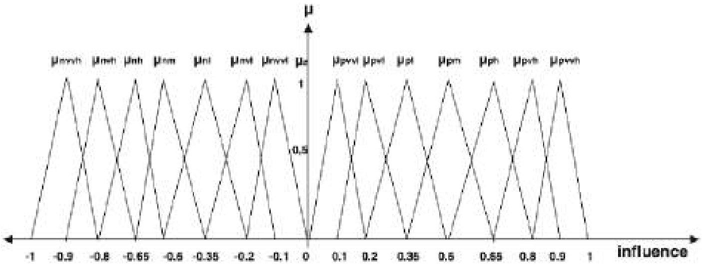


Fig. 2. Membership functions of linguistic variables [5]

Table 1. The aggregated adjacency matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0.00	0.76	0.74	0.68	0.67	0.69	0.67	0.00	0.75	0.63
C2	0.60	0.00	0.71	0.35	0.00	0.73	0.35	0.00	0.74	0.68
C3	0.61	0.53	0.00	0.78	0.59	0.00	0.00	0.58	0.57	0.77
C4	0.75	0.00	0.45	0.00	0.72	0.00	0.00	0.72	0.35	0.35
C5	0.78	0.00	0.71	0.73	0.00	0.00	0.30	0.70	0.65	0.50
C6	0.70	0.78	0.50	0.45	0.00	0.00	0.75	0.00	0.72	0.61
C7	0.70	0.64	0.58	0.00	0.00	0.64	0.00	0.00	0.43	0.60
C8	0.58	0.00	0.00	0.75	0.71	0.00	0.00	0.00	0.00	0.00
C9	0.75	0.72	0.00	0.60	0.61	0.43	0.00	0.00	0.00	0.65
C10	0.70	0.65	0.71	0.00	0.58	0.00	0.00	0.45	0.61	0.00

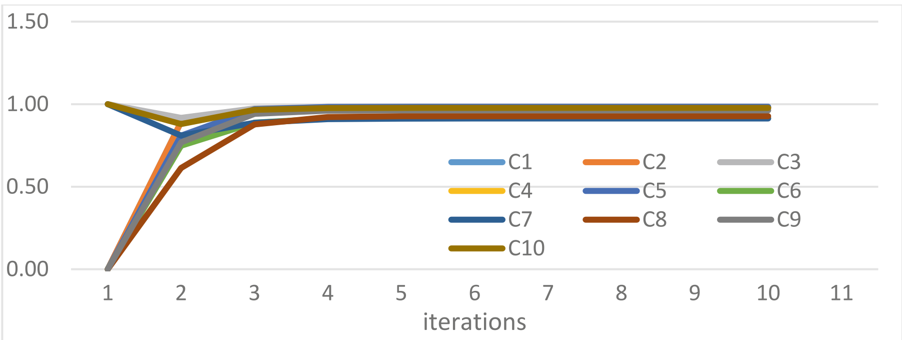


Fig. 3. The behavior of the system during iterations (aggregated)

According to the study's findings, the most important variables that impact the value of an agency to an airline are; annual reservation volume (C1), the ratio of the airline to the total business volume of the agency (C9), and the agency's sensitivity to the fare offers/being very negotiable (C10). The study suggests that airlines should focus on these variables while working on discount rates and fare rules for agencies to improve their value proposition to the agencies.

4 Conclusion

In conclusion, the civil aviation industry has been growing rapidly, but it has always been struggling with low-profit margins due to various external factors, including competition, fuel prices, currency fluctuations, interest rates, and global crises. The recent Covid-19 pandemic has further worsened the situation, causing significant financial losses and layoffs in the global aviation industry. In such a challenging environment, revenue management has become more important than ever for aviation companies to increase revenue and reduce costs. Group reservations are an essential aspect of revenue management that has not been extensively studied in the literature. This study aimed to investigate the value of agencies to airlines and explore the dynamics that agencies create while doing business with airlines by using Fuzzy Cognitive Mapping. The results showed that the sensitivity of agencies to fare offers and their annual reservation volume are crucial factors affecting the value of the agency to the airline. The findings of this study can provide insights into pricing strategies for airlines and help them to maximize their revenue.

For further researches, it is indicated to construct a performance evaluation system for the airline agencies using the most important variables that are found as annual reservation volume, the ratio of the airline to the total business volume of the agency, and the agency's sensitivity to the fare offers/being very negotiable. That performance system could be used as a basis for determining discount strategies.

References

1. Agrawal, A.: Sustainability of airlines in India with Covid-19: challenges ahead and possible way-outs. J. Reven. Pricing Manage. **20**(4), 457–472 (2021)
2. Belobaba, P.P.: Did LCCs save airline revenue management? J. Reven. Pricing Manage. **10**, 19–22 (2011)
3. Chandana, S., Leung, H., Levy, J.: Disaster management model based on modified fuzzy cognitive maps. In: 2007 IEEE International Conference on Systems, Man and Cybernetics, pp. 392–397. IEEE (2007)
4. Eurocontrol: Think Paper No. 8: Impact of COVID-19 on European Aviation in 2020 and Outlook 2021, pp. 1–12 (2021). https://www.eurocontrol.int/publication/what-covid19-did-european-aviation-2020-outlook-2021
5. Kadaifci, C., Sari, I.U.: Examination of Factors Affecting Customer Oriented Product Design in Automotive Industry. Customer Oriented Product Design: Intelligent and Fuzzy Techniques, pp. 451–467 (2020)
6. Kosko, B.: Fuzzy cognitive maps. Int. J. Man Mach. Stud. **24**, 65–75 (1986)
7. Lee, J.: Effects of operational performance on financial performance. Manage. Sci. Lett. **9**(1), 25–32 (2019)
8. O'Connell, J.F.: The Routledge Companion to Air Transport Management. In: Halpern, N., Graham, A. (eds.) (2018)

9. Raliff, R., Vinod, B.: Airline princing and revenue management: a future outlook. J. Reven. Pricing Manage. **4**(3), 302 (2005)
10. Venkataraman, S.V., Kaushik, A., Mishra, S.: A network group booking model for airline revenue management. J. Model. Manage. (2021)
11. Wildenberg, M., et al.: Linking thoughts to flows-Fuzzy cognitive mapping as tool for integrated landscape modelling. In: Proceedings of the 2010 International Conference on Integrative Landscape Modelling—Linking Environmental, Social and Computer Sciences, Montpellier, February 3–5 (2010)
12. Yang, H.: Airlines' futures. J. Reven. Pricing Manage. **6**, 309–311 (2007)

Procurement Performance Metric Selection Using SCOR Modelling and Scenario-Based Fuzzy Cognitive Map

Ayça Maden[1(✉)] and S. Emre Alptekin[2]

[1] Management Engineering Department, Beykent University, Istanbul, Turkey
aycamaden@beykent.com
[2] Industrial Engineering Department, Galatasaray University, 34357 Ortaköy–Istanbul, Turkey
ealptekin@gsu.edu.tr

Abstract. The present study investigates a company's procurement process with a weak performance measurement and management system in its supply chain. The challenge in evaluating the customer-specific engineer-to-order projects in the case company prompted the use of SCOR modelling as a standardization tool to analyze the supply chain processes and metrics. The selection of performance metrics is a dynamic process that needs to be updated based on changing company needs. However, the complex relationships between metrics make it difficult for stakeholders to predict their impact on results. To address this, we utilized scenario-based Fuzzy Cognitive Maps to prioritize important SCOR metrics and fuzzy numbers in our evaluations to account for the inherent fuzziness in procurement processes. To the best of our knowledge, this is the first study to evaluate the trade-offs between procurement performance metrics using SCOR modelling.

Keywords: Procurement · Fuzzy Cognitive Map · SCOR · Performance Measurement

1 Introduction

One of the open issues of procurement processes is the ambiguity in the relationships between KPIs (Key Performance Indicators) of the procurement process (Abolbashari et al. 2018). SCOR (Supply Chain Operations Reference) ensures methodology, diagnostic and benchmarking instruments that support businesses to make striking and fast developments in supply chain processes (APICS 2015). However, SCOR KPIs are complex to manage without a systematic approach. There is not enough methodology to manage SCOR KPIs in literature, which is one of the most significant troubles in supply chain management (Akkawuttiwanich & Yenradee 2018; Estampe et al. 2013). A company cannot satisfy all of its metrics, so they have to trade off among the KPIs (Akkawuttiwanich & Yenradee 2018). In this work, we focused on using a standardized process, metric and quality definitions. Specifically, we used SCOR modelling since the case company has a vague and imprecise process and metric definitions.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 465–471, 2023.
https://doi.org/10.1007/978-3-031-39777-6_55

SCOR modelling enabled us to analyse the case company both for the as-is and to-be processes of the case company. In addition, we used Fuzzy Cognitive Maps to focus on important SCOR metrics since there are trade-offs between SCOR performance metrics.

In this paper, Sect. 1 holds the Introduction, Sect. 2 holds SCOR Metrics, Sect. 3 holds Methodology, Sect. 4 holds the Case study, and Sect. 5 holds the Conclusion.

2 SCOR Metrics

Metrics of the SCOR are arranged in order of rank (APICS 2015). For example, level 1 metrics can help measure a company's accomplishments in satisfying its aims regarding the competitive area (APICS 2015). The SCOR Level-1 metrics can be seen in Table 1,

Table 1. The SCOR Level-1 metrics

Performance Attribute	Level-1 Strategic Metrics
Reliability	Perfect Order Fulfillment
Responsiveness	Order Fulfillment Cycle Time
Agility	Upside Supply Chain Flexibility
	Upside Supply Chain Adaptability
	Downside Supply Chain Adaptability
	Overall Value at Risk
Cost	Total Cost to Serve
Asset Management Efficiency	Cash-to-Cash Cycle Time
	Return on Supply Chain Fixed Assets
	Return on Working Capital

Level 2 metrics can clarify performance situations for Level 1 (APICS 2015). Level 3 performance metrics are concerned with SCOR Level 3 processes (Cheng et al. 2010). Hence, we can choose the supply chain performance metrics in a process-based approach after the SCOR Level 3 modelling (Cheng et al. 2010).

3 Methodology

3.1 Establishing FCM Structure

1. Initial concept values are set up. The normalized weight vector is determined as the initial concept values, as in Eq. (1) (Baykasoğlu & Gölcük 2015):

$$C_{MA_1}^{t=0} \quad \ldots \quad C_{MA_p}^{t=0} \quad \ldots \quad C_{MA_n}^{t=0} = (w_1 \quad \ldots \quad w_p \quad \ldots \quad w_n) \tag{1}$$

where $C_{MA_i}^{t=0}$ is the concept value of the ith main attribute when t = 0.

2. Fuzzy influence matrices are elicited.
3. Fuzzy influence matrices, which are taken from a group of decision-makers are aggregated by utilizing Eq. (2) (Baykasoğlu & Gölcük 2015):

$$\tilde{e}_{ji}^{MA} = \frac{\sum_{u=1}^{s} \tilde{e}_{jiu}^{MA}}{S} \tag{2}$$

where $\tilde{e}_{ji}^{MA}$ shows the aggregated fuzzy influence matrix of the main attributes
4. Finally, the aggregated crisp influence matrix e_{ji}^{MA} is achieved.

3.2 Simulating FCMs and Obtaining the Final Weights

After establishing the FCM structure, the steps of the FCMs to obtain final weights in the methodology can be followed (Baykasoğlu & Gölcük 2015);

1. Specifying the parameters of the activation function,
2. Running of FCMs. FCM simulation is conducted to obtain dynamic behaviour based on Eq. (3):

$$C_{MA_i}^{t+1} = \mathrm{f}(C_{MA_i}^{t} + \sum\nolimits_{j\,=\,1}^{n} e_{ji}^{MA} \times C_{MA_j}^{t}) \tag{3}$$
$$j = 1$$

where $\mathrm{C}_{\mathrm{MA}_i}^{\mathrm{t}}$ is the concept value of the *i*th main attribute at time *t*,
3. Normalizing the steady state weights.

$$w_i = \frac{C_{MA_i}}{\sum_{i=1}^{n} C_{MA_i}}, \tag{4}$$

The final crisp weights are represented by,

$$\mathrm{I}_{\mathrm{MA}}\,(MA_1 MA_2 \ldots MA_n), \quad (w_1 \;\; w_{2\ldots} \;\; w_n), \;\; \text{where} \; \sum\nolimits_{i=1}^{n} w_i = 1.$$

In this paper, we used Triangular Fuzzy Numbers (TFN). For the FCM method, we preferred the sigmoidal FCMs. We used FCMapper to make the Fuzzy Cognitive Map analysis. Developed by Michael Bachhofer and Martin Wildenberg, FCMapper is an open-licensed excel macro (FCMapper). FCM_Scenarios screen allows simulation of various scenarios. The GUI interface of the FCMapper excel tool can be seen in Fig. 1.

4 Case Study

We made a case study application in a company producing customer-specific project-based products. In the case company, different departments use different performance calculations for the same processes. In addition, project-based working environments necessitated keeping pace with the changing working conditions of the company, especially in the procurement process. According to the company strategy, this study selected supply chain performance metrics from the SCOR Level 3 source engineer-to-order product process. The case company has 4 Level 3 sub-processes of source engineer-to-order

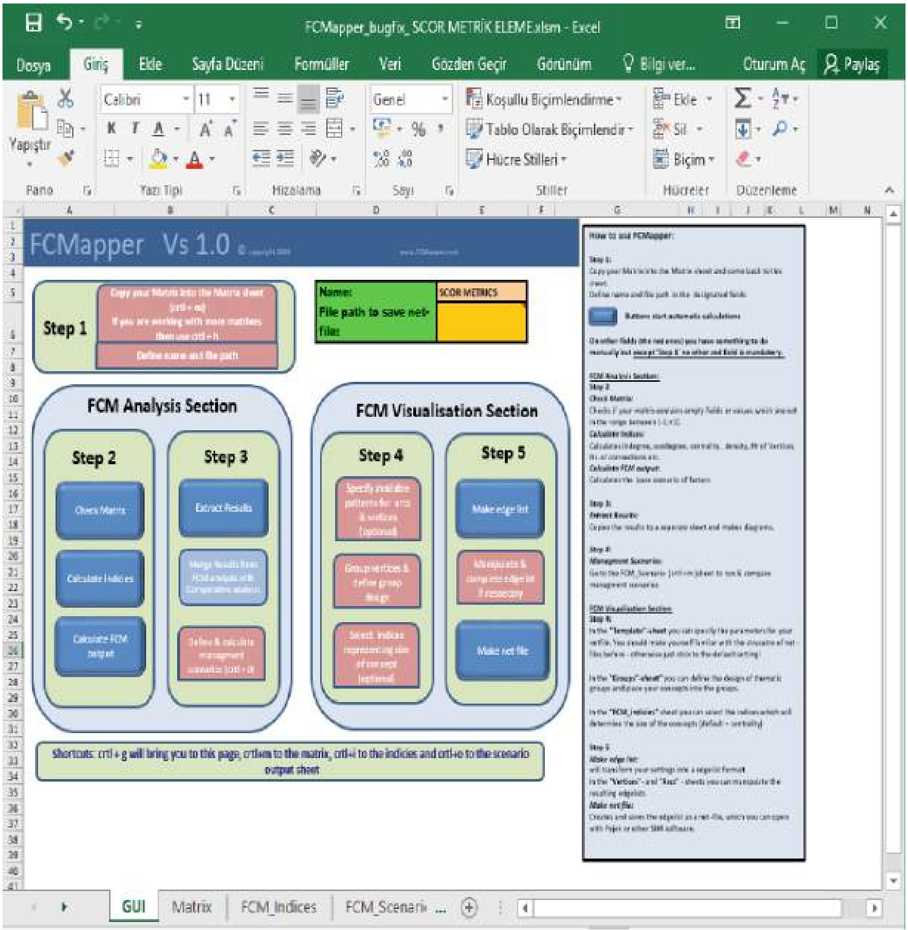


Fig. 1. GUI interface of the FCMapper (http://www.fcmappers.net/joomla/)

product: Identify Sources of Supply (sS3.1), Select Final Supplier, Negotiate (sS3.2), Receive Product (sS3.4) and Verify Product (sS3.5). Related to the company strategy of cost, responsiveness, reliability, and agility, there are 24 SCOR Level 3 performance metrics in the sS3, sS3.1, sS3.2, sS3.4, and sS3.5 processes of the SCOR model. The top priority SCOR metrics of the case company were determined after prioritizing and eliminating these metrics using Fuzzy Cognitive Maps. Depending on the company's specific processes and strategy (cost, time, agility, and reliability), 24 SCOR performance metrics in the source engineer-to-order product can be seen in Table 2,

According to the company stakeholders, these 24 metrics are measurable, understandable, and unambiguous. We reduced 24 metrics using the Fuzzy Cognitive Maps method, which enabled us to have a sufficient number of metrics without loss of information. We chose the average value as the threshold value to focus on more critical metrics

Table 2. SCOR Level 3 performance metrics

sS3	AG.3.9 Additional source volumes obtained in 30 days
	AG.3.40 Current Purchase Order Cycle Times
	AG.3.42 Current Source Volume
	AG.3.46 Demand sourcing-supplier constraints
	CO.3.005 Sourcing Labour Cost
	CO.3.006 Sourcing Automation Cost
	CO.3.007 Sourcing Property, Plant and Equipment Cost
	CO.3.008 Sourcing GRC, Inventory, and Overhead Cost
	CO.3.009 Purchased Materials Cost
	CO.3.010 Material Transportation Cost
	CO.3.011 Material Customs, Duties, Taxes and Tariffs Cost
	CO.3.012 Material Risk and Compliance Cost
sS3.1	RS.3.35 Identify Sources of Supply Cycle Time
sS3.2	RL.3.17% of suppliers with an EMS or ISO 14001 certification
	RS.3.125 Select Supplier and Negotiate Cycle Time
	RL.3.18% Orders/ Lines Processed Complete
	RL.3.20% Orders/ Lines Received On-Time to Demand Requirement
sS3.4	RL.3.22% Orders/ lines received with correct packaging
	RL.3.23% Orders/ Lines Received with Correct Shipping Documents
	RS.3.113 Receiving Product Cycle Time
	RL.3.19% Orders/ Lines Received Defect Free
sS3.5	RL.3.21% Orders/ lines received with the correct content
	RL.3.24% Orders/lines received damage-free
	RS.3.140 Verify Product Cycle Time

according to their weights. However, scenario analysis in Fuzzy Cognitive Map allowed us to include different metrics below the average threshold value in our calculations.

For initial concept values of the Fuzzy Cognitive Map method, we used aggregated evaluations of three experts from the case company. Here, the concepts are SCOR Level 3 metrics. Then, we defuzzified the fuzzy influence matrix of the metrics using the centroid formula, and we found aggregated crisp values of the decision makers' evaluations using the average operator. In this study, the sigmoid function was used since it gives values between 0 and 1. To reach the long-term influences among the SCOR metrics, aggregated crisp values of the experts were entered into FCMapper software, and steady-state values were obtained in 10 iterations. After finding the steady-state values, we calculated the normalized weights of the metrics by dividing each metric value by the summation of the steady-state values.

Then, we chose the metrics equal to or above the average of the normalized 24 metrics; 0,0416 value (1/24); to be used for analysis. The determined metrics are AG.3.46, CO.3.006, CO.3.012, RS.3.35, RS.3.125, RL.3.19. Using the same aggregated values at the beginning of the FCM method, selected six metrics were activated by setting the other metrics' values to 0 in FCMapper software. Thus, relations between these six metrics were the same as in the previously applied Fuzzy Cognitive Map. To reach the long-term influences between the six SCOR Level 3 metrics, steady-state values are calculated, and we reached the steady-state values in 13 iterations as in Table 3. In Table 3, columns indicate SCOR metrics while rows indicate iterations.

Table 3. Steady-state concept weights

	1	2	3	4	5	6
1	0,599	0,786	0,731	0,731	0,832	0,832
2	0,525	0,732	0,664	0,675	0,781	0,767
3	0,514	0,717	0,650	0,663	0,766	0,747
4	0,514	0,713	0,647	0,660	0,762	0,742
5	0,514	0,713	0,647	0,659	0,761	0,741
6	0,514	0,712	0,647	0,659	0,761	0,741
7	0,514	0,712	0,647	0,659	0,761	0,741
8	0,515	0,712	0,647	0,659	0,761	0,741
9	0,515	0,712	0,647	0,659	0,761	0,741
10	0,515	0,712	0,647	0,659	0,761	0,741
11	0,515	0,712	0,647	0,659	0,761	0,741
12	0,515	0,712	0,647	0,659	0,761	0,741
13	0,515	0,712	0,647	0,659	0,761	0,741

After finding the steady-state values, we found the normalized weights of the metrics divided by the summation of the steady-state values. The final normalized weights of the six metrics can be seen in Table 4,

Table 4. Normalized weights of the concepts

1	2	3	4	5	6
0,128	0,177	0,160	0,163	0,189	0,184

Finally, normalized weights of the concepts enabled us the importance ranking of the concepts, which are RS.3.125, RL.3.19, CO.3.006, RS.3.35, CO.3.012, and AG.3.46.

5 Conclusion

Businesses and supply chains that are weak in terms of performance measurement and management fail to achieve their goals in their business processes, which is among the topic of our case supply chain. Because of huge complications in customer-specific engineer-to-order projects, it is difficult to find project-based performance metrics and focus on the most problematic ones. Procurement stakeholders tend to evaluate performance metrics individually, and they cannot simultaneously analyse all the problematic metrics. In our study, there is uncertainty because of the subjectivity in human judgments. The increase or decrease in the calculated SCOR performance metrics was determined by subjective judgments. Using the Fuzzy Cognitive Map, we reached the significance ranking of the concepts, which are RS.3.125-Select Supplier and Negotiate Cycle Time, RL.3.19-Percent of the Orders/Lines Received Defect Free, CO.3.006-Sourcing Automation Cost, RS.3.35-Identify Sources of Supply Cycle Time, CO.3.012-Material Risk, and Compliance Cost, and AG.3.46-Demand sourcing-supplier constraints. In our future study, the degree of increase or decrease of the metrics from their target values will be examined in detail. Our future work will use data-driven modelling of the performance metrics focusing on more accurate tools to find problematic metrics.

Acknowledgements. This research has been financially supported by Galatasaray University Research Fund, with the project number FBA-2022-1091.

References

Abolbashari, M.H., Chang, E., Hussain, O.K., Saberi, M.: Smart buyer: a Bayesian network modelling approach for measuring and improving procurement performance in organizations. Knowl. Based Syst. **142**, 127–148 (2018)

APICS. Supply Chain Operations Reference Model, Revision 11.0 (2015)

Akkawuttiwanich, P., Yenradee, P.: Fuzzy QFD approach for managing SCOR performance indicators. Comput. Ind. Eng. **122**, 189–201 (2018)

Estampe, D., Lamouri, S., Paris, J.L., Brahim-Djelloul, S.: A framework for analysing supply chain performance evaluation models. Int. J. Prod. Econ. **142**(2), 247–258 (2013)

Baykasoğlu, A., Gölcük, İ: Development of a novel multiple-attribute decision making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS. Inf. Sci. **301**, 75–98 (2015)

Cheng, J.C., Law, K.H., Bjornsson, H., Jones, A., Sriram, R.D.: Modeling and monitoring of construction supply chains. Adv. Eng. Inform. **24**(4), 435–455 (2010)

FCMapper, FCMapper - Fuzzy Cognitive Mapping Software Solution. http://www.fcmappers.net/joomla/

Joint Application of the FCM and Structural Shift Detection Algorithms in Monitoring

Zinaida K. Avdeeva(✉), Elena A. Grebenyuk, and Svetlana V. Kovriga

V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Russia
zinaida.avdeeva@gmail.com

Abstract. The paper discusses features of joint application of fuzzy cognitive maps (FCM) and algorithms for detecting structural shifts in the combined monitoring of prices in commodity markets, including (1) digital monitoring to detect structural shifts in the time series of the object of observation, (2) situational monitoring of the environment of its functioning to structure the current situation and generate signals about its changes, (3) information exchange control and formation of output aggregated signals. Integration of such methods is based on digital processing of analytical and expert information, which makes it possible to take into account information about changes in the environment of object functioning (occurred, expected or possible) and improve the quality of detection, form variants of possible changes in object state; and thus, ultimately, to improve the efficiency of solving target tasks of analysis and forecasting of controlled object.

Keywords: Commodity markets · Monitoring · Structural shifts · Nonstationary time series · Sequential analysis · Fuzzy cognitive map · Scenario simulation

1 Introduction

The current rate and frequency of changes in the conditions of functioning and development of socio-economic, organizational, technical and other objects are being pointedly put the problem of medium- and long-term forecasting in conditions of the emergence of significant events of the external environment that determine the nature and direction of process changes. Accordingly, monitoring and forecasting of the object and its functioning conditions should be based on methods that take into account all available qualitative and quantitative information. In the literature one can find such integrated methods and models for specific areas in the field of financial market, for some objects of commodity markets [1–4]. In addition, there are various approaches to quantitative forecasting based on the addition to time series of data derived from qualitative information [5], such as expert opinions on the predicted change in value, assessment of market sentiment through analysis of tone, news activity and interest in topics. Nevertheless, predicting several steps ahead over long time horizons with the changing external environment remains a challenge, as the level of forecast uncertainty increases as the horizon increases. Due to changes in the macro- and business- environment leading to structural shifts on the forecast horizon (within a year), the forecast model built on the historical

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 472–479, 2023.
https://doi.org/10.1007/978-3-031-39777-6_56

interval ceases to describe the dynamics of the forecast price and becomes unsuitable. The relevance of the identified problem lies in the increasing need for forecasts of the considered class of objects (macroeconomic indicators, for example, GDP, prices of commodities, such as scrap metal, oil and others, shares, exchange rates), the quality and reliability of which can be improved by including information on the impact of external environment events on the forecast indicator in a quantitative model.

Previously, we developed a scheme of combined target indicator forecasting using the example of commodity market price [6], which consists in the construction of a fuzzy cognitive map (FCM) (see an example on Fig. 1) of the influence of the external environment on the dynamics of the forecast object, the construction of a forecast model for the year ahead, the monitoring of the external environment and time series in the current observation mode, the correction of the model on the forecast horizon based on monitoring results. The main task of combined monitoring of non-stationary processes is to detect proactive signals of structural shifts by situational monitoring algorithms from analytical information. For timely replacement and adjustment of the prediction model, the combined monitoring algorithm [7] was developed, which includes three units: 1) unit to detect structural shifts in time series based on [12, 13], 2) unit to track changes in the external environment (situational monitoring) based on FCM-modeling of incoming information, 3) control unit, performing information interaction between digital and situation monitoring units, formation of input data for them and aggregated output signal.

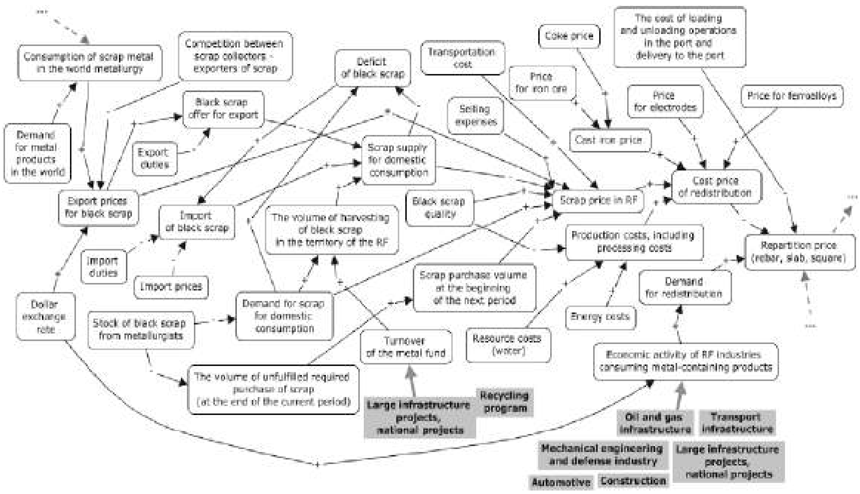


Fig. 1. The sign FCM (fragment) of mutual influences of factors associated with the price of black scrap (gray dotted arrows with ellipsis indicate the boundaries of the selected fragment).

This paper discusses the peculiarities of joint application of fuzzy cognitive maps to model the incoming analytical and news information [7] and the method of change detection [12] in the mode of current observations on the example of price monitoring in the commodity market of metals. One of the important features of commodity markets that we take into account when building a monitoring scheme is the fact that prices for various price goods can follow general trends. Such price trends are called co-movement [3, 4]. Changes in the price of commodities are due to the co-movement of group prices and prices of commodities that are part of the same value chain, the impact of financial and economic indicators, and the influence of environmental events. The Fig. 1 shows a fragment simplify version (without weights) of the fuzzy cognitive map (FCM) of the situation on the scrap metal market, respectively the model includes factors and causal relationships between object parameters and external factors that determine the dynamics of prices in the market. FCM include the key parameters of the value chain: prices for primary and secondary raw materials– the cost of production – demand for final products. The keywords associated with the FCM factors are applied to collecting information and data for monitoring. In particular, two sets are generated for the case study: analytical reports on market conditions from chosen information sources and time series of scrap metal market prices and related markets.

2 Problem Statement. Methods and Algorithms

We applied the developed algorithms to solve the following problem. The objects of monitoring are series Y and groups of time series $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$, $\hat{Y}4$, which, under the influence of an unstable external environment and the activities of market participants, changes (structural shifts) of one of the following types occur: changes in the drift of the process; changes in the type of process: from a process with a stochastic trend to a stationary one or from a stationary one to a process with a stochastic trend; changes in the connections (Granger causality) between the series of the group $\hat{Y}i$ and the series Y.

For the considering case Y are series of prices for raw materials, changes in which may occur under the influence of the following groups of time series $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$, $\hat{Y}4$ (system-forming factors): factors influencing the dynamics of the monitored group of goods and the groups of goods included in the value chain of the controlled group; factors of the world economy and factors reflecting the state policy of influencing commodity markets;

The level of influence of each of the groups $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$, $\hat{Y}4$ on the monitoring object Y is set by its weight, which characterizes the contribution of the group to the change of Y.

It is required to find for each group $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$, $\hat{Y}4$ the moments of the points of changes in the totality of the series included in it and the weight of these groups influence on the monitoring object Y. It is required to build a monitoring system that includes digital monitoring of price time series and situation monitoring to detect the changes listed above in a number of Y and in groups $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$, $\hat{Y}4$.

2.1 Methods of Processing and Generating Signals About the State Of the External Environment

To describe the situation, for example for the commodity market in question, we have defined the FCM of the form K_f (X, A, f), in which $X = (x_1, \ldots, x_n)$ is a set of factors of the situation S; $A = [a_{ij}]$ is the $N \times N$ matrix of factors mutual influence, where $a_{ij} \in$

$[-1; 1]$ is the weight of influence of factor x_i on the factor x_j at a linguistic scale; f is a function that defines the rule of factor value change. The set of factors X are divided into classes according to the belonging of factors to the system-forming (SF) groups $\{X^{Y_i}\}$, For our example, it is accordingly: $\hat{Y}1$- scrap metal, $\hat{Y}2$- secondary raw materials, $\hat{Y}3$- finished metal products for metal products;$\hat{Y}4$ -macroeconomic indicators.

At least once a month, analytical messages on market conditions are received at the input of the situation-monitoring unit. Each message is processed by means of FCM- keywords and accordingly presented as a vector $X^d_{isor1}(x^d_i(0); oc^d)$, where $x^d_i(0) \in [-1; 1]$- the observed value on the linguistic scale for the factor mentioned in the message, for the factor, xi; oc^d- associated event (if any). From the information obtained, a vector $X^d(0)$ is constructed at time d and for a given target parameter y. Thus, the state of the situation at any discrete point of time t ≥ 0 [6] expressed as follows

$$X(t+1) = Q(t)X(0) + Q(t)G(0), \tag{1}$$

where $Q(t) = E^N + A + A^2 + \ldots + A^t = (E_N - A)^{-1}$ - the matrix of integral influences, in which q_{ij}-elements estimate all direct and indirect influences between factor i and j. [6].

Then for each observed situation $X^d(0)$ we determine the prediction y* using the formula (1) and obtain the following estimates: an estimate of the dynamics $r_y = sign\left(y^* - y^0\right)$ and vector of significance of the SF-groups $C^d = \{q^{Y_i}_{S^d_{x_k}}\}$, where $q^{Y_i}_{S^d_{x_k}}$ are calculated as a significance of the Y_i for some scenario S^d_y defined as

$$q^{Y_i}_{S^d_y} = \sum_{k1=1}^{m1}\sum_{j1=1}^{l1} r_{x_{k1}} \times q^{x^{Y_i}_{j1}}_{x_{k1}} + \sum_{j2=1}^{l2} r_{x^{Y_i}_{j2}} \times q^{y^{V_i}}_{x^{M_i}_{j2}}, \tag{2}$$

where the first addend is the cumulative influence of the factors from X^{Y_i} on causal factors related to a model Yi on target indicator, y, and the second addend is the cumulative influence of the causal factors related to $\hat{Y}$i on target indicator y.

The FCM generates trend change signals $In^d(y) = (\Delta^d; \overrightarrow{C}^d_j)$, $\Delta \in \{-1, 0, 1\}$–the trend change $\Delta = sign(y^d - y^d(0))$, $\overrightarrow{C}^d = [c_j]$– weights of significance of the influence of SF- groups of systemic factors $\hat{Y}$j. These estimates are transmitted to the control unit to confirm the detected qualitative signals at the quantitative process data level. Another scenario for the unit is to run a procedure to explain from the observed information the signals detected by the digital monitoring of changes in process properties.

2.2 Methods of Processing and Generating Signals About Changes in Object Properties

We solve the problem of detecting structural shifts both for an individual time series and for a group, all the series of which are subject to the influence of the same factors, using the methods of sequential analysis [8–11].

Upon receipt of the next observation, the sequential algorithm decides on the presence or absence of a change in the process or determines its type. The criteria for the quality of detection are the minimization of the delay, subject to the restriction on the probability of a false alarm. This approach makes it possible to detect deviations from a stable state at an early stage, and prevent the object from going into "emergency mode".

The digital monitoring unit performs the following functions: monitoring of time series Y, and monitoring of groups $\hat{Y}i$ and $\hat{Y}j$. The input of the digital monitoring unit receives the following signals: the current values of Y and values of time series of group $\hat{Y}i$: $Y \in \hat{Y}i$, and time series of group $\hat{Y}j$ with the maximum weight.

In the algorithm we use for monitoring the time series Y, the input information is an independent random sequence [12, 13]. The series of prices in the commodity market are not independent sequences; their current values depend on their past values. Therefore, we approximate these series with parametric models and feed the "residuals" – the difference between the values of the series and its model - to the input of the algorithm. In the model of a series, we include its past values and the past values of the series from Ŷi, which are Granger causal for the series Y, and construct it by time series of differences.

The monitoring algorithm of time series Y is an ensemble of detection algorithms, the number of which coincides with the number of models constructed and depends on the number of series in the $\hat{Y}$i group, which are causal by Grainger for the series Y.

If for a given series in the group $\hat{Y}$i there are no series that are Granger causal for it, then we describe it with a parametric model.

The monitoring algorithm of the group $\hat{Y}$i, i = 1, 2, 3 … k we build from the residuals of the time series. To detect points of change, we use a Bayesian Online Changepoint Detection [14] and an algorithm for detecting changes in the level of a non-stationary process [12, 13].

2.3 Control Unit

The control unit organizes the exchange of signals between the situation and digital monitoring units: sends a request for confirmation of the situation monitoring signal to the digital block; sends a request for confirmation of information about a change detected by digital monitoring to the situation monitoring block; generates an aggregated signal about the detected changes, and an assessment of the degree of confidence in the correctness of the signal.

Before starting monitoring, we feed the algorithm as input 1) a list of groups of system-forming factors and weights that reflect the significance of the influence of each SF-group on the monitoring object Y, 2) time series distributed over groups $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$,…, $\hat{Y}n$, when distributing time series from the database into SF-groups.

The situation monitoring unit receives information about expert-significant events and changes in the system-forming factors of the external environment, based on which it determines the qualitative dynamics of object Y, $In^d(y) = (\Delta^d;\ \overrightarrow{C}_j^d)$, such as signals about a change in the Y trend or a change in the weights of the SF-groups influence.

The digital monitoring unit receives time series values distributed across SF- groups, requests from the control unit to change search parameters in order to refine signals of situation monitoring.

The block generates two types of output signals: about changes in the monitoring object Y and about the occurrence of points of change in groups $\hat{Y}1$, $\hat{Y}2$, $\hat{Y}3$, $\hat{Y}n$.

The control unit checks the consistency of signals about changes in the monitoring object Y and in the group $\hat{Y}I$ with the maximum weight and generates an output signal or sends a request to the digital monitoring unit about changes in search parameters or to the situation monitoring unit. In addition to aggregated signals, the control unit also sends the scenarios for the development of the situation formed as a result of situation monitoring to the unit for solving target problems of process analysis, for example, forming long-term forecasts of its behaviour.

3 Description of the Experiment

Figure 2 shows signals about process changes in 2019 generated by the joint processing of the information and time series of commodity market parameters. The *Trend change signal* line indicates the expected trend change directions (rising, falling) of the Y price range and its intensity in conventional units. The *Signal of significance change* lines indicate the weights of SF-groups. The *Task of confirmation* line contains information about the direction of communication between digital and situation monitoring signals. Directional vertical arrows indicate the direction of the signal trend.

SITUATION MONITORING							
	1	2	3	4	5	6	7
Observation period	01'19	02'19	03'19	04'19	05'19	06'19	07'19
Trend change signal In($\tilde{y}$)	1 Δ=0.3	0	−1 Δ= −0.1	0	1 Δ=0.1	−1 Δ= −0.1	1 Δ=0.3
Adittional signal				Inf (Yune, decrese, if <agreem. Rus-Ger=0>			
Signal of significance chance		1	1	0	1	1	1
M_i; C_i^0	C_i^1	C_i^2	C_i^3	C_i^4	C_i^5	C_i^6	C_i^7
SF1; 0,4	0,7	0,6	0,3		0,3	0,7	0,4
SF2; 0,3		0,4	0,2		0,2		0,6
SF3; 0,2			0,5		0,5	0,3	
SF4; 0,1	0,3						
M_i - model	C_i - weight of the model						
SUPERVISOR							
Observation period	01'19	02'19	03'19	04'19	05'19	06'19	07'19
Task of confirmation	**SM→DM**		**SM→DM**	**DM→SM**	**SM→DM**	**DM→SM**	**DM→SM**
Monitoring scenario						Inf (Yune, decrese, if <agreem. Rus-Ger=0>	
Result	confirm		confirm	confirm	not	confirm	confirm
DIGITAL MONITORING DIGITAL MONITORING of SF1-GROUP							
Observation period	01'19	02'19	03'19	04'19	05'19	06'19	07'19
Data confirmation	25.01		29.03	12.04		07.06	05.07
Digital signal	End of decreasing		Begin of decreasing	Stabilization		Begin of decreasing	End of decreasing

Fig. 2. Summary of monitoring results

Figure 3 shows a snapshot of the series dynamics of the $Y3$-product group and the $Y1$-scrap metal group. At the bottom of the figure, there is a graph of the monitored indicator. Moments of SF-groups trend changes are marked by arrows: downward trend - downward arrow, upward trend - upward arrow. The digital monitoring algorithm revealed a change for the product group from stable to downtrend on 02.08.2019, for the SF- scrap metal group a change from rising to falling on 22.03.2019 and from falling to rising on 27.09.2019. The bottom part of the figure shows the moments of change of the monitored indicator, whose dynamics were influenced by the dynamics of the products, the explanation of which is formed in the situational monitoring unit. Moments of trend change of the monitored parameter indicated by the double arrows (June 7 and July 26). It should be noted that it did not react to the trend change to the rising 27.09.2023 in the SF scrap metal group.

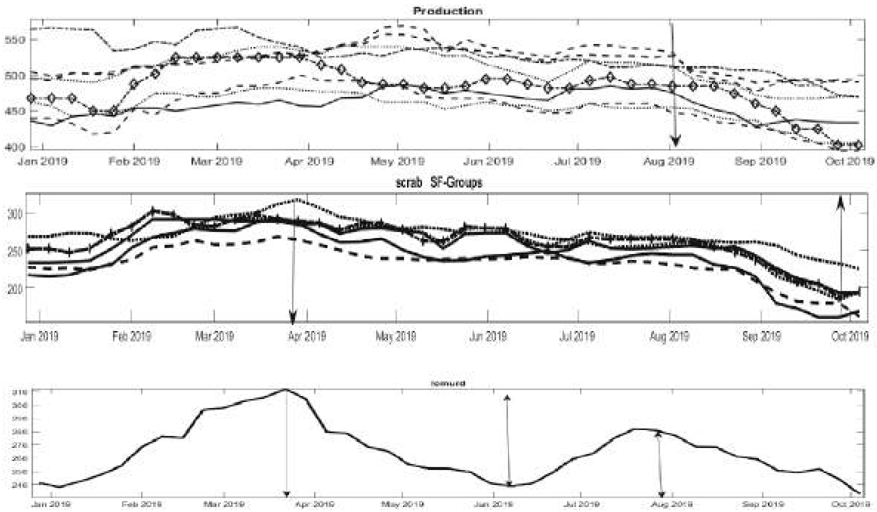


Fig. 3. Summary of digital monitoring results

As can be seen from the results shown in Fig. 2, the situation-monitoring signal is ahead of the digital monitoring signal. Out of 4 signals detected at the beginning of the period by market information, 3 trend change signals were confirmed by the digital monitoring algorithm one week later. С другой стороны, цифровые сигналы об изменении обнаруженные в апреле, июне и июле были объяснены информацией в блоке ситуационного мониторинга.

4 Conclusions

In a simulation using available data and market information for the period 2019, we have confirmed the hypothesis that situational monitoring algorithms using qualitative information can detect pro-active signals of structural shifts. The joint applying of the two algorithms for detecting changes in processes represented analytical information and time series can be ensured primarily by the correct representation of the system time in the control unit supported the exchange of signals between the situation and digital monitoring. However, the bottleneck of this integration remains the digital processing of analytical information, for both updating the FCM and entering information into the situation-monitoring unit.

Acknowledgements. The work was supported by the Russian Science Foundation (grant 23-21-00455).

References

1. Filimonov, V., Bicchetti, D., Maystre, N., Sornette, D.: Quantification of the high level of endogeneity and of structural regime shifts in commodity markets. J. Int. Money Financ. **42**, 174–192 (2014)
2. Zhang, D., Broadstock, D.C.: Global financial crisis and rising connectedness in the international commodity markets. Int. Rev. Finan. Anal. **68** (2020) https://doi.org/10.1016/j.irfa.2018.08.003
3. Byrne, J.P., Fazio, G., Fiess, N.: Primary commodity prices: co-movements, common factors and fundamentals. J. Dev. Econ. **101**, 16–26 (2013)
4. Liu, P., Qiu, Z., Xu, D.X.: Financial investments and commodity prices. Int. Rev. Financ. (2021). https://doi.org/10.1111/irfi.12361
5. Petropoulos, F., Apiletti, D., et al.: Forecasting: theory and practice. Int. J. Forecast. **38**(3), 705–871 (2022). https://doi.org/10.1016/j.ijforecast.2021.11.001
6. Avdeeva, Z., Grebenyuk, E., Kovriga, S.: Raw material price forecasting on commodity markets: application of expert and quantitative information. Adv. Syst. Sci. Appl. **22**(4), 126–143 (2022)
7. Avdeeva, Z., Grebenyuk, E., Kovriga, S.: Combined monitoring of heterogeneous information for forecasting of the non-stationary process dynamics. In: Proceedings of 15th Multiconferense on Control Problem «IT in Control». SPb, pp. 181–185 (2022)
8. Zeileis, A., Leisch, F., Kleiber, C., Hornik, K.: Monitoring structural change in dynamic econometric models. J. Appl. Economet. **20**(1), 99–121 (2005)
9. Nikiforov, I.V.: A simple recursive algorithm for diagnosis of abrupt changes in random signals. IEEE Trans. Inf. Theory **46**(7), 2740–2746 (2000)
10. Gösmann, J., Kley, T., Dette, H.: A new approach for open-end sequential change point monitoring. J. Time Series Anal. **42**(1) (2019). https://doi.org/10.1111/jtsa.12555
11. Page, E.S.: Continuous inspection schemes. Biometrika **41**(1–2), 100–115 (1954)
12. Grebenyuk, E.A.: Monitoring and identification of structural shifts in processes with a unit root. In: Proceedings of the 13th International Conference Management of large-scale system development (MLSD) (2020). https://doi.org/10.1109/MLSD49919.2020.9247829
13. Grebenyuk, E.A.: Algorithms for detecting changes in the properties of non-stationary processes in the monitoring mode. In: Proceedings of the 14th International Conference Management of Large-Scale System Development (MLSD) (2021). https://doi.org/10.1109/MLSD52249.2021.9600109
14. Adams, R.P., MacKay, D.J.C.: Bayesian online change point detection. ArXiv. University of Cambridge, UK (2007). https://doi.org/10.48550/arXiv.0710.3742. Accessed 15 Jan 2023

Evaluating the Impact of Risk Component in Supply Chain Management Using Intuitionistic Fuzzy Cognitive Map

Hasibe Berfu Öztürk(✉) and Şule Önsel Ekici

Department of Industrial Engineering, Istanbul-Cerrahpasa University, Istanbul, Turkey
berfudemir@beykent.edu.tr, sule.ekici@iuc.edu.tr

Abstract. The growing complexity of supply chains is attributed to a range of factors, including globalization, competitive conditions, growing relationships among supply chain partners, and evolving technology. This phenomenon is associated with increased uncertainties and new risks that adversely affect many firms. Therefore, identifying, assessing, and managing these risks is essential for establishing a profitable, competitive, and sustainable supply chain over the long term. Supply chain risk management (SCRM) has emerged as a critical area of inquiry and practice for managing these risks. The primary objective of this study is to develop a model that predicts supply chain risks. To achieve this objective, the study commences with a comprehensive literature review to identify relevant articles related to the topic of interest. Based on the document coding rules using the identified articles, variables were determined. The identified variables were subsequently transformed into linguistic variables and triangular membership functions using a fuzzy set approach. An intuitionistic fuzzy cognitive map was created using the adjacency matrix obtained from the normalization process. Furthermore, scenario analyses were conducted to identify measures. The resulting model provides a comprehensive framework for predicting supply chain risks and identifying measures to mitigate them. This study advances the state of the art in SCRM research and practice.

Keywords: Supply Chain Risk Management (SCRM) · Document Coding · Intuitionistic Fuzzy Cognitive Map (iFCM) · Scenario Analysis

1 Introduction

The prevailing business trends have resulted in the formation of intricate and constantly changing supply chains. The complexity of goods and services, outsourcing, and globalization has all led to an escalation in the level of supply chain risk [10]. In essence, a supply chain refers to a collection of companies and their associated activities, engaged in the production, delivery, and consumption of a product or service. Consequently, businesses depend on their supply chains to operate efficiently for their continued existence, and each company participates in one or more supply chains [8]. To manage these risks, Supply Chain Risk Management (SCRM) has emerged as a structured and phased

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 480–487, 2023.
https://doi.org/10.1007/978-3-031-39777-6_57

methodology aimed at identifying, evaluating, ranking, minimizing, and monitoring potential supply chain disruptions [5]. With these perspectives in mind, the objective of this study is to develop a model that predicts supply chain risks. Unlike previous studies, this study utilizes the documentary coding technique to determine the SCRM related variables and the interactions among these variables. Furthermore, the key contribution of this study is the creation of a fuzzy cognitive map using document coding, which distinguishes it from other supply chain risk studies in the literature by analyzing the SCRM from a system perspective. The primary aim of this study is to evaluate the impact of risks on the supply chain. For this reason, document coding was initially conducted to identify the concepts that affect the supply chain in this study. Afterwards, the final concepts were determined with the help of expert opinions. In the last stage, an intuitionistic fuzzy cognitive map was generated with the final variables obtained. With scenario analysis, it is aimed to make suggestions regarding the effects of these risks on the supply chain. The structure of this paper is as follows: Sect. 3 presents the methodology and its application. Section 4 presents the experimental results. Finally, in Sect. 5, the conclusions are discussed, along with suggestions for future research.

2 Literature Review

The fuzzy cognitive maps (FCMs) allow for malleable and intricate depictions of the connections between ideas, employing fuzzy logic to denote the degree and direction of these relationships [2]. FCMs are fuzzy directed graphs where nodes represent concepts and edges represent strengths of relationships. The utilization of fuzzy cognitive maps has been observed in various fields within the literature. FCMs have been widely applied to analyze complicated, causal-based systems in terms of modeling, decision making, analysis, prediction, classification, and so on [1]. Dursun and Gümüş [9] proposed a comprehensive supply chain management process configuration that enables an understanding of the relationships between supply chain integration, supply chain strategies, supply chain risk factors, and performance criteria. To take the mutual relationships between the criteria into account, the intuitive fuzzy cognitive map methodology was utilized. In 2020, Purnomo et al. [11] obtained sustainable SCM concepts in higher education laboratories through in-depth interviews and arranged them using the Delphi method. The intuitive fuzzy cognitive map was used to identify the relationships between the concepts. In 2018, Rezaee et al. [12] attempted to pinpoint and prioritize potential failures in the manufacturing process using a multi-stage FCM method and Process Failure Model and Effects Analysis. The aim of the study by Rezaee et al. [12] is to present a model for service supply chain performance evaluation (SSCPE) based on FCMs. Panderi ve Azar [13] recommended a model for service supply chain performance evaluation (SSCPE) based on fuzzy cognitive mapping (FCM). Shokouhyar et al. [23] presented a smart, sustainable supply chain practices structure on the basis of the relational view. Kosko [2] provides extensive research on FCM methodology offering a comprehensive analysis of the literature.

3 Methodology and Application

This research aims to address a fundamental research question pertaining to the impact of risks on the supply chain and the interconnectedness among these risks. By identifying the risk factors that significantly affect the supply chain, the objective is to provide recommendations based on these factors. The first step in creating a cognitive map involves identifying causal concepts and linking them to effect concepts [6]. Documentary coding is a technique used in this process, which allows for the analysis of ideas advocated by decision-makers [7]. Therefore, a comprehensive literature review was conducted in this study to identify relevant articles for the purpose of determining variables and their interconnections by considering document coding rules with the help of the identified articles. Upon completion of the document coding process for this study, it was observed that the number of variables was greater than expected, which could potentially complicate the creation of the fuzzy cognitive map. To accomplish this, the opinions of two experts were sought. The experts reviewed the accuracy of the coding and provided their insights for the classification of the variables. Consequently, fifteen key variables (C1: Inventory Risk, C2: Strategy Risk, C3: Supply and Demand Risk, C4: Production Risk, C5: Product Risk, C6: Distribution Risk, C7: Financial Risk, C8: Employee Risk, C9: Political Risk, C10: Operational Risk, C11: Information Risk, C12: Environmental Risk, C13: Supply Chain Performance, C14: Supply Risk, C15: Supply Chain Risk) were identified. The sub-risk groups and indicators of key risk concepts have been identified. For instance, there are three sub-risks related to distribution risk, namely delay risk, logistics risk, and facility risk. As for logistics risk, there are indicators such as delivery delay, losses in distribution, and lack of transportation equipment capacity.

The relations between the variables are calculated by using intuitionistic fuzzy sets. They have an advantage in allowing for modeling of certain uncertainties on the membership. Intuitionistic fuzzy sets offer an advantage in modeling certain uncertainties on the membership function [16]. The intuitionistic fuzzy set assigns a membership degree, a non-membership degree, and a hesitancy degree to each element [14]. Atanassov [15] provides a detailed explanation on intuitionistic fuzzy set. Fuzzy Set Theory [3] provides a means for describing data using imprecise, linguistic terms. Fuzzy Set Theory enables the formal definition of such terms within mathematical logic [17]. Table 1 presents the linguistic variables determined for this study.

The center of gravity method (COG) was used to determine the weights of the variables. To determine the final weight of the relationships belonging to multiple effects and conceptual variables, the weighted average method was used to calculate the degrees of membership, non-membership, and hesitancy separately. The center of gravity method (COG) is the most preferred approach in meeting this requirement. One of the benefits of this method is that the degree of overlap of some membership functions does not affect the result of clarification if the membership functions in the output linguistic categories are symetric [18]. In this method, each membership function is weighted according to the maximum membership value [19].

The adjacency matrix was obtained as a result of the calculations performed for all variables. With the help of the adjacency matrix, a fuzzy cognitive map was generated.

Table 1. Linguistic Scale.

Linguistic Scala	Triangular Fuzyy Number
Very Strong	(0.90, 0.05, 0.05)
Strong	(0.75, 0.20, 0.05)
Medium	(0.50, 0.40, 0.10)
Low	(0.25, 0.60, 0.15)
Very Low	(0.10, 0.80, 0.10)

The simulation process was halted after 20 iterations, when equilibrium was achieved. To interpret the variables, the input, output, and centrality degrees of all variables were computed using FCMMapper.

4 Results

The resulting FCM is shown in Fig. 1 and it is analyzed in terms of centrality and density. Centrality denotes all the links between a variable and the rest of the system, revealing the contribution of the concept to the system [21]. In this study, financial risk is the most critical variable with a total rating of 2.79. Subsequently, the supply chain performance and environmental risk variables follow the financial risk variable in terms of centrality values. The density (clustering coefficient) of a fuzzy cognitive map is a correlation index [20]. The density index calculated for this study was 0.173. However, a high-density index indicates the existence of more likely causal relationships between the variables. Variables are categorized based on their upper degree and degree values. Transmitter variables control the system, but the system does not affect them. Thus, a lower number of transmitter variables is expected in a cognitive map with well-executed causal relationships [20]. In the developed model, there are five transmitter variables: strategic risk, production risk, employee risk, political risk, and information risk. These variables affect other variables, but they are not affected by any variable. The receiver variables in the system express the purpose of the system. Other variables affect the receiver variable, but the receiver variable does not affect any variable. Receiver variables are the strategic concepts of the system. In the developed model, there are two receiver variables: supply risk and supply chain performance variables. In this context, the output of the model in terms of the receiver variable is highly meaningful for the study. The remaining variables, namely inventory risk, supply risk, product risk, distribution risk, financial risk, operational risk, and supply chain risk, are classified as ordinary variables. These variables affect and are affected by each other. In the final analysis of FCM, the complexity score of the model is measured by taking the ratio of the number of receiver variables to transmitter variable [20]. Accordingly, the complexity score was calculated as 0.4.

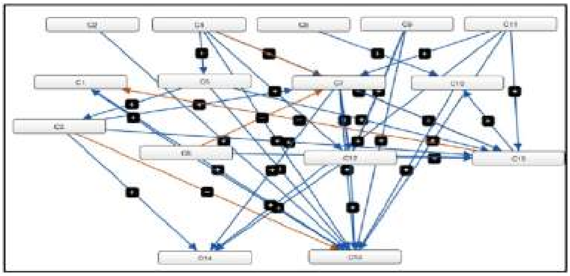


Fig. 1. Network Diagram of iFCM.

Scenarios

Scenarios provide a comprehensive depiction of a plausible future situation based on a complex network of influencing factors [23]. The primary aim of scenario analysis was to examine the compatibility of the obtained analysis results with the data obtained from document coding.

a. Scenario 1

The initial scenario was conducted based on the central focus of the study, namely the concept of supply chain risk. As shown in Fig. 2, an increase in the concept of the supply chain risk causes a strong negative increase in inventory risk and a positive increase in supply chain performance. Parallel with this finding, [26] highlighted a negative relationship between supply chain risk and inventory risk, while in their study [4] emphasized a positive significant relationship between supply chain risk and operational risk.

One reason for a negative relationship between supply chain risk and inventory risk could be due to the fact that firm's large inventories can help mitigate supply chain risks such as stockouts or delays, but also creates additional inventory costs and increased risk of inventory obsolescence. Furthermore, the positive significant relationship between supply chain risk and operational risk highlights the importance of effective supply chain risk management in maintaining operational performance and delivering high-quality products or services to customers.

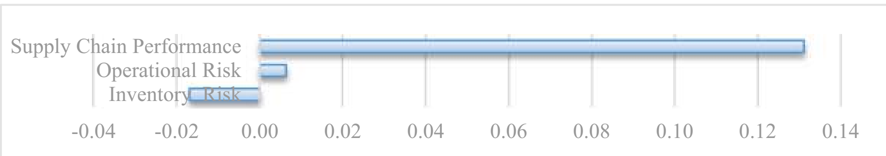


Fig. 2. Result of Scenario 1

b. Scenario 2

The second scenario was created by considering the centrality scores. The most central concept of the model is financial risk (with a centrality value of 2.79). To monitor the effects on the current system, the percentage change in the remaining variables when the financial risk variable increases is given in Fig. 3. As a result of this change, distribution risk and environmental risk increase in a very strong and positive manner. Parallel with this finding, in their study [26] highlighted a positive significant relationship between financial risk and environmental risk, while [25] emphasized a positive significant relationship between financial risk and supply chain risk. According to the study by reference [24], supply chain disruptions - which can be identified as indicators of supply chain risk- can have an impact on the stock market indicators associated with financial risk.

The existence of a positive association between supply chain risk and financial risk is a foreseeable outcome, given that disruptions within the supply chain can exert an influence on a company's financial performance. Consequently, the effective management of supply chain risk assumes paramount significance in mitigating financial risk and enhancing the overall financial performance of a firm. Additionally, the management of environmental risk holds a pivotal role in the context of sustainable supply chain risk management, as it enables companies to attenuate the adverse effects of risks on their financial performance.

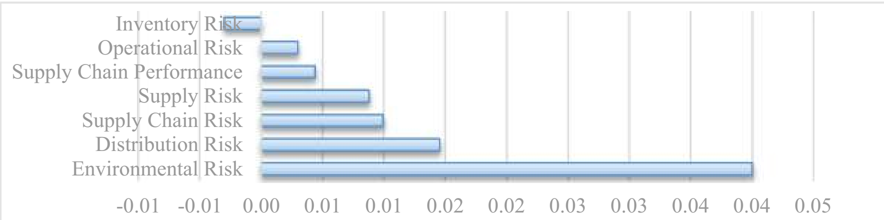


Fig. 3. Result of Scenario 2

c. Scenario 3

The third scenario was constructed by employing transmitter variables that possess zero indegree. As shown in Fig. 4, an increase in the strategic risk, production risk, employee risk, political risk, and information risk cause a strong positive increase in financial risk. This finding aligns with the assertions made by [22] in their respective investigations, wherein they establish a positive association between information risk and financial risk. Furthermore, as per the study cited in reference [26], environmental risks are predominantly interconnected with governmental, economic, and social factors.

Environmental risks can be influenced by various factors, including government policies, economic conditions, social factors. Government policies and regulations can affect the environmental impact of production processes, while economic conditions can influence the demand for products or services. Effective management of these risks is crucial for maintaining financial stability and achieving long-term sustainability in the supply chain. By focusing on these risks, companies can effectively distribute their resources and efforts, resulting in improved risk mitigation and overall performance of the supply chain.

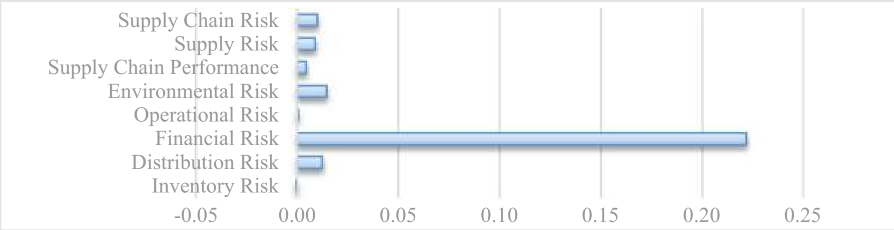


Fig. 4. Result of Scenario 3

5 Conclusion

As supply chains have become more complex and globalized, there are a multitude of risks that can disrupt the delivery of goods and services. As such, it is essential for organizations to identify evaluate their supply chain risks, and take appropriate actions to reduce or eliminate those risks. Therefore, it is important for organizations to identify and evaluate their supply chain risks, and take appropriate measures to reduce or eliminate those risks. As a further suggestion, the data collection period can be extended for future studies so that researchers can enhance data size to observe more complex causal interactions of various factors.

References

1. Bakhtavar, E., Valipour, M., Yousefi, S., Sadiq, R., Hewage, K.: Fuzzy cognitive maps in systems risk analysis: a comprehensive review. Complex Intell. Syst. **7**(2), 621–637 (2020). https://doi.org/10.1007/s40747-020-00228-2
2. Kosko, B.: Fuzzy cognitive maps. Int. J. Man. Mach. Stud. **24**(1), 65–75 (1986)
3. Zadeh, L.A.: Fuzzy sets. Inform. Control **8**(3), 338–353 (1965)
4. Rahmen, M.: The impact of social customer relationship management on operational performance in the service sector. J. Serv. Manage. Res. **6**(1), 23–36 (2022). https://doi.org/10.1016/jsmr.2022.01.004
5. Gurtu, A., Johny, J.: Supply chain risk management: literature review. Risks **9**(1), 16 (2021)
6. Karaali, F. Ç., & Ülengin, F.: Yapay Sinir Ağları ve bilişsel haritalar kullanılarak işsizlik oranı öngörü çalışması. İTÜ DERGİSİ/d, 7(3) (2008)
7. Axelrod, R. (ed.): Structure of Decision: The Cognitive Maps of Political Elites. Princeton University Press (2015)
8. Hugos, M.H.: Essentials of Supply Chain Management. Wiley (2018)
9. Dursun, M., Gumus, G.: Intuitionistic fuzzy cognitive map approach for the evaluation of supply chain configuration criteria. Math. Meth. Appl. Sci. **43**(13), 7788–7801 (2020)
10. Singh A.R., Jain R., Mishra P.K.: Risk in Supply Chain Management, National Conference on Advances in Mechanical Engineering (2009)
11. Purnomo, M.R.A., Anugerah, A.R., Dewipramesti, B.T.: Sustainable supply chain management framework in a higher education laboratory using intuitionistic fuzzy cognitive map. J. Indust. Eng. Manage. **13**(2), 417–429 (2020)
12. Rezaee, M.J., Yousefi, S., Valipour, M., Dehdar, M.M.: Risk analysis of sequential processes in food industry integrating multi-stage fuzzy cognitive map and process failure model and effects analysis. Comput. Ind. Eng. **123**, 325–337 (2018)

13. Rezaei Pandari, A., Azar, A.: A fuzzy cognitive mapping model for service supply chains performance. Meas. Bus. Excell. **21**(4), 388–404 (2017)
14. Xu, Z.: Intuitionistic fuzzy set theory: a survey and recent applications. Comput. Intell. Neurosci. **2017**, 1–21 (2017)
15. Atanassov, K.: Intuitionistic fuzzy implications and Modus Ponens, Notes on Intuitionistic Fuzzy Sets, vol. 11, 2005, no. 1, 1–5. R. Feys, Modal logics. Gauthier-Villars, Paris (1965)
16. Torra, V.: Hesitant fuzzy sets. Int. J. Intell. Syst. **25**(6), 529–539 (2010)
17. Nieto-Morote, A., Ruz-Vila, F.: A fuzzy approach to construction project risk assessment. Int. J. Project Manage. **29**(2), 220–231 (2011)
18. Užga-Rebrovs, O., Kuļešova, G.: Comparative analysis of fuzzy set defuzzification methods in the context of ecological risk assessment. Inform. Technol. Manage. Sci. **20**(1), 25–29 (2017)
19. Mogharreban, N., Dilalla, L.F.: Comparison of defuzzification techniques for analysis of non-interval data. In: NAFIPS 2006–2006 Annual Meeting of the North American Fuzzy Information Processing Society, pp. 257–260 (2006)
20. Özesmi, U., Özesmi, S.L.: Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. Ecol. Model. **176**, 43–64 (2004)
21. Ülengin, F., Işık, M., Ekici, ŞÖ., Özaydın, Ö., Kabak, Ö., Topçu, Y.İ: Policy developments for the reduction of climate change impacts by the transportation sector. Transp. Policy **61**, 36–50 (2018)
22. Li, J., Zhou, Y., Zhu, Q., Huang, W.: The impact of risk information sharing on supply chain financial performance: evidence from China. Sustainability **11**(17), 4672 (2019)
23. Shokouhyar, S., Pahlevani, N., Mir Mohammad Sadeghi, F.: Scenario analysis of smart, sustainable supply chain on the basis of a fuzzy cognitive map. Manage. Res. Rev. **43**(4), 463–496 (2019)
24. Wan, C., Yan, X., Zhang, D., Qu, Z., Yang, Z.: An advanced fuzzy Bayesian-based FMEA approach for assessing maritime supply chain risks. Transport. Res. Part E: Logist. Transport. Rev. **125**, 222–240 (2019)
25. Korucuk, A., Memiş, S.: Investigating the effects of financial risk on supply chain risk. J. Account. Finan. Audit. Stud. **4**(1), 55–68 (2018)
26. Abdel-Basset, M., Mohamed, R.: A novel plithogenic TOPSIS-CRITIC model for sustainable supply chain risk management. J. Clean. Prod. **247**, 119586 (2020)

Neural Networks

Enhancing E-Commerce Query Expansion Using Generative Adversarial Networks (GANs)

Altan Cakir[1] and Mert Gurkan[2(✉)]

[1] Istanbul Technical University, Sarıyer, 34467 Istanbul, Turkey
altan.cakir@itu.edu.tr
[2] Insider, Kağıthane, 34415 Istanbul, Turkey
mert.gurkan@useinsider.com

Abstract. In this study, we propose an innovative approach to query expansion (QE) in e-commerce, aiming to enhance the effectiveness of information search. Our method utilizes a generative adversarial network (GAN) called modified QE conditional GAN (mQE-CGAN) to expand queries by generating synthetic queries that incorporate semantic information from textual input. The (mQE-CGAN) framework consists of a generator and a discriminator. The generator is a sequence-to-sequence transformer model trained to produce keywords, while the discriminator is a recurrent neural network model used to classify the generator's output in an adversarial manner. By incorporating a modified CGAN framework, we introduce various forms of semantic insights from the query-document corpus into the generation process. These insights serve as conditions for the generator model and are instrumental in improving the query expansion task. Through various preliminary experiments, we demonstrate that the utilization of condition structures within the mQE-CGAN framework significantly enhances the semantic similarity between the generated sequences and reference documents. Compared to baseline models, our approach achieves an impressive increase of approximately 5–10% in semantic similarity.

Keywords: Generative Adversarial Networks · Information Retrieval · E-Commerce

1 Introduction

In e-commerce, search queries are typically matched with appropriate keywords, categories, and frequencies provided by business partners. This matching process often relies on straightforward rules such as exact, similarity, and phrase matching, which are based on the presence of shared tokens. However, to achieve more intelligent matching, AI-based techniques become necessary. These techniques enable the association of a query with relevant keywords, even when there is a limited overlap in their shared tokens.

The mathematical concept of the high frequency approach in power series provides a framework for understanding the nature of search queries. According

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 491–498, 2023.
https://doi.org/10.1007/978-3-031-39777-6_58

to this distribution, the majority of search queries fall within the main center, represented by the most frequently occurring queries. Conversely, the less common queries with low frequencies form the tail end of the distribution curve. However, the presence of rare queries presents a challenge for systems, as they tend to receive minimal query volume traffic. As a result, there is a lack of sufficient data related to these rare queries, necessitating the generation of synthetic data to address this data deficiency.

It is widely recognized that search queries typically adhere to a power law distribution, where the bulk of queries reside within the main center of the curve, while rare queries with low frequency form the tail. However, the limited occurrence of low-frequency queries poses challenges for systems as they are often excluded from the overall query volume traffic. This exclusion results in a data deficiency that necessitates synthetic generation to address the lack of sufficient data for these infrequent queries.

Matching rare queries with relevant documents in a search engine poses a significant challenge due to the distribution of incoming queries. Often, additional processes are necessary to facilitate the effective matching between these infrequent queries and the existing database of documents. To tackle this issue, researchers have explored various methodologies discussed in the literature, including relevance feedback methods, similarity-based approaches for query-document matching, machine translation models for query transformation, and query expansion methods. Among these approaches, query expansion methods offer an effective solution for search engines. This is because alternative methodologies mentioned tend to introduce additional complexities in terms of time and memory requirements. They either increase the number of steps involved in query-document matching or expand the search space, which can significantly impact the performance of the search engine. In contrast, query expansion methods provide an efficient means to improve matching without imposing excessive burdens on the search engine's computational resources.

Query expansion is a critical area of research in the field of Information Retrieval (IR) with broad applications spanning question answering, information filtering, and multimedia document matching tasks [1]. It involves enhancing the performance of matching input sequences with the document corpus of an IR system by reformulating the given input sequences [2]. The primary challenge in this task stems from the inherent characteristics of the input data commonly utilized in such studies.

2 Related Studies and Timeline

The integration of deep learning applications into natural language processing has paved the way for the emergence of word embeddings, providing effective means to capture semantic information within textual data [3]. Word embeddings enable the assessment of semantic relationships among words, offering valuable insights. Leveraging this capability, query expansion problems have been tackled through diverse approaches that assess the similarity between words in a query and potential expansion candidate terms.

The common use of word embedding techniques in numerous natural language processing and information retrieval tasks has spurred research endeavors aimed at enhancing the precision of word representations, particularly in specific scenarios. In this regard, various approaches have been proposed to generate alternative embeddings of tokens for query expansion [4]. Furthermore, investigations have explored the utilization of task-specific trained word embeddings for query expansion [5]. This approach facilitates the creation of word representations that are better equipped to capture the contextual nuances and semantic characteristics of the training corpus.

Based on recent studies [6,7], we introduce a ad-hoc query expansion technique for search engine optimization. Our approach leverages a prefix tree as a look-ahead strategy to generate expansion terms for a given query. By employing this method, we aim to enhance the effectiveness of search engine optimization efforts.

3 GAN Model Design

Initially, GAN models were trained using noise as input for the generation process. However, the advent of conditional GAN models introduced the ability to incorporate chosen conditions to aid in query generation. Similar to previous research in the field of query expansion, GAN-based architectures also employ the approach of augmenting user queries with relevant existing information. GAN models utilize various conditions, such as parts of the text data, class labels available during training, or extracted properties of the query and documents, to increase the probability of matching queries with desired documents. In the study by Lee et al. [8], a conditional GAN structure is proposed for query expansion, specifically focusing on enriching rare queries within search engines. On the other hand, Huang et al. [9] leverage a widely-used pseudo-relevance feedback method in the query expansion domain as the condition for generating expansion terms in their research.

The integration of conditions within the GAN framework serves as a focal point for researchers seeking to increase the efficacy of query expansion and elevate the relevance of retrieved documents in response to user queries.

- **Model:** In the context of e-commerce applications, we introduce a promising conditional generative adversarial network (GAN) model that incorporates the semantic correlation between query and document pairs as input. Our model features a generator, which is a sequence-to-sequence encoder-decoder model, and a discriminator that utilizes an LSTM-based binary classifier. We delve into the specifics of the model framework and present a comprehensive evaluation of the training process, highlighting the effectiveness of our conditional approach.
- **Conditional Query Expansion:** In this study, we present innovative approaches for conditioning generative adversarial networks (GANs). Our focus lies in designing conditioning structures that effectively capture the semantic relationships between query-document pairs.

- **Datasets:** We evaluate the performance of our generative model using user query and document pairs obtained from customers of Insider.[1]. To assess the efficacy of our models across diverse data characteristics, we conduct evaluations by testing the proposed models on various customer datasets.

4 System Architecture

4.1 *m*QE-CGAN Framework

The framework proposed for adversarial training using the mQE-CGAN model is illustrated in Fig. 1. In this framework, the generator model receives input queries along with assigned condition vectors specific to each query. It employs a sequence-to-sequence structure to generate expansion terms based on the provided queries. On the other hand, the discriminator model within the adversarial setup performs binary classification on the expanded synthetic queries and the documents that correspond to the original user queries.

The condition mechanisms reviewed in this study aim to optimize the available data. As query-document pairs within the datasets represent user searches and the corresponding matching documents, the condition approaches concentrate on the semantic and similarity metrics of the given queries and their associated documents, as determined by the search engine.

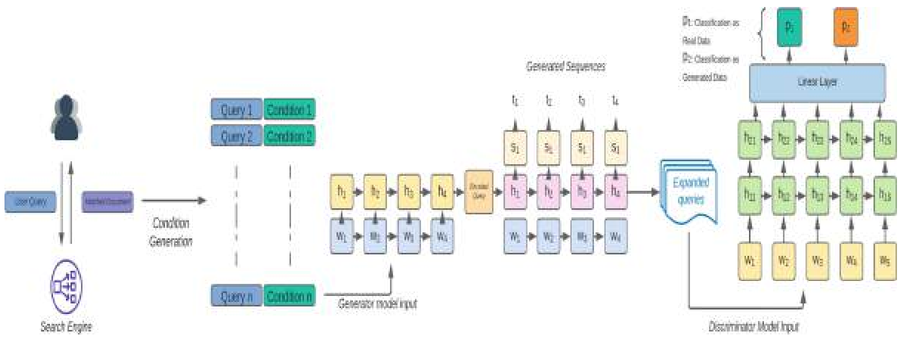


Fig. 1. Diagram of the *m*QE-CGAN framework.

Generator Model. The architecture's generator model utilizes an encoder-decoder sequence-to-sequence framework, which takes FastText word embedding representations [10] of user search queries and their associated condition vectors as input. To enable backpropagation with discrete input sequences, the generator's decoder employs Monte Carlo rollouts, a technique utilized in prior studies [8,11]. This approach facilitates the transfer of rewards generated by the discriminator to the generator at each generation step (Fig. 2).

[1] https://useinsider.com.

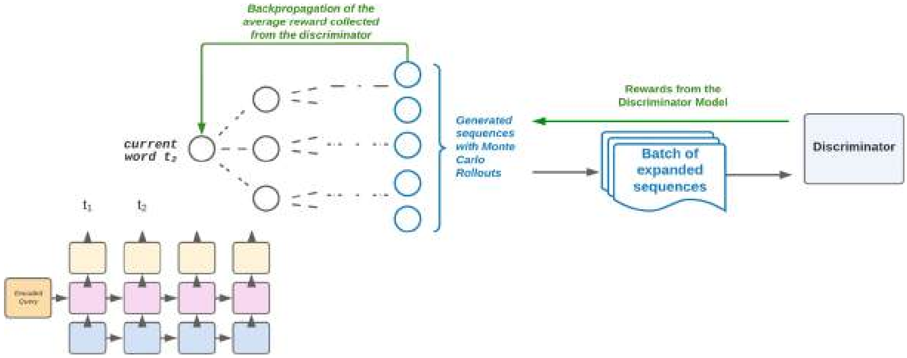


Fig. 2. The Monte Carlo rollouts diagram. The generation of a batch of sequences by the decoder of the network at each step. These generated batches are then assessed by the discriminator to provide guidance during the generation process of the generator model.

5 Experiments

5.1 Datasets

The study's datasets were derived from an analysis of user behavior within Insider's search engine product. These datasets specifically comprise user search queries and the top-ranked products retrieved from Insider's customers' platforms. Importantly, it should be emphasized that the datasets employed in this study do not contain any personally identifiable information pertaining to the users. Throughout the data collection process, any data that could potentially identify users was meticulously removed and discarded.

In general, users tend to use fewer words when entering queries in search engines to find the desired documents. Consequently, queries in search engines often contain fewer words compared to the corresponding documents. This observation holds true for the datasets utilized in our study as well. We firmly believe that accounting for the disparity in word count between queries and documents is crucial in the design of search engines and other information retrieval systems. It is imperative that our systems possess the ability to comprehend the intent behind user queries, even when they are brief and succinct.

The difference in word count between queries and documents presents various challenges for search engines. One such challenge is the cold start problem, which is common problem from frequency count rating. In the context of rare query inputs from users, search engines are more susceptible to this issue. The datasets generated for this study specifically aim to test the capabilities of the mQE-CGAN framework in tackling this challenge. The cold start problem occurs when a search engine or recommendation system lacks sufficient information about a user or item. This can happen when a user is new to the system or when an item is newly introduced. In the case of rare query inputs, the search engine may lack the necessary information to comprehend the user's query intent, resulting in irrelevant search results. By employing rare query inputs in the generated datasets, the study seeks to push the limits of the mQE-CGAN framework and enable it to effectively address the cold start problem, thereby returning relevant results even for uncommon query inputs.

5.2 Generator Evaluation Metrics

The results of the evaluation metrics after integrating the condition generation strategies into the generator model can be found in the table below.

Table 1. The results indicate that the generator model successfully produces high-quality expanded queries for all the companies included in the dataset. Among the companies, C1, a cosmetics company based in Turkey, achieves the best performance. These findings serve as preliminary results presented at the conference, and further in-depth results will be shared in a forthcoming research article.

Dataset	Condition	CE Loss	Perplexity	WC	SS (μ, ϵ)
C1	Baseline Generator	1.266	3.650	1.07	(0.602, 0.173)
	Word Sim	1.328	3.792	1.02	**(0.696, 0.169)**
	Document Sim	1.258	3.536	0.99	(0.659, 0.178)
	TF-IDF	1.288	3.644	1.15	(0.606, 0.176)

The table presents the results of evaluation metrics, namely Semantic Similarity (SS), BLEU, and ROUGE, for different condition generation strategies. The semantic strategy that utilizes semantic information from the query-document corpus to generate conditions outperforms other strategies in all metrics. The results indicate that integrating the condition generation strategies into the generator model improves its performance. The SS metric shows that the generated sequences are more similar to the reference documents when the semantic strategy is employed. Similarly, the BLEU and ROUGE metrics reveal higher n-gram similarity and word overlap, respectively, between the generated sequences and reference documents for the semantic strategy.

6 Results

The best performing generator models, as shown in Table 1, are those conditioned with the Word Similarity method, which achieves the highest semantic evaluation metrics. This method provides precise embedding vectors of words that are most similar in meaning to the words in the query, enabling the generator models to better understand the context and generate more relevant expanded queries. Similar to pseudo-relevance feedback methods, the Word Similarity method enhances the query with documents that initially matched with it, helping the generator models to better understand the user's intent. However, the Document Similarity and TF-IDF Weighting conditions yield slightly worse semantic evaluation metrics than the Word Similarity condition since they do not provide as much context information about the query. In some cases, these conditions can even generate irrelevant expanded queries. Our analysis revealed that the Document Similarity conditions tend to guide the generation process in inaccurate ways as the most similar documents to given input queries can be differentiating from the reference documents, leading to the generation of irrelevant expanded queries. The results of this study are preliminary, and further research will be published in an upcoming research article.

According to our model performances, it is evident that the model exhibits the least success with the dataset of Company 1 (C1). This outcome aligns with our expectations following a thorough analysis of the dataset properties. The C1 dataset proves to be the most challenging, primarily due to its large vocabulary size in comparison to the other utilized datasets. We are currently conducting further investigations within the scope of adversarial learning models to address this issue. We remain optimistic that our future endeavors will yield improvements in the model's performance specifically on the C1 dataset, thereby enabling the generation of more relevant expanded queries across a wider range of datasets.

7 Conclusion

Our research focused on the integration of generative adversarial networks (GANs), query expansion, and condition structures derived from query-document relationships. Our mQE-CGAN framework demonstrated that query expansion can effectively enrich user queries with limited information and generate semantically similar sequences to the documents in the datasets. The framework has shown promising evaluation metrics, especially on processed datasets, making it a viable option for optimizing search engines in the e-commerce domain. We believe that the mQE-CGAN framework has great potential for enhancing the performance of search engines in the e-commerce domain. We anticipate that future work will lead to further advancements and improvements in the framework.

References

1. Carpineto, C., Romano, G.: A survey of automatic query expansion in information retrieval. ACM Comput. Surv. (2012). https://doi.org/10.1145/2071389.2071390
2. Azad, H.K., Deepak, A.: Query expansion techniques for information retrieval: a survey. Inf. Process. Manag. **56**, 1698–1735 (2019). https://doi.org/10.1016/j.ipm.2019.05.009
3. Roy, D., Paul, D., Mitra, M., Garain, U.: Using word embeddings for automatic query expansion. Computing Research Repository (CoRR) (2016)

4. Sordoni, A., Bengio, Y., Nie, J.Y.: Learning concept embeddings for query expansion by quantum entropy minimization. In: Proceedings of the AAAI Conference on Artificial Intelligence (2014). https://doi.org/10.1609/aaai.v28i1.8933
5. Diaz, F., Mitra, B., Craswell, N.: Query expansion with locally-trained word embeddings. Computing Research Repository (CoRR) (2016). https://doi.org/10.48550/arXiv.1605.07891
6. Qi, W., et al.: ProphetNet-Ads: a looking ahead strategy for generative retrieval models in sponsored search engine. In: Zhu, X., Zhang, M., Hong, Yu., He, R. (eds.) NLPCC 2020. LNCS (LNAI), vol. 12431, pp. 305–317. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60457-8_25 ISBN 978-3-030-60457-8
7. Lian, Y., et al.: An end-to-end generative retrieval method for sponsored search. Computing Research Repository (CoRR) (2021). https://doi.org/10.48550/arXiv.1902.00592
8. Lee, M.C., Gao, B., Zhang, R.: Rare query expansion through generative adversarial networks in search advertising. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2018). https://doi.org/10.1145/3219819.3219850
9. Huang, M., Wang, D., Liu, S., Ding, M.: GQE-PRF: generative query expansion with pseudo-relevance feedback. In 2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT) (2021). https://doi.org/10.48550/arXiv.2108.06010
10. Bojanowski, P., Grave, E., Joulin, A., Mikolov, T.: Enriching word vectors with subword information. Computing Research Repository (CoRR) (2016). https://doi.org/10.48550/arXiv.1607.04606
11. Yu, L., Zhang, W., Wang, J., Yong, Yu.: Seqgan: sequence generative adversarial nets with policy gradient. In: Proceedings of the AAAI Conference on Artificial Intelligence (2017). https://doi.org/10.48550/arXiv.1609.05473

MEConvNN-Designing Memory Efficient Convolution Neural Network for Visual Recognition of Aerial Emergency Situations

Unse Fatima(✉), Junbom Pyo, YeongMin Ko, and Moongu Jeon

School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, South Korea
unse.fatima@gm.gist.ac.kr, {xenonluv,KoYeongMin,mgjeon}@gist.ac.kr

Abstract. Unmanned aerial vehicles (UAVs) play a vital role in calamity and natural disaster management due to their remote sensing capabilities. Specifically, UAVs/drones equipped with visual sensors can have remote access to confined areas wherein human access is limited. Growing inventions in deep learning incubate the efficacy of such UAVs/drones in terms of computational capability with limited resources and lead us to effectively utilize this technology in visual recognition of emergency situations, like floods in urban areas, earthquakes or fires in forests, and traffic accidents on busy highways. This can be beneficial in mitigating the consequences of such events on the environment and people more rapidly with minimum men and material loss. However, most deep learning architectures used in this domain with high accuracy are costly regarding memory and computational resources. This motivates us to propose a framework that can be computationally efficient and can be utilized on an embedded system suitable for smaller platforms. In this work, we formalize and investigate that problem and design a memory-efficient neural network for visual recognition of emergency situations named MEConvNN. To this end, we have effectively used dilated convolutions to extract the spatial representation. The proposed method is experimentally evaluated using Aerial Image Database for Emergency Response (AIDER), showing comparative efficacy with the state-of-the-art methods. Specifically, the proposed method achieves accuracy with less than a 2% drop compared to state-of-art methods but is more memory efficient in contrast to state-of-art methods.

Keywords: Unmanned Aerial Vehicles (UAVs) · Remote sensing · Convolutional neural networks · Image Classification

1 Introduction

Unmanned aerial vehicles (UAVs) have surfaced as an effective resource for remote sensing applications recently, offering numerous advantages that make them a tempting option compared to traditional remote sensing platforms such

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 499–506, 2023.
https://doi.org/10.1007/978-3-031-39777-6_59

as satellites or manned aircraft. One of the main advantage of UAVs is the maneuverability that enables them to be launched and operated from almost any location. This would allow them to remotely access confined spaced that are inaccessible to other platforms. This is particularly useful in applications like search and rescue (SAR) situations monitoring, where UAVs can be dispatched to the scene of disaster to capture images and videos. The captured images/videos can be transmitted to central command center in real-time for damage assessment. The earliest access to this visual data can assist in evaluating the magnitude of the damage and pinpoint areas necessitating urgent intervention to enhance the efficiency of rescue and relief operations.

During SAR monitoring, the UAVs are required to be capable of identifying and recognizing the objects on the ground. In this regards, Deep learning provides numerous framework for accurate objection detection and recognition. Operating such algorithms necessitates ample processing power onboard UAVs. The successful detection and recognition of objects would allow us to quickly and accurately identify areas of damage, locate survivors, and assess the overall extent of the disaster. By leveraging the power of machine learning algorithms and deep learning models, image classification can enable more effective and efficient emergency responses.

Utilizing Convoultional Neural Network (CNN) architectures in UAVs offers several advantages for remote sensing applications like real-time detection and classification. While CNNs offer many advantages, their on-board implementation in UAV's may face constraints due to limited computing power, memory, power consumption and storage capacity. These limitations need to be considered carefully while designing and implementing CNNs for UAV-based applications.

This work explores the visual recognition of emergency situation using remote sensing data by designing a memory efficient dilated convolution (MEDC) architecture. It comprises of a dilated convolutional block. The proposed frameworks's performance is validated on publicly available dataset. The proposed model renders better performance with respect to memory consumption than the state-of-the-art methods. The organization of the remaining paper is as such: Sect. 2 delves into relevant research. The suggested approach is detailed in Sect. 3. Section 4 presents the experiments and their outcomes. Finally, the paper's conclusion is outlined in Sect. 5.

2 Literature Review

Compared to aircrafts and satellites, remote sensing data acquisition through UAVs is cost-effective [1,2] and their flexibility allows to acquire high-resolution images from relatively low altitudes. Owing to this attribute, recently they have been used effectively in classification and analysis of disaster stricken areas to facilitate in search and rescue operations [3–6]. According to [7], image classification tasks are well performed by deep convolutional neural networks architectures because of their ability to identify significant features from two-dimensional images. Numerous CNN architectures have been developed over the years that achieved higher accuracy than human level. These architectures include VGG16 [8], ResNet50 [9], Inception [10], Xception [11] and many others.

Recently, multiple studies have used CNN to classify images captured by UAVs for emergency responses. In [12], a system was designed to identify avalanches victims in mountainous region during winter. Usually rescue teams face challenges due to difficult evacuation routes. However, unmanned Aerial Vehicles (UAVs) equipped with sensors like cameras can help in locating victims for their earliest evacuation. It employs a CNN approach with Support Vector Machine (SVM) for object classification. This paper comparatively analyzed the performance of CNN-SVM and HOG-SVM techniques using image resolution as a key parameter. The CNN-SVM method proved to be superior, with an accuracy of 94% compared to 85.82%. In [13], the researchers suggested utilizing a fundamental CNN framework for categorizing images of forest fires. This approach employed a structure resembling that of AlexNet [14], incorporating a sequence of convolutional and pooling layers, flattening and being connected to two dense layers. The final layer utilized a sigmoid activation function for binary classification, ultimately yielding an accuracy rate of 95%. However, in both of these studies, a single class is targeted. Most relevant work is presented in [15] which utilizes same dataset, however our architecture outperforms in terms of computational efficiency and mean F1 score.

In this work, we focus on the computational efficiency of the architecture for drone dataset using an algorithm which allows to consume less memory while having a good F1 score.

3 Methodology

The proposed methodology as shown in Fig. 1 is comprised of memory-efficient dilated convolutions (MEDC) block. The dilated convolution module of the proposed method makes use of dilation rate to capture and process images at varying resolutions. MEDC block as shown in Fig. 2, captures the visual information at various scales from the input image by applying depth-wise convolutions. In each convolution different dilation rates is employed. Subsequently, it combines the resulting output feature maps through addition and then reduces the number of channels with a 1×1 convolution. The dilation rate expands the receptive field by using spacing between kernel points without augmenting the quantity of parameters. Suppose I_n represents the feature map of input with a kernel K and a dilation rate l. The output feature map O for dilated convolution block is determined as:

$$O(i,j) = \sum_{p=1}^{I} \sum_{q=1}^{J} I_n(i + l \times p, j + l \times q) \in K(p,q) \tag{1}$$

whereas, i and j represent indices of the output feature map, p and q are the indices of the kernel. The formula iterates over all valid positions of the kernel on the feature map of input, implementing the dilated convolution at each position.

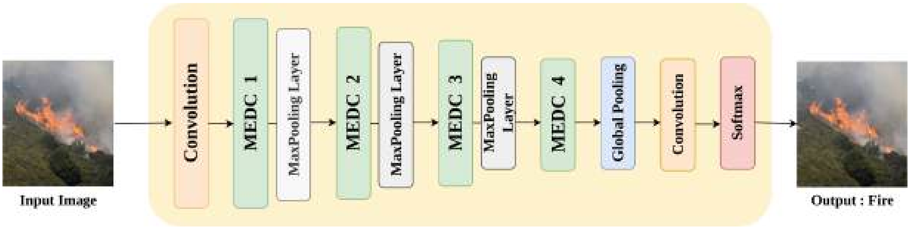


Fig. 1. The structure of Memory-Efficient Convolutional Neural Network

The dilated convolution module leads to lower computational cost while maintaining good F1 score. In standard convolution, the convolutional layer's parameter count rises as the kernel size increases. In order to capture larger receptive field using conventional convolutions, stacking of multiple layers or a larger kernel size is required. Consequently increasing total number of parameters and hence memory usage. However, with dilated convolutions a 3×3 kernel with $l = 2$ will have same receptive field as a 5×5 kernel of a conventional convolution method but with less than half number of parameters. Thus dilated convolution provides the better output to that of the conventional convolution operation with much lesser memory usage. Therefore, it reduces the computation cost of executing the proposed framework.

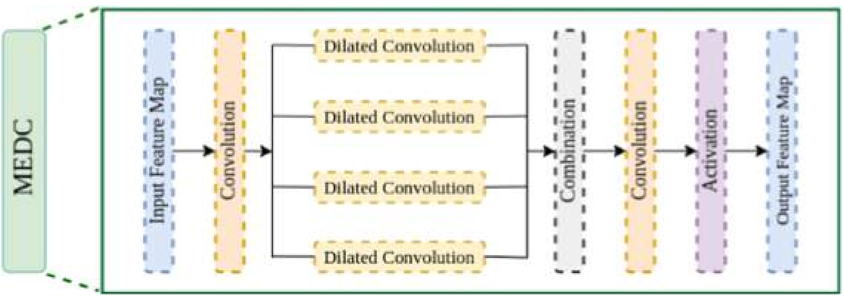


Fig. 2. Memory-Efficient Dilated Convolution Block architecture

4 Experiment and Result Analysis

This section numerates the experimental setup followed by the outcomes of the presented approach. The performance of the suggested technique is assessed using a publicly available dataset named Aerial Image Dataset for Emergency Response (AIDER) [15] . To evaluate performance, we employed the standard metric: mean F1 Score which is the hormonic mean of precision and recall.

4.1 Dataset Description

We employed a publicly accessible dataset designed for emergency response applications, called the Aerial Image Dataset for Emergency Response [15]. This dataset is collected using images from four distinct disaster events, namely Collapsed Building, Fire/Smoke, Traffic Incidents and Flood, also a Normal class. Images are captured at varying resolutions, under diverse illumination conditions, and from different views. The dataset is intentionally imbalanced, normal class features more images to better resemble real world scenarios. Table 1 displays the quantity of images in each category, illustrating the images distribution in train and test subsets of the dataset.

Table 1. The images distribution in the train and test subsets.

Multi-Class Dataset	Train set	Test set
Collapsed Building	408	103
Fire	416	105
Flooded Areas	420	106
Traffic Incidents	388	97
Normal Class	3512	878

4.2 Training Methodology

The suggested ME-ConvNN method is implemented using the Keras deep learning library, with TensorFlow as the back-end. All networks use an image size of 240×240 pixels. The proposed network training is done for 200 epochs using an Nvidia GTX 1080 with memory of 8GB. The loss function utilized for the training network is categorical cross-entropy loss as we have many classes available in the dataset. The SGD optimizer is employed during training of proposed framework with weight decay of 1×10^{-3}. To further enhance the dataset, a range of techniques like illumination changes, blurring, sharpening, and shadowing have been applied. Before adding each image to the training batch, random geometric transformations like rotations, translations and others are also adopted with probability-based approach. The aim of these transformations is to mitigate over-fitting and boost the diversity of the training set, resulting in enhanced generalization of proposed framework.

4.3 Evaluation Criterion

To evaluate the performance of the presented network, mean F1 Score is chosen for five classes to effectively address the imbalanced dataset. This metric demonstrates a low count of false positives and false negatives, reflecting a balance between precision and recall.

$$F1 = \frac{2}{N} \times \sum_{i=1}^{N} \frac{\mathrm{Pre}_i \times \mathrm{Rec}_i}{\mathrm{Pre}_i + \mathrm{Rec}_i} \quad (2)$$

where N is the total number of classes present in the dataset, Pre_i and Rec_i represent the precision and recall of the i^{th} class.

4.4 Results

As part of the result, we present both the quantitative as well as qualitative outcomes of the proposed framework.

Table 2. A quantitative analysis of ME-ConvNN is conducted in comparison to state-of-the-art methods relative to memory consumption. The evaluation utilizes the mean F1 score as the performance metric.

Model	Parameters	F1 Score (%)	Memory (MB)
VGG16	14,840,133	95.83	59.36
Inception	22,171,429	96.13	88.68
Xception	21,516,845	96.45	86.06
EmergencyNet	90,892	94.16	0.36
ME-ConvNN	57,204	94.87	0.22

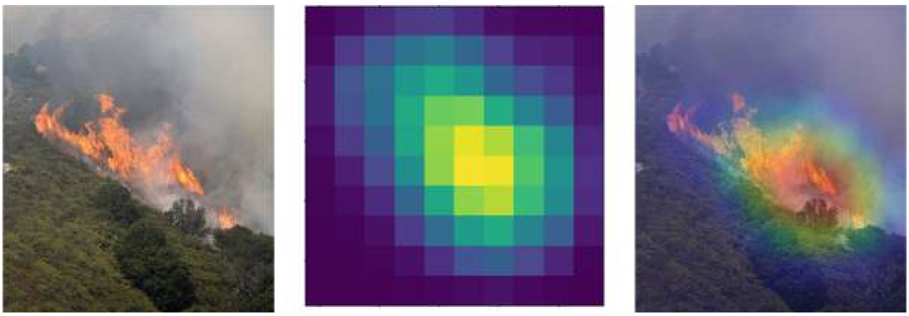


Fig. 3. Qualitative evaluation of ME-ConvNN architecture. Taking fire class image as a sample

We illustrated the proposed model's performance utilizing emergency response dataset using existing techniques, including VGG16 [8], Inception [10], Xception [11] and EmergencyNet [15].

ME-ConvNN has outperformed all mentioned techniques in terms of memory consumption as compared to VGG16, Inception and Xception model while marginally compromising on F1 score drop. Our network is almost 38.8% more computationally efficient in comparison with EmergencyNet and has a greater F1 score as well. Results are shown in Table 2.

In addition, the qualitative results of the proposed ME-ConvNN are demonstrated in Fig. 3, indicating which part of the image led ME-ConvNN to its final classification decision. This also validates the fact that the network has been successfully converged and optimally trained.

5 Conclusion and Future Work

In this article, we presented ME-ConvNN, a memory-efficient architecture for multi-class image classification. The ME-ConvNN utilizes a dilated convolution block that leverages varying dilation rates to increase the receptive field while minimizing the parameter counts. The ME-ConvNN's performance is assessed on a public dataset using the mean F1 Score as the evaluation criterion. Compared to other state-of-the-art methods to the best of our knowledge, ME-ConvNN offers greater computational efficiency without significantly sacrificing F1 score.

The combination of enhanced memory efficiency and a fair F1 score enables the potential for real-time implementation of this architecture in search and rescue operations. In future research, we plan to collect more aerial images for different classes to enrich diversity and further increase the efficiency of ME-ConvNN in terms of memory.

Acknowledgement. This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2014-3-00077, Development of Global Multi-target Tracking and Event Prediction Techniques Based on Real-time Large-Scale Video Analysis), and Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism (No. R2022060001, Development of service robot and contents supporting children's reading activities based on artificial intelligence)

References

1. Alvarez-Vanhard, E., Corpetti, T., Houet, T.: UAV & satellite synergies for optical remote sensing applications: a literature review. Sci. Remote Sens. **3**, 100019 (2021)
2. Matese, A., et al.: Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. Remote Sens. **7**(3), 2971–2990 (2015)
3. Valsan, A., Parvathy, B., GH, V.D., Unnikrishnan, R.S., Reddy, P.K., Vivek, A.: Unmanned aerial vehicle for search and rescue mission. In: International Conference on Trends in Electronics and Informatics (ICOEI) (48184), pp. 684–687. IEEE (2020)
4. Alsamhi, S.H., et al.: Multi-drone edge intelligence and SAR smart wearable devices for emergency communication. Wirel. Commun. Mob. Comput. **2021**, 1–12 (2021)

5. Ribeiro, R.G., Cota, L.P., Euzébio, T.A., Ramírez, J.A., Guimarães, F.G.: Unmanned-aerial-vehicle routing problem with mobile charging stations for assisting search and rescue missions in postdisaster scenarios. IEEE Trans. Syst. Man Cybern. Syst. **52**, 6682–6896 (2021)
6. Schedl, D.C., Kurmi, I., Bimber, O.: An autonomous drone for search and rescue in forests using airborne optical sectioning. Sci. Rob. **6**(55), eabg1188 (2021)
7. Zong, M., Wang, R., Chen, X., Chen, Z., Gong, Y.: Motion saliency based multi-stream multiplier ResNets for action recognition. Image Vision Comput. **107**, 104108 (2021)
8. Simonyan, K. and Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014). arXiv preprint arXiv:1409.1556
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
10. Szegedy, C., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
11. Chollet, F.: Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1251–1258 (2017)
12. Bejiga, M.B., Zeggada, A., Nouffidj, A., Melgani, F.: A convolutional neural network approach for assisting avalanche search and rescue operations with UAV imagery. Remote Sens. **9**(2), 100 (2017)
13. Srinivas, K., Dua, M.: Fog computing and deep CNN based efficient approach to early forest fire detection with unmanned aerial vehicles. In: Smys, S., Bestak, R., Rocha, Á. (eds.) ICICIT 2019. LNNS, vol. 98, pp. 646–652. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-33846-6_69
14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Commun. ACM **60**, 84–90 (2017)
15. Kyrkou, C., Theocharides, T.: EmergencyNet: efficient aerial image classification for drone-based emergency monitoring using atrous convolutional feature fusion. IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens. **13**, 1687–1699 (2020)

Multiple Neural Networks for Clustering and Prediction of the Particulate Matter (PM2.5): A Case Study of Beijing

Martha Ramírez and Patricia Melin(✉)

Tijuana Institute of Technology, Tijuana, Mexico
pmelin@tectijuana.mx

Abstract. Air pollution represents a world health problem for decades. The monitoring of pollutants in the air is a process that allows using the information recorded in a period to notify what the current measurement is and the future measurement of the polluting indices or variables, such as particulate matter (PM2.5), which can be cause of people to have respiratory problems among others health concerns. So, it is necessary to inform to the population the levels of the contaminants in the air. Thus, in this work we proposed a method to predict the particulate matter (PM2.5) time series using multiple unsupervised artificial neural networks to first perform information clustering tasks. Subsequently, multiple supervised neural networks are used to carry out the prediction tasks for each of the formed information segments. In other words, the outputs obtained by the first set of unsupervised neural networks are used to prepare the inputs of the set of supervised neural networks. The advantages of the method for clustering and prediction of time series are shown in the results of the simulation carried out. Also, the utilization of neural networks to find similarity of the data makes it possible to cluster and subsequently make the prediction of the particulate matter (PM2.5) based on each one of the identified clusters, which allow to have helpful and valuable information to plan and carry out preventive health actions specifically for certain locations or people groups.

Keywords: Neural Networks · Time Series · Prediction · Competitive Clustering · Air Quality

1 Introduction

As a support in the decision-making process [1], the analysis of historical information and monitoring of pollutants in the air makes it possible to use the collected information to issue early warnings on the current and future measurement of pollutant variables, such as particulate matter (PM2.5).

Therefore, considering that one of the objectives of the World Health Organization (WHO) is achieving global air quality guidelines for countries to achieve recommended air quality levels, helping to prevent respiratory diseases in the population [2], in this paper a method for time series of particulate matter (PM2.5) prediction is presented.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 507–514, 2023.
https://doi.org/10.1007/978-3-031-39777-6_60

During the last decades, many technological advances have contributed to the utilization of Artificial Neural Networks (ANNs) to solve complex problems, this is due to the ability of ANNs to learn non-linear relationships [3, 4]. ANNs are mathematical models inspired in the connection of the biological neurons in the brain [5], so they can learn directly from the data and be trained for pattern recognition [6, 7]. Depending on the type of problem to solve and the quality of the available dataset, we can use supervised [8, 9] or unsupervised [10–12] learning algorithms to train the ANNs.

Thus, the main objective is to create a model for time series prediction using unsupervised neural networks to divide the data based on how similar they are between them, and later use supervised neural networks (NNs) to predict the future values for each data cluster. Therefore, with our proposal the main contribution is a method that uses multiple neural NNs to predict time series, seeking to take advantage of the two types of training (supervised and unsupervised) offered by these computational models. First, a set of competitive neural networks is used to classify inputs based on a given number of classes [13], and secondly, the outputs obtained (data clusters) by a set of unsupervised artificial neural networks are used to prepare the inputs of the set of Nonlinear autoregressive neural networks. This set of supervised neural networks is used to predict future values for each data cluster [14, 15]. It is important to mention that our proposal differs from most existing methods, by performing a combination of supervised and unsupervised training algorithms, unlike the computational models in the literature that mostly use supervised training algorithms for time series prediction; which represents a great advantage for our method, since the prediction can be made by specific segments of the data, which can be divided by geographic region, frequency or specific characteristics.

Beneath we briefly describe the contents of this paper. In Sect. 2, an overview of the problem is presented. We show the used methods in Sect. 3. The experimental and discussion of results are shown in Sects. 4 and 5, respectively. Finally, in Sect. 6, the general conclusions are presented.

2 Problem Description

Despite the fact that clean air is a human right, about 7 million people die every year due to air pollution. There is also a huge burden of sickness, hospitalization, reduced life expectancy and the associated social and economic impacts of lost productivity and health care costs [16].

Part of the civil and governmental actions is to reduce the level of pollutants in the environment. However, is a significant challenge that requires long-term action. Meanwhile it is necessary to have tools that make it possible to predict the level of air pollution, in order to inform the population in a timely manner so that direct exposure can be avoided as far as possible when there are high concentrations of pollutants in the environment.

For this work we selected the Beijing Multi-Site Air-Quality Data dataset from the UCI Machine Learning Repository. This dataset consists of 420768 instances for 18 attributes (row number, year, month day, hour, PM2.5, PM10, SO2, NO2, CO,O3, temperature, pressure, dew point temperature, precipitation, wind direction, wind speed, station name), corresponding to hourly air pollutants data from 12 nationally controlled air-quality monitoring sites from the Beijing Municipal Environmental Monitoring Center. The period is from March 1st, 2013, to February 28th, 2017 [17].

In order to carry out the clustering tasks corresponding to the concentration of the particulate matter ($PM_{2.5}$), we obtained the monthly average of the variables PM25, PM10, SO2, NO2, CO and O3. For the prediction tasks we also obtained the monthly and daily average of the variable PM25.

3 Proposed method

In this work, we propose a model that consists of two phases, the first one is to use a competitive neural network [18] for clustering tasks because of their performance to find similarities and to label each element of the dataset with the class that belongs to it, that means, all inputs units ***i*** are connected to all output units ***j*** with weight $\mathbf{w_{ij.}}$ The input dimension is the number of inputs, and the number of outputs is equal to the number of clusters that the data has. The activation value a_{j} of the output unit is calculated by the inner product of the input and weight vectors and returns the output unit with the highest activation ***k***, the weights leading to this unit are updated by the competitive function or the so-called winner-take-all learning rule as Eq. (1) and the weights of the output unit with the smallest activation are updated to be even closer, and can be written as Eq. (2) [19]:

$$\mathrm{w}_k(t+1) = \frac{\mathrm{w}_k(t) + \lambda(\mathrm{x}(t) - \mathrm{w}_k(t))}{\|\mathrm{w}_k(t) + \lambda(\mathrm{x}(t) - \mathrm{w}_k(t))\|} \tag{1}$$

$$\mathrm{w}_k(t+1) = \mathrm{w}_k(t) + \lambda(\mathrm{x}(t) - \mathrm{w}_k(t)) \tag{2}$$

where w is an input weight vector, k is the winner output unit, t is a vector position, η is the learning rate and x is an input vector. Thus, when an input x is presented to the network, the weight vector closest to x rotates toward it.

The main objective is to detect the similarities in the dataset between the variables or attributes, seeking to classify the information based on the average monthly concentration. So, we have used the competitive NN to separate the dataset into four classes: the first Cluster 1(C1), the second class: Cluster 2 (C2), the third class: Cluster 3 (C3) and the last class: Cluster 4 (C4).

For the prediction tasks in the second phase, we used a Nonlinear Autoregressive (NAR) neural network [20], because of the good results obtained previously.

By using a NAR Network, the future values of a time series y(t) are predicted only from past values of that series and can be written as Eq. (3) [21]:

$$y(t) \; = f(y(\mathrm{t}-1), y(\mathrm{t}-2), \ldots, y(t-d)) \tag{3}$$

where $y(t)$ represents a value of the time series y at time t, d is a time delay parameter and f is an activation function.

We can observe (Fig. 1) for each one of the competitive neural networks the groups formed corresponding to each class. When the attributes belonging to each cluster are identified, a subset of the time series data is integrated, which becomes the input to a NAR neural network. This applies to each of the four classes.

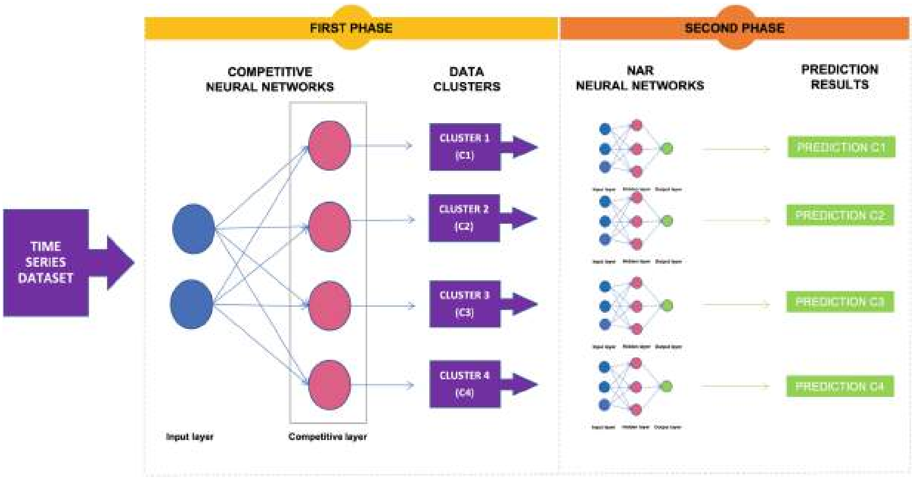


Fig. 1. Illustration of the proposed method.

To perform the clustering tasks, multiple competitive neural networks of four classes are used, and for the prediction tasks of the information segments of each of the classes, a NAR network with Levenberg-Marquardt algorithm for training is used. The hidden layer consists of 10 neurons.

4 Experimental Results

During the experimentation for prediction (Tables 1–2), we use the entire data set, 30 executions of the NAR neural network (the hidden layer consists of 10 neurons). For the 12 monitoring stations we calculated the monthly average of the variables (576 data points per variable) to make the prediction of the six polluting variables. In the Tables we show the Average (A), Best (B) and Worst (W) relative root mean squared error (%RMSE), as a percentage. So, based on the monthly average concentration of the six pollutant variables PM25, PM10, SO2, NO2, CO, O3 (2013–2017), the monitoring sites that recorded similar concentrations were grouped into four classes (C1, C2, C3, C4) by using competitive neural networks, where each experiment consists of performing 30 executions of the competitive neural network.

Subsequently, we made the prediction of the PM25 variable for each of the clusters formed (Table 3).

Based on the average monthly concentration of the variable PM25 by using competitive NNs we classify the monitoring sites into four classes. Subsequently, we made the prediction of the PM25 variable for each of the clusters formed (Table 4).

Table 1. Prediction of average daily value for six pollution variables

Attribute	A %RMSE	B %RMSE	W %RMSE
PM25	0.00002079	0.00002051	0.00002092
PM10	0.00001861	0.00001846	0.00001888
SO2	0.00003233	0.00003203	0.00003371
NO2	0.00002248	0.00002238	0.00002277
CO	0.00002412	0.00002374	0.00002434
O3	0.00001709	0.00001703	0.00001717

Table 2. Prediction of average monthly value for six pollution variables

Attribute	A %RMSE	B %RMSE	W %RMSE
PM25	0.000307320	0.000293227	0.000341818
PM10	0.000305745	0.000298708	0.000329358
SO2	0.000626912	0.000533747	0.002499245
NO2	0.000523055	0.000501873	0.000581513
CO	0.000412227	0.000398324	0.000439063
O3	0.000328529	0.000316985	0.000395431

Table 3. Prediction of average daily PM2.5 concentration by clusters of multiple variables

Cluster	Monitoring sites	Sample	A %RMSE	B %RMSE	W %RMSE
C1	7	10227	0.00003446	0.00003407	0.00003527
C2	2	2922	0.00020571	0.00020172	0.00021161
C3	2	2922	0.00021293	0.00020702	0.00022527
C4	1	1461	0.00059876	0.00047015	0.00369635

Table 4. Prediction of average daily PM2.5 concentration by clusters of individual variables

Cluster	Monitoring sites	Sample	A %RMSE	B %RMSE	W %RMSE
C1	2	2922	0.00019933	0.00018991	0.00024125
C2	3	4383	0.00012389	0.00012092	0.00013315
C3	4	5844	0.00010090	0.00009914	0.00010343
C4	3	4383	0.00010871	0.00010702	0.00011159

Subsequently, we made the prediction of the average monthly variable PM25 variable for each of the clusters formed from the classification of monitoring sites based on six polluting variables (Table 5).

Table 5. Prediction of monthly PM2.5 concentration by clusters of multiple variables

Cluster	Monitoring sites	Sample	A %RMSE	B %RMSE	W %RMSE
C1	7	336	0.00050199	0.00046543	0.00074532
C2	2	96	0.00342633	0.00230991	0.00632787
C3	2	96	0.00377913	0.00307685	0.00633747
C4	1	48	0.00678047	0.00508235	0.01450819

Based on the clusters formed using competitive neural networks we made the prediction of the average monthly variable PM25 variable for each cluster (Table 6).

Table 6. Prediction of monthly PM2.5 concentration by clusters of individual variables

Cluster	Monitoring sites	Sample	A %RMSE	B %RMSE	W %RMSE
C1	2	96	0.00273402	0.00234010	0.00472917
C2	3	144	0.00202317	0.00157779	0.00410329
C3	4	192	0.00110263	0.00095638	0.00161259
C4	3	144	0.00183348	0.00146168	0.00321809

5 Discussion of Results

In this work we aim at separating the results for each of the clusters identified using competitive neural networks, with the idea of carrying out an appropriate decision-making process by data segments. Besides, we noted that for some data clusters, it is sought to almost reach the prediction of the values in comparison with the prediction made considering all the values of the time series, with the advantage that the prediction made would be specific to a particular area or place.

For the case of the clusters formed using the time series of the daily average of the multiple variables, we noted that 58% of the number of elements belong to cluster 1, with which, for the NAR neural network used with this class were obtained similar results to the prediction made with the complete data set, so that for the rest of clusters 2, 3 and 4, the %RMSE obtained was slightly higher compared to the results obtained with the complete data set. Similar results were obtained in the experiments carried out with the time series of the monthly average of the monthly variable PM25.

6 Conclusions

We have presented a model to test the combination of unsupervised and supervised neural networks for the clustering and the prediction of the particulate matter ($PM_{2.5}$) respectively. We identified by experimentation the monitoring sites with similar data (clusters), after the prediction of the particulate matter ($PM_{2.5}$) based on each one of the clusters identified was obtained.

Also, we identified the advantages of using multiple neural networks to predict time series, aiming out taking advantage of the two types of training (supervised and unsupervised) offered by these computational models, which allows us to analyze results for a particular place or area. It is important to consider the complexity of the problem to be solved and select the phase of the method that is required to be used since it is possible to operate each one individually or simultaneously.

As future works, we will consider the use of a self-organizing map (SOM) with other fuzzy logic techniques, looking for hybrid classification techniques, with the idea that time-series predictions can be generated for each of the geographical areas, communities, or industrial sectors.

References

1. Tsai, Y., Zeng, Y., Chang, Y.: Air Pollution Forecasting Using RNN with LSTM. In: 2018 IEEE 16th International Conference on Dependable, Autonomic and Secure Computing, 16th International Conference on Pervasive Intelligence and Computing, 4th International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), pp. 1074–1079 (2018). https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00178
2. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Executive summary, Geneva: World Health Organization, pp. 1–16 (2021). https://apps.who.int/iris/handle/10665/345334
3. Carreon-Ortiz, H., Valdez, F., Melin, P., Castillo, O.: Architecture optimization of a non-linear autoregressive neural networks for mackey-glass time series prediction using discrete mycorrhiza optimization algorithm. Micromachines **14**(1), 149 (2023). https://doi.org/10.3390/mi14010149
4. Castillo, O., Melin, P.: Forecasting of COVID-19 time series for countries in the world based on a hybrid approach combining the fractal dimension and fuzzy logic, Chaos, Solitons & Fractals, vol 140, (2020). https://doi.org/10.1016/j.chaos.2020.110242
5. Pulido, M., Melin, P.: Comparison of Genetic Algorithm and Particle Swarm Optimization of Ensemble Neural Networks for Complex Time Series Prediction. In: Melin, P., Castillo, O., Kacprzyk, J. (eds.) Recent Advances of Hybrid Intelligent Systems Based on Soft Computing. SCI, vol. 915, pp. 51–77. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-58728-4_3
6. Zhelev, S., Avresky, D.R.: Using LSTM neural network for time series predictions in financial markets. In: 2019 IEEE 18th International Symposium on Network Computing and Applications (NCA), pp. 1–5 (2019). https://doi.org/10.1109/NCA.2019.8935009
7. Ramirez, E., Melin, P., Prado-Arechiga, G.: Hybrid model based on neural networks, type-1 and type-2 fuzzy systems for 2-lead cardiac arrhythmia classification. Expert Syst. Appl. **126**, 295–307 (2019). https://doi.org/10.1016/j.eswa.2019.02.035

8. Moghar, A., Hamiche, M.: Stock market prediction using LSTM recurrent neural network. Proc. Comput. Sci. **170**, 1168–1173 (2020). https://doi.org/10.1016/j.procs.2020.03.049
9. Wei, D.: Prediction of stock price based on LSTM neural network. In: 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), pp. 544–547 (2019). https://doi.org/10.1109/AIAM48774.2019.00113
10. Patricia Melin,: Introduction to Type-2 Fuzzy Logic in Neural Pattern Recognition Systems. In: Patricia Melin, (ed.) Modular Neural Networks and Type-2 Fuzzy Systems for Pattern Recognition, pp. 3–6. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-24139-0_1
11. Cherif, A., Cardot, H., Boné, R.: SOM time series clustering and prediction with recurrent neural networks. Neurocomputing **74**(11), 1936–1944 (2011). https://doi.org/10.1016/j.neucom.2010.11.026
12. Melin, P., Castillo, O.: Spatial and temporal spread of the COVID-19 pandemic using self organizing neural networks and a fuzzy fractal approach. Sustainability **13**, 8295 (2021). https://doi.org/10.3390/su13158295
13. Ding, X., Hao, K., Cai, X., Tang, S., Chen, L., Zhang, H.: A novel similarity measurement and clustering framework for time series based on convolution neural networks. IEEE Access **8**, 173158–173168 (2020). https://doi.org/10.1109/ACCESS.2020.3025048
14. Hu, Y., Sun, X., Nie, X., Li, Y., Liu, L.: An enhanced LSTM for trend following of time series. IEEE Access **7**, 34020–34030 (2019). https://doi.org/10.1109/ACCESS.2019.2896621
15. Sánchez, D., Melin, P.: Modular Neural Networks for Time Series Prediction Using Type-1 Fuzzy Logic Integration. In: Melin, P., Castillo, O., Kacprzyk, J. (eds.) Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization. SCI, vol. 601, pp. 141–154. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-17747-2_11
16. Personal interventions and risk communication on air pollution: summary report of WHO expert consultation, 12–14 February 2019, Geneva, Switzerland. World Health Organization, pp. 9–10 (2020). https://apps.who.int/iris/handle/10665/333781
17. Zhang, S., Guo, B., Dong, A., He, J., Xu, Z., Chen, S.X.: 2017 Cautionary tales on air-quality improvement in Beijing. In: Proceedings of the Royal Society, pp. 1–14. A 473:20170457 (2017). https://doi.org/10.1098/rspa.2017.0457
18. Méndez, E., Lugo, O., Melin, P.: A Competitive Modular Neural Network for Long-Term Time Series Forecasting. In: Melin, P., Castillo, O., Kacprzyk, J. (eds.) Nature-Inspired Design of Hybrid Intelligent Systems. SCI, vol. 667, pp. 243–254. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-47054-2_16
19. Jang, J.-S., R, Mizutani E.: Unsupervised learning and other neural networks. In: Neuro-fuzzy and soft computing: a computational approach to learning and machine Intelligence, Prentice Hall, Upper Saddle River, pp. 301–310 (1997)
20. Melin, P., Monica, J.C., Sanchez, D., Castillo, O.: A new prediction approach of the COVID-19 virus pandemic behavior with a hybrid ensemble modular nonlinear autoregressive neural network. Soft Comput **27**, 2685–2694 (2023).https://doi.org/10.1007/s00500-020-05452-z
21. Mónica, J.C., Melin, P., Sánchez, D.: Genetic Optimization of Ensemble Neural Network Architectures for Prediction of COVID-19 Confirmed and Death Cases. In: Castillo, O., Melin, P. (eds.) Fuzzy Logic Hybrid Extensions of Neural and Optimization Algorithms: Theory and Applications. SCI, vol. 940, pp. 85–98. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68776-2_5

Overcoming the Challenges of Uncertainty in Forecasting Economic Time Series Through Convolutional Neural Networks and Other Intelligent Approaches

Dmytro Kovalenko(✉) and Dusan Marcek

Department of Applied Informatics, Faculty of Economics, VŠB– Technical University of Ostrava, Sokolská tř. 33, 70200 Ostrava, Czech Republic
kov0387@vsb.cz

Abstract. This article provides insights into the use of artificial neural networks (ANNs) and convolutional neural networks (CNNs) as the tools for forecasting economic time series, where uncertainty refers to incomplete information about the future. To improve the forecasting ability of CNN architectures and capture long-term dependencies in the input sequence we used the WaveNet models which dilate convolutions with skip connections in the input sequence. The residual blocks with skip connections are defined in a specific way that allows for easier information flow through the network while avoiding the vanishing gradient problem, making it a potential innovation in the field of deep learning. Another innovative aspect is the use one-hot encoding for the target sequences using categorical cross-entropy loss function.

Keywords: Neural networks · RNN · LSTM · CNN · WaveNet

1 Introduction

Neural networks have already found their place among other approaches, models or paradigms of artificial intelligence. Unlike more conventional approaches based on the symbolic representation of the information, the knowledge of a neural network is spread units constituting elements, similarly as in biological neural networks. On the other hand, it is sometimes difficult to assess what knowledge has been acquired by the network since the analysis of the internal representation is difficult.

To process data with spatiotemporal structure recurrent neural networks (RNNs) were suggested. RNNs were successfully applied in many real-life applications where processing time dependent information was necessary. Unlike feedforward neural networks, units in RNNs are fed by activities from previous time steps recurrent through connections. In this way, contextual information can be kept in units' activities, enabling RNNs to process time series. Common algorithms used for RNNs training are based on gradient minimization of the output error.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 515–522, 2023.
https://doi.org/10.1007/978-3-031-39777-6_61

The main goal of this research is to investigate the ability of alternative CNN architectures to enhance the accuracy of economic time series forecasting using WaveNet as a baseline model. The paper aims to demonstrate the potential of the WaveNet model in recognizing complex relationships within time series data, handling patterns across different time scales, and processing long-term dependencies with large time lags between dependent variables.

The paper also presents several novel approaches to improve the performance of the WaveNet model for forecasting economic time series. First, the use of logarithmic returns for normalization, which is suitable for statistical analysis, and the adoption of the WaveNet model to overcome the limitations of traditional CNNs, are unique contributions. Second, the employment of one-hot encoding for target sequences and the categorical cross-entropy loss function to predict multiple categories of logarithmic returns instead of a single continuous value is a novel approach, which is beneficial in financial time series forecasting. Third, the optimization of the model's hyperparameters, such as the number of filters, filter width, expansion coefficients, and dropout coefficient, and replacement of the categorical loss function with MSE, significantly improved the forecasting accuracy.

Section 2 provides a literature review. Section 3 provides the basic methodology of the deep learning for CNNs. Section 4 proposes innovative components of the WaveNet-based model. Section 5 introduces the data for our research. The methods for implementing learning algorithms, as well as several commercial procedures and software are used and described in Sect. 6. Experimental results and discussion are given in Sect. 7. Concluding remarks are given in Sect. 8.

2 The Literature Review

Standard RNNs trained by common gradient descend techniques have problems with processing time series containing long-time dependencies. To overcome this limitation, a novel architecture called long short-term memory (LSTM) networks was suggested, which can learn long-term dependencies in data and has been used in practical applications.

When propagating teaching signal, error tends to vanish or blow up [1], see also [2]. To overcome this limitation and to solve tasks traditional RNNs cannot solve, novel architecture was suggested. Long short-term memory (LSTM) networks are equipped with special units called constant error carrousels with self- recurrent feedback connection of constant weight set to 1 [3]. In this way, error signal can span theoretically infinite time distances. LSTM architecture has further developed: forgetting mechanism was introduced [4]. The problem of time series forecasting is solved in LSTM with the help of a memory cell referred to as the LSTM cell. This type of network possesses the capability to learn long-term dependencies in data, which shows its applicability to a time series problem. Practical applications of LSTM networks in robotics is discussed in [5].

CNN is another deep learning algorithm applied in stock market prediction after LSTM. In the study [6] Di Persio and Honchar's demonstrated the advantages of using CNN over other methods in forecasting stock market price movements of the S&P 500 index. The study established the effectiveness of CNNs in forecasting economic time series, particularly in the stock market.

In [7] the authors proposed a convolutional deep-learning-based network for foreign exchange rate, crude oil price and gold price predictions. The explanatory variables for this were exchange rate changes among PKR/USD, GBD/USD and HKD/USD currencies. The daily data of exchange rates ranging from early 2008 to late 2018 were utilized. The authors also added sentiment analysis values. Mean absolute error (MAE) and root mean squared error (RMSE) were used as evaluation metrics. The results were obtained through the proposed deep learning methodology with linear and support vector regression models. Their results show that deep learning-based methods perform better than other models.

3 The Proposed Methodology of CNNs

Our proposed system is based on a deep learning model. According to [8] all configurations of CNNs are based on the classical feedforward artificial neural network and therefore adopt most of its basic principles of structure, training and inference. These configurations rely on weight sharing, local receptive fields, and subsampling to extract features from input feature maps and reduce the number of free parameters, thereby enhancing the network's ability.

In general, CNN configurations consist of several types of layers, including the Input Layer that accepts data, such as images or time series data. The Convolutional Layer extracts various features from the input feature map by creating a convolved kernel with the input layer and forwarding the response to the next layers in the architecture. The Pooling Layer performs the merge operation, leading to a reduction in the dimensions of maps on other layers and a decrease in the number of synapses and free parameters. The Fully-Connected Layer serves as a classifier, performing the inner product of the input data vector and the transpose weight vector plus bias. The Rectified Linear Unit Layer increases the decision function's nonlinear properties by removing negative values from the activation map and converting them to zero, without activating the convolutional layers. The Dropout Layer helps prevent overfitting by randomly dropping out a certain percentage of neurons in the network during training. The Batch Normalization Layer normalizes the activations of the previous layer, enhancing the network's generalization performance and training speed. Finally, the Output Layer produces the final predictions.

Although CNNs are powerful tools for forecasting, they may encounter limitations when applied to economic time series. This is because CNNs assume that data points are independent, which does not hold for time series as they have temporal dependencies and order. As a result, processing long-term dependencies, especially with large time lags between dependent variables, can be challenging. Additionally, fixed-size convolution kernels may not be capable of recognizing patterns presented in different data time scales, resulting in information loss.

4 Innovative Components of the WaveNet-Based Model

Initially, WaveNet was represented by van den Oord et al. [9] (see also [10]) as a deep neural network originally designed for generating raw audio waveforms. Despite being a fully probabilistic and autoregressive model, the paper demonstrates that it can be efficiently trained on data with high temporal resolution. The model's ability to capture the characteristics of many different speakers and switch between them by conditioning on speaker identity has also potential applications in time series analysis. Additionally, the paper explores the use of WaveNet as a discriminative model, that can be applied to forecasting time series by predicting the class labels of the target variable, based on the input features.

WaveNet utilizes causal convolutions as its primary component, as depicted in Fig. 1. Causal convolutions guarantee that the model only utilizes past information to make predictions and is incapable of utilizing future information. Essentially, the prediction at time $t + 1$ can solely rely on the past information x_1 to x_t and cannot be based on any future information x_{t+1} to x_t. This is a crucial aspect when modeling time series data, where the future is unknown and cannot be used for making predictions.

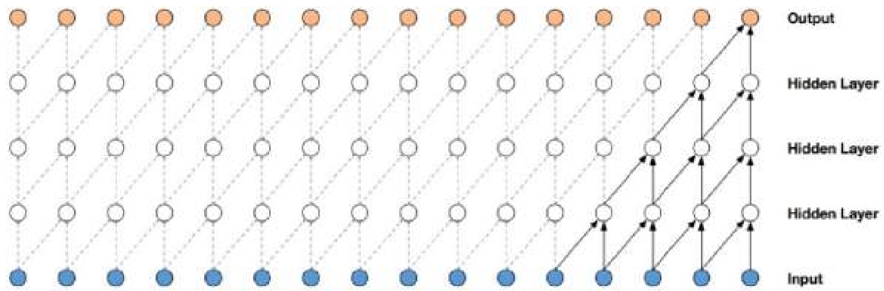


Fig. 1. Casual convolutional layer (Source [9]).

In WaveNet, the input historical time series data are transformed into a matrix that is equivalent to a causal masked convolution produced by multiplying the convolutional kernel with the mask tensor before application. The method of predicting the next value $\hat{x}(t + 1)$ of one-dimensional time series $x_{tt=0}^{N-1}$ conditional on the series' history, $x(0), \ldots., x(t)$ using convolutional neural network of the WaveNet architecture is based on maximizing the likelihood function (1) [10]

$$(x|\alpha)p = \prod_{t=0}^{N-1} p(x(t+1)|x(0), \ldots., x(t), \alpha) \tag{1}$$

The WaveNet network utilizes stacked layers of dilated convolutions to learn the likelihood function. Dilated convolutions apply a filter to every d_{th} element in the input vector, enabling the model to learn long-term dependencies between distant data points. The network consists of L layers of dilated convolutions with dilations increasing by a factor of two and filters of size 1×2. Each layer is followed by a non-linearity, resulting in output feature maps f^l, where $l = 1, \ldots., L$. The final layer is succeeded by a 1×1 convolution to decrease the number of channels back to one. The model is trained to predict a forecasted time series $\hat{x}_t$ based on the input time series x_t.

5 Used Data

We solve the problem of forecasting stock prices based on historical data. The input data for this problem consist of historical stock price data for Apple Inc. (AAPL), which was obtained from Yahoo Finance (Yahoo Finance, n.d.). Figure 2 demonstrate the historical daily prices for Apple Inc. Stocks by "Close" on NASDAQ from December 12, 1980, to December 17, 2022.

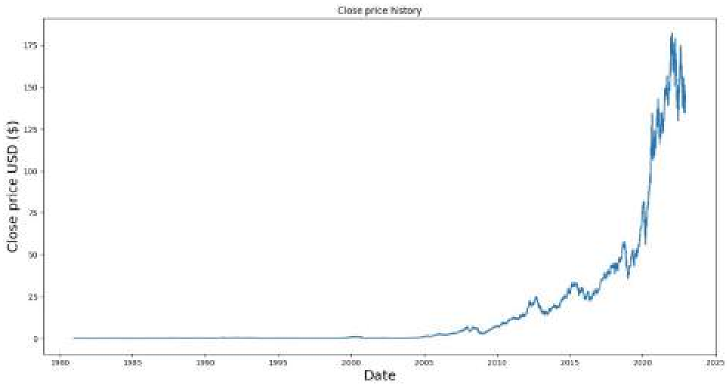


Fig. 2. The time plot of daily prices for Apple Inc. Stocks by "Close" from December 12, 1980 till December 17, 2022 (Source: own processing).

6 Proposed Solution

The basic model was built using the Keras and TensorFlow toolkit. According to Keras (n.d.), Keras is a deep learning API written in Python and runs on TensorFlow. Keras was developed with a focus on enabling fast experimentation. TensorFlow is an end-to-end platform that makes it easy for users to build and deploy Machine Learning models (TensorFlow, n.d.). Additionally, we used the following Python libraries: Yfinance, Pandas, Numpy, Scikit-Learn and Malplotlib.

The proposed solution uses the WaveNet model architecture for time series forecasting. It fetches historical stock data and drops any negative or zero closing prices from the dataset.

The log-returns for each day are calculated, which is the natural logarithm of the closing price of a day divided by the closing price of the previous day.

The input and label sequences of a certain length are generated using the function, which consists of consecutive log-returns. The input and label sequences are reshaped to have a 4D shape required by the WaveNet model, and the label sequences are one-hot encoded into binary classes.

The WaveNet model (see Fig. 3) is defined and uses dilated convolutions, skip connections, and residual connections to capture long-term temporal dependencies in time-series data. The model is compiled using Cross-entropy as the loss function and accuracy as the evaluation metric and trained on the training data. The model is evaluated on the test data using MAE and RMSE metrics. The accuracy of the model is calculated for each sequence and the overall accuracy is reported. The WaveNet model as shown in Fig. 3 takes an input sequence of log-returns and passes it through an input layer. The input sequence is then processed by a series of convolutional layers with dilated convolutions. The dilation rate is increased exponentially in each layer to allow the model to capture long-term temporal dependencies in the input sequence. Skip connections are added between the convolutional layers to help the model learn both short-term and long-term dependencies. The residual connections allow the model to retain information from previous layers and avoid the vanishing gradient problem. The output of the convolutional layers and skip connections is summed together to form the final activations. Dropout is applied to the final activations to prevent overfitting. The final activations are then passed through a convolutional layer and another activation function. The output layer uses a sigmoid activation function to predict the next log-return in the sequence. The model is optimized using the Adam optimizer.

7 Experimental Results and Discussion

The purpose of this section is to compare the prediction accuracy of applied CNN models with the traditional LSTM. In the experimental calculations we used LSTM, CNN, and their hybrid ConvLSTM models and evaluated their effectiveness using identical metrics, namely RMSE and MAE. As expected, the LSTM model showed the highest result in forecasting, with a RMSE value of 7.948 and MAE of 4.603.

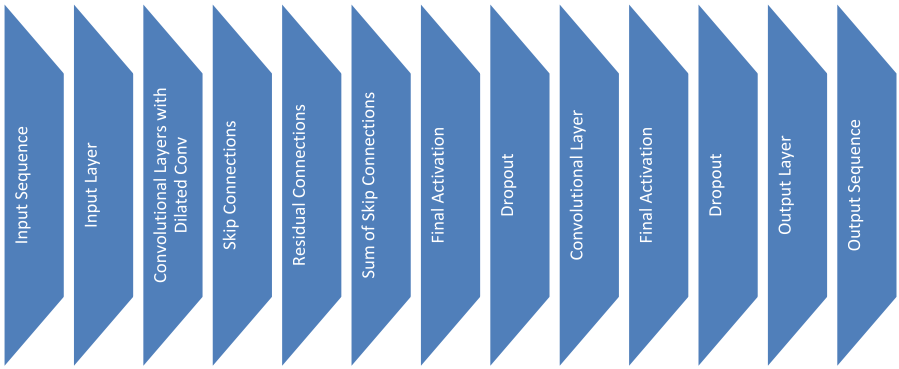


Fig. 3. Schematic diagram of the architecture WaveNet model (Source: own processing).

However, the basic CNN and ConvLSTM models with 2D convolution did not perform satisfactorily, as shown in Table 1. Despite this, we were able to reduce the value of the RMSE and MAE metrics by half by adjusting hyperparameters during previous experiments.

Table 1. The empirical statistical forecasting accuracy assessment of the presented CNN architectures (models)

Indicators / Metrics	RMSE	MAE
Indicators of the basic CNN model	86.461	71.445
Indicator of the best CNN model	43.842	37.265
Indicators of the basic ConvLSTM model	74.082	55.682
Indicator of the best ConvLSTM model	53.604	39.330
Indicators of the basic WaveNet model	0.160	0.148
Indicator of the best WaveNet model	0.002	0.001

The obtained RMSE and MAE values in Table 1 demonstrated the ability of CNN to predict daily time series and improve forecasting results. This provided the impetus to search for new solutions to further improve the performance of convolutional models.

To improve the performance of convolutional models, we used logarithmic returns to normalize the distribution of returns and make them more suitable for statistical analysis. We also choose to use WaveNet because it can handle patterns present in different time scales of the data, including long-term relationships with large time delays between dependent variables, while avoiding the vanishing gradient problem.

Furthermore, we used single coding for target sequences and a categorical cross-entropy loss function to enable the model to predict multiple categories of logarithmic returns instead of a single continuous value. This approach is particularly useful for financial time series, where changes in the direction or magnitude of returns can have a significant impact on investment decisions. The base model showed RMSE = 0.160, MAE = 0.1482.

To optimize the performance of the model, we tuned hyperparameters such as the number of filters in the convolutional layers, filter width, expansion coefficients, and dropout coefficient. We also replaced the categorical loss function with MSE, which significantly improved the metrics of RMSE = 0.0017 and of MAE = 0.0013.

8 Conclusion

This paper demonstrates the potential of the WaveNet model to improve the efficiency of forecasting economic time series using convolutional neural networks. The WaveNet architecture is capable of overcoming the limitations of traditional CNNs, specifically in recognizing complex relationships within time series data, handling patterns across different time scales, and processing long-term dependencies with large time lags between dependent variables. The use of dilated convolutions with skip connections, unique to WaveNet, enables the model to capture long-term dependencies while avoiding the vanishing gradient problem.

Moreover, employing one-hot encoding for target sequences and the categorical cross-entropy loss function allows the model to predict multiple categories of log-returns, which is particularly beneficial in financial time series forecasting. Our experiments indicate that WaveNet models can outperform traditional LSTM, CNN, and ConvLSTM models in terms of forecasting accuracy.

Future research will investigate the possibility of enhancing basic WaveNet models by adding a module that can predict multiple future time points. Additionally, incorporating relevant information such as economic indicators or sentiment analysis data may improve the model's forecast accuracy. Transfer learning will also be explored, whereby the model is pre-trained on a large dataset of similar time series data and fine-tuned for a specific forecasting task. Finally, WaveNet models require interpretation due to their complex architecture and numerous parameters.

References

1. Hochreiter, S., Bengio, Y., Frasconi, P., Schmidhuber, J.: Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In: Kolen, J. and Kremer, S., editors, Fielg Guide to Dynamic Recurrent Networks, pp. 237–243. Wiley-IEEE Press (2001)
2. Čerňanský, M.: Comparison of Recurrent Neural networks with Markov Models on Complex Symbolic Sequences. Disertation. Slovak University of Technology in Bratislava Faculty of Informatics and Information Technology, Bratislava (2006)
3. Hochreiter, J., Smidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
4. Gers, F.A., Smidhuber, J., Cummins, F.A.: Learning to forget: continual prediction with LSTM. Neural Comput. **12**(10) 2451–2471 (2000)
5. Bakker, B.: Reinforcement learning with LSTM. In: Advancens in Neural Information Processing Systems 13, pp. 1475–1482. MIT Press (2002)
6. Di Perisio, L., Honchar, O.: Artificial Neural Networks architectures for stock price prediction: comparisons and applications. Int. J. Circ., Syst. Signal Process. **10**, 403–413 (2016)
7. Yasir, M., et al.: An intelligent event-sentiment-based daily foreign exchange rate forecasting systém. Appl. Sci. **9**, 2980 (2019). https://doi.org/10.3390/app9152980
8. Madera, M., Marcek, D.: Intelligence in Finance and Economics for Predicting High-Frequency Data. https://www.mdpi.com/journal/mathematics
9. Van den Oord, A., et al.: WaveNet: A Generative Model for Raw Audio. (2016). https://arxiv.org/abs/1609.03499
10. Borovykh, A., Bohte, S., Oosterlee, C.W.: Conditional time series forecasting with convolutional neural networks (2017). https://doi.org/10.48550/arXiv.1703.04691

Statistics

InterCriteria Analysis as an Intelligent Tool for Intuitionistic Fuzzy Decision Making: Case Study of Statistics for Science, Technology and Information Society of Turkish Statistical Institute

Veselina Bureva(✉) and Sotir Sotirov

Laboratory of Intelligent Systems, "Prof. Dr. Assen Zlatarov" University, "Prof. Yakimov" Blvd., Burgas 8010, Bulgaria
{vbureva,ssotirov}@btu.bg

Abstract. In the current paper Science, Technology and Information Society datasets are analyzed. They are downloaded from Turkish Statistical Institute. The aim of the research is to determine unseen relationships between the enterprises, individuals and their usage of ICT facilities. InterCriteria analysis (ICA) is applied over the datasets to determine dependencies and independencies between the data. ICA is performed using ICrAData software. The results of ICA application are visualized into intuitionistic fuzzy triangle and those of the research determine dependencies between enterprises. The relationships between facilities provided via web pages are also analyzed.

Keywords: InterCriteria Analysis · Intuitionistic Fuzzy Pairs · Information and Communication Technologies

1 Introduction to InterCriteria Analysis (ICA) and Literature Review

InterCriteria Analysis (ICA) was introduced in 2014 by the Bulgarian scientists Krassimir Atanassov, Deyan Mavrov and Vassia Atanassova [3, 6, 14]. The technique is proposed as a decision making method based on the theories of index matrices [1] and intuitionistic fuzzy sets [2, 5]. ICA works on relations between objects on the base of certain criteria. The relationships between the objects are received in the form of intuitionistic fuzzy pairs. Depending on the degree of membership, various dependencies are determined: similarity, independence or uncertainty. In the theory of ICA the similarity between the objects is determined as positive consonance, the independence between the objects is called negative consonance and the uncertainty is considered as dissonance. Popular extensions of ICA are proposed in the time: ICA for triples [8], ICA over intuitionistic fuzzy data [10], Three Dimensional ICA [15], ICA using special type intuitionistic fuzzy implications [4]. Different algorithms for InterCriteria relations calculation, threshold values and intuitionistic fuzzy triangle interpretations are published in [7, 9].

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 525–531, 2023.
https://doi.org/10.1007/978-3-031-39777-6_62

ICA is successfully applied in the fields of artificial intelligence, medicine and healthcare, QSAR and molecular modelling, biotechnology, ecology, economics and finance, industry, education and sports. Depending on the field of application ICA determines the tendencies between the selected objects or the behavior between the criteria. In series of papers university rankings of different countries are made. Relationships between Bulgarian indicators are determined. Dependencies between Bulgarian universities in the area of communication and computer technologies are analyzed. Afterwards, ICA is applied to university rankings of Poland, Turkey, Slovakia, Australia, India and United Kingdom. The behavior of the indicators in the different countries is observed. A universities segmentation in the selected countries is made [11].

ICA is also used to analyze the air quality, the quality of live, quality of water. In these research works parameters for the quality of air, quality of water and live quality are studied. Dependency between different groups of data or relations between different towns, countries is observed. Accurate parameters are selected after the ICA applications. Study of bulk properties and fraction properties of crude oils are analyzed using ICA [11].

ICA is also applied to study over various EU enterprises. The results are already published in [12]. In the current paper web facilities usages in the enterprises are analyzed for the first time. The aim is to present the degree of application of the new ICT technologies. The connection between appropriate fields of the industry is studied. The usage of different types of technologies is compared.

The paper has the following structure. Section 1 presents a brief literature review of ICA. Section 2 contains two applications of the ICA. The Datasets for Science, Technology and Information Society in Turkey are discussed. The first application determines dependencies between enterprises. The second application of ICA finds unseen relationships between the ICT technologies used in the enterprises. The research is made for the years 2020 and 2021. In Sect. 3 some conclusion remarks are discussed.

2 InterCriteria Analysis Applied to the Datasets for Science, Technology and Information Society in Turkey

Science, Technology and Information Society datasets (2020, 2021) are downloaded from the website of Turkish Statistical Institute [16]. Thereafter, ICA is applied to find possible relationships between facilities and websites of enterprises. In the current research the following abbreviations of the enterprises are used:

- E_1 - Electricity, gas and steam, water supply, sewerage and waste management;
- E_2 – Construction;
- E_3 - Wholesale and retail trade; repair of motor vehicles and motorcycles;
- E_4 - Transportation and storage;
- E_5 - Accommodation and food service activities;
- E_6 - Information and communication;
- E_7 - Real estate activities;
- E_8 - Professional, scientific and support activities;
- E_9 - Administrative and support activities;
- E_{10} - Repair of computers.

The websites and facilities provided via web pages have the following abbreviations:

- F_1 - Enterprises having websites;
- F_2 - Description of goods or services and price lists;
- F_3 - Online ordering or reservation or booking;
- F_4 - Possibility for visitors to customize or design online goods or services;
- F_5 - Tracking or status of orders placed;
- F_6 - Personalized content in the website for regular/recurrent visitors;
- F_7 - Links or references to the enterprise's social media profiles.

The datasets for enterprises having websites and facilities provided via web pages by economic activity and size group are used for the current work. The aim of the ICA application is to find possible correlations or independencies between the enterprises. The ICA applications are made using ICrAData software [13]. The outcomes of applying ICA over ICT data are presented in the form of intuitionistic fuzzy pairs with degrees of membership and degrees of non-membership. The intuitionistic fuzzy pairs, received from ICA application over datasets for websites and facilities, are presented in Fig. 1.

$\langle\mu,\nu\rangle$	F_1	F_2	F_3	F_4	F_5	F_6	F_7
F_1	$\langle 1.00, 0.00\rangle$	$\langle 0.62, 0.36\rangle$	$\langle 0.49, 0.51\rangle$	$\langle 0.36, 0.60\rangle$	$\langle 0.44, 0.56\rangle$	$\langle 0.36, 0.64\rangle$	$\langle 0.53, 0.47\rangle$
F_2	$\langle 0.62, 0.36\rangle$	$\langle 1.00, 0.00\rangle$	$\langle 0.62, 0.36\rangle$	$\langle 0.47, 0.47\rangle$	$\langle 0.62, 0.36\rangle$	$\langle 0.44, 0.53\rangle$	$\langle 0.44, 0.53\rangle$
F_3	$\langle 0.49, 0.51\rangle$	$\langle 0.62, 0.36\rangle$	$\langle 1.00, 0.00\rangle$	$\langle 0.80, 0.16\rangle$	$\langle 0.96, 0.04\rangle$	$\langle 0.69, 0.31\rangle$	$\langle 0.78, 0.22\rangle$
F_4	$\langle 0.36, 0.60\rangle$	$\langle 0.47, 0.47\rangle$	$\langle 0.80, 0.16\rangle$	$\langle 1.00, 0.00\rangle$	$\langle 0.78, 0.18\rangle$	$\langle 0.64, 0.31\rangle$	$\langle 0.82, 0.13\rangle$
F_5	$\langle 0.44, 0.56\rangle$	$\langle 0.62, 0.36\rangle$	$\langle 0.96, 0.04\rangle$	$\langle 0.78, 0.18\rangle$	$\langle 1.00, 0.00\rangle$	$\langle 0.69, 0.31\rangle$	$\langle 0.73, 0.27\rangle$
F_6	$\langle 0.36, 0.64\rangle$	$\langle 0.44, 0.53\rangle$	$\langle 0.69, 0.31\rangle$	$\langle 0.64, 0.31\rangle$	$\langle 0.69, 0.31\rangle$	$\langle 1.00, 0.00\rangle$	$\langle 0.60, 0.40\rangle$
F_7	$\langle 0.53, 0.47\rangle$	$\langle 0.44, 0.53\rangle$	$\langle 0.78, 0.22\rangle$	$\langle 0.82, 0.13\rangle$	$\langle 0.73, 0.27\rangle$	$\langle 0.60, 0.40\rangle$	$\langle 1.00, 0.00\rangle$

Fig. 1. Result of ICA application for websites and facilities relationships study, 2020 - Intuitionistic Fuzzy Pairs between the Websites and Facilities

ICA identifies the following relationships between the websites and facilities provided via web pages for 2021 (Fig. 2 and Fig. 4): 1 pair of websites and facilities provided via web pages in strong positive consonance, 2 pairs of websites and facilities provided via web pages in weak positive consonance, 6 pairs of websites and facilities provided via web pages in weak dissonance, 8 pairs of websites and facilities provided via web pages in dissonance, 3 pairs of websites and facilities provided via web pages in strong dissonance and 1 pair of websites and facilities provided via web pages in weak negative consonance in the ICT dataset for 2021.

ICA identifies the following relationships between the websites and facilities provided via web pages for 2020 (Fig. 2 and Fig. 4): 1 pair of websites and facilities provided via web pages in strong positive consonance, 4 pairs of websites and facilities provided via web pages in weak positive consonance, 3 pairs of websites and facilities provided via web pages in weak dissonance, 7 pairs of websites and facilities provided via web pages in dissonance and 6 pairs of websites and facilities provided via web pages in strong dissonance.

The results are interpreted into the intuitionistic fuzzy triangle. The left IF-triangle presents results for websites and facilities provided via web pages in 2020. The right

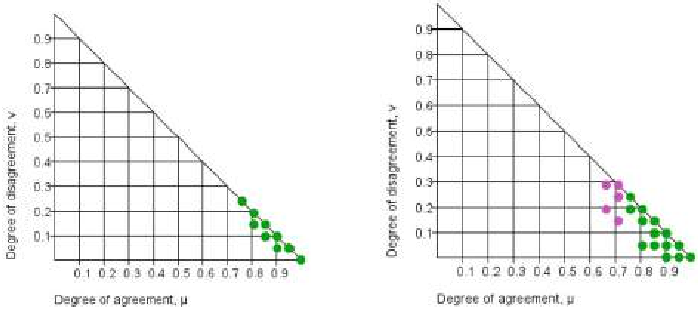


Fig. 2. Websites and Facilities Relationships interpreted into the Intuitionistic Fuzzy Triangle – 2020 (left) and 2021 (right)

IF-triangle presents results for websites and facilities provided via web pages in 2021 (Fig. 3).

Obviously, the behavior between the facilities is mostly concentrated in the area of weak dissonance, dissonance and strong dissonance. We can conclude that most facilities are independent. The pairs with dependencies (2020–2021) have the following form:

- The pair of facilities "*Online ordering or reservation or booking - Tracking or status of orders placed*" is in strong positive consonance in the years 2020 and 2021. The relation between these two facilities is strong and they can be merged in a generalized facility.
- The pairs "*Online ordering or reservation or booking - Possibility for visitors to customize or design online goods or services*" and "*Possibility for visitors to customize or design online goods or services - Tracking or status of orders placed*" are in weak positive consonance in years 2020 and 2021. The dependency between the pairs is weak and constant in time.

The second application of ICA determines the type of the relations between different types of enterprises. The intuitionistic fuzzy pairs of ICA application for 2020 are presented in Fig. 3. ICA is applied using ICrAData software [12]. The following results are obtained (Fig. 4 and Fig. 5):

- 9 pairs of enterprises in strong positive consonance, 17 pairs of enterprises in positive consonance, 10 pairs of enterprises in weak positive consonance and 9 pairs of enterprises in weak dissonance for 2021.
- 21 pairs of enterprises in strong positive consonance, 12 pairs of enterprises in positive consonance and 12 pairs of enterprises in weak positive for 2020.

Obviously, the behavior between enterprises is mostly concentrated in the area of strong positive consonance, positive consonance and weak positive consonance. We can conclude that most enterprises have similar websites and facilities provided via web pages.

The pairs in strong positive consonance (2020 and 2021) have the following form: "*Electricity, gas and steam, water supply, sewerage and waste management – Construction*", "*Electricity, gas and steam, water supply, sewerage and waste management - Real*

$\langle\mu, \nu\rangle$	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
E_1	⟨1.00, 0.00⟩	⟨0.95, 0.05⟩	⟨0.71, 0.24⟩	⟨0.90, 0.05⟩	⟨0.67, 0.29⟩	⟨0.90, 0.10⟩	⟨0.95, 0.05⟩	⟨1.00, 0.00⟩	⟨0.86,0.10⟩	⟨0.90, 0.10⟩
E_2	⟨0.95, 0.05⟩	⟨1.00, 0.00⟩	⟨0.76, 0.19⟩	⟨0.95, 0.00⟩	⟨0.71, 0.24⟩	⟨0.86, 0.14⟩	⟨1.00, 0.00⟩	⟨0.95, 0.05⟩	⟨0.90, 0.05⟩	⟨0.86,0.05⟩
E_3	⟨0.71, 0.24⟩	⟨0.76, 0.19⟩	⟨1.00, 0.00⟩	⟨0.76, 0.24⟩	⟨0.95, 0.05⟩	⟨0.81, 0.14⟩	⟨0.76, 0.19⟩	⟨0.71, 0.24⟩	⟨0.86, 0.14⟩	⟨0.71, 0.14⟩
E_4	⟨0.90, 0.05⟩	⟨0.95, 0.00⟩	⟨0.76, 0.24⟩	⟨1.00, 0.00⟩	⟨0.71, 0.29⟩	⟨0.81, 0.14⟩	⟨0.95, 0.00⟩	⟨0.90, 0.05⟩	⟨0.90, 0.10⟩	⟨0.81, 0.05⟩
E_5	⟨0.67, 0.29⟩	⟨0.71, 0.24⟩	⟨0.95, 0.05⟩	⟨0.71, 0.29⟩	⟨1.00, 0.00⟩	⟨0.76, 0.19⟩	⟨0.71, 0.24⟩	⟨0.67, 0.29⟩	⟨0.81, 0.19⟩	⟨0.67, 0.19⟩
E_6	⟨0.90, 0.10⟩	⟨0.86, 0.14⟩	⟨0.81, 0.14⟩	⟨0.81, 0.14⟩	⟨0.76, 0.19⟩	⟨1.00, 0.00⟩	⟨0.86, 0.14⟩	⟨0.90, 0.10⟩	⟨0.86, 0.10⟩	⟨0.90, 0.00⟩
E_7	⟨0.95, 0.05⟩	⟨1.00, 0.00⟩	⟨0.76, 0.19⟩	⟨0.95, 0.00⟩	⟨0.71, 0.24⟩	⟨0.86, 0.14⟩	⟨1.00, 0.00⟩	⟨0.95, 0.05⟩	⟨0.90, 0.05⟩	⟨0.86, 0.05⟩
E_8	⟨1.00, 0.00⟩	⟨0.95, 0.05⟩	⟨0.71, 0.24⟩	⟨0.90, 0.05⟩	⟨0.67, 0.29⟩	⟨0.90, 0.10⟩	⟨0.95, 0.05⟩	⟨1.00, 0.00⟩	⟨0.86, 0.10⟩	⟨0.90, 0.00⟩
E_9	⟨0.86,0.10⟩	⟨0.90, 0.05⟩	⟨0.86, 0.14⟩	⟨0.90, 0.10⟩	⟨0.81, 0.19⟩	⟨0.86, 0.10⟩	⟨0.90, 0.05⟩	⟨0.86, 0.10⟩	⟨1.00, 0.00⟩	⟨0.81,0.05⟩
E_{10}	⟨0.90, 0.10⟩	⟨0.86,0.05⟩	⟨0.71, 0.14⟩	⟨0.81, 0.05⟩	⟨0.67, 0.19⟩	⟨0.90, 0.00⟩	⟨0.86, 0.05⟩	⟨0.90, 0.00⟩	⟨0.81,0.05⟩	⟨1.00, 0.00⟩

Fig. 3. Result of ICA application for finding enterprises relationships, 2020 - Intuitionistic Fuzzy Pairs between the Enterprises

estate activities", "*Transportation and storage - Real estate activitie*s", "*Electricity, gas and steam, water supply, sewerage and waste management -Professional, scientific and support activities*","*Construction - Professional, scientific and support activities*", "*Real estate activities - Professional, scientific and support activities*".

The pairs in positive consonance (2020 and 2021) have the following form: "*Wholesale and retail trade; repair of motor vehicles and motorcycles - Administrative and support activities*", "*Electricity, gas and steam, water supply, sewerage and waste management - Repair of Computers*", "*Information and communication - Repair of computers*", "*Professional, scientific and support activities - Repair of computers*".

The pairs in positive consonance (2020 and 2021) have the following form: "*Accommodation and food service activities - Information and communication*", "*Accommodation and food service activities - Administrative and support activities*", "*Transportation and storage - Repair of Computers*", "*Administrative and support activities - Repair of Computers*".

The dependencies of the pairs of enterprises in strong positive consonance, positive consonance and weak positive consonance in the years 2020 and 2021 are constant. The relationships between the types of enterprises are visible in each of the observed years.

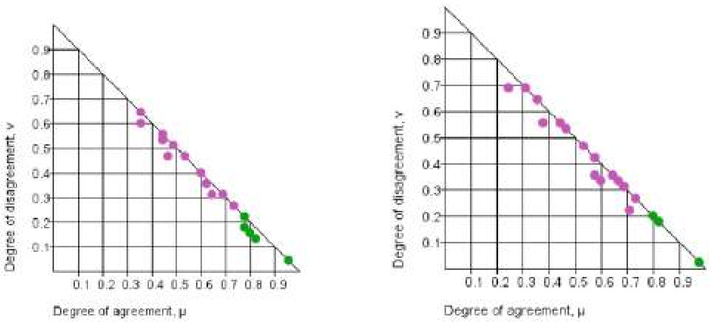


Fig. 4. Results of ICA application over the relationships between the enterprises – 2020 (left) and 2021 (right)

The number of intuitionistic fuzzy pairs from ICA application over the datasets for enterprises having websites and facilities provided via web pages by economic activity are presented in Fig. 5. The data for two years is observed: 2020 and 2021.

Type of correlations	*Number of pairs of criteria/facilities in 2021*	*Number of pairs of criteria/facilities in 2020*	*Number of pairs of enterprises in 2021*	*Number of pairs of enterprises in 2020*
***strong positive consonance* [0,95; 1,00]**	1	1	9	21
***positive consonance* [0,85; 0,95)**	-	-	17	12
***weak positive consonance* [0,75; 0,85)**	2	4	10	12
***weak dissonance* [0,67; 0,75)**	5	3	9	-
***dissonance* [0,57; 0,67)**	6	5	-	-
***strong dissonance* [0,43; 0,57)**	3	6	-	-
***dissonance* [0,33; 0,43)**	2	2	-	-
***weak dissonance* [0,25; 0,33)**	1	-	-	-
***weak negative consonance* [0,15;0,25)**	1	-	-	-
***negative consonance* [0,05;0,15)**	-	-	-	-
***strong negative consonance* [0,00;0,05)**	-	-	-	-

Fig. 5. Results from ICA applied to the datasets to study the relationships between the facilities and the relationships between the enterprises

3 Conclusion

In the current paper a study of the datasets for enterprises having websites and facilities provided via web pages (years: 2020, 2021) is presented. The ICA is applied over the data to determine the relations between the types of enterprises that use websites and facilities provided via web pages. A second application of ICA determines dependencies between the websites and facilities provided via web pages (years: 2020, 2021). In the future research works web technologies used in the enterprises in different countries will be a subject of research. The obtained results will be compared.

Acknowledgments. The authors are thankful for the support provided by the European Regional Development Fund and the Operational Program "Science and Education for Smart Growth" under contract UNITe No. BG05M2OP001–1.001–0004-C01 (2018–2023).

References

1. Atanassov, K.T.: Index Matrices: Towards an Augmented Matrix Calculus. Springer International Publishing, Cham (2014)
2. Atanassov, K.: On Intuitionistic Fuzzy Sets Theory. Springer, Berlin (2012)
3. Atanassov, K., Mavrov, D., Atanassova, V.: InterCriteria decision making: a new approach for multicriteria decision making, based on index matrices and intuitionistic fuzzy sets. Issues Intuitionistic Fuzzy Sets Generalized Nets **11**, 1–8 (2014)

4. Atanassov, K., Ribagin, S., Sotirova, E., Bureva, V., Atanassova, V., Angelova, N.: Intercriteria analysis using special type of intuitionistic fuzzy implications. Notes on Intuitionistic Fuzzy Sets **23**(5), 61–65 (2017)
5. Atanassov, K., Szmidt, E., Kacprzyk, J.: On intuitionistic fuzzy pairs. Notes Intuitionistic Fuzzy Sets **19**(3), 1–13 (2013)
6. Atanassov, K., Atanassova, V., Gluhchev, G.: InterCriteria analysis: ideas and problems, Notes Intuitionistic Fuzzy Sets **21**(1), 81–88 (2015)
7. Atanassova, V., Mavrov, D., Doukovska, L., Atanassov, K.: Discussion on the threshold values in the InterCriteria Decision Making approach. Int. J. Notes on Intuitionistic Fuzzy Sets **20**(2), 94–99, ISSN 1310–4926 (2014)
8. Atanassova, V., Doukovska, L., Michalikova, A., Radeva, I.: Intercriteria analysis: From pairs to triples. Notes Intuitionistic Fuzzy Sets, **22**(5), ISSN:Print ISSN 1310–4926; Online ISSN 2367–8283, 98–110 (2016)
9. Atanassova, V.: Interpretation in the Intuitionistic Fuzzy Triangle of the Results, Obtained by the InterCriteria Analysis, 16th World Congress of the International Fuzzy Systems Association (IFSA), 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), 30. 06–03. 07: Gijon. Spain **2015**, 1369–1374 (2015). https://doi.org/10.2991/ifsa-eusflat-15.2015.193
10. Bureva, V., Sotirova, E., Atanassova, V., Angelova, N., Atanassov, K.: Intercriteria Analysis over Intuitionistic Fuzzy Data. In: Lirkov, I., Margenov, S. (eds.) LSSC 2017. LNCS, vol. 10665, pp. 333–340. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73441-5_35
11. Chorukova, E., Marinov, P., Umlenski, I.: Survey on Theory and Applications of InterCriteria Analysis Approach. In: Atanassov, K.T. (ed.) Research in Computer Science in the Bulgarian Academy of Sciences. SCI, vol. 934, pp. 453–469. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72284-5_20
12. Doukovska, L., Atanassova, V., Shahpazov, G., Capkovic, F.: InterCriteria analysis applied to variuos EU enterprises. In: Proceedings of the Fifth International Symposium on Business Modeling and Software Design, Milan, Italy, 2015, 284–291 (2015)
13. Ikonomov N., Vassilev, P., Roeva, O.: ICrAData – software for intercriteria analysis. INT.J.BIOAUTOMATION **22**(1), 1–10 (2018)
14. InterCriteria Analysis Portal, Information. Available online at: https://intercriteria.net/dfni-i-02-5-2014-en/info/
15. Traneva, V., Tranev, S., Szmidt, E., Atanassov, K.: Three Dimensional Intercriteria Analysis over Intuitionistic Fuzzy Data. In: Kacprzyk, J., Szmidt, E., Zadrożny, S., Atanassov, K.T., Krawczak, M. (eds.) IWIFSGN/EUSFLAT -2017. AISC, vol. 643, pp. 442–449. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-66827-7_40
16. Turkish Statistical Institute. Science, Technology and Information Society. Retrieved March 31, 2022 from: https://data.tuik.gov.tr/Kategori/GetKategori?p=Science,-Technology-and-Information-Society-102

Generating High-Quality Prediction Intervals for Regression Tasks via Fuzzy C-Means Clustering-Based Conformal Prediction

Saleh Msaddi and Tufan Kumbasar(✉)

Control and Automation Engineering Department, Istanbul Technical University, Istanbul, Turkey
{msaddi18,kumbasart}@itu.edu.tr

Abstract. Accurately assessing uncertainty and prediction of a regression model is essential for making informed decisions, especially in high-risk tasks. Conformal Prediction (CP) is a powerful distribution-free uncertainty quantification framework for building such models as it is capable to transform a single-point prediction of any machine learning model into a Prediction Interval (PI) with a guarantee of encompassing the true value for specified levels of confidence. On the other hand, to generate high-quality PIs, the PIs should be as narrow as possible while enveloping a certain amount of uncertainty (i.e. confidence level). The generated width of the PIs mainly depends on the nonconformity measure used within the CP. In this study, we propose two novel Fuzzy c-Means Clustering (FCM) based nonconformity measures for CP with nearest neighbors to learn distribution-free and high-quality PIs for regression. The proposed approach generates tight PIs by evaluating the degree of nonconformity of a new data point compared to the so-called calibration points via Fuzzy Sets (FSs). From the calibration dataset, we extract representative FSs via FCM and assign every test point alongside the nearest neighbors within the calibration dataset with membership grades to adapt the nonconformity measure. To evaluate the performance, we present statistical comparisons and demonstrate that the proposed FCM-based nonconformity measures result in high-quality PIs.

Keywords: Conformal Prediction · Uncertainty Quantification · Fuzzy Clustering

1 Introduction

In recent years, we have witnessed the huge success of Machine learning (ML) in various applications. However, most of the studies focus on point-wise estimation which might not be sufficient in high-risk tasks as there is a need for assessment of the reliability of the prediction [1].

This work was supported by the BAGEP Award of the Science Academy.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 532–539, 2023.
https://doi.org/10.1007/978-3-031-39777-6_63

Conformal Prediction (CP) is a distribution-free method that is capable of transforming a single-point prediction of any ML model into a Prediction Interval (PI) with a guarantee of enveloping the true value for specified levels of confidence [4,5,8]. In the CP literature, the Inductive CP (ICP) is mainly preferred due to its computational efficiency in comparison to the transductive CP [9]. ICP requires a training set to train an ML model, and a disjoint calibration set to compute the nonconformity scores to define the width of PIs. In ICP, a nonconformity measure assesses the degree of difference of a new example from a set of the seen points. Given that all conformal predictors are valid, the main research focus is on defining nonconformity measures that generate high-quality PIs, i.e. informative and tight PIs, [3,7,10–12].

In this paper, we introduce two novel nonconformity measures for ICP Regression (ICPR) based on Fuzzy c-Means Clustering (FCM) and the residual information of the $k-$Nearest Neighbors ($k-$NNs). We use FCM to cluster the attributes of the calibration set and extract representative Fuzzy Sets (FSs). We modify the nonconformity measure in [3] to introduce a notion of the importance of $k-$NNs using FSs. The first nonconformity measure compares the similarity of the $k-$NNs to each cluster center. Since FCM is a distance-based clustering algorithm, this provides more importance to the $k-$NNs near the cluster center and less importance to those farther. The second nonconformity measure takes into account the similarity of the $k-$NNs to the cluster centers and the test point. Thus, giving more attention to the $k-$NNs that share the same properties of the test point. An empirical investigation on publicly available datasets is performed to illustrate the efficiency of the proposed approach compared to the standard nonconformity measure used in ICPR. The results indicate that both FCM-based nonconformity measures result in high-quality PIs, guaranteeing coverage and tightness.

The rest of this paper is organized as follows. Section 2 introduces the proposed FCM-Based Conformal Regression. Section 3 presents comparative results and discussions. Section 4 gives the driven conclusions and future work.

2 Inductive Conformal Regression via FCM

In this section, we first provide a brief introduction to FCM and an explanation of the ICPR framework. Then, we discuss the construction and implementation of two novel nonconformity measures for ICPR based on FCM.

2.1 Brief Overview on FCM

FCM [2], is a popular method for clustering data into groups based on their similarity, allowing for partial membership of a data point to multiple clusters. Given a finite set $X \subset \mathbb{R}^p$, $X = \{x_1, \dots, x_N\}$ and an integer c, such that $2 \leq c \leq N$, a fuzzy c partition of X is a $c \times N$ fuzzy partition matrix $U = [u_{ij}]$ whose entries satisfy $u_{ij} \in [0, 1], \forall i, j, \sum_{i=1}^{c} u_{ij} = 1$, and $\sum_{j=1}^{N} u_{ij} > 0, \forall i$.

Given $U \in M_{fc}$, where M_{fc} denotes the set of fuzzy c partitions of X, and the vector of cluster centers $V = (v_1, v_2, \ldots, v_c)$ such that $v_i \in \mathbb{R}^p$, the fuzzy c-means functional $J_m : M_{fc} \times \mathbb{R}^{cp} \to \mathbb{R}^+$ is defined as:

$$J_m(U, v) = \sum_{j=1}^{N} \sum_{i=1}^{c} (u_{ij})^m (d_{ij}^2) \tag{1}$$

where $m \in [1, \infty)$ is the fuzziness parameter and $d_{ij}^2 = \|x_j - v_i\|^2$. The FCM algorithm performs iterative optimization of the functional J_m to produce a fuzzy c partition of X and computes u_{ij} as follows:

$$u_{ij} = 1/\sum_{n=1}^{c} \left(\frac{d_{ij}}{d_{in}}\right)^{2/m-1} \tag{2}$$

2.2 Preliminaries on ICPR

Here, we provide the outline of the ICPR framework; a thorough discussion and analysis of the validity of CP and ICPR can be found in [9,13]. We are given an exchangeable dataset consisting of a training set $\mathcal{D} = \{(x_1, y_1), \ldots, (x_l, y_l)\}$ and a test set $\mathcal{D}_{ts} = \{x_{l+1}, \ldots, x_{l+r}\}$, where $x_j \in \mathbb{R}^p$ is feature vector and y_i is the target output. Given a test point $x_{l+1} \in \mathcal{D}_{ts}$, the task is to generate a PI that includes y_{l+1} with a confidence level κ.

The ICPR framework is presented in Algorithm 1, where q is the number of calibration samples. The common and standard common nonconformity measure for ICPR is the absolute error (residual) between the true output value y_j and the predicted output value $\hat{y}_j = f(x_j), \forall (x_j, y_j) \in \mathcal{D}_{cal}$ defined as [9]:

$$\psi_j = |y_j - \hat{y}_j| \tag{3}$$

While the PIs generated by (3) are guaranteed to provide the desired level of coverage, they do not adapt to the difficulty level of predicting a sample. In practical settings, we prefer having wider PIs for challenging samples and tighter PIs for easier samples. To generate adaptive PIs, the nonconformity measure (3) is normalized with a term expressing the difficulty level of predicting a sample. A general normalized nonconformity measure is expressed as follows:

$$\psi_j = \frac{|y_j - \hat{y_j}|}{\phi_j + \theta} \tag{4}$$

where ϕ_j is the difficulty estimator and $\theta \geq 0$ is a parameter used to control the sensitivity of the nonconformity measure [8].

2.3 FCM-Based ICPR: Fuzzified Nonconformity Measures

To generate high-quality PIs, we propose two novel normalized nonconformity measures that weight the residuals of the $k-$NNs via FSs generated via FCM. For both measures:

Algorithm 1. Standard ICPR

Input: $\mathcal{D} = \{(x_1, y_1), \ldots, (x_l, y_l)\}$, $\mathcal{D}_{ts} = \{x_{l+1}, \ldots, x_{l+r}\}$, and q

Initialization:

$h := l - q$

$S := \{\}$

Split $\mathcal{D}$ into two disjoint sets: proper training set, $\mathcal{D}_{tr} = \{(x_1, y_1), \ldots, (x_h, y_h)\}$ calibration set, $\mathcal{D}_{cal} = \{(x_{h+1}, y_{h+1}), \ldots, (x_l, y_l)\}$

Training Phase:

Train $f(.)$ on $\mathcal{D}_{tr}$

Calibration Phase:

for $j = 1$ `to` q **do**

    calculate $\hat{y}_{h+j} = f(x_{h+j})$

    calculate α_{h+j} using (3)

    store α_{h+j} in S

end for

Sort S in descending order to obtain $S = (\alpha_1, \ldots, \alpha_q)$

get α_s^δ where $s = \lfloor \delta(q+1) \rfloor$

Testing Phase:

for $g = 1$ `to` r **do**

    calculate $\hat{y}_{l+g} = f(x_{l+g})$

    Output: the PI using $(\hat{y}_{l+g} - \alpha_s^\delta, \hat{y}_{l+g} + \alpha_s^\delta)$

end for

- we extract FSs by clustering the feature space $X_{cal} \subset \mathcal{D}_{cal}$. Thus, through learned V_{cal}, we are capable to calculate $U_{cal} = [u_{ij}]$ of X_{cal}.
- we define the $k-$NNs of a test point x_{l+1} as the k closest points to x_{l+1} in X_{cal} based on the Euclidian distance, denoted as $\{x_{l+1}^1, \ldots, x_{l+1}^k\}$.
- we provide a notion of importance to the residuals of the $k-$NNs by assessing the degree of similarity of x_{l+1} and its $k-$NNs with respect to V_{cal} via two novel fuzzified nonconformity measures.

The steps for implementing the fuzzified nonconformity measures are explained below and summarized in Algorithm 2.

Calibration Phase; given that $f(.)$ is trained on $\mathcal{D}_{tr}$, the absolute residuals are calculated $\forall (x_j, y_j) \in \mathcal{D}_{cal}$ as follows:

$$r_j = |y_j - \hat{y}_j| \tag{5}$$

- The first nonconformity measure that we propose is based on the similarity of the $k-$NNs with respect to V_{cal}. For a given v_i, the residuals are weighted by the membership grades of the $k-$NNs to give more importance to the neighbors closer to v_i, and less importance to those farther. The nonconformity measure is then taken as the average similarity with respect to V_{cal} as follows:

$$\phi_j = \frac{1}{c} \sum_{i=1}^{c} \frac{\sum_{n=1}^{k} u_{ij}^n r_j^n}{\sum_{n=1}^{k} u_{ij}^n} \tag{6}$$

For example, given two cluster centers v_1 and v_2, a point z and its $k-$NNs $\{e, f, g\}$. The nonconformity score of z is calculated as:

$$\phi_z = \frac{1}{2}\left(\frac{u_{1e}r_e + u_{1f}r_f + u_{1g}r_g}{u_{1e} + u_{1f} + u_{1g}} + \frac{u_{2e}r_e + u_{2f}r_f + u_{2g}r_g}{u_{2e} + u_{2f} + u_{2g}}\right)$$

– The second nonconformity measure is a modification of (6). Besides evaluating the similarity of the $k-$NNs with respect to V_{cal}, we suggested accounting for the similarity of the $k-$NNs with respect to x_j. In this case, we draw more attention to the $k-$NNs that are more likely to share the same characteristics of x_j. The nonconformity measure is then calculated as the total weighted average as follows:

$$\phi_j = \frac{\sum_{i=1}^{c} u_{ij} \sum_{n=1}^{k} u_{ij}^{n} r_j^{n}}{\sum_{i=1}^{c} u_{ij} \sum_{n=1}^{k} u_{ij}^{n}} \tag{7}$$

Again, using the same illustrative example, the nonconformity score of z can be calculated using (7) as:

$$\phi_z = \frac{u_{1z}(u_{1e}r_e + u_{1f}r_f + u_{1g}r_g) + u_{2z}(u_{2e}r_e + u_{2f}r_f + u_{2g}r_g)}{u_{1z}(u_{1e} + u_{1f} + u_{1g}) + u_{2z}(u_{2e} + u_{2f} + u_{2g})}$$

Assuming that $u_{1z} > u_{2z}$, more emphasis is then given to the $k-$NNs that are closer to v_1.

Now that we have defined our proposed nonconformity measures, we can calculate ψ_j from (4) using the definition of ϕ_j as in (6) or (7) to find ψ_s^{δ}.

Testing Phase; $\forall x_{l+g} \in \mathcal{D}_{ts}$, $f(.)$ is used to predict $\hat{y}_{l+g}$. To construct PIs around $\hat{y}_{l+g}$, we first find the $k-$NNs of x_{l+g} in $\mathcal{D}_{cal}$. At this point, we can directly use (6) to calculate the nonconformity score of x_{l+g}. Yet, we need to find the membership grades of x_{l+g} to use (7). Therefore, the test fuzzy partition matrix $U_{ts} = [u'_{ij}]$ of $\mathcal{D}_{ts}$ is calculated via V_{cal}. The PIs are defined as follows:

$$\left(\hat{y}_{l+g} - \psi_s^{\delta}(\phi_{l+g} + \theta), \hat{y}_{l+g} + \psi_s^{\delta}(\phi_{l+g} + \theta)\right) \tag{8}$$

3 Results and Discussions

The efficiency of the proposed nonconformity measures is demonstrated on datasets tabulated in Table 1. We normalized the target outputs into the range of 0 to 1 to comprehensively compare interval widths across all datasets. The number of calibration samples is set as suggested in [6].

The experiments comprised a 10×10-fold cross-validation scheme. The regression model was selected as Random Forest Regressor with 500 random trees for all experiments. We set $k = 25$ and $m = 2$ in all the experiments, while c is selected according to the size of the datasets.

Algorithm 2. FCM-based ICPR

Input: $\mathcal{D} = \{(x_1, y_1), \ldots, (x_l, y_l)\}$, $\mathcal{D}_{ts} = \{x_{l+1}, \ldots, x_{l+r}\}$, q, k, c, m, β, and δ.
Initialization:
$h := l - q$
$R := \{\}$
$S := \{\}$
Split $\mathcal{D}$ into $\mathcal{D}_{tr} = \{(x_1, y_1), \ldots, (x_h, y_h)\}$ and $\mathcal{D}_{cal} = \{(x_{h+1}, y_{h+1}), \ldots, (x_l, y_l)\}$
Training Phase:
Train $f(.)$ on $\mathcal{D}_{tr}$
Calibration Phase:
Using FCM, cluster $X_{cal} \subset \mathcal{D}_{cal}$ into c clusters to get V_{cal} and U_{cal}
for $j = 1$ `to` q **do**
    calculate $\hat{y}_{h+j} = f(x_{h+j})$
    calculate r_{h+j} from (5)
    store r_{h+j} in R
end for
for $j = 1$ `to` q **do**
    find the k-NNs of x_{h+j} in $\mathcal{D}_{cal}$
    calculate σ_{h+j} from (6) or (7)
    calculate α_{h+j} from (4)
    store α_{h+j} in S
end for
Sort S in descending order to obtain $S = (\alpha_1, \ldots, \alpha_q)$
get α_s^δ where $s = \lfloor \delta(q+1) \rfloor$
Testing Phase:
Calculate U_{ts} of $X_{ts} \subseteq \mathcal{D}_{ts}$ with respect to V_{cal}
for $g = 1$ `to` r **do**
    calculate $\hat{y}_{l+g} = f(x_{l+g})$
    find the $k-$NNs of x_{l+g} in $\mathcal{D}_{cal}$
    calculate σ_{l+g} from (6) or (7)
    Output: PI using (8)
end for

Table 1. Description of the Datasets and Hyperparameters

Dataset	Size	Features	q	c
Boston	506	13	99	2
Abalone	4177	8	899	3
deltaA	7129	5	1599	3
kin8nm	8192	8	1799	4

Table 2 reports the mean coverage and the mean interval width for 90%, 95%, and 99% confidence levels produced by the nonconformity measures (3), (6), (7). We also report the effect of the sensitivity parameter θ on the coverage and intervals' width with $\theta = 0$ and $\theta = 0.1$.

In almost all cases, all the nonconformity measures achieve the desired coverage; which is expected since CP provides probabilistic predictions with provable

Table 2. Uncertainty Quantification Performance

		Mean Coverage			Mean Interval Width		
	Measure	90%	95%	99%	90%	95%	99%
Boston	(3)	0.913	0.951	0.992	0.233	0.322	0.716
	(6) - $\beta = 0$	0.910	0.951	0.993	0.236	0.312	0.622
	(7) - $\beta = 0$	0.910	0.955	0.992	0.235	0.315	0.624
	(6) - $\beta = 0.1$	0.908	0.951	0.992	0.227	0.309	0.655
	(7) - $\beta = 0.1$	0.908	0.951	0.992	0.227	0.311	0.657
Abalone	(3)	0.900	0.950	0.990	0.248	0.329	0.542
	(6) - $\beta = 0$	0.903	0.950	0.990	0.237	0.305	0.505
	(7) - $\beta = 0$	0.902	0.950	0.990	0.236	0.305	0.504
	(6) - $\beta = 0.1$	0.901	0.950	0.990	0.235	0.308	0.504
	(7) - $\beta = 0.1$	0.901	0.950	0.990	0.235	0.308	0.506
deltaA	(3)	0.902	0.951	0.990	0.12	0.158	0.26
	(6) - $\beta = 0$	0.900	0.951	0.990	0.114	0.148	0.229
	(7) - $\beta = 0$	0.899	0.951	0.990	0.114	0.148	0.228
	(6) - $\beta = 0.1$	0.902	0.950	0.990	0.117	0.150	0.239
	(7) - $\beta = 0.1$	0.902	0.950	0.990	0.117	0.150	0.239
kin8nm	(3)	0.901	0.952	0.990	0.331	0.401	0.552
	(6) - $\beta = 0$	0.902	0.951	0.990	0.315	0.371	0.489
	(7) - $\beta = 0$	0.902	0.951	0.990	0.315	0.371	0.489
	(6) - $\beta = 0.1$	0.902	0.952	0.990	0.319	0.377	0.501
	(7) - $\beta = 0.1$	0.902	0.952	0.990	0.319	0.377	0.501

guarantees of accuracy. Given a set of valid conformal regressors, it is worth investigating the tightness of the PIs by comparing the mean interval width. Across all datasets, we can clearly observe that our proposed nonconformity measures result in significantly tighter PIs compared to the standard measure. Comparing the mean interval widths across all levels of confidence, (6) and (7) result in high-quality PIs guaranteeing both coverage and tightness. A significant improvement is observed in high confidence levels, e.g. 95% 99%. On the other hand, we can see that the PIs obtained in the Boston dataset are relatively large; this is partially due to the small size of the dataset. Besides, the guarantee of producing high-quality PIs was preserved for both values of θ in (6) and (7).

4 Conclusion and Future Work

In this work, we presented two nonconformity measures for ICPR based on FCM. Based on the membership grades and $k-$NNs residuals, both fuzzy nonconformity measures produce adaptive PIs enveloping the target output with a given confidence level. Experiments show that our proposed measures produce reliable and tighter PIs compared to the standard nonconformity measure (3). It is important to note that the selection of k, m, c, and θ impacts the width of

PIs but not their validity. Tighter PIs can be obtained by adjusting the values of these parameters according to the data. For instance, one might consider an exhaustive search to find the combination of m and c that yields the highest fuzzy partition coefficient.

Our future plan includes creating more informative nonconformity measures based on FSs and extending the work to Type-2 FSs.

References

1. Beke, A., Kumbasar, T.: More than accuracy: a composite learning framework for interval type-2 fuzzy logic systems. IEEE Trans. Fuzzy Syst. **31**(3), 734–744 (2023). https://doi.org/10.1109/TFUZZ.2022.3188920
2. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: The fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2–3), 191–203 (1984)
3. Boström, H.: crepes: a Python package for generating conformal regressors and predictive systems. In: Conformal and Probabilistic Prediction with Applications, pp. 24–41. PMLR (2022)
4. Chernozhukov, V., Wüthrich, K., Zhu, Y.: Distributional conformal prediction. Proc. Natl. Acad. Sci. **118**(48), e2107794118 (2021)
5. Devetyarov, D., et al.: Conformal predictors in early diagnostics of ovarian and breast cancers. Progress Artif. Intell. **1**(3), 245–257 (2012)
6. Johansson, U., Boström, H., Löfström, T., Linusson, H.: Regression conformal prediction with random forests. Mach. Learn., 155–176 (2014). https://doi.org/10.1007/s10994-014-5453-0
7. Papadopoulos, H., Gammerman, A., Vovk, V.: Normalized nonconformity measures for regression conformal prediction. In: Proceedings of the IASTED International Conference on Artificial Intelligence and Applications, pp. 64–69 (2008)
8. Papadopoulos, H., Haralambous, H.: Reliable prediction intervals with regression neural networks. Neural Netw. **24**(8), 842–851 (2011)
9. Papadopoulos, H., Proedrou, K., Vovk, V., Gammerman, A.: Inductive confidence machines for regression. In: Elomaa, T., Mannila, H., Toivonen, H. (eds.) ECML 2002. LNCS (LNAI), vol. 2430, pp. 345–356. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-36755-1_29
10. Romano, Y., Patterson, E., Candes, E.: Conformalized quantile regression. In: Advances in Neural Information Processing Systems, vol. 32 (2019)
11. Seedat, N., Jeffares, A., Imrie, F., van der Schaar, M.: Improving adaptive conformal prediction using self-supervised learning. arXiv preprint arXiv:2302.12238 (2023)
12. Sesia, M., Romano, Y.: Conformal prediction using conditional histograms. Adv. Neural. Inf. Process. Syst. **34**, 6304–6315 (2021)
13. Vovk, V., Gammerman, A., Shafer, G.: Algorithmic Learning in a Random World, vol. 29. Springer, New York (2005). https://doi.org/10.1007/b106715

Robust Parameter Estimation with Orthogonal Regression Using Interval Constraint Satisfaction Technique

Sejal K. Raval[1](✉), Kamlesh N. Raval[2], Dipesh Makwana[3], and Paresh P. Kotak[4]

[1] Instrument and Control Department, Government Polytechnic, Ahmedabad, India
sejalr73@gmail.com
[2] Computer Engineering Department, R.C.Technical Institute, Ahmedabad, India
[3] I.C. Department, L.D. College of Engineering, Ahmedabad, India
[4] S. and S. Gandhi College, Surat, India

Abstract. Parameter estimation is a first and foremost task to design a proper mathematical model. Outliers in a data can often lead to improper parameter estimation. In present work a novel technique using interval constraint satisfaction technique is suggested for getting robust parameter estimation of a system using linear orthogonal regression. In this proposed method we used M estimator with tukey's biweight function and modified it for orthogonal regression. Tukey's biweight is generally computed using IRLS which has computational issues of convergence and local minima. Here, benefits of interval analysis based method is utilized to develop the proposed method. Two well-known examples of datasets are used to validate the proposed method.

Keywords: Orthogonal regression · Interval constraint satisfaction technique with Branch and Prune algorithm (ICSTBP) · Interval Global optimization (IGO) · Vectorized IGO (VIGO)

1 Introduction

The major issue faced in regression analysis, is the appearance of outliers in the data set that can effectively distort the classical least squares estimator and lead to unreliable results. Outlierness of any observation can be evaluated by finding its distance from the central tendency of the observations. The central tendency of the data points can be represented either by sample mean, standard deviation (SD), median or median of absolute deviation (MAD). The purpose of an estimator is to reduce the average error between estimated set of parameters $\hat{\beta}$ and actual parameter set β. there are several types of robust estimators like L estimator, RM estimator, LS estimator and many more to list. In 1964 Huber generalized median regression to a wider class of estimators, called M estimators, by considering other functions than the absolute value in equation. From which we are concentrating on tukey's biweight function which has higher efficiency but weight is calculated using iterative method and available numeric iterative algorithms are

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 540–547, 2023.
https://doi.org/10.1007/978-3-031-39777-6_64

heavily influenced by starting point provided by the user for the possible solution. There are some computational issues with existing methods like unguaranteed convergence of the iterative methods, and even if the method converges, the speed and accuracy of the convergence is still dependent on initial guess [17]. There exists another problem of local minima. If the method converges depending on starting point it may get stuck at local minima. So the final solution may or may not be the global solution. This has inspired us for our present work. In which we propose a novel method of robust parameter estimation based on interval arithmetic concept, which is not dependent on initial guess and guarantees global minima.

2 Outliers and Robust Estimators

In regression computation, major three types of outliers influence the LS estimator. Rousseeuw and Leroy (1987) define them as vertical outliers, bad leverage points and good leverage points. To understand this terminology, consider a simple linear regression as shown in Fig. 1 [7].

Vertical outliers are those data points that have outlying values for the corresponding error term (the y-dimension) but are not outlying in the space of explanatory variables (the x-dimension). Their presence affects the LS-estimation and affect the estimated intercept.

Good Leverage points are those informations that are outlying in the space of explanatory variables but that are located close to the regression line. Their presence does not affect the LS-estimation but it affects statistical inference since they do deflate the estimated standard errors.

Bad Leverage points are observations that are both outlying in the space of explanatory variables and located far from the true regression line. Their presence affects significantly the LS-estimation of both the intercept and the slope.

To deal with the problem caused by outliers there exists two ways.

1. Regression diagnostics [21]

- Detect outliers
- Handle outliers, generally remove them from the data set [21, 22]
- Use classical methods to estimate regression parameters

2. Robust regression [21]

- Fit a regression to the bulk of the data.
- Outliers now stand out having large residuals.

M estimator [20]

M-estimators are robust estimators that are constructed from MLE (or MAP) assuming a non-normally-distributed noise**.** The principle consists in taking a distribution that has a PDF with ‘heavy tails’. Having such tails means that large errors are less improbable than it would be with the normal distribution. Many different M-estimators based upon several noise distributions have been proposed. Note that some of these M-estimators are

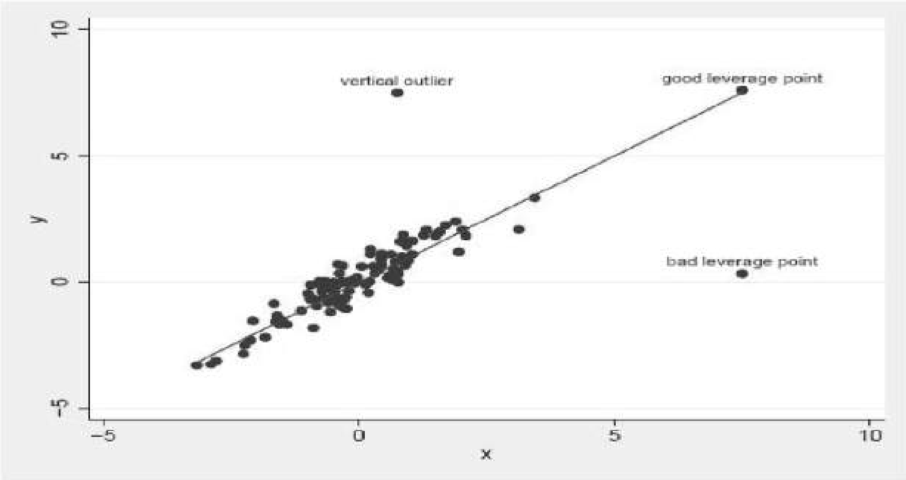


Fig. 1. Types of Outlier

statistically grounded in the sense that the considered distributions are related to actual facts. Other M-estimators are a little bit more artificial in the sense that their underlying probability distributions have been made up for the need of robustness. Sometimes, the underlying probability distribution of an M-estimator is not even an actual probability distribution since it does not sums up to 1. In this case, the 'probability distribution' is to be considered as some kind of 'score function' which takes high values for probable inputs and lower values for less probable inputs. [20].

We now give some notation and terms related with M-estimators.

An M-estimator is typically defined using what is called a *ρ-function*. The ρ function of an M-estimator is a function from $\mathbb{R}$ to $\mathbb{R}_+$ that defines the underlying probability distribution of the M-estimator:

$$\mathrm{P(e)} = \exp(-\rho(\mathrm{e})) \tag{1}$$

The cost function minimized by the M-estimator is obtained by

$$\min_{\mathrm{P}} \sum\nolimits_{i=1}^{d} \rho(\mathrm{e}_i) \tag{2}$$

Note that the least-squares cost function is just a particular case of M-estimator with $\rho(x) = x^2$. Two other interesting functions can be associated to an M-estimator: the *influence* function and the *weight* function. The influence function, denoted ψ, is defined as:

$$\psi(x) \stackrel{\text{def}}{\equiv} \frac{\partial \rho}{\partial x}(x) \tag{3}$$

The weight function w, also known as the attenuation function, is defined as:

$$wx \stackrel{def}{\equiv} \psi xx \tag{4}$$

It measures the weight given to a particular measurement in the cost function according to the error it produces. With M-estimators, the principle is to give large weights to small errors and small weights to large errors (which are likely to come from outliers). Note that, as defined in Eq. (4), the weight function w is not defined at zero. It is generally possible to define it at this particular point with a continuous extension [20].

3 Interval Constraint Satisfaction Technique

Interval Analysis. Interval analysis is a less explored branch of applied mathematics in which a real number is visualized as a pair of real number denoting two end points. A closed interval x $= [\underline{x}, \overline{x}]$ with $\underline{x}, \overline{x} \in$ R, can be regarded as the set of all real numbers enclosing the lower and upper bound, or as an approximation of some real numbers lying within that set [6]. In the present work, a problem i.e. orthogonal regression has been posed as constrain satisfaction problem (CSP) and has been solved by interval constraint satisfaction technique. The power of interval arithmetic and its algorithms to find all possible solution has been utilized. In addition, interval computations provide an appropriate framework to deal with uncertain data [5].

Interval Constraint Solving Techniques
A recent powerful tool is the class of methods called as interval constraint satisfaction techniques (ICST) for solving constraint satisfaction problems [14]. ICST possesses three useful properties of completeness, reliability and efficiency [1, 2, 14]:

There are many solvers available which make use of ICST. In our present work we used RealPaver 0.4 as a constraints solver because it is coded using linux platform which is an open source available for researchers. The constraint solver engine of RealPaver uses the famous branch and prune algorithm. The algorithm is well explained in many resources [12, 15].

4 Proposed Method

Our approach to robust regression estimate is to use residuals as the perpendicular distance from fitted line and estimate the parameter $\overset{\gtreqless}{\beta}$ using M estimator. In M estimator $\overset{\gtreqless}{\beta}$ defined by

$$\sum_{i=1}^{n} \rho\left(\frac{r_i \ \hat{\beta})}{\hat{\sigma}}\right) = min \tag{5}$$

With a bounded ρ and a high breakdown point preliminary scale $\overset{\gtreqless}{\sigma}$. If ρ has a derivative ψ it follows that

$$\sum\nolimits_{i=1}^{n} \psi \ \ \frac{r_i}{\hat{\sigma}}\Big) x_i = 0 \tag{6}$$

where, ρ (.) is a loss function which is even, non-decreasing for positive values and less increasing than the square function. To guarantee scale equivariance (i.e. independence with respect to the measurement units of the dependent variable), residuals are standardized by a measurement of dispersion $\hat{\hat{\sigma}}$. The ψ is redescending, consequently the estimating equation may have multiple solutions corresponding to multiple local minima of the function on the left hand side of ().

a. M-Estimates with Known Scale

i. Numerical Computation and issues

One can use any of the general methods for solving Eq. 5.6, such as the Newton-Raphson procedure [23]. If the procedure converges, the convergence is fast, but there is no guarantee that it converges. This happens in the case of a location M – estimate, where the iterations are required to solve ψ function. If ψ is bounded, its derivative $\psi\prime$ tends to zero at infinity, and hence the denominator is not bounded away from zero, which makes the procedure unreliable.

Another method is based on representing Eq. 5.6 in the form of a weighting function as

$$\sum_{i=1}^{n} \psi \left. \frac{r_i}{\hat{\sigma}} \right) r_i x_i = 0 \tag{7}$$

For Tukey's biweight function, the weighting function becomes

$$W(t) = \begin{cases} 1 - \left(\left(\frac{t}{k}\right)^2\right)^2, if \left|\frac{t}{k}\right| \leq 1 \\ 0, otherwise \end{cases} \tag{8}$$

Equation 7 is solved using an iterative process known as "iteratively Reweighted Least Squares" (IRLS or IRWLS) which involves the following steps [23]:

I. Start with a robust initial estimates of $\hat{\beta}_0$ (using any basic estimator) and compute $\hat{\sigma}$ as MADN by taking the median of the non-absolute residuals.
II. For k = 0, 1, 2,…
 a. Given $\hat{\beta}_k$, for i = 1,…, n compute the residuals $r_{i,k} = y_i - x_i'\hat{\beta}_k$ and $w_{i,k} = W\left(r_{i,k}/\hat{\sigma}\right)$.
 b. Compute $\hat{\beta}_{k+1}$, by solving $\sum_{i=1}^{n} w_{i,k} x_i \ \left. y_i - x_i'\hat{\beta}\right) = 0$.
III. Stop when $max_i\left(\left|r_{i,k} - r_{i,k+1}\right|\right)/\hat{\sigma} < \epsilon$

The choice of starting point influences the convergence of the algorithm and the final result (for Tukey's biweight ρ function).

A third method of solving equation 5.6 is based on iterative pseudo-observations However; the convergence has the same difficulties as the IRLS, and is much slower.

In general, computational methods available in the literature for M-estimation are iterative in nature. Especially if ψ is redescending, these methods have the following computational issues [4, 15].

1. The methods are not robust to the initialization; the choice of the initial estimate has a significant influence on the quality of the obtained M estimate. Since the commonly used methods are gradient-based, the computations can get stuck at local solutions.
2. The convergence of the methods is not guaranteed. If the method converges, the solution may not be a global solution.

a. Outlier dataset [15]:

Table 1. Outlier dataset

	1	2	3	4	5	6	7	8
Y	2	4	60	7	9	12	14	15
X	9	10						
Y	18	5						

b. Chaterjee-hadi dataset [10]:
This dataset (Table 2) shows x axis and y axis is a data without outlier.

Table 2. Chaterjee-hadi dataset

x	0	5	15	16	17	18
y	8.1	11	8.20	8.30	9.4	9.3
x	19	20	21	22	23	24
y	9.6	10.3	11.3	11.4	12.2	12.9

The obtained values of m and c of the proposed algorithm are available in form of interval values using Outer boxes. ICSTBP gives all the solutions available in the initial search region. The lines fitted using the midpoint of the interval answers are plotted in Figs. 2 and 3.

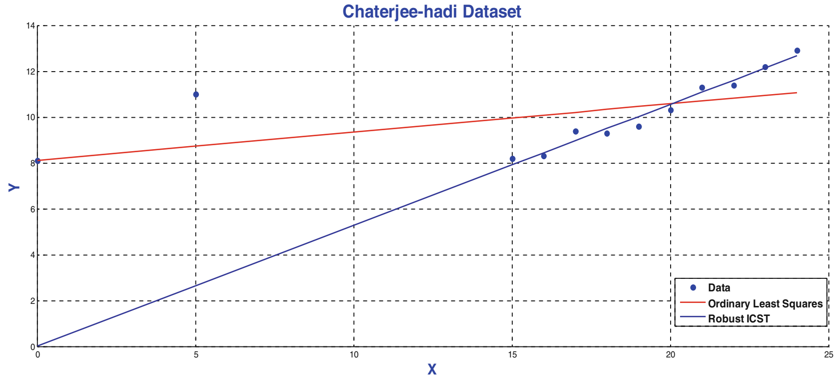


Fig. 2. Estimated parameters

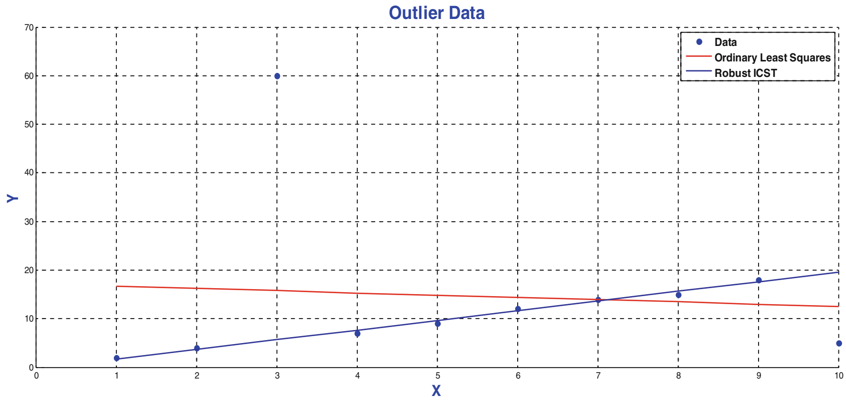


Fig. 3. Estimated parameters

5 Conclusion and Discussion

In the present work, a novel method for parameter estimation of orthogonal regression is proposed and validated. The simulation results show that the proposed method efficiently reduces impact of outliers and gives more precise solution than obtained in currently available mathematical tools like matlab. This method utilizes interval based algorithms where it need not to give perfect initial guess of parameters rather to give a possible range of solution is given which can be extended up to infinity [1]. So from the results it can be concluded that to suppress the effect of outliers is a nonlinear part of the regression, for which the method is best suited as interval analysis solves the problem of local minima. With simple modification in program we can show that the proposed method can be extended for large data set hence the scalability of method can be justified. For multidimensional orthogonal regression where the problem of local solution is of major consideration [18], for which this method can be extended as future work. In future, IGO and VIGO algorithm can be applied for faster computation.

References

1. Granvilliers, L.: RealPaver User's Manual: Solving Nonlinear Constraints by Interval Computations, edition 0.4, for realPaver version 0.4 edition (2004)
2. Granvilliers, L., Benhamou, F.: RealPaver: an interval solver using constraint satisfaction techniques. ACM Trans. Math. Soft. **32**(1), 138–156 (2006)
3. Hansen, E., Walster, G.: Global Optimization using Interval Analysis, second edition
4. Maronna, R.A., Martin, R.D., Yohai, V.J.: Robust Statistics: Theory and Methods. John Wiley-Sons, Chichester, England (2006)
5. Moore, R.E.: Interval Analysis. Prentice-Hall, Englewood Cliffs, New Jersey (1966)
6. Natraj, P.S.V., Kalla, R.: Computation of limit cycles for uncertain nonlinear fractional –order systems IOP publishing, T1 36/014021 (2009)
7. VanHentenryck, P., Michel, L., Deville, Y.: Numerica: A Modeling Language forGlobal Optimization. The MIT Press (1997)
8. Puget, J.F.: A C++ implementation of CLP (1994)
9. VanHentenryck, P., Michel, L., Benhamou, F.: Newton-constraint programming over nonlinear constraints. Sci. Comput. Program. **30**(1–2), 83–118 (1998)
10. Chatterjee, S., Hadi, A.S.: Regression Analysis by Example, 4th Edition. John Wiley& Sons, Inc., (2006)
11. DeMuth, J.E.: Basic Statistics and Pharmaceutical Statistical Applications (Biostatistics(New York, N. Y.), 2. Marcel Dekker., New York (1988)
12. Raval, S.K: An alternative method for parameter estimation using orthogonal regression with Interval Constraint satisfaction Technique 978–0–7695–4692–6/12 2012 IEEE https://doi.org/10.1109/CSNT2012.147
13. www.iopscience.iop.org/jpconf7_78_012084.pdf
14. www.nlreg.com/orthogonal.htm
15. Modi, C.K.: Interval Analysis based Technique for Robust estimation and Regression. PhD dissertation, Systems and Control Engineering, IIT Bombay (2010)
16. Dissanaike, G., Wang, S.: A critical examination of orthogonal regression. https://doi.org/10.2139/ssrn.407560
17. http://www.mathworks.com/matlabcentral/fileexchange/6716-orthogonal-linear-regression
18. Helfrich, H.P., Zwick, D.: A trust region method for implicit orthogonal distance regression. Num. Algorith. **5**, 535–545 (1993)
19. Calafiore, G.C.: Outliers robustness in multivariate orthogonal regression. IEEE Trans. Syst. Man Cybern.-Part A: Syst. Humans. **30**(6), 674−679 (2000)
20. http://www.brnt.eu/phd/node13.html
21. Rousseeuw, P.J., Leroy, A.M.: Robust Regression and Outlier Detection. JohnWiley-Sons, New York (1987)
22. Barnett, V., Lewis, T.: Outliers in statistical Data, 3rd edn. JohnWiley and Sons Inc, NewYork (1998)
23. Maronna, R.A., Martin, R.D.,Yohai, V.J.: Robust Statistics : Theory and Methods. John Wiley-Sons, Chichester, England (2006)

Two-Way ANOVA for Fuzzy Observations and an Implication

Zeynep Gökkuş[1(✉)], Sevil Şentürk[2], Taner Yildiz[3], and Elçin Yeşiloğlu Cevher[3]

[1] Vocational School, Computer Technologies Department, Kastamonu University, Kastamonu, Turkey
z.gokkus@gmail.com
[2] Faculty of Science, Department of Statistics, Eskişehir Technical University, Eskişehir, Turkey
[3] Faculty of Agriculture, Department of Agricultural Machinery and Technologies Engineering, Ondokuz Mayıs University, Samsun, Turkey
t.yildiz@omu.edu.tr

Abstract. Analysis of variance (ANOVA) is an effective and powerful method among studies investigating how the continuous dependent variable is affected by categorical independent variables by measurement. It is difficult to talk about the precision of measurements in real life problems. For this reason, many statistical methods have been extended to fuzzy numbers. The most commonly used ANOVA method is one-way-ANOVA, and there are methods developed for fuzzy one-way-ANOVA in the literature. In this study, the theoretical structure of the two-way-ANOVA method is constructed for the first time using non-symmetric triangular fuzzy numbers and its analysis is performed on a sample.

Keywords: Analysis of variance · arithmetic fuzzy numbers · agricultural machinery design

1 Introduction

It is certain that traditional analysis of variance (ANOVA) methods will be insufficient in resolving observations containing uncertainty. Therefore, it is appropriate to analyze data with uncertainty using fuzzy methods. For this reason, several researchers have proposed different ANOVA models for different fuzzy number types in different fields in the literature. These are as follows: surveying the trace of one-way fuzzy analysis of variance (FANOVA) and comparing it's results with regression model [14], whole testing of one-way ANOVA under normal fuzzy observations [15], bootstrap method for approximating the asymptotic one sample tests by fuzzy observations [16], on a given Hilbert space, developing a one-way ANOVA approach for the functional data [17], using the moment correction for uncertain data for processing ANOVA model [18], bootstrap asymptotic multi-sample testing of means for fuzzy observations [19, 20], regarding the cuts of fuzzy observations for one-way ANOVA problem on fuzzy data for optimization approach [21], developing one-way ANOVA for fuzzy observations basis of

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 548–559, 2023.
https://doi.org/10.1007/978-3-031-39777-6_65

extension principle approach [22], comparing several groups with FANOVA [23], generalizing PCIs are able to measure the capability of a fuzzy-valued process in producing products based on a triangular fuzzy observations (TFNs) [24], taking an interest in the Latin Square Design problem (LSD) test using Trapezoidal Fuzzy Numbers [27], proposed testing is analyzed under varieties of trapezoidal fuzzy models [28], a new method for testing intuitionistic fuzzy hypotheses of one-factor and two-factor ANOVA model using intuitionistic fuzzy data samples [25], applying an intuitionistic fuzzy two-factor ANOVA (2-D IFANOVA) through intuitionistic fuzzy sets (IFSs) and index matrices (IMs), investigating how the number of COVID-19 cases depend on the "density" and "climate zone" factors for Europe [26], introducing the neutrosophic analysis of variance (NANONA) [29], presenting a novel Neutrosophic-Principal Component Analysis and Two-way Neutrosophic ANOVA to analyzing the reasons of Performance Gap between private and public school students [30], and generalizing Two-Way ANOVA to neutrosophic two-way ANOVA solving neutrosophic data [31]. Besides the FANOVA models and applications mentioned above, there is extending one-way ANOVA to a case where observations are non-symmetric TFNs rather than real numbers [1–3]. FANOVA model obtained by the method is similar to the classical one-way ANOVA model and is very easy to implement.

The aim of this paper is constructed the theoretical base of a simple fuzzy two-way ANOVA model, as an extension of ANOVA, is firstly explained for non-symmetric triangular fuzzy observations. Preliminary information about fuzzy concepts and some arithmetic operations is given in Section 2. The proposed two-way FANOVA is explained in Section 3. The proposed two-way FANOVA model is applied to an example that depends on biological-technical properties from four different chestnut cultivars for the design and development of agricultural machinery in Section 4. And conclusion and further research works are in final section

2 Preliminaries

Let X be universal set and $F(X) = \{A|A : X \to [0, 1]\}$. Any $A \in F(X)$ is called a fuzzy set on X. Most of all, let R be set of all real numbers. It is used [1, 2] $F_T(R) = \left\{\tilde{T}\middle|a, s_a^l, s_a^r \in R; s_a^l, s_a^r\rangle 0\right\}$, where

$$\tilde{T}(x) = T\left(a, s_a^l, s_a^r\right)(x) = \begin{cases} \left(x - a + s_a^l\right)/s_a^l \; if \; a - s_a^l \le x < a \\ \left(a + s_a^r - x\right)/s_a^r \; if \; a \le x < a + s_a^r \\ 0, \; elsewhere\,. \end{cases} \tag{1}$$

Any $\tilde{T} \in F_T(R)$ is called a triangular fuzzy number (TFN). Here, a is center point, s_a^l and s_a^r are the left and right widths of a and it can be written as $T\left(a, s_a^l, s_a^r\right)$ [1, 2].

Definition 2.1. [1–3] Let $\tilde{A}, \tilde{B} \in F(R)$, then.

$$\tilde{A}\ominus\tilde{B} = \left\{\int_0^1 g(\pi)\left[\tilde{A}_\pi(-)\tilde{B}_\pi\right]^2 d\pi\right\}^{1/2} \tag{2}$$

is called distance between $\tilde{A}$ and $\tilde{B}$, in which for any $\pi \in [0, 1]$,

$$\tilde{A}_\pi(-)\tilde{B}_\pi = \left\{[a_1(\pi) - b_1(\pi)]^2 + [a_2(\pi) - b_2(\pi)]^2\right\}^{1/2} \tag{3}$$

measured the distance between $\tilde{A}_\pi = [a_1(\pi), a_2(\pi)]$ and $\tilde{B}_\pi = [b_1(\pi), b_2(\pi)]$, $a_1(0)$, $a_2(0)$, $b_1(0)$, $b_2(0)$ are taken as finite real numbers and g is a real valued non-decreasing function on [0, 1] with $g(0) = 0$ and $\int_0^1 g(\pi)d\pi = \frac{1}{2}$ (for instance $g(\pi) = \frac{m+1}{2}\pi^m$ where $m = 1, 2, 3, \ldots$).

Remark 2.1. [1, 2] In Definition 2.1, $a_1(\pi) - b_1(\pi)$ and $a_2(\pi) - b_2(\pi)$ are the distances between the left and right end points of the π-cut of $\tilde{A}$ and $\tilde{B}$, respectively. The weight of $\left[\tilde{A}_\pi(-)\tilde{B}_\pi\right]^2$ is $g(\pi)$, and the non-decreasing characteristic of g is the higher membership of π-cut, and it is adjusted the distance among $\tilde{A}$ and $\tilde{B}$. This operation synthetically represents the knowledge on all membership degree. The favor of this operation is that fuzzy numbers may allow different π-cuts have different weights.

Remark 2.2. [1, 2] The distance specified in definition 2.1 corresponds to the whole deviation of a and b if the two numbers $\tilde{A}$ and $\tilde{B}$ are real numbers a and b, respectively.

Theorem 2.1. [1, 2] The distance between TFNs $\tilde{A} = T\left(a, s_a^l, s_a^r\right)$ and $\tilde{B} = T\left(a, s_a^l, s_a^r\right)$ is.

$$\begin{aligned}\left(\tilde{A}\ominus\tilde{B}\right)^2 &= (a-b)^2 + \tfrac{1}{(m+2)(m+3)}\left[\left(s_a^l - s_b^l\right)^2 + \left(s_a^r - s_b^r\right)^2\right] \\ &\quad + \tfrac{a-b}{(m+1)}\left[\left(s_a^r - s_b^r\right)^2 - \left(s_a^l - s_b^l\right)^2\right],\end{aligned} \tag{4}$$

where the weighted function $g(\pi) = \frac{m+1}{2}\pi^m$ for $m = 1, 2, 3, \ldots$.

3 Two-Way Fuzzy Analysis of Variance (FANOVA)

Classical two-way ANOVA models are available in several books and articles [4–6]. In this study, classical two-way ANOVA models are not mentioned in terms of simplicity. In this study, the classical two-way ANOVA model is extended to fuzzy two-way ANOVA, similar to the fuzzy one-way ANOVA method described in [1, 2, 3].

Suppose the data or observations are TFNs of the form $\tilde{y}_{ijk} = T\left(y_{ijk}, s_{y_{ijk}}^l, s_{y_{ijk}}^r\right)$ to represent $\tilde{y}_{ijk}$ about y_{ijk}. All theoretical model elements such as statistical hypotheses, random variables and population parameters are crisp, and the model is $Y_{ijk} = \mu + \lambda_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$. Here, μ represents the grand mean i, λ_i the effect of Factor A, β_j the effect of Factor B, γ_{ij} the interaction effect of Factor A*Factor B, and $\varepsilon_{ijk} \sim N\left(0, \sigma^2\right)$ the error. All hypotheses, Y_{ijk}'s being ordinary random variables, are as in the classical two-way ANOVA model and are shown as follows.

Hypotheses for Factor A;

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_r$$

$$H_1 : not\ all\ \mu_i\ are\ equal.$$

Hypotheses for Factor B;

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_s$$

$$H_1 : not\ all\ \mu_j\ are\ equal.$$

Hypotheses for interaction of Factor A*Factor B;

$$H_0 : \mu_{11} = \mu_{12} = \cdots = \mu_{rs}$$

$$H_1 : not\ all\ \mu_{ij}\ are\ equal.$$

However, unlike the classical two-way ANOVA model, the observations are TFNs, not crisp numbers. According to Definition 2.1, while $\tilde{y}_{ijk}$'s are TFNs, the equations for SST, MSE, SSTR, SSE, MSTR and F statistics for each of Factor A-Factor B-interaction of Factor A*Factor B are as follows.

$\widetilde{SST} = \sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{n_{ij}}\left(\tilde{y}_{ijk} \ominus \overline{\tilde{y}_{\ldots}}\right)^2$ for $\overline{\tilde{y}_{\ldots}} = \frac{1}{n_t}\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{n_{ij}}\tilde{y}_{ijk} = T\left(\overline{y_{\ldots}}, \overline{s^l_{y_{\ldots}}}, \overline{s^r_{y_{\ldots}}}\right)$;

factor A, $\widetilde{SSTR}_A = \sum_{i=1}^{r}\sum_{k=1}^{n_i} n_i\left(\overline{\tilde{y}_{i..}} \ominus \overline{\tilde{y}_{\ldots}}\right)^2$, $\widetilde{MSTR}_A = \frac{\widetilde{SSTR}_A}{r-1}$, $\tilde{f}_A = \frac{\widetilde{MSTR}_A}{\widetilde{MSE}}$ for $\overline{\tilde{y}_{i..}} = \frac{1}{n_i}\sum_{i=1}^{n_i}\sum_{j=1}^{s}\tilde{y}_{ijk} = T\left(\overline{y_{i..}}, \overline{s^l_{y_{i..}}}, \overline{s^r_{y_{i..}}}\right)$;

factor B, $\widetilde{SSTR}_B = \sum_{j=1}^{s}\sum_{k=1}^{n_j} n_j\left(\overline{\tilde{y}_{.j.}} \ominus \overline{\tilde{y}_{\ldots}}\right)^2$, $\widetilde{MSTR}_B = \frac{\widetilde{SSTR}_B}{s-1}$, $\tilde{f}_B = \frac{\widetilde{MSTR}_B}{\widetilde{MSE}}$ for $\overline{\tilde{y}_{.j.}} = \frac{1}{n_j}\sum_{i=1}^{r}\sum_{j=1}^{n_j}\tilde{y}_{ijk} = T\left(\overline{y_{.j.}}, \overline{s^l_{y_{.j.}}}, \overline{s^r_{y_{.j.}}}\right)$;

interaction factor of Factor A*Factor B, $\widetilde{SSTR}_{A*B} = \sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{k=1}^{rs}\left(\overline{\tilde{y}_{ij.}} \ominus \overline{\tilde{y}_{i..}} \ominus \overline{\tilde{y}_{.j.}} \oplus \overline{\tilde{y}_{\ldots}}\right)^2$, $\widetilde{MSTR}_{A*B} = \frac{\widetilde{SSTR}_{A*B}}{(r-1)(s-1)}$, $\tilde{f}_{A*B} = \frac{\widetilde{MSTR}_{A*B}}{\widetilde{MSE}}$ for $\overline{\tilde{y}_{ij.}} = \frac{1}{n_j}\sum_{j=1}^{n_j}\tilde{y}_{ijk} = T\left(\overline{y_{ij.}}, \overline{s^l_{y_{ij.}}}, \overline{s^r_{y_{ij.}}}\right)$;

error, $\widetilde{SSE} = \frac{1}{r \mathrm{x} s}\left[\left(\sum_{i=1}^{r}\left(\overline{\tilde{y}_{i..}} \ominus\left(\sum_{j=1}^{s}\left(y_{ijk} \ominus \overline{\tilde{y}_{i..}}\right)^2\right)\right)\right)+\left(\sum_{j=1}^{s}\left(\overline{\tilde{y}_{.j.}} \ominus\left(\sum_{i=1}^{r}\left(y_{ijk} \ominus \overline{\tilde{y}_{.j.}}\right)^2\right)\right)\right)+\left(\sum_{i=1}^{r}\sum_{j=1}^{s}\left(\overline{\tilde{y}_{ij.}} \ominus\left(\sum_{i=1}^{r}\left(y_{ijk} \ominus \overline{\tilde{y}_{ij.}}\right)^2\right)\right)\right)\right]$, $\widetilde{MSE} = \frac{\widetilde{SSE}}{n_t - rs}$.

The Decision Rule: [1, 2] Let $\tilde{f}_A$, $\tilde{f}_B$ and $\tilde{f}_{A*B}$ be the observed values of the test statistics. $F_{1-\pi,r-1,n_t-r}$, $F_{1-\pi,s-1,n_t-s}$ and $F_{1-\pi,rs-1,n_t-rs}$ be the π th quartiles of the fisher distribution with $r-1$ and $n_t - r$, $s-1$ and $n_t - s$, $rs - 1$ and $n_t - rs$ degrees of freedom, respectively. At π significance level, H_0 hypotheses are accepted if $\tilde{f}_A \le F_{1-\pi,r-1,n_t-r}$, $\tilde{f}_B \le F_{1-\pi,s-1,n_t-s}$ and $\tilde{f}_{A*B} \le F_{1-\pi,rs-1,n_t-rs}$; otherwise the H_1 hypotheses are accepted.

In testing two-way FANOVA, the p-values can be calculated by $p - \text{value} = P\left(F > \tilde{f}\right)$ in which $\tilde{f}$ s are the observed values of the test statistics [1, 2].

Remark 3.1. According to Remark 2.1, if the observations are precise numbers y_{ijk} they are indicator functions $I_{\{y_{ijk}\}}$ for $i = 1, \ldots, r, j = 1, \ldots, s$ and $k = 1, \ldots, n_{ij}$, then all the established extended statistics for FANOVA correspond to ANOVA [1, 2].

Theorem 3.1 [1, 2]. In two-way FANOVA model, supposed the observations are $\tilde{y}_{ijk} = T\left(y_{ijk}, s^l_{y_{ijk}}, s^r_{y_{ijk}}\right) \in F_T(R), i = 1, \ldots, r, j = 1, \ldots, s, k = 1, \ldots, n_{ij}$, then the observed crisp values of $\widetilde{sst}$, $\widetilde{sstr}$ and $\widetilde{sse}$ are as follows:

$$\widetilde{SST} = sst_y + \frac{1}{(m+2)(m+3)}\left(sst_{s^l_y} + sst_{s^r_y}\right), \tag{5}$$

$$\begin{cases} \widetilde{SSTR}_A = sstr_{(A)y} + \frac{1}{(m+2)(m+3)}\left(sstr_{(A)s^l_y} + sstr_{(A)s^r_y}\right) \\ \widetilde{SSTR}_B = sstr_{(B)y} + \frac{1}{(m+2)(m+3)}\left(sstr_{(B)s^l_y} + sstr_{(B)s^r_y}\right) \\ \widetilde{SSTR}_{A*B} = sstr_{(A*B)y} + \frac{1}{(m+2)(m+3)}\left(sstr_{(A*B)s^l_y} + sstr_{(A*B)s^r_y}\right) \end{cases}, \tag{6}$$

$$\widetilde{SSE} = sse_y + \frac{1}{(m+2)(m+3)}\left(sse_{s^l_y} + sse_{s^r_y}\right), \tag{7}$$

where sst_y, $sstr_y$ and sse_y are real values of sst, $sstr$ and sse for the centre points of $T\left(y_{ijk}, s^l_{y_{ijk}}, s^r_{y_{ijk}}\right)$, and $sst_{s^l_y}$, $sstr_{s^l_y}$ and $sse_{s^l_y}$ the real values of sst, $sstr$ and sse for the left widths of $T\left(y_{ijk}, s^l_{y_{ijk}}, s^r_{y_{ijk}}\right)$, and $sst_{s^r_y}$, $sstr_{s^r_y}$ and $sse_{s^r_y}$ the real values of sst, $sstr$ and sse for the right widths of $T\left(y_{ijk}, s^l_{y_{ijk}}, s^r_{y_{ijk}}\right)$, respectively.

Result 3.1. [1, 2] Likewise Theorem 3.1, the observed values of $\widetilde{mstr}$s, $\widetilde{mse}$ and the test statistics $\tilde{f}$ s are as follows:

$$\begin{cases} \widetilde{MSTR}_A = \frac{\widetilde{SSTR}_A}{r-1} = mstr_{(A)y} + \frac{1}{(m+2)(m+3)}\left(mstr_{(A)s^l_y} + mstr_{(A)s^r_y}\right) \\ \widetilde{MSTR}_B = \frac{\widetilde{SSTR}_B}{s-1} = mstr_{(B)y} + \frac{1}{(m+2)(m+3)}\left(mstr_{(B)s^l_y} + mstr_{(B)s^r_y}\right) \\ \widetilde{MSTR}_{A*B} = \frac{\widetilde{SSTR}_{A*B}}{rs-1} = mstr_{(A*B)y} + \frac{1}{(m+2)(m+3)}\left(mstr_{(A*B)s^l_y} + mstr_{(A*B)s^r_y}\right) \end{cases}, \tag{8}$$

$$\widetilde{MSE} = \frac{\widetilde{SSE}}{n_t - rs} = sse_y + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right), \tag{9}$$

$$\begin{cases} \tilde{f}_A = \dfrac{\widetilde{MSTR_A}}{\widetilde{MSE}} = \dfrac{mstr_A + \frac{1}{(m+2)(m+3)}\left(mstr_{(A)s_y^l} + mstr_{(A)s_y^r}\right)}{sse + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right)} \\ \tilde{f}_B = \dfrac{\widetilde{MSTR_B}}{\widetilde{MSE}} = \dfrac{mstr_B + \frac{1}{(m+2)(m+3)}\left(mstr_{(B)s_y^l} + mstr_{(B)s_y^r}\right)}{sse + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right)} \\ \tilde{f}_{A*B} = \dfrac{\widetilde{MSTR_{A*B}}}{\widetilde{MSE}} = \dfrac{mstr_{A*B} + \frac{1}{(m+2)(m+3)}\left(mstr_{(A*B)s_y^l} + mstr_{(A*B)s_y^r}\right)}{sse + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right)} \end{cases}. \tag{10}$$

In this case, the two-way FANOVA table is constructed as in Table 1.

Table 1. Two-way FANOVA table

Source of variation	$\widetilde{SS}$	df	$\widetilde{MS}$	$\tilde{f}$
Factor A	$\widetilde{SSTR_A}$	$(r-1)$	$\widetilde{MSTR_A} = \widetilde{SSTR_A}/(r-1)$	$\tilde{f}_A = \widetilde{MSTR_A}/\widetilde{MSE}$
Factor B	$\widetilde{SSTR_B}$	$(s-1)$	$\widetilde{MSTR_B} = \widetilde{SSTR_B}/(s-1)$	$\tilde{f}_B = \widetilde{MSTR_B}/\widetilde{MSE}$
Interaction A*B	$\widetilde{SSTR_{A*B}}$	$(r-1)(s-1)$	$\widetilde{MSTR_{A*B}} = \widetilde{SSTR_{A*B}}/((r-1)(s-1))$	$\tilde{f}_{A*B} = \widetilde{MSTR_{A*B}}/\widetilde{MSE}$
Error	$\widetilde{SSE}$	$(n_t - rs)$	$\widetilde{MSE} = \widetilde{SSE}/(n_t - rs)$	
Total	$\widetilde{SST}$	$(n_t - 1)$		

4 Implication and Results

Knowing the biological and technical properties of grainy products are important in the design and development of agricultural machinery [7, 9]. The one of the most important factors in design of agricultural machinery separation systems, conveying of agricultural product, processing and packaging systems are mechanical properties [10]. Damages in seed and grainy products mostly occur during harvesting-threshing, cleaning, sorting and transmission. The effect of external forces that cause mechanical damage to grainy products can be caused by static or dynamic impact loads [11]. Mechanical effects can damage the crops during, harvest and post-harvest. In addition, it can adversely affect the storage and shelf life of products. Therefore, it is necessary to determine the maximum loads of grainy products depending on their size under mechanical loads [13].

In this study, rupture energy measurements of four different chestnut cultivars (Macit 55, Akyüz, Ali Nihat and Bouche de Betizac) are determined at the same loading rate and in three different axes. The X-axis is the length, the Y-axis is the width, and the Z-axis is the thickness dimension of the chestnuts. The data in the study are at a loading speed of 10 mm min^{-1} of chestnut cultivars; It was obtained by static loading in the x-x (length), y-y (width), and z-z (thickness) axes at a humidity level of 75–85% and represented by Fig. 1.

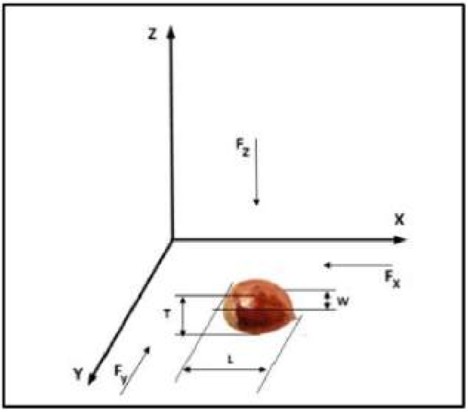


Fig. 1. Chesnut dimensions and load axes [7]

In this study, it is aimed to check whether the rupture energy measurements for four different chestnut cultivars (factor B) and three different dimensions (factor A) differ from each other. For this reason, rupture energy measurements are considered as non-symmetric fuzzy triangular numbers. It is investigated with two-way FANOVA whether there is a difference in rupture energy between chestnut cultivars and between dimensions, and whether there is an interaction between chestnut cultivars and sizes in terms of rupture energy. In other words, it is going to test for dimensions "$H_0 : \mu_1 = \mu_2 = \mu_3$" versus "$H_1$ not all μ_i's are equal, for i = 1, 2, 3"; for chestnut cultivars "$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$" versus "$H_1$ not all μ_i's are equal, for i = 1, 2, 3, 4"; and for interaction of dimensions*chestnut cultivars "$H_0 : \mu_{11} = \ldots = \mu_{34}$" versus "$H_1$ not all μ_{ij}'s are equal, for i = 1, 2, 3 and j = 1, 2, 3, 4". For this tests, 61 non-symmetric triangular rupture energies are selected as the experimental units in two-way FANOVA with sample sizes for dimensions $n_{1j} = 20$, $n_{2j} = 20$ and $n_{3j} = 21$, and for chesnut cultivars $n_{i1} = 13$, $n_{i2} = 15$, $n_{i3} = 18$ and $n_{i4} = 15$, respectively as seen Fig. 2–3 and Table 2.

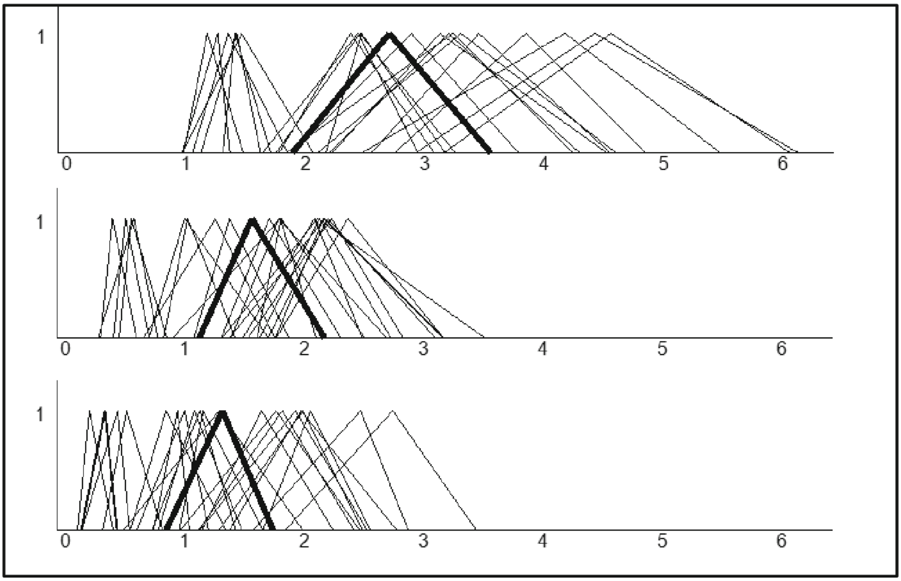


Fig. 2. Triangular fuzzy observations for Factor A (for r = 1, 2, 3, respectively)

As Theorem 3.1 and Result 3.1, two-way FANOVA statistics for the non-symmetric TFN in Table 2 can be calculated. For instance, for $m = 1$, the total sum of squares is calculated as follows

$$\widetilde{SST} = \sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{n_{ij}}\left(y_{ijk} - \overline{y_{...}}\right)^2 + \frac{1}{3x4}\left[\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{n_{ij}}\left(s^l_{y_{ijk}} - \overline{s^l_{y_{...}}}\right)^2 + \sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{n_{ij}}\left(s^r_{y_{ijk}} - \overline{s^r_{y_{...}}}\right)^2\right]$$
$$= \left[(1.33577 - 1.97246)^2 + \ldots\right] + \frac{1}{3x4}\left[\left((0.3 - 0.57213)^2 + \ldots + (0.1 - 0.62459)^2 + \ldots\right)\right] = 66.231.$$

By Eq. 10, the two-way FANOVA test statistics is calculated as $\tilde{f}_A = 53.5145$, $\tilde{f}_B = 44.3465$ and $\tilde{f}_{A*B} = 4.4887$ for $m = 1$. All computations are done by MS Excel.

Comparing the two-way FANOVA test statistics in Table 3, all three H_1 hypotheses are accepted at 0.05 significance level. The critical values of ANOVA tests are $F_{(A)1-\pi,r-1,n_t-rs} = F_{0.95;2;49} = 3.15$ and p-value $= 2.0455\text{x}10^{-4}$, $F_{(B)1-\pi,s-1,n_t-rs} = F_{0.95;3;49} = 2.76$ and p-value $= 1.4692\text{x}10^{-16}$, $F_{(A*B)1-\pi,rs-1,n_t-rs} = F_{0.95;6;49} = 2.25$ and p-value $1.9278\text{x}10^{-4}$. These results strongly show accuracy of H_1 hypotheses. Therefore, there are significant differences between chestnut dimensions (factor A) and chestnut cultivars (factor B) in terms of rupture energy, and the interaction of chestnut sizes and chestnut cultivars (interaction of A*B) is significant, based on the recorded fuzzy data in Table 2.

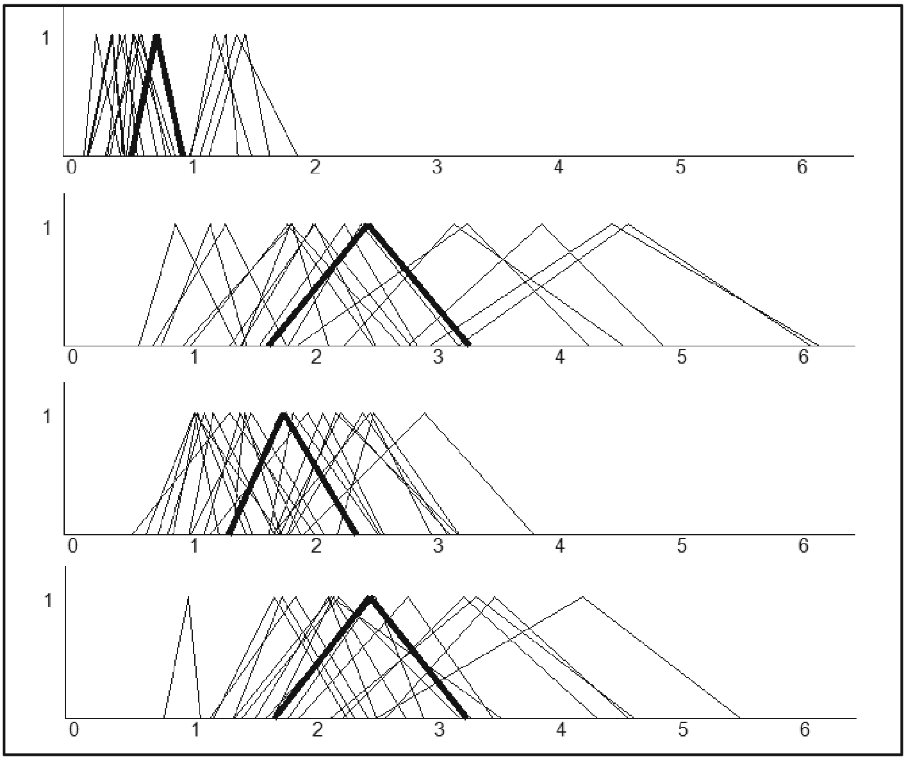


Fig. 3. Triangular fuzzy observations for Factor B (for s = 1, 2, 3, 4, respectively)

Table 2. Fuzzy observations for rupture energy of chestnut dimensions and chestnut cultivars

		Factor A (Chestnut dimensions)		
		Length	Width	Thickness
Factor B (Chestnut cultivars)	**Macit 55**	(1,33577,0,3,0,1), (1,2477,0,2,0,3), (1,4939,0,3,0,2), (1,42266,0,3,0,5)	(0,57185,0,1,0,2), (0,61978,0,1,0,3), (0,64316,0,3,0,2), (0,46361,0,1,0,2)	(0,40446,0,2,0,1), (0,5074,0,3,0,1), (0,5815,0,2,0,3), (0,39659,0,2,0,1), (0,26987,0,1,0,2)
	Akyüz	(4,4979,1,5,1,7), (3,9172,1,1,1), (3,3059,1,4,1), (3,1953,0,9,1,4), (4,6303,1,4,1,5)	(1,8693,0,9,0,3), (1,8577,0,4,0,9), (1,3247,0,6,0,5), (2,4365,0,6,0,8), (2,2977,0,7,0,6)	(1,1962,0,4,0,3), (2,0433,0,6,0,8), (1,8336,0,8,0,7), (2,0572,0,7,0,5), (0,91171,0,3,0,5)

(*continued*)

Table 2. (*continued*)

		Factor A (Chestnut dimensions)		
		Length	Width	Thickness
	Ali Nihat	(1,5344,0,5,0,6), (2,9564,1,0,9), (1,4789,0,1,0,3), (2,4487,0,6,0,7), (2,5106,0,8,0,5), (2,5401,0,3,0,7)	(1,09119,0,2,0,4), (1,06824,0,3,0,7), (1,87566,0,2,0,7), (2,2293,0,4,1), (2,2675,0,5,0,9), (1,4444,0,3,0,5)	(1,223,0,2,0,5), (1,1487,0,3,0,4), (1,0676,0,4,0,2), (1,9966,0,8,0,6), (1,3568,0,8,0,7), (2,1223,0,4,0,5)
	Bouche de Bectizac	(3,2674,1,1,1), (3,3726,1,2,1,3), (2,5254,0,7,0,8), (4,245,1,7,1,3), (3,5231,0,9,1,1)	(2,1968,0,5,0,6), (2,2432,0,8,1), (1,7754,0,4,0,6), (2,1552,0,6,0,4), (2,1818,0,8,1,4)	(1,8852,0,7,0,6), (1,0083,0,2,0,1), (1,7116,0,5,0,6), (2,8114,0,9,0,7), (2,5425,0,9,0,4)

Table 3. Details of two-way FANOVA for the rupture energy of chestnut dimensions and chestnut cultivars

Source of variation	$\widetilde{SS}$	df	$\widetilde{MS}$	$\tilde{f}$
Factor A	2	25,76487	11,34578	69,39778
Factor B	3	28,20611	9,402038	57,50867
Interaction A*B	6	5,709964	0,951661	5,820944
Error	49	7,805453	0,163489	
Total	60	66,2309		

5 Conclusion

In applied science such as agricultural machinery and technology, it can be encountered with vague situations such as determination of mechanical properties robustness [12]. In such problems, ANOVA can not solve the uncertain test so the problems have to generalize basis of fuzzy observations [1–3, 8]. The proposed two-way FANOVA model is a generalized form of the classical two-way ANOVA based on non-symmetric TFNs. In other words, if all observations are real numbers, two-way FANOVA reduces to two-way ANOVA because vagueness of the fuzzy statistics are removed and where is only the centre point of them [1, 2].

For further research, the proposed two-way FANOVA approach can be adapted to different types of fuzzy numbers and fuzzy sets, such as trapezoidal fuzzy number, intuitionistic fuzzy sets.

References

1. Parchami, A., Mashinchi, M., Kahraman, C.: A case study on vehicle battery manufacturing using fuzzy analysis of variance. In: Kahraman, C., Cevik Onar, S., Oztaysi, B., Sari, I. U., Cebi, S., Tolga, A. C. (eds.) Intelligent and Fuzzy Techniques: Smart and Innovative Solutions, pp. 916–923. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-51156-2_106
2. Parchami, A., Mashinchi, M., Kahraman, C.: An implication of fuzzy ANOVA in vehicle battery manufacturing. J. Mahani Math. Res. **10**(2), 33–47 (2021)
3. Parchami, A., Nourbakhsh, M.R., Mashinchi, M.: Analysis of variance in uncertain environments. Complex Intell. Syst. **3**, 189–196 (2017)
4. Şenoğlu, B., Acıtaş, Ş.: İstatistiksel deney tasarımı: sabit etkili modeller. Nobel (2011). (Turkish)
5. Hocking, R.R.: Methods and Applications of Linear Models: Regression and the Analysis of Variance. Wiley, New York (1996)
6. Armstrong, R.A., Eperjesi, F., Gilmartin, B.: The application of analysis of variance (ANOVA) to different experimental designs in optometry. Ophthalmic Physiol. Opt. **22**(3), 248–256 (2002)
7. Yıldız, T., Cevher, E.Y.: Some mechanical properties of chestnut in relation to product processing and equipment design. Turkish J. Agric.-Food Sci. Technol. **10**(8), 1565–1570 (2022)
8. Kahraman, C., Kabak, Ö.: Fuzzy statistical decision-making. Fuzzy Stat. Decision-Making: Theory Appl., 315–327 (2016)
9. Yurtlu, Y.B., Yeşiloğlu, E.: Mechanical behaviour and split resistance of chestnut under compressive loading (2011)
10. Öztekin, Y.B., Taner, A., Duran, H.: Chestnut (Castanea sativa Mill.) cultivar classification: an artificial neural network approach. Notulae Botanicae Horti Agrobotanici Cluj-Napoca **48**(1), 366–377 (2020)
11. Massantini, R., Moscetti, R., Frangipane, M.T.: Evaluating progress of chestnut quality: a review of recent developments. Trends Food Sci. Technol. **113**, 245–254 (2021)
12. Mancuso, S., Ferrini, F., Nicese, F.P.: Chestnut (Castanea sativa Mill.) genotype identification: an artificial neural network approach. J. Hortic. Sci. Biotechnol. **74**(6), 777–784 (1999)
13. Serdar, Ü., et al.: Horticultural characteristics of chestnut growing in Turkey. Erwerbs-obstbau **60**(3), 239–245 (2018)
14. 14. De Garibay, V.G.: Behaviour of fuzzy ANOVA. Kybernetes **16**(2), 107–112 (1987)
15. López-Díaz, M., et al.: Introduction to ANOVA with fuzzy random variables. In: Soft Methodology and Random Information Systems, vol. 26, pp. 487–494. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-44465-7_60
16. Montenegro, M., Colubi, A., Casals, M.R., Gil, M.A.: Asymptotic and bootstrap techniques for testing the expected value of a fuzzy random variable. Metrika **59**, 31–49 (2004)
17. Cuevas, A., Febrero, M., Fraiman, R.: An anova test for functional data. Comput. Stat. Data Anal. **47**(1), 111–122 (2004)
18. Konishi, M., Okuda, T., Asai, K.: Analysis of variance based on fuzzy interval data using moment correction method. Int. J. Innov. Comput. Inf. Control **2**(1), 83–99 (2006)
19. Gil, M.A., Montenegro, M., González-Rodríguez, G., Colubi, A., Casals, M.R.: Bootstrap approach to the multi-sample test of means with imprecise data. Comput. Stat. Data Anal. **51**(1), 148–162 (2006)
20. González-Rodríguez, G., Colubi, A., Gil, M.Á.: Fuzzy data treated as functional data: a one-way ANOVA test approach. Comput. Stat. Data Anal. **56**(4), 943–955 (2012)
21. Wu, H.C.: Analysis of variance for fuzzy data. Int. J. Syst. Sci. **38**(3), 235–246 (2007)

22. Nourbakhsh, M., Mashinchi, M., Parchami, A.: Analysis of variance based on fuzzy observations. Int. J. Syst. Sci. **44**(4), 714–726 (2013)
23. Parchami, A., Ivani, R., Mashinchi, M., Kaya, İ.: An implication of fuzzy ANOVA: metal uptake and transport by corn grown on a contaminated soil. Chemom. Intell. Lab. Syst. **164**, 56–63 (2017)
24. Parchami, A., Sadeghpour-Gildeh, B., Nourbakhsh, M., Mashinchi, M.: A new generation of process capability indices based on fuzzy measurements. J. Appl. Stat. **41**(5), 1122–1136 (2014)
25. Anuradha, D., Kalpanapriya, D.: Intuitionistic fuzzy ANOVA and its application in medical diagnosis. Res. J. Pharm. Technol. **11**(2), 653–656 (2018)
26. Traneva, V., Mavrov, D., Tranev, S.: Fuzzy two-factor analysis of COVID-19 cases in Europe. In: 2020 IEEE 10th International Conference on Intelligent Systems (IS), pp. 533–538. IEEE, August 2020
27. Parthiban, S., Gajivaradhan, P.: A comparative study of LSD under fuzzy environments using trapezoidal fuzzy numbers. IOSR J. Math. (IOSR-JM) **12**(1), 57–75 (2016)
28. Parthiban, S., Gajivaradhan, P.: A comparative study of two factor ANOVA model under fuzzy environments using trapezoidal fuzzy numbers. Int. J. Fuzzy Math. Arch. **10**(1), 1–25 (2016)
29. Aslam, M.: Neutrosophic analysis of variance: application to university students. Complex Intell. Syst. **5**(4), 403–407 (2019). https://doi.org/10.1007/s40747-019-0107-2
30. Nortey, E.N., Eric, W.N., Eunice, O.A.: Neutrosophic-principal component analysis of causes of performance gap among private and public school students in the basic education certificate examination. Asian J. Probab. Stat. **20**(3), 132–149 (2022)
31. Miari, M., Anan, M.T., Zeina, M.B.: Neutrosophic two way ANOVA. Int. J. Neutrosophic Sci. **18**(3), 73–83 (2022)

An Application of Fuzzy ANOVA on Field of Agricultural Machinery

Zeynep Gökkuş[1](✉), Sevil Şentürk[2], Taner Yildiz[3], and Elçin Yeşiloğlu Cevher[3]

[1] Vocational School, Computer Technologies Department, Kastamonu University, Kastamonu, Turkey
z.gokkus@gmail.com

[2] Faculty of Science, Department of Statistics, Eskişehir Technical University, Eskişehir, Turkey
sdeligoz@eskisehir.edu.tr

[3] Faculty of Agriculture, Department of Agricultural Machinery and Technologies Engineering, Ondokuz Mayıs University, Samsun, Turkey
t.yildiz@omu.edu.tr

Abstract. ANOVA is a statistical analysis test that addresses the significance of variances in the means of different populations. If the data includes imprecise or vagueness in the applied sciences, the classical ANOVA cannot be the solution. In which cases, fuzzy ANOVA based on fuzzy observations ought to be used and are particularly more suitable for such problems in contrast to ANOVA, Fuzzy ANOVA (FANOVA) uses fuzzy numbers and is a generalization of ANOVA. In recent years, however; particularly, fuzzy statistical methods have been put into use for the designing process of agricultural machinery. ANOVA for the studies that have been put out regarding literature, there is not a study present that uses the FANOVA method. Therefore, in this study, span, which is one of the foremost factors to be taken into consideration for the design of an agricultural machine to be developed for harvesting grainy products, on the max force that can be applied to a chestnut was evaluated and the results were assessed.

Keywords: Analysis of variance · arithmetic fuzzy numbers · design of agricultural machinery

1 Introduction

As in the whole world, agricultural areas have reached their final limit in our country as well. In order to nourish the increasing population, it is inevitable to produce from these limited areas with the highest quality and efficiency. It is one of the most important works of humanity to maximize agricultural products in terms of quality and quantity. This is only possible with the use of modern agricultural technologies. It is important to know the technical properties of biological material in the design and manufacture of agricultural machinery, product processing, product quality control and product quality improvement. The content of these studies can be listed such as selecting quality products, removing foreign materials from the products and classifying the

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 560–571, 2023.
https://doi.org/10.1007/978-3-031-39777-6_66

products in accordance with their purpose [32]. Knowing the physical and mechanical properties of biological material such as shape, volume, surface area and surface roughness, breaking forces and breaking energies, deformation and hardness contributes to the design and development of agricultural machinery related to the transportation, cleaning-classification, processing and storage of the products [33].

In addition to the design and manufacture of agricultural machinery, knowing the biological and technical characteristics of agricultural products is important in their operation, control, determination of work efficiency, analysis, evaluation of the quality of products and presentation of new vegetable or animal products to consumers. Besides, knowing these properties is important not only for engineers, but also for food scientists, food industrialists and plant breeders. Moreover, it provides benefits for the other designers and experts engaged in animal production [34–37].

Therefore, it is necessary to determine the physico-mechanical properties of agricultural products in biological material testing laboratories and to design and develop them according to the results before the design and manufacture of agricultural machinery. When the studies on agricultural machinery design are examined, it is noteworthy that applications of analysis of variance (ANOVA) and fuzzy methods are frequently used.

2 Literature Review

ANOVA models applied in the field of agricultural machinery design can be listed as: in terms of physical and technical characteristics of three different almond varieties [36]; in terms of elasticity coefficient properties of pea, chickpea and soybean seeds [33]; the breaking forces and breaking energies measured in cotton varieties [36]; and some physico-mechanical properties of three different chestnut cultivars [38] whether they are similar or not was modeled with ANOVA. Moreover, using the ANOVA method to compare the surface quality of the mulberry branch and measure the cutting forces under different conditions [39], comparing of the lateral overturning and backward rollover characteristics of a multipurpose agricultural machine developed in South Korea, with a validated simulation model and the equations obtained from the literature review [40], applying ANOVA on determining the comfort problems of mobile agricultural machinery workers in the working environment and improving the comfort [41], determining the levels of soil compaction in cultivated fields by ANOVA [42], investigation of the relationship between shear parameters such as thrust force and delamination of fiber-pull out, edge chipping, uncut fibers and others damages in composite materials by ANOVA [43], determining of the effect of agricultural machinery on both physical and hydraulic properties of agricultural soils used for rice cultivation in northern Ghana by ANOVA [44], demonstration of the feasibility of maintenance and repair of agricultural machines while they are in operation with ANOVA [45], investigation of the relationship between the adoption of lean production by companies in the agricultural machinery and applications sector in Brazil and their market share and value creation with ANOVA [46], using a single-wheeled test device to evaluation of the effects of three-level speed, three-level slippage and three-wheel load on net traction by ANOVA at the soil bin facility of the Department of Agricultural Machinery of Urmia University [47], applying ANOVA and regression models for 19 indicators related to agricultural machinery systems and agricultural production in QiXing farm in Heilongjiang reclamation area [48], using one-way ANOVA, PCA, k-NN, LDA and SVM methods while determining geographical origins with fingerprint technique on 20 different elements in 135 peanut samples taken from China's Jilin Province, Jiangsu Province, and Shandong Province [49], determining the positioning accuracy of four GNSS receivers with different specifications and operating modes by ANOVA to select a Global Navigation Satellite System (GNSS) according to farm needs [50], and comparing two conceptual interface designs for use in an agricultural air seeder with an existing user interface by ANOVA [51].

In addition to above references, fuzzy methods, which have been applied in almost every field in recent years, have also started to be applied in the field of agricultural machinery design. Such as: choosing the right agricultural machinery service facility with fuzzy AHP [52], considering the real motion situation of agricultural machines in operation, the fuzzy adaptive finite impulse response Kalman filter (FA-FIR-KF) algorithm to integrate position and state information and some necessary to make innovative improvements suggestion of auxiliary optimization algorithms [53], classification of agricultural products according to fruit shape and size using Support Vector Machines (SVMs) and determination of grade using Fuzzy Logic (FL) approach [54], using fuzzy logic to assess the sustainability of rice production in Mazandaran province, Iran [55], combining fuzzy AHP method with global and machine learning-based sensitivity analysis by integrating with geographic information system for agricultural suitability mapping [56], determining the effectiveness of three tractors in the same category operating in climate and soil conditions in the Belgrade (Serbia) region according to their reliability, sustainability and functionality impact indicators with fuzzy set theory and min-max composition [57], joint selection of service providers in network nodes with ANFIS and ANN to design the maintenance service network in agriculture [58], developing autonomous navigation for 4WS agricultural machinery with a path tracking algorithm based on fuzzy control pure tracking model [59], simulation of hydraulic transport robot control system based on fuzzy PID controller [60], and proposing an improved fuzzy Stanley model (SM) based on particle swarm optimization (PSO) in which the control gain is adaptively varied according to tracking error, velocity and orientation, to increase the adaptability of the path tracking algorithm [61].

When the studies in the literature are examined, it is seen that fuzzy ANOVA methods are also studied. Some of these are: surveying the trace of one-way fuzzy analysis of variance (FANOVA) and comparing it's results with regression model [14], whole testing of one-way ANOVA under normal fuzzy observations [15], bootstrap method for approximating the asymptotic one sample tests by fuzzy observations [16], on a given Hilbert space, developing a one-way ANOVA approach for the functional data [17], using the moment correction for uncertain data for processing ANOVA model [18], bootstrap asymptotic multi-sample testing of means for fuzzy observations [19, 20], regarding the cuts of fuzzy observations for one-way ANOVA problem on fuzzy data for optimization approach [21], developing one-way ANOVA for fuzzy observations basis of extension principle approach [22], comparing several groups with FANOVA [23], generalizing PCIs are able to measure the capability of a fuzzy-valued process in producing products based on a triangular fuzzy observations (TFNs) [24], taking an interest in the

Latin Square Design problem (LSD) test using Trapezoidal Fuzzy Numbers [27], proposed testing is analyzed under varieties of trapezoidal fuzzy models [28], a new method for testing intuitionistic fuzzy hypotheses of one-factor and two-factor ANOVA model using intuitionistic fuzzy data samples [25], applying an intuitionistic fuzzy two-factor ANOVA (2-D IFANOVA), through intuitionistic fuzzy sets (IFSs) and index matrices (IMs), investigating how the number of COVID-19 cases depend on the "density" and "climate zone" factors for Europe [26], introducing the neutrosophic analysis of variance (NANONA) [29], presenting a novel Neutrosophic-Principal Component Analysis and Two-way Neutrosophic ANOVA to analyzing the reasons of Performance Gap between private and public school students [30], and generalizing Two-Way ANOVA to neutrosophic two-way ANOVA solving neutrosophic data [31]. Besides the FANOVA models and applications mentioned above, there is extending one-way ANOVA to a case where observations are non-symmetric TFNs rather than real numbers [1–3].

The studies mentioned above show that both ANOVA and fuzzy methods are used intensively and up-to-date in the field of agricultural machinery. However, although there are FANOVA methods available in the literature, no FANOVA application has been found in the field of agricultural machinery. The aim of this study is to use the one-way FANOVA method for cases where there are non-symmetric triangular observations in the field of agricultural machinery design.

The aim of this study is to use the one-way FANOVA method for cases where there are non-symmetric triangular observations in the field of design of agricultural machinery. For this reason, first of all, the one-way FANOVA method for unsymmetrical triangular observations are briefly explained. The method is applied to a sample data in the field of design of the agricultural machinery in chapter 4, and finally the results are discussed and future work plans are mentioned.

3 One Way Analysis of Variance for Non-symmetric Triangular Fuzzy Observations

In the classical one-way ANOVA model, the variability within and between groups according to one-factor is examined. There are many sources in the literature for the details of these processes and can be examined [4–6]. Here, we are interested in a fuzzy ANOVA (FANOVA) model, which is very similar to the ANOVA model and easy to implement [1, 2].

Let X be universal set and $F(X) = \{A|A : X \to [0, 1]\}$. Any $A \in F(X)$ is called a fuzzy set on X. Most of all, let R be set of all real numbers. It is used [1, 2] $F_T(R) = \left\{\tilde{T}|a, s_a^l, s_a^r \in R; s_a^l, s_a^r > 0\right\}$, where

$$\tilde{T}(x) = T\left(a, s_a^l, s_a^r\right)(x) = \begin{cases} \left(x - a + s_a^l\right)/s_a^l \ if \ a - s_a^l \le x < a \\ \left(a + s_a^r - x\right)/s_a^r \ if \ a \le x < a + s_a^r \\ 0, \ elsewhere\,. \end{cases} \tag{1}$$

Any $\tilde{T} \in F_T(R)$ is called a triangular fuzzy number (TFN). Here, a is center point, s_a^l and s_a^r are the left and right widths of a and it can be written as $T\left(a, s_a^l, s_a^r\right)$ [1, 2].

There are extensive studies for arithmetic operations on non-symmetric triangular numbers expressed as $\tilde{T}$. For detailed information [1, 2] and [3] can be examined.

Suppose the data or observations are TFNs of the form $\tilde{y}_{ijk} = T\left(y_{ijk}, s^l_{y_{ijk}}, s^r_{y_{ijk}}\right)$ to represent $\tilde{y}_{ij}$ about y_{ij}. All theoretical model elements such as statistical hypotheses, random variables and population parameters are crisp, and the model is $Y_{ij} = \mu + \theta_i + \varepsilon_{ij}$, with $\varepsilon_{ij} \sim N\left(0, \sigma^2\right)$. Here, Y_{ij}'s are random variables and statictical hypotheses for FANOVA are similar to those for ANOVA, are written as follows,

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_r$
$H_1 : notall \mu_i areequal$ (there is at least one pair with unequal means).

However, there is only one point in this model that differs from the classical ANOVA assumptions that observations are non-symmetric TFNs rather than real numbers. When $\tilde{y}_{ij}$'s are non-symmetric TFNs, the observed $\overline{\tilde{y}_{..}}$ and $\overline{\tilde{y}_{i.}}$ values can be calculated as follows [1, 2],

$$\overline{\tilde{y}_{..}} = \frac{1}{n_t} \sum\nolimits_{i=1}^{r} \sum\nolimits_{j=1}^{n_i} \tilde{y}_{ij} = T\left(\overline{y_{..}}, \overline{s^l_{y_{..}}}, \overline{s^r_{y_{..}}}\right) \tag{2}$$

$$\overline{\tilde{y}_{i.}} = \frac{1}{n_i} \sum\nolimits_{j=1}^{n_i} \tilde{y}_{ij} = T\left(\overline{y_{i.}}, \overline{s^l_{y_{i.}}}, \overline{s^r_{y_{i.}}}\right). \tag{3}$$

Observed SST, SSTR, SSE, MSTR, MSE and $\tilde{f}$ statistical values can be obtained as follows [1, 2],

$$\tilde{sst} = \sum\nolimits_{i=1}^{r} \sum\nolimits_{j=1}^{n_i} \left(\tilde{y}_{ij} \ominus \overline{\tilde{y}_{..}}\right)^2 \tag{4}$$

$$\tilde{sstr} = \sum\nolimits_{i=1}^{r} n_i \left(\overline{\tilde{y}_{i.}} \ominus \overline{\tilde{y}_{..}}\right)^2 \tag{5}$$

$$\tilde{sse} = \sum\nolimits_{i=1}^{r} \sum\nolimits_{j=1}^{n_i} \left(\tilde{y}_{ij} \ominus \overline{\tilde{y}_{i.}}\right)^2 \tag{6}$$

$$\tilde{mstr} = \frac{\tilde{sstr}}{r-1} \tag{7}$$

$$\tilde{mse} = \frac{\tilde{sse}}{n_t - r} \tag{8}$$

$$\tilde{f} = \frac{\tilde{mstr}}{\tilde{mse}} = \frac{n_t - r}{r-1} \frac{\tilde{sstr}}{\tilde{sse}}. \tag{9}$$

The decision rule of FANOVA is similar to ANOVA. If the FANOVA test statistic $\tilde{f}$ is greater than, the α th quantile of fisher distribution with $r-1$ and $n_t - r$ degrees of freedom, $F_{1-\alpha;r-1;n_t-r}$, the H_0 hypothesis is rejected otherwise H_1 is rejected [1, 2].

When the observations are the crisp numbers y_{ij}, that is they are indicator functions $I_{\{y_{ij}\}}$ for $i = 1, \ldots, r$ and $j = 1, \ldots, n_i$, then all the extended statistics for FANOVA coincide with ANOVA statistics [1, 2]. In this study, m = 1 is taken to ensure that the α-cut membership value of the TFNs is 1. And, observed crisp values of $s\tilde{s}t$, $s\tilde{s}tr$, $s\tilde{s}e$, $M\tilde{S}TR$, $M\tilde{S}E$ and $\tilde{f}$ are as follows:

$$S\tilde{S}T = sst_y + \frac{1}{(m+2)(m+3)}\left(sst_{s_y^l} + sst_{s_y^r}\right) \tag{10}$$

$$S\tilde{S}TR = sstr_y + \frac{1}{(m+2)(m+3)}\left(sstr_{s_y^l} + sstr_{s_y^r}\right) \tag{11}$$

$$S\tilde{S}E = sse_y + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right) \tag{12}$$

$$M\tilde{S}TR = \frac{S\tilde{S}TR}{r-1} = mstr_{y} + \frac{1}{(m+2)(m+3)}\left(mstr_{s_y^l} + mstr_{s_y^r}\right) \tag{13}$$

$$M\tilde{S}E = \frac{S\tilde{S}E}{n_t - r} = sse_{y} + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right) \tag{14}$$

$$\tilde{f} = \frac{M\tilde{S}TR}{M\tilde{S}E} = \frac{mstr + \frac{1}{(m+2)(m+3)}\left(mstr_{s_y^l} + mstr_{s_y^r}\right)}{sse + \frac{1}{(m+2)(m+3)}\left(sse_{s_y^l} + sse_{s_y^r}\right)}. \tag{15}$$

where sst_y, $sstr_y$ and sse_y are real values of sst, $sstr$ and sse for the centre points of $T\left(y_{ijk}, s_{y_{ijk}}^l, s_{y_{ijk}}^r\right)$, and $sst_{s_y^l}$, $sstr_{s_y^l}$ and $sse_{s_y^l}$ the real values of sst, $sstr$ and sse for the left widths of $T\left(y_{ijk}, s_{y_{ijk}}^l, s_{y_{ijk}}^r\right)$, and $sst_{s_y^r}$, $sstr_{s_y^r}$ and $sse_{s_y^r}$ the real values of sst, $sstr$ and sse for the right widths of $T\left(y_{ijk}, s_{y_{ijk}}^l, s_{y_{ijk}}^r\right)$, respectively.

The values obtained by Eq. 10–15 can be summarized in Table 1.

Table 1. One-way FANOVA table

Source of variation	$\tilde{ss}$	Degrees of freedom	$\tilde{ms}$	$\tilde{f}$
Between treatments	$S\tilde{S}TR$	$r-1$	$M\tilde{S}TR = \frac{S\tilde{S}TR}{r-1}$	$\tilde{f} = \frac{M\tilde{S}TR}{M\tilde{S}E}$
Within treatments (error)	$S\tilde{S}E$	$n_t - r$	$M\tilde{S}E = \frac{S\tilde{S}E}{n_t-r}$	
Total	$S\tilde{S}T$	$n_t - 1$		

4 Results of FANOVA on Design of the Agricultural Machinery for Grainy Products

Knowing the biological and technical properties of grainy products are important in the design and development of agricultural machinery [7, 9]. The one of the most important factors in design of agricultural machinery separation systems, conveying of agricultural product, processing and packaging systems are mechanical properties [10]. Damages in seed and grainy products mostly occur during harvesting-threshing, cleaning, sorting and transmission. The effect of external forces that cause mechanical damage to grainy products can be caused by static or dynamic impact loads [11]. Mechanical effects can damage the crops during, harvest and post-harvest. In addition, it can adversely affect the storage and shelf life of products. Therefore, it is necessary to determine the maximum loads of grainy products depending on their size under mechanical loads [13].

Rupture energy measurements of four different chestnut cultyvars (Macit 55, Akyüz, Ali Nihat and Bouche de Betizac) are determined at the same loading rate and in three different axes. The X-axis is the length, the Y-axis is the width, and the Z-axis is the thickness dimension of the chestnuts. The data in the study are at a loading speed of 10 mm min^{-1} of chestnut cultivars; It was obtained by static loading in the x-x (length), y-y (width), and z-z (thickness) axes at a humidity level of 75–85%.

In this study, data that can be used in the design of an agricultural machinery that can be used in the stages of harvesting, transportation, cleaning-classification, processing and storage of chestnuts without causing any damage and product loss are taken into account. The largest dimension of the chestnut is its length [7]. For this reason, in this study, the amount of rupture energy to be applied by the machine to be designed according to the length dimension of the chestnut is measured. As a result, it is determined whether the rupture energy measurements are similar according to the length dimension of different chestnut cultivars. Therefore, "$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$" versus "$H_1 : notall \mu_i' sareequal, fori = 1, 2, 3, 4$" is tested. In this test, whether the rupture energy measurements according to the length dimension are same on four different chestnut cultivars is investigated by one-way FANOVA. Here, observations for four different chestnut cultivars measured for length dimension are considered in four columns and the observations are non-symmetric TFNs.

For this rupture energy test, a total of 20 non-symmetric triangular rupture energy measures are selected in the one-way FANOVA test, with sample sizes $n_1 = 4, n_2 = 5, n_3 = 6,$ and$n_4 = 5$, respectively (see Table 2) (Fig. 1).

According to the methodology given in Sect. 3, the FANOVA statistics reported in Table 3 can be obtained by performing the necessary arithmetic operations on the non-symmetric TFNs given in Table 2. All computations are performed with MS Excel software. The observed $S\tilde{S}T$, $SS\tilde{T}R$ and $S\tilde{S}E$ values are calculated as follows,

$$\begin{aligned}
\widetilde{SST} &= \sum_{i=1}^{4}\sum_{j=1}^{n_i}\left(y_{ij}-\overline{y_{..}}\right)^2 + \tfrac{1}{3\times 4}\left[\sum_{i=1}^{4}\sum_{j=1}^{n_i}\left(s^l_{y_{ij}}-\overline{s^l_{y_{..}}}\right)^2 + \sum_{i=1}^{4}\sum_{j=1}^{n_i}\left(s^r_{y_{ij}}-\overline{s^r_{y_{..}}}\right)^2\right] \\
&= \left[(1.33577-2.7725)^2+\ldots+(3.5231-2.7725)^2\right] + \tfrac{1}{3\times 4}\left[(0-0.79)^2+\ldots\right. \\
&\quad \left.(1.5-0.79)^2+(1-0.81)^2+\ldots+(0.5-0.81)^2\right] = 23.597.
\end{aligned}$$

$$\begin{aligned}
\widetilde{SSTR} &= \sum_{i=1}^{4} n_i\left(\overline{y_{i.}}-\overline{y_{..}}\right)^2 + \tfrac{1}{3x4}\left[\sum_{i=1}^{4} n_i\left(s^l_{y_{i.}}-\overline{s^l_{y_{..}}}\right)^2 + \sum_{i=1}^{4} n_i\left(s^r_{y_{i.}}-\overline{s^r_{y_{..}}}\right)^2\right] \\
&= \left[4(1.375-2.77)^2+5(3.91-2.77)^2+\ldots\right] + \tfrac{1}{3\times 4}\left[4(0.525-0.79)^2+5(1.02-0.79)^2+\ldots\right. \\
&\quad \left.+4(0.825-0.81)^2+5(1.16-0.81)^2+\ldots\right] = 17.989.
\end{aligned}$$

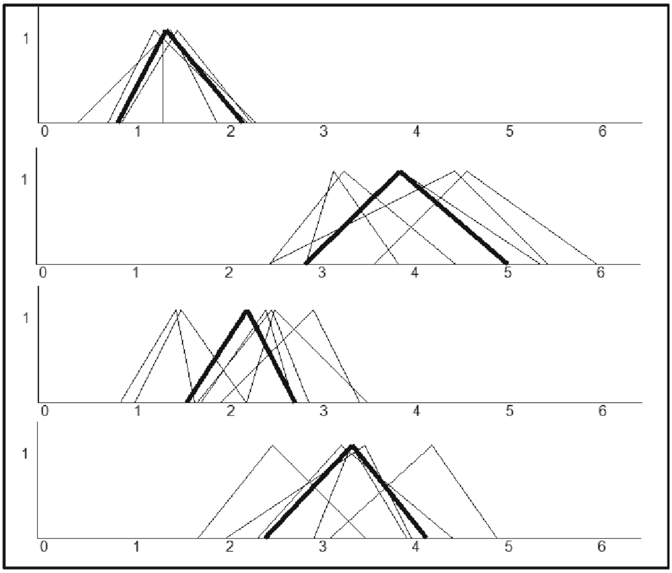


Fig. 1. Non-symmetric triangular observations for rupture energy according to length dimension of chestnut cultivars (i = 1, 2, 3, 4 for Macit 55, Akyüz, Ali Nihat and Bouche de Bectizac, respectively)

Table 2. Non-symmetric triangular observations for rupture energy according to length dimension of chestnut cultivars

Observation #	Macit 55	Akyüz	Ali Nihat	Bouche de Bectizac
1	(1.33577, 0, 1)	(4.4979, 2, 1)	(1.5344, 0.5, 0.7)	(3.2674, 0.9, 1.2)
2	(1.2477, 0.5, 1)	(3.9172, 1, 1.5)	(2.9564, 1, 0.5)	(3.3726, 0.4, 0.6)
3	(1.4939, 0.6, 0.8)	(3.3059, 0.8, 1.2)	(1.4789, 0.6, 0.2)	(2.5254, 0.8, 1)
4	(1.42266, 1, 0.5)	(3.1953, 0.3, 0.7)	(2.4487, 0.7, 0.3)	(4.245, 1.1, 0.7)
5		(4.6303, 1, 1.4)	(2.5106, 0.8, 0.4)	(3.5231, 1.5, 0.5)
6			(2.5401, 0.3, 1)	

$$\widetilde{SSE} = \sum_{i=1}^{4}\sum_{j=1}^{n_i}\left(y_{ij} - \overline{y_{i.}}\right)^2 + \frac{1}{3\times 4}\left[\sum_{i=1}^{4}\sum_{j=1}^{n_i}\left(s^l_{y_{ij}} - \overline{s^l_{y_{i.}}}\right)^2 + \sum_{i=1}^{4}\sum_{j=1}^{n_i}\left(s^r_{y_{ij}} - \overline{s^r_{y_{i.}}}\right)^2\right]$$
$$= \left[(1.33577 - 1.375)^2 + \ldots + (4.4979 - 3.91)^2 + \ldots\right] + \frac{1}{3\times 4}\left[(0 - 0.525)^2 + \ldots + (2 - 1.02)^2 + \ldots \right.$$
$$\left. +(1 - 0.825)^2 + \ldots + (1 - 1.16)^2 + \ldots\right] = 5.444.$$

Finally, the other observed values and FANOVA test statistic $\tilde{f}$ can be seen from Table 3.

Critical value of ANOVA test is $F_{1-\alpha,r-1,n_t-r} = F_{0.05;3;16} = 3.24.$. Comparing the obtained FANOVA test statistic ($\tilde{f}$) with the F, H0 can be rejected at a significance level

Table 3. One-way FANOVA results for non-symmetric triangular rupture energy measurments

Source of variation	$\tilde{ss}$	Degrees of freedom	$\tilde{ms}$	$\tilde{f}$
Between treatments	17.989	3	5.996	265,87
Within treatments (error)	5.444	16	0.023	
Total	23.597	19		

of 0.05. Also, the p-value $=1.86875 \times 10^{-6}$ can be calculated. This strongly indicates the accuracy of H1. Therefore, based on the fuzzy data recorded in Table 2, it is clear that chestnut cultivars differ according to rupture energy measurements.

5 Conclusions

It can be encountered with vague situations the determination of robustness of the mechanical properties in applied sciences such as agricultural machinery and technologies [12]. In such problems, ANOVA could not solve the uncertain test. Because of this, the problems have to generalize basis of fuzzy observations [1–8].

The one-way FANOVA model is a generalized form of the classical one-way ANOVA based on non-symmetric TFNs. In other words, if all observations are real numbers, one-way FANOVA reduces to one-way ANOVA because vagueness of the fuzzy statistics are removed and where is only the centre point of them [1, 2].

The results showed that the measurements of rupture energy for the four chestnut cultivars differ for the length dimension. This situation reveals the necessity of being more careful in the design and development of the agricultural machinery, technologies and systems. Damages occur mostly during harvesting-threshing, cleaning, sorting and transmission in the grainy products. For example, damaged grains in wheat reduce the mill quality. If it is to be used as a seed, its germination ability decreases. If cracks occur in grainy products, the microbiological activity increases. This causes deterioration of fatty acids, accumulation of oil in the oil seeds, and plunging of the machine elements. In addition, the cooking quality of grainy products such as beans decreases.

For further research, the one-way FANOVA approach can be adapted to different types of fuzzy numbers and fuzzy sets, such as trapezoidal fuzzy number, intuitionistic fuzzy sets.

References

1. Parchami, A., Mashinchi, M., Kahraman, C.: A case study on vehicle battery manufacturing using fuzzy analysis of variance. In: Kahraman, C., Cevik Onar, S., Oztaysi, B., Sari, I.U., Cebi, S., Tolga, A.C. (eds.) INFUS 2020. AISC, vol. 1197, pp. 916–923. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-51156-2_106
2. Parchami, A., Mashinchi, M., Kahraman, C.: An implication of fuzzy ANOVA in vehicle battery manufacturing. J. Mahani Math. Res. **10**(2), 33–47 (2021)
3. Parchami, A., Nourbakhsh, M.R., Mashinchi, M.: Analysis of variance in uncertain environments. Complex Intell. Syst. **3**, 189–196 (2017)

4. Şenoğlu, B., Acıtaş, Ş.: İstatistiksel deney tasarımı: sabit etkili modeller. Nobel (2011). (Turkish)
5. Hocking, R.R.: Methods and Applications of Linear Models: Regression and the Analysis of Variance. Wiley, New York (1996)
6. Armstrong, R.A., Eperjesi, F., Gilmartin, B.: The application of analysis of variance (ANOVA) to different experimental designs in optometry. Ophthalmic Physiol. Opt. **22**(3), 248–256 (2002)
7. Yıldız, T., Cevher, E.Y.: Some mechanical properties of chestnut in relation to product processing and equipment design. Turk. J. Agric.-Food Sci. Technol. **10**(8), 1565–1570 (2022)
8. Kahraman, C., Kabak, Ö.: Fuzzy statistical decision-making. In: Kahraman, C., Kabak, Ö. (eds.) Fuzzy statistical decision-making. SFSC, vol. 343, pp. 1–12. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39014-7_1
9. Yurtlu, Y.B., Yeşiloğlu, E.: Mechanical behaviour and split resistance of chestnut under compressive loading (2011)
10. Öztekin, Y.B., Taner, A., Duran, H.: Chestnut (Castanea sativa Mill.) cultivar classification: an artificial neural network approach. Notulae Botanicae Horti Agrobotanici Cluj-Napoca **48**(1), 366–377 (2020)
11. Massantini, R., Moscetti, R., Frangipane, M.T.: Evaluating progress of chestnut quality: a review of recent developments. Trends Food Sci. Technol. **113**, 245–254 (2021)
12. Mancuso, S., Ferrini, F., Nicese, F.P.: Chestnut (Castanea sativa Mill.) genotype identification: an artificial neural network approach. J. Horticult. Sci. Biotechnol. **74**(6), 777–784 (1999)
13. Serdar, Ü., et al.: Horticultural characteristics of chestnut growing in Turkey. Erwerbs-obstbau **60**(3), 239–245 (2018)
14. de Garibay, V.G.: Behaviour of fuzzy ANOVA. Kybernetes **16**(2), 107–112 (1987)
15. Montenegro, M., González-Rodríguez, G., Gil, M.A., Colubi, A., Casals, M.R.: Introduction to ANOVA with fuzzy random variables. In: López-Díaz, M., Gil, M.Á., Grzegorzewski, P., Hryniewicz, O., Lawry, J. (eds.) Soft methodology and random information systems, pp. 487–494. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-44465-7_60
16. Montenegro, M., Colubi, A., Casals, M.R., Gil, M.A.: Asymptotic and bootstrap techniques for testing the expected value of a fuzzy random variable. Metrika **59**, 31–49 (2004)
17. Cuevas, A., Febrero, M., Fraiman, R.: An anova test for functional data. Comput. Stat. Data Anal. **47**(1), 111–122 (2004)
18. Konishi, M., Okuda, T., Asai, K.: Analysis of variance based on fuzzy interval data using moment correction method. Int. J. Innov. Comput. Inf. Control **2**(1), 83–99 (2006)
19. Gil, M.A., Montenegro, M., González-Rodríguez, G., Colubi, A., Casals, M.R.: Bootstrap approach to the multi-sample test of means with imprecise data. Comput. Stat. Data Anal. **51**(1), 148–162 (2006)
20. González-Rodríguez, G., Colubi, A., Gil, M.Á.: Fuzzy data treated as functional data: a one-way ANOVA test approach. Comput. Stat. Data Anal. **56**(4), 943–955 (2012)
21. Wu, H.C.: Analysis of variance for fuzzy data. Int. J. Syst. Sci. **38**(3), 235–246 (2007)
22. Nourbakhsh, M., Mashinchi, M., Parchami, A.: Analysis of variance based on fuzzy observations. Int. J. Syst. Sci. **44**(4), 714–726 (2013)
23. Parchami, A., Ivani, R., Mashinchi, M., Kaya, İ: An implication of fuzzy ANOVA: metal uptake and transport by corn grown on a contaminated soil. Chemom. Intell. Lab. Syst. **164**, 56–63 (2017)
24. Parchami, A., Sadeghpour-Gildeh, B., Nourbakhsh, M., Mashinchi, M.: A new generation of process capability indices based on fuzzy measurements. J. Appl. Stat. **41**(5), 1122–1136 (2014)
25. Anuradha, D., Kalpanapriya, D.: Intuitionistic fuzzy ANOVA and its application in medical diagnosis. Res. J. Pharm. Technol. **11**(2), 653–656 (2018)

26. Traneva, V., Mavrov, D., Tranev, S.: Fuzzy two-factor analysis of COVID-19 cases in Europe. In: 2020 IEEE 10th international conference on intelligent systems (IS), pp. 533–538. IEEE (2020)
27. Parthiban, S., Gajivaradhan, P.: A comparative study of LSD under fuzzy environments using trapezoidal fuzzy numbers. IOSR J. Math. (IOSR-JM) **12**(1), 57–75 (2016)
28. Parthiban, S., Gajivaradhan, P.: A comparative study of two factor ANOVA model under fuzzy environments using trapezoidal fuzzy numbers. Int. J. Fuzzy Math. Arch. **10**(1), 1–25 (2016)
29. Aslam, M.: Neutrosophic analysis of variance: application to university students. Complex Intell. Syst. **5**(4), 403–407 (2019). https://doi.org/10.1007/s40747-019-0107-2
30. Nortey, E.N., Eric, W.N., Eunice, O.A.: Neutrosophic-principal component analysis of causes of performance gap among private and public school students in the basic education certificate examination. Asian J. Prob. Stat. **20**(3), 132–149 (2022)
31. Miari, M., Anan, M.T., Zeina, M.B.: Neutrosophic two way ANOVA. Int. J. Neutrosophic Sci. **18**(3), 73–83 (2022)
32. Yağcıoğlu, A.: Tarımsal Elektrifikasyon (Genişletilmiş 2. Basım). EÜZF Yayınları, Turkish, p 488 (19966)
33. Saraçoğlu, T., Şengün, G.N.: Bazı Tohumların İki Farklı Yöntemle Esneklik Katsayılarının Belirlenmesi ve Karşılaştırılması. Tarım Makinaları Bilimi Dergisi **14**(2), 135–142 (2018)
34. Mohsenin, N.N.: Application of engineering techniques to evaluation of texture of solid food materials. J. Texture Stud. **1**(2), 133–154 (1970)
35. Isık, A.: A research on determining agricultural structure and mechanization properties of the farms in Çukurova region. In: 6th International Congress on Agricultural Mechanization and Energy, Ankara (1996)
36. Kiliçkan, A., Coşkun, M.B.: Ege bölgesinde yetiştirilen farklı pamuk çeşitlerinin makinalı hasada ilişkin bazı fiziksel özelliklerinin belirlenmesi. adnan menderes üniversitesi ziraat fakültesi dergisi **6**(1), 87–90 (2009). (Turkish)
37. Sezer, S. A., Çetin, M.: Erik Meyvesinin Farklı Hasat Dönemlerindeki Bazı Fiziksel ve Mekanik Özelliklerinin Belirlenmesi. Anadolu Tarım Bilimleri Dergisi **36**(1), 73–79 (2021). (Turkish)
38. Hamleci, B., Güner, M.: Kestanenin Sıkıştırma Yükü Altındaki Mekanik Davranışlarının Belirlenmesi. Tarım Makinaları Bilimi Dergisi **11**(4), 301–307 (2015). (Turkish)
39. Meng, Y., Wei, J., Wei, J., Chen, H., Cui, Y.: An ANSYS/LS-DYNA simulation and experimental study of circular saw blade cutting system of mulberry cutting machine. Comput. Electron. Agric. **157**, 38–48 (2019)
40. Hwang, S.J., Jang, M.K., Nam, J.S.: Application of lateral overturning and backward rollover analysis in a multi-purpose agricultural machine developed in south korea. Agronomy **11**(2), 297 (2021)
41. Hostens, I., Papaioannou, G., Spaepen, A., Ramon, H.: Buttock and back pressure distribution tests on seats of mobile agricultural machinery. Appl. Ergon. **32**(4), 347–355 (2001)
42. Tonui, W., Ndiema, C., Mutai, E.K.: Determination of soil compaction levels by agricultural machinery in cultivated fields using dynamic cone penetrometer. Int. J. Res. Appl. Sci. Eng. Technol. **4**, 494–500 (2016)
43. Sakthivel, M., Vijayakumar, S., Prasad, N.K.: Drilling analysis on basalt/sisal reinforced polymer composites using ANOVA and regression model. Appl. Math. Sci. **9**(66), 3285–3290 (2015)
44. Abagale, F.K.: Effect of agricultural machinery on physical and hydraulic properties of agricultural soils. J. Soil Sci. Environ. Manag. **12**(2), 58–65 (2021)
45. Skobtsov, I.G., Shilovsky, V.N., Konanov, D.G.: Integrated system of criteria for determining operational maintenance of forest and agricultural machinery. In: IOP Conference Series: Earth and Environmental Science, vol. 954, no. 1, p. 012072. IOP Publishing (2022)

46. Forrester, P. L., Shimizu, U.K., Soriano-Meier, H., Arturo Garza-Reyes, J., Basso, L.F.C.: Lean production, market share and value creation in the agricultural machinery sector in Brazil. J. Manuf. Technol. Manag. **21**(7), 853–871 (2010)
47. Taghavifar, H., Mardani, A.: Net traction of a driven wheel as affected by slippage, velocity and wheel load. J. Saudi Soc. Agric. Sci. **14**(2), 167–171 (2015)
48. Yu, X., Leng, Z., Zhang, H.: Optimal models for impact of agricultural machinery system on agricultural production in Heilongjiang agricultural reclamation area. In: 2012 24th Chinese Control and Decision Conference (CCDC), pp. 832–835. IEEE (2012)
49. Zhao, H., Wang, F., Yang, Q.: Origin traceability of peanut kernels based on multi-element fingerprinting combined with multivariate data analysis. J. Sci. Food Agric. **100**(10), 4040–4048 (2020)
50. Catania, P., et al.: Positioning accuracy comparison of GNSS receivers used for mapping and guidance of agricultural machines. Agronomy **10**(7), 924 (2020)
51. Rakhra, A.K., Green, M.K., Mann, D.D.: The influence of a user-centred design focus on the effectiveness of a user interface for an agricultural machine. Agric. Sci. **11**(11), 947 (2020)
52. Osuch, A., Osuch, E., Rybacki, P., Przygodziński, P., Kozłowski, R., Przybylak, A.: A decision support method for choosing an agricultural machinery service workshop based on fuzzy logic. Agriculture **10**(3), 76 (2020)
53. Li, S., et al.: Agricultural machinery GNSS/IMU-integrated navigation based on fuzzy adaptive finite impulse response Kalman filtering algorithm. Comput. Electron. Agric. **191**, 106524 (2021)
54. Mustafa, N.B.A., Ahmed, S.K., Ali, Z., Yit, W.B., Abidin, A.A.Z., Sharrif, Z.A.M.: Agricultural produce sorting and grading using support vector machines and fuzzy logic. In: 2009 IEEE İnternational Conference on Signal and İmage Processing Applications, pp. 391–396. IEEE (2009)
55. Amini, S., Rohani, A., Aghkhani, M.H., Abbaspour-Fard, M.H., Asgharipour, M.R.: Assessment of land suitability and agricultural production sustainability using a combined approach (Fuzzy-AHP-GIS): a case study of Mazandaran province, Iran. Inf. Process. Agric. **7**(3), 384–402 (2020)
56. Talukdar, S., et al.: Coupling geographic information system integrated fuzzy logic-analytical hierarchy process with global and machine learning based sensitivity analysis for agricultural suitability mapping. Agric. Syst. **196**, 103343 (2022)
57. Miodragović, R., Tanasijević, M., Mileusnić, Z., Jovančić, P.: Effectiveness assessment of agricultural machinery based on fuzzy sets theory. Expert Syst. Appl. **39**(10), 8940–8946 (2012)
58. Ren, W., Wu, K., Gu, Q., Hu, Y.: Intelligent decision making for service providers selection in maintenance service network: an adaptive fuzzy-neuro approach. Knowl.-Based Syst. **190**, 105263 (2020)
59. Zhang, C., Gao, G., Zhao, C., Li, L., Li, C., Chen, X., Research on 4WS agricultural machine path tracking algorithm based on fuzzy control pure tracking model. Machines **10**(7), 597 (2022)
60. Jin, X., Chen, K., Zhao, Y., Ji, J., Jing, P.: Simulation of hydraulic transplanting robot control system based on fuzzy PID controller. Measurement **164**, 108023 (2020)
61. Sun, Y., Cui, B., Ji, F., Wei, X., Zhu, Y.: The full-field path tracking of agricultural machinery based on PSO-enhanced fuzzy stanley model. Appl. Sci. **12**(15), 7683 (2022)

Applications

New Product Development Methodology with Using Sustainability Driven QFD and Neutrosophic Sets: A Case Study of the Aircraft Stabilizer

Sezen Ayber(✉) and Nihal Erginel

Eskisehir Technical University, İki Eylul Campus, 26555 Eskişehir, Turkey
sezenayber@gmail.com, nerginel@eskisehir.edu.tr

Abstract. Today, an increasing number of businesses are realizing that sustainability is an integral part of their business strategy, while meeting customer needs. For this reason, in order to meet social, economic and environmental problems, businesses must consider sustainability requirements as well as customer expectations in the New Product Development (NPD) process. Quality Function Deployment (QFD) is a concept used for transforming customer needs (CNs) into engineering characteristics (ECs) at each phase of NPD and production. In this study, the first matrix of QFD called house of quality (HoQ) is extended with defining sustainability requirements (SRs) besides CNs. In classical QFD studies, exact numbers are often used as a result of human evaluations when determining the importance of CNs and the relationships between CNs and ECs. These evaluations are generally realized by using linguistic terms which causes uncertainty and vagueness. To overcome this problem, the neutrosophic fuzzy set could be applied. In this study, while determining the importance of CNs and SRs, relationships between CNs, SRs and ECs, single valued neutrosophic set (SVNS) is used. The case study is implemented on the horizontal stabilizer of the F-16 aircraft.

Keywords: New Product Development · QFD · Sustainability · Single Valued Neutrosophic Sets · Horizontal Stabilizer

1 Introduction

Companies want to develop new products due to the use of resources, market strategy, desire for growth, competition, technological developments, changes in consumer preferences and globalization [1]. NPD is a process to transform technology estimations, market opportunities and gaps into a product/service [2]. NPD is vital to the survival and profitability of companies, so it has become the focus of businesses in today's competitive business environment.

Customer requirements and wants are very important in the development process of products. In this context, companies must listen to the voice of their customers in order to make the right product and service design decisions. Companies that use the

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 575–583, 2023.
https://doi.org/10.1007/978-3-031-39777-6_67

customer-oriented NPD process will develop economically. On the other hand, there is a need to balance this economic development with environmental protection and social needs. For this reason, sustainability requirements aiming at social, environmental, and economic improvements, as well as customer needs, should be considered together in NPD processes.

QFD is a methodology for transforming customer needs into engineering characteristics at each phase of NPD and production. In this study, the first matrix of QFD is extended with defining sustainability requirements besides customer needs. While defining the importance of CNs and SRs, relationships between CNs, SRs and ECs, single valued neutrosophic sets will be used. The aim of this study is to propose NPD methodology by using sustainability driven QFD and neutrosophic sets. The case study will be implemented on the horizontal stabilizer of the F-16 aircraft.

The rest of this study is arranged as: Section 2 presents the QFD method and the related works with QFD, fuzzy theory and sustainability. Section 3 introduces SVNS. Section 4 proposes NPD methodology with using sustainability driven QFD and SVNS. Section 5 demonstrates the case study for the horizontal stabilizer of the F-16 aircraft. Section 6 presents the conclusions.

2 Literature Review

QFD is a useful method for transforming CNs into ECs conceptualized by Dr. Yoji Akao [3]. It has been commonly used and performed in several areas, such as NPD, quality management, decision sciences, all over the world [4]. QFD uses four matrices: HoQ, part deployment, process planning and operation planning [5]. HoQ occurs at the core of QFD and shows how the ECs satisfy CNs. Inputs of HoQ are CNs, ECs, importance of each CNs, correlations among ECs, relationships between CNs and ECs. The major output of HoQ is the prioritized ECs. Determining the importance of CNs and the relationships between CNs and ECs is generally realized by using linguistic terms which causes uncertainty and vagueness. To overcome this problem, fuzzy set theory could be applied. Kwong and Bai [6] used fuzzy QFD to design the bicycle. Zhang et al. [7] applied fuzzy QFD for ergonomic product design. Onar et al. [8] proposed hesitant fuzzy QFD. Yu et al. [9] suggested intuitionistic fuzzy QFD to design the steering wheel. Van et al. [10] developed neutrosophic QFD to select a green supplier. Haber et al. [11] used fuzzy QFD to design haemodialysis equipment. Haktanır and Kahraman [12] proposed an interval-valued Pythagorean fuzzy QFD. Wang et al. [13] applied fuzzy QFD to design a training device. Gündoğdu and Kahraman [14] developed spherical fuzzy QFD for designing a delta robot technology.

In today's business environment, there is a considerable attention to green and sustainable product design and development. Sustainability aims to manage risks by adapting economic, environmental, and social factors to the operations to create long-term value in businesses. Masui et al. [15] proposed QFDE method considering environmental issues for NPD process. Kuo et al. [16] suggested the Eco-QFD method with using fuzzy theory. Vinodh and Chintha [17] applied fuzzy QFD to achieve sustainable production. Fargnoli et al. [18] proposed a method by using QFD and life cycle analysis to handle environment and cost aspects of sustainability. Finger and Lima-Junior [19] applied hesitant fuzzy QFD to make supplier programs. In the literature, one of aspects of sustainability generally studied with QFD. Studies handling all aspects of sustainability (economic, environmental, and social) with QFD are rare.

3 Single Valued Neutrosophic Sets

Zadeh [20] first put forward the fuzzy set theory in 1965 to model uncertainty. Traditional fuzzy sets have only membership functions. Atanassov [21] developed the intuitionistic fuzzy set, which is an extension of fuzzy sets, by adding the non-membership function. The sum of the degrees of membership and non-membership must be less than 1. Yager [22] developed the pythagorean fuzzy set, allowing the definition of the sum of the degrees of membership and non-membership to be greater than 1, but does not allow the sum of their squares to be greater than 1. Smarandache [23] introduced the neutrosophic set theory to reflect the human way of thinking more accurately. Neutrosophic set has an important place in the modeling of uncertainty as it is defined by "truth-membership degree", "falsity-membership degree", and "indeterminacy-membership degree". It provides more flexibility compared to other extensions of fuzzy sets. It is also used in solving real-life problems. Wang et al. [24] developed a special case of neutrosophic set, single-valued neutrosophic set. Basic definitions of neutrosophic set are as follows:

Definition 1. [22] A neutrosophic set N in X is defined by:

$$N = \Big\{ < x, (T_N(x), I_N(x), F_N(x)) > \ : x \in X, (T_N(x), I_N(x), F_N(x) \in \Big\} 0^-, 1^+ \Big(\Big) \quad 0^- \le T_N(x) + I_N(x) + F_N(x) \le 3^+ \tag{1}$$

Definition 2. [24] A single valued neutrosophic set N in X is defined by:

$$\begin{aligned} & N = \{< x, (T_N(x), I_N(x), F_N(x)) > \ : x \in X, (T_N(x), I_N(x), F_N(x) \in [0, 1]\} \\ & 0 \le T_N(x) + I_N(x) + F_N(x) \le 3 \end{aligned} \tag{2}$$

Definition 3. [25] A single valued triangular neutrosophic number (SVTNN).

$\tilde{a} = (< (a_1, b_1, c_1); \chi_{\tilde{a}}, \sigma_a, \omega_{\tilde{a}}) >)$ is a neutrosophic set on the set of real numbers, whose truth-membership, indeterminacy-membership and falsity-membership functions are respectively defined by,

$$\mu_{\tilde{a}}(x) = \begin{cases} \frac{(x-a_1)\chi_{\tilde{a}}}{b_1-a_1}, a_1 \le x \le b_1 \\ \frac{(c_1-x)\chi_{\tilde{a}}}{c_1-b_1}, b_1 \le x \le c_1 \\ 0, \textit{otherwise} \end{cases} \tag{3}$$

$$v_{\tilde{a}}(x) = \begin{cases} \frac{(b_1-x+\sigma_{\tilde{a}}(x-a_1))}{b_1-a_1}, a_1 \le x \le b_1 \\ \frac{x-b_1+\sigma_{\tilde{a}}(c_1-x)}{c_1-b_1}, b_1 \le x \le c_1 \\ 0, \textit{otherwise} \end{cases} \tag{4}$$

$$\lambda_{\tilde{a}}(x) = \begin{cases} \frac{(b_1-x+\omega_{\tilde{a}}(x-a_1))}{b_1-a_1}, a_1 \le x \le b_1 \\ \frac{x-b_1+\omega_{\tilde{a}}(c_1-x)}{c_1-b_1}, b_1 \le x \le c_1 \\ 0, \textit{otherwise} \end{cases} \tag{5}$$

Definition 4. [25] Let $\tilde{a} = (<(a_1, b_1, c_1); \chi_{\tilde{a}}, \sigma_a, \omega_{\tilde{a}}) >)$ and $\tilde{b} = (<(a_2, b_2, c_2); \chi_{\tilde{b}}, \sigma_b, \omega_{\tilde{b}}) >)$ be two SVTNN. Then,

$$\tilde{a} + \tilde{b} = < (a_1 + a_2, b_1 + b_2, c_1 + c_2); \chi_{\tilde{a}} \wedge \chi_{\tilde{a}}, \sigma_a \vee \sigma_b, \omega_{\tilde{a}} \vee \omega_{\tilde{b}}) > \tag{6}$$

$$\tilde{a}\tilde{b} = \begin{cases} < (a_1 a_2, b_1 b_2, c_1 c_2); \chi_{\tilde{a}} \wedge \chi_{\tilde{a}}, \sigma_a \vee \sigma_b, \omega_{\tilde{a}} \vee \omega_{\tilde{b}}) >, & (c_1 > 0, c_2 > 0) \\ < (a_1 c_2, b_1 b_2, c_1 a_2); \chi_{\tilde{a}} \wedge \chi_{\tilde{a}}, \sigma_a \vee \sigma_b, \omega_{\tilde{a}} \vee \omega_{\tilde{b}}) >, & (c_1 \langle 0, c_2 \rangle 0) \\ < (c_1 c_2, b_1 b_2, a_1 a_2); \chi_{\tilde{a}} \wedge \chi_{\tilde{a}}, \sigma_a \vee \sigma_b, \omega_{\tilde{a}} \vee \omega_{\tilde{b}}) >, & (c_1 < 0, c_2 < 0) \end{cases} \tag{7}$$

Definition 5. [26] Convert the SVTNN into crisp values by:

$$S = \frac{1}{8}(a_1 + b_1 + c_1)(2 + a_{\tilde{a}} - \sigma_{\tilde{a}} - \omega_{\tilde{a}}) \tag{8}$$

4 NPD Methodology with Using Sustainability Driven QFD and Neutrosophic Sets

Today, companies should focus not only on production in accordance with customer needs but also on protecting nature and providing economic and social sustainability in NPD processes. In this study, NPD methodology with using sustainability driven QFD and neutrosophic sets is proposed. The steps of the proposed methodology are as follows:

Step 1: Defining goals of the company and the product, determining CRs, defining linguistic importance degrees of each CRs and assigning SVTNN by using Table 2.

Step 2: Determining SRs, defining linguistic importance degrees of each SRs and assigning SVTNN by using Table 2.

Step 3: Determining ECs and defining linguistic correlations among ECs by using Table 2.

Step 4: Determining relationships between CRs-ECs, SRs-ECs using linguistic terms and finding SVTNN corresponding to them. Empty cell shows no correlation.

Step 5: Finding importance degrees of each ECs by multiplying the linguistic importance degrees of CRs & SRs and the linguistic relationship degrees (CRs-ECs, SRs-ECs). Converting them to crisp values.

Step 6: Finding relative importance degree of each EC by dividing the degree of importance degree of each EC to the cumulative degree of all importance.

Step 7: Prioritizing ECs and recommending actions to ECs with the high importance.

5 A Case Study of the Aircraft Stabilizer

The proposed NPD methodology with using sustainability driven QFD and neutrosophic sets is applied on the horizontal stabilizer of the F-16 aircraft. The horizontal stabilizer is one of the control surfaces that is a composite-based structural part. The use of composite materials in aircraft construction is increasing. It is very important to consider and examine how these materials are produced. With the production of the horizontal stabilizer, the locality and nationality rate of the Turkish Air Force will increase. Also, there is a potential for handling NPD of the horizontal stabilizer considering sustainability. A team of company experts conducts the evaluation and defines CRs, SRs and ECs. These are listed in Table 1. Linguistic terms and corresponding SVTNN scale for importance degrees of CRs & SRs and relationship degrees (CRs-ECs, SRs-ECs) are given in Table 2 [27].

Table 1. Customer needs, sustainability requirements and engineering characteristics [26].

CN1	Ease of installation and integrity
CN2	It must be performed flight control mission
CN3	No corrosion
CN4	Resistant to environmental conditions while performing the flight mission
CN5	Strength to flight loads
CN6	Easy to repair
CN7	No surface damage and edge delamination
CN8	Electrical property suitability
CN9	Resistant to shock loads and blast
CN10	Interchangeability
CN11	Accuracy in horizontal stabilizer width & height & thickness
SR1	Long life cycle
SR2	It can be provided fuel savings
SR3	Low manufacturing cost
SR4	Suitable for flight safety
EC1	Mechanical strength
EC2	Fatigue strength
EC3	Corrosion resistance
EC4	Weight
EC5	Dimensional stability
EC6	Thermal insulation
EC7	Electrical conductivity
EC8	Thermal expansivity
EC9	Stiffness
EC10	Moisture resistance
EC11	Aeroelasticity
EC12	Surface geometry

(*continued*)

Table 1. (*continued*)

EC13	Wear tolerance
EC14	Abrasion resistance
EC15	Drag resistance
EC16	Projection area

Table 2. Linguistic terms and corresponding SVTNN scale [27].

Low importance (LI) / Low relation (LR)	((4.6; 5.5; 8.6); 0.4; 0.7; 0.2)
Not low importance (NLI) / Not low relation (NLR)	((4.7; 6.9; 8.5); 0.7; 0.2; 0.6)
Very low importance (VLI) / Very low relation (VLR)	((6.2; 7.6; 8.2); 0.4; 0.1; 0.3)
Completely low imp. (CLI) / Completely low rel. (CLR)	((7.1; 7.7; 8.3); 0.5; 0.2; 0.4)
More or less low imp. (MLLI) / More or less low rel. (MLLR)	((5.8; 6.9; 8.5); 0.6; 0.2; 0.3)
Fairly low imp. (FLI) / Fairly low rel. (FLR)	((5.5; 6.2; 7.3); 0.8; 0.1; 0.2)
Essentially low imp. (ELI) / Essentially low rel. (ELR)	((5.3; 6.7; 9.9); 0.3; 0.5; 0.2)
High imp. (HI) / High rel. (HR)	((6.2; 8.9; 9.1); 0.6; 0.3; 0.5)
Not high imp. (NHI) / Not high rel. (NHR)	((4.4; 5.9; 7.2); 0.7; 0.2; 0.3)
Very high imp. (VHI) / Very high rel. (VHR)	((6.6; 8.8; 10); 0.6; 0.2; 0.2)
Completely high imp. (CHI) / Completely high rel. (CHR)	((6.3; 7.5; 8.9); 0.7; 0.4; 0.6)
More or less high imp. (MLHI) / More or less high rel. (MLHR)	((5.3; 7.3; 8.7); 0.7; 0.2; 0.8)
Fairly high imp. (FHI) / Fairly high rel. (FHR)	((6.5; 6.9; 8.5); 0.6; 0.8; 0.1)
Essentially high imp. (EHI) / Essentially high rel. (EHR)	((7.5; 7.9; 8.5); 0.8; 0.5; 0.4)

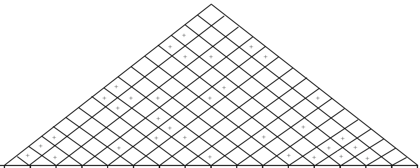
Linguistic importance degrees of CRs & SRs, linguistic relationship degrees (CRs-ECs, SRs-ECs), crisp absolute & relative importance values are demonstrated in HoQ in Fig. 1. These values are presented for EC6 as follows;

$$\begin{aligned}&[(<(4.4, 5.9, 7.2); 0.7, 0.2, 0.3)>)(<(6.3, 7.5, 8.9); 0.7, 0.4, 0.6)>)]\\&+[(<(6.6, 8.8, 10); 0.6, 0.2, 0.2)>)(<(6.2, 8.9, 9.1); 0.6, 0.3, 0.5)>)]\\&=[(<(68.6, 122.9, 155.1); 0.6, 0.4, 0.6)>)]\end{aligned}$$

Then the output is converted to crisp value by using Eq. (8).

$$S = \frac{1}{8}(68.6 + 122.6 + 155,1)(2 + 0.6 - 0,4 - 0.6) = 69.3$$

Relative importance degree of each EC is found by dividing the degree of importance degree of each EC to the cumulative degree of all importance. For EC6, relative absolute importance value is found as (69.3/3000) $\times$ 100 = 2.31. According to relative absolute importance values, mechanical strength with 9.51, abrasion resistance with 8.13 and wear tolerance with 8.11 are found as the most important technical requirements.



WHATS \ HOWS		Importance Evaluations	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	EC9	EC10	EC11	EC12	EC13	EC14	EC15	EC16
Customer needs	CN1	CLI					HR				HR		VLR	HR	HR			VLR
	CN2	EHI	HR	HR		HR				HR			EHR	CHR	HR		EHR	HR
	CN3	HI	HR		EHR							CHR				HR		
	CN4	NHI	CHR	CHR	EHR			CHR		CHR		EHR				HR		
	CN5	CHI	HR	HR		HR				HR			EHR	CHR	HR		EHR	HR
	CN6	NHI	HR	HR	HR				HR			HR			HR	CHR		
	CN7	NHI	HR	HR	HR				HR			HR			HR	CHR		
	CN8	HI							EHR									
	CN9	CHI	CHR	CHR	CHR				HR	HR		HR		HR	HR	HR	HR	
	CN10	VHI					HR		HR				HR	HR				HR
	CN11	CLI					EHR			CHR	CHR							HR
Sustainability requirements	SR1	HI	EHR	EHR	EHR					HR		HR			HR	HR		
	SR2	CLI				EHR	HR							EHR			EHR	
	SR3	VHI	HR		HR		HR	HR	HR			HR	CHR	CHR	CHR	HR	HR	HR
	SR4	MLHI	CHR	CHR	CHR				HR	HR		HR			HR	HR		
Absolute importance			284.9	218.2	222.7	99.0	184.0	69.3	206.2	188.6	68.3	226.3	154.5	231.6	243.1	243.7	167.1	188.5
Relative absolute importance (%)			9.51	7.28	7.43	3.30	6.14	2.31	6.88	6.30	2.28	7.55	5.15	7.73	8.11	8.13	5.58	6.29

Fig. 1. Linguistic HoQ and relative absolute importance values.

6 Conclusion

Companies need a new methodology that integrates sustainability requirements besides the customer expectations during new product development in order to meet the social, economic and environmental issues. Because the market has a deep sensitivity to environmental protection of the earth. Also, nowadays customers consider the social and economic dimensions of the product. For this reason, NPD methodology with using sustainability driven QFD and neutrosophic sets to meet customer and sustainability requirements is presented in this study. Linguistic terms can be integrated while constructing HoQ in the methodology. Single valued triangular neutrosophic number are employed for importance degrees of CRs & SRs and relationship degrees (CRs-ECs, SRs-ECs). The methodology has been successfully used in the design of a horizontal stabilizer. In the future studies, other extensions of ordinary fuzzy sets can be used while modeling uncertainty and vagueness in HoQ and a comparative analysis through these QFD analyses can be executed.

References

1. Dver, A.: Software Product Management Essentials, 2003rd edn. Anclote Pres, USA (2003)
2. Krishnan, V., Ulrich, K.T.: Product development decisions: a review of the literature. Manage. Sci. **47**(1), 1–21 (2001)
3. Mizuno, S., Akao, Y.: QFD: Customer-Driven Approach to Quality Planning and Deployment. 1st edn. Asian Productivity Organization, Tokyo, Japan (1994)
4. Sivasamy, K., Arumugam, C., Devadasan, S.R., Murugesh, R., Thilak, V.M.M.: Advanced models of quality function deployment: a literature review. Qual. Quant. **50**(3), 1399–1414 (2016)
5. Liu, H., Wang, C.: An advanced quality function deployment model using fuzzy analytic network process. Appl. Math. Model. **34**(11), 3333–3351 (2010)

6. Kwong, C.K., Bai, H.: A fuzzy AHP approach to the determination of importance weights of customer requirements in quality function deployment. J. Intell. Manuf. **13**(5), 367–377 (2002)
7. Zhang, F., Yang, M., Liu, W.: Using integrated quality function deployment and theory of innova-tion problem solving approach for ergonomic product design. Comput. Ind. Eng. **76**(1), 60–74 (2014)
8. Onar, S.Ç., Büyüközkan, G., Öztayşi, B., Kahraman, C.: A new hesitant fuzzy QFD approach: an application to computer workstation selection. Appl. Soft Comput. **46**, 1–16 (2016)
9. Yu, L., Wang, L., Bao, Y.: Technical attributes ratings in fuzzy QFD by integrating interval-valued intuitionistic fuzzy sets and Choquet integral. Soft. Comput. **22**(6), 2015–2024 (2016). https://doi.org/10.1007/s00500-016-2464-8
10. Van, L., Yu, V., Dat, L., Dung, C., Chou, S.Y., Loc, N.: New integrated quality function deployment approach based on interval neutrosophic set for green supplier evaluation and selection. Sustainability **10**(3), 838 (2018)
11. Haber, N., Fargnoli, M., Sakao, T.: Integrating QFD for product-service systems with the Kano model and fuzzy AHP. Total Qual. Manage. Bus. Excell. **31**(4), 1–26 (2018)
12. Haktanır, E., Kahraman, C.: A novel interval-valued pythagorean fuzzy QFD method and its application to solar photovoltaic technology development. Comput. Ind. Eng. **132**, 361–372 (2019)
13. Wang, D., Yu, H., Wu, J., Meng, Q., Lin, Q.: Integrating fuzzy based QFD and AHP for the design and implementation of a hand training device. J. Intell. Fuzzy Syst. **36**(4), 3317–3331 (2019)
14. Gündoğdu, F.K., Kahraman, C.: A novel spherical fuzzy QFD method and its application to the linear delta robot technology development. Eng. Appl. Artif. Intel. **87**, 103348 (2020)
15. Masui, K., Sakao, T., Kobayashi, M., Inaba, A.: Applying quality function deployment to environmentally conscious design. Int. J. Q. Reliab. Manag. **20**(1), 90–106 (2003)
16. Kuo, T.C., Wub, H.H., Shieh, J.I.: Integration of environmental considerations in quality function deployment by using fuzzy logic. Expert Syst. Appl. **36**, 7148–7156 (2009)
17. Vinodh, S., Chintha, K.S.: Application of fuzzy QFD for enabling sustainability. Int. J. Sustain. Eng. **4**(4), 313–322 (2011)
18. Fargnoli, M., Costantino, F., Gravio, G.V., Tronci, M.: Product service-systems implementation: a customized framework to enhance sustainability and customer satisfaction. J. Clean. Prod. **188**, 387–401 (2018)
19. Finger, G.S.W., Lima-Junior, F.R.: A hesitant fuzzy linguistic QFD approach for formulating sustainable supplier development programs. Int. J. Prod. Econ. **247**, 108428 (2022)
20. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
21. Atanassov, K.T.: Intuitionistic fuzzy sets. In: Intuitionistic Fuzzy Sets. Studies in Fuzziness and Soft Computing, pp. 1–137. Physica, Heidelberg (1999)
22. Yager, R.R.: Pythagorean fuzzy subsets. In: Proceedings of the 2013 Joint IFSA World Congress and NAFIPS Annual Meeting, IFSA/NAFIPS 2013, pp. 57–61. Edmonton, AB, Canada (2013)
23. Smarandache, F.: A Unifying Field in Logics Neutrosophy: Neutrosophic Probability, Set and Logic. American Research Press, Rehoboth (1998)
24. Wang, H., Smarandache, F., Zhang, Y., Sunderraman, R.: Single valued neutrosophic sets. Multispace Multistructure **4**, 410–413 (2010)
25. Deli, I., Şubaş, Y.: A ranking method of single valued neutrosophic numbers and its applications to multi-attribute decision making problems. Int. J. Mach. Learn. Cybern. **8**(4), 1309–1322 (2016). https://doi.org/10.1007/s13042-016-0505-3

26. Ayber, S., Erginel, N.: Combining customer and sustainability requirements in quality function deployment for a composite aircraft structure. In: International Conference on Industrial Engineering and Operations Management, pp. 1935–1943. IEOM Society, Istanbul (2022)
27. Abdel-Baseta, M., Changb, V., Gamala, A., Smarandache, F.: An integrated neutrosophic ANP and VIKOR method for achieving sustainable supplier selection: a case study in importing field. Comput. Ind. **106**, 94–110 (2019)

Design of 2 DOF Fuzzy Control System for Fuel Cell Systems

Fatih Kendir[1,2](✉) and Tufan Kumbasar[1]

[1] Control and Automation Engineering Department, Istanbul Technical University, İstanbul, Turkey
{kendir,kumbasart}@itu.edu.tr
[2] AVL Research and Engineering, İstanbul, Turkey
fatih.kendir@avl.com

Abstract. Fuel Cell Systems (FCSs) are a promising energy source for various applications due to their high efficiency and low environmental impact. However, the commercialization of FCSs has been hindered by their limited lifespan, which is largely affected by cell degradation. To address this issue, we propose a 2 DOF fuzzy control system that is capable to maximize the net stack output power while considering cell degradation. The proposed control system is composed of a data-driven Reference Generator (RG), a data-driven Feedforward Controller (FFC), and a feedback PI controller which are designed to regulate the Oxygen Excess Ratio (OER) to achieve maximum net power output. To handle the dominant nonlinearity, we learn and deploy Fuzzy Models (FMs) as the RG and FFC. In this context, we analyzed the stack current-OER relationship at optimal stack temperature and learned the nonlinear characteristic with FMs. Then, the PI controller is tuned to track the generated reference to maximize the net stack power for different stack currents. To analyze the impact of improvements on cell lifetime, we developed a degradation model based on electrochemical impedance spectroscopy. The results show that the fuzzy control system can maximize the net stack output power while simultaneously prolonging fuel cell life as a result of the high-performing OER control loop. We believe that the proposed 2 DOF fuzzy control system is a step towards the commercialization of fuel cells and enables their widespread adoption as a sustainable energy source for various applications.

Keywords: Fuel Cell Systems · Data-driven design · Fuzzy Modeling

1 Introduction

Fuel Cell Systems (FCSs) have several advantages over traditional energy sources, such as high efficiency and low environmental impact. However, cell degradation has restricted their commercialization, resulting in a limited lifespan. In FCSs, the chemical energy of oxygen and hydrogen reactants is converted directly into electricity as shown in Fig. 1. The oxygen excess ratio (OER) control

This work was supported by the BAGEP Award of the Science Academy.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 584–591, 2023.
https://doi.org/10.1007/978-3-031-39777-6_68

loop that adjusts the air mass flow is the most important among other control loops such as cooling and humidification in the air supply subsystem since it has a heavy effect on the net power output due to slow dynamics compared to the hydrogen supply subsystem. To shape the OER control loop, a FeedForward Controller (FFC) and PI FeedBack Controller (FBC) are designed in [1]. Also, to maximize net power output, an extremum-seeking controller was proposed to regulate OER around an optimal value in [2]. An adaptive control system with asymmetric OER constraints to avoid oxygen starvation is also designed [3]. A sliding mode controller was also tuned to handle nonlinearities and uncertainties [4]. A fuzzy controller to improve net power output was studied in [5]. Another study has focused on designing a fuzzy controller combined with a feed-forward decoupled controller for an air supply system to track load demand [6]. To increase oxygen utilization without recirculation, a fuzzy controller based on stack current and voltage variation was proposed in [7]. In all these studies, either static or lookup-based OER control systems are designed without considering cell degradation.

In this paper, we present the design of a 2DOF fuzzy control system for OER tracking of FCSs. As illustrated in Fig. 1, the main components of the proposed fuzzy control systems are the FBC, FFC, and Reference Generator (RG). In this paper, we present a systematic data-driven approach to represent the FFC and RG via Fuzzy Models (FMs). In this context, the FCS model was simulated for pre-defined operating points to obtain the necessary training data for FFC and RG. To examine the control performance, we presented comparative simulation results on a realistic FCS model. To show the superiority of FMs, we conducted comparative simulations where FM and Polynomial Regression (PR) model-based RGs and FFCs were employed. Moreover, Electrochemical Impedance Spectrum (EIS) analysis was performed to obtain the impedance of FCS which is used as an indicator of cell lifetime. The results showed that the proposed 2DOF fuzzy control system improves net power output while reducing cell impedance which could lead to a longer cell lifetime.

The paper is organized as follows: Sect. 2 presents the proposed 2DOF fuzzy control structure and its data-driven design. Section 3 presents comparative results. Finally, the driven conclusions and future work are given in Sect. 4.

2 Proposed 2DOF Fuzzy Control Structure

The proposed 2DOF fuzzy control system is given in Fig. 1. Here, the acquisition of maximum net power output (P_{net}) is converted into a reference tracking problem where desired OER ($\lambda_{O_2,des}$) depends on stack current (I_{st}). In contrast to lookup-based solutions in literature, we propose a data-driven FM-based RG for $\lambda_{O_2,des}$. Then, an FBC regulates λ_{O_2} to track the commanded reference and calculates the closed-loop control input, compressor voltage ($V_{cm,FB}$). We also propose a data-driven FM-based FFC involved in the control system which calculates open-loop control input ($V_{cm,FF}$) to improve the transient performance of λ_{O_2} tracking.

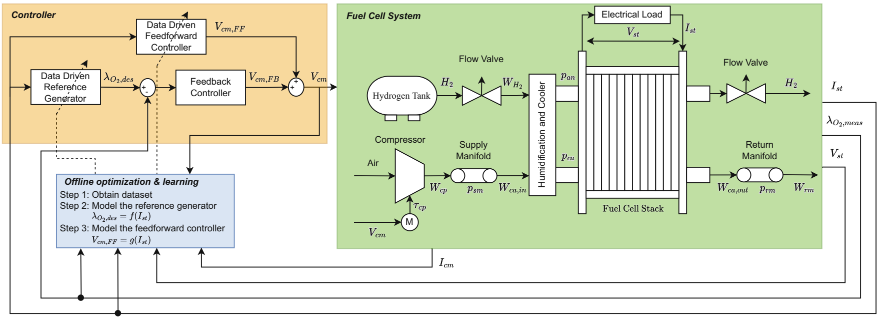


Fig. 1. Fuel cell system schematic with proposed 2DOF fuzzy control structure

2.1 Problem Description

In FCS, the dynamics of P_{net} is dominated by airpath as the hydrogen path dynamics is relatively faster due to the pressurized H_2 in the tank. Thus, improvements in OER tracking control could bring better P_{net} output of FCS. P_{net} is defined as the difference between the produced power by stack and consumed power by auxiliary devices such as air compressor, coolant pump, etc. The power consumed by auxiliary devices is dominated by the air compressor, thus the rest can be neglected. As a result, P_{net} is defined as follows [1]:

$$P_{net} = V_{st} I_{st} - V_{cm} I_{cm} \tag{1}$$

The amount of oxygen supplied to the FCS has to be more than the required amount of oxygen for chemical reactions to prevent oxygen starvation during load changes. The OER (λ_{O_2}) is defined as the ratio between mass flow rates of supplied oxygen and consumed oxygen to the FCS. The OER has to be controlled to maximize P_{net} and to prevent oxygen starvation and thus cell degradation.

Let us first analyze the steady state (SS) behavior of P_{net} at the optimal stack temperature $T_{st} = 80'\text{C}$ for various I_{st} and λ_{O_2} values. The results presented in Fig. 2(a) show that P_{net} is maximized for specific λ_{O_2} values for each I_{st}. At high OER, P_{net} slightly deteriorates; on the other hand, starvation occurs in the cells that yield to degradation at low OER. Therefore, OER has to be accurately controlled to obtain maximum P_{net} from FCS and decelerate degradation.

2.2 Data Driven Design of FM-Based RG and FFC

To design RG, we need to model the nonlinear relationship between λ_{O_2} and I_{st} presented in Fig. 2(b). In this context, we define:

$$\lambda_{O_2,des} = f(I_{st}) \tag{2}$$

In a similar manner, to design FFC, we define the following function between V_{cm} and I_{st} to model the nonlinearity in Fig. 2(c):

$$V_{cm,FF} = g(I_{st}) \tag{3}$$

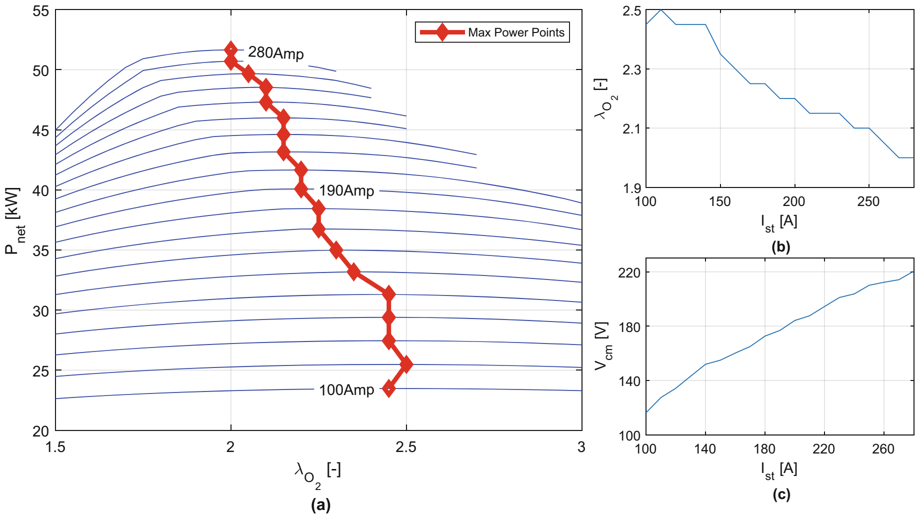


Fig. 2. FCS characteristics at stack temperature $T_{st} = 80\,^\circ$C: (a) The behaviour of P_{net} for various λ_{O_2} and I_{st}, (b) optimal λ_{O_2} that maximizes P_{net} vs. I_{st}, (c) V_{cm} at optimal λ_{O_2} vs. I_{st}.

To model $f(I_{st})$ and $g(I_{st})$, the training data was acquired by simulating the FCS model [1] for SS analysis. For this purpose, the FCS is simulated for different I_{st} and λ_{O_2} inputs. By considering the FCS operational limitations, I_{st} is gridded by 10 in range of [100–280]A and λ_{O_2} is gridded by 0.05 in range of [1.5–3.0] and simulations were performed. Then, the λ_{O_2} values that maximize SS value P_{net} for each I_{st} and corresponding V_{cm} values were obtained. By this way, training data sets with $N = 19$ samples $[I_{st}, \lambda_{O_2}]$ and $[I_{st}, V_{cm}]$ were collected. The collected datasets are standardized prior to training. In the training of the models, we used the RMSE as the loss function which is defined as:

$$\mathcal{L} = \sqrt{\sum_{i=1}^{N} \frac{(\hat{y}_i - y_i)^2}{N}} \tag{4}$$

Here, y_i is the actual output, and $\hat{y}_i$ is the predicted output.

Polynomial Regression: To obtain a baseline performance for comparison, we also learned PR models. For RG, the PR model is defined as follows:

$$f(I_{st}) = p_1 I_{st}^n + p_2 I_{st}^{n-1} + ... + p_n I_{st} + p_{n+1} \tag{5}$$

where n is the order of the polynomial that needs to be set. Similarly, a PR model is also constructed for $g(I_{st})$. The PR models are trained via the recursive least square approach. The resulting RMSE values of each training with different orders of PR models are evaluated and provided in Table 1. As expected, the fitting performance is improved when the order of the polynomial is increased.

Fuzzy Modelling: We modeled the RG and the FFC via Takagi-Sugeno FMs. The rule structure of FMs with K rules is defined as follows [8]:

$$R_k: \text{ If } I_{st} \text{ is } A_k \text{ then } y \text{ is } Y_k, k = 1, 2, ..., K \tag{6}$$

where Y_k is the consequent Membership Function (MF) that is defined as follows:

$$Y_k = a_k I_{st} + b_k \tag{7}$$

and A_k is the antecedent MF defined with a Gaussian fuzzy set. The membership grade is defined as:

$$\mu_{A_k}(I_{st}) = \exp\left(\frac{-(I_{st} - c_k)^2}{2\sigma_k^2}\right) \tag{8}$$

We employed the ANFIS learning approach to learn the characteristics of the RG given in (2) and the FFC given in (3) via $[I_{st}, \lambda_{O_2}]$ and $[I_{st}, V_{cm}]$, respectively. To obtain the FMs for the RG and FFC, we used two fuzzy modeling techniques which are learning an FM via MFs and rules derived from data clusters found using Subtractive Clustering (SCM) and Fuzzy C-Means (FCM). For the FCM approach, the number of clusters is set as 5 with a fuzziness parameter value of 1.8 for both RG and FFC. For the SCM, the number of clusters is computed according to a defined cluster influence range value of 0.25. The resulting number of clusters for RG is 5 while 6 for FFC. Note that the number of clusters is equal to the number of fuzzy rules K.

The training performances of FMs are tabulated in Table 1. The performance results of FMs are the same for the RG. Yet, the FM-SCM is performing better in the training of FFC. Overall, it appears that the FMs have better results over PR models according to the calculated $\mathcal{L}$. However, it is necessary to reflect these improvements in the fuel cell net power output and cell lifetime.

Table 1. RMSE performance comparison

Design Method	RG	FFC
PR (n=6)	0.0165	1.003
PR (n=7)	0.0164	0.9963
FM-SCM	0.0094	0.4503
FM-FCM	0.0094	0.6277

3 Comparative Results and Analysis

Here, we evaluate the performance of the proposed 2DOF fuzzy controller in comparison to its PR model counterpart. For both 2DOF controllers, a PI controller is employed as FBC with the coefficients $K_p = 36$ and $K_i = 216$. Moreover, to show the contribution of the 2DOF controller, we designed a 1DOF controller as well. In the 1DOF control structure, the same learned PR model and FM are employed for the RG alongside a PI FBC with $K_p = 50$ and $K_i = 600$.

To illustrate how the controllers perform in different operating points, we conducted two sets of simulations ($J = 2$) with respect to I_{st}. In the first set, I_{st} is applied as a step input from 135 to 265 [A] while in the second set as a step input from 145 to 195 [A]. To analyze the control performances, the Integral Absolute Error (IAE) and settling time T_s metrics were used. For P_{net}, the sum of normalized IAE was evaluated which is defined as:

$$e_{P_{net}} = \sum_{j=1}^{J} \frac{\sum_{s=1}^{S} |P(s)^j_{net,des} - P(s)^j_{net,sim}|}{max(P(s)^j_{net,des}) - min(P(s)^j_{net,des})} \tag{9}$$

while for λ_{O_2} the sum of IAE was used as follows:

$$e_\lambda = \sum_{j=1}^{J}\sum_{s=1}^{S} |\lambda(s)^j_{O_2,des} - \lambda(s)^j_{O_2,sim}| \tag{10}$$

In Table 2, the resulting performance measures are tabulated. It can be observed that the 2DOF structures are superior to their 1DOF counterparts in terms of $e_{P_{net}}$, $T_{s,P_{net}}$ and e_λ. Therefore, the results of 2DOF control structure are depicted in Fig. 3(a) and (b) for P_{net} and λ_{O_2} respectively. Here, to make the contribution of the 2DOF structure illustrative, a simulation result of the 1DOF control system is also given. The results of 2DOF controllers are very close to each other in terms of P_{net}, yet; the transient response of λ_{O_2} is better.

Table 2. Comparative control performance measures.

Structure	Model	$e_{P_{net}}\,[W]$	$T_{s,P_{net}}\,[s]$	$e_\delta\,[-]$	$T_{s,\delta}\,[s]$
1 DOF	PR (n=6)	306	1.257	0.157	0.79
	PR (n=7)	306	1.257	0.157	0.79
	FM-FCM	294	1.25	0.153	0.783
	FM-SCM	291	1.249	0.153	0.783
2 DOF	PR (n=6)	292	0.894	0.129	1.013
	PR (n=7)	292	0.892	0.129	1.015
	FM-FCM	281	0.888	0.123	0.972
	FM-SCM	279	0.885	0.123	0.969

To investigate the impact of different transient responses, impedance diagrams based on EIS were obtained by using the degradation model for min-max values of λ_{O_2}. The difference in the transient response of λ_{O_2} control brings differences in impedance diagrams as shown in Fig. 4(a) at the current density 0.94643 A.cm^{-2}. In [9,10], it is shown that the degradation is decelerated when the arc size in the impedance diagram is minimized. According to the results in

Fig. 4(a), the arc size of impedance obtained from 2DOF fuzzy controller simulations is minimized which decelerates the degradation and results in a prolonged lifetime of FCS. Moreover, the voltage value corresponding to the same current density is higher when the FMs are utilized in 2DOF design as presented by polarization curves in Fig. 4(b). The results indicate that the proposed fuzzy control structure provides better power output and decelerates degradation.

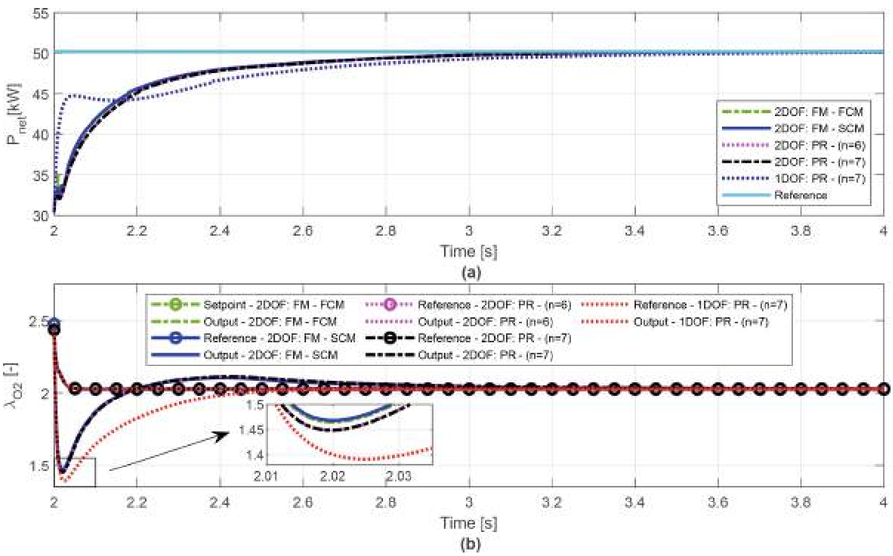


Fig. 3. Simulation results for the first operating point: (a) P_{net} (b) OER

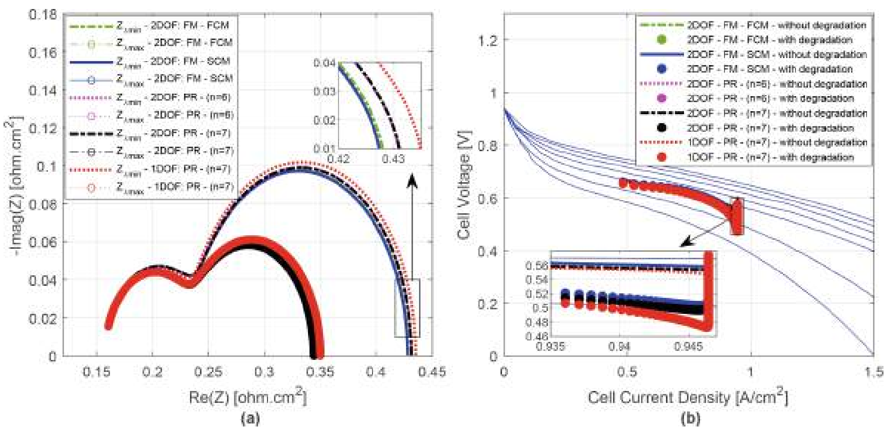


Fig. 4. Degradation analysis for the first operating point: (a) EIS, (b) Polarization curve

4 Conclusion and Future Work

We proposed a 2DOF fuzzy control structure for OER tracking of FCSs to improve net power output while reducing cell degradation. The results clearly showed that the 2DOF fuzzy control structure brings improvements over its PR-based 1DOF and 2DOF counterparts in terms of net power output and degradation. In brief, the 2DOF fuzzy control system is a promising development in the commercialization of fuel cells while reducing cell degradation which is a barrier to making FCS a feasible and sustainable energy source.

In future work, we plan to investigate FMs by exploring various MFs in the 2DOF control structure. We also aim to extend our study by involving neural network-based approaches.

References

1. Pukrushpan, J., Peng, H., Stefanopoulou, A.: Control-oriented modeling and analysis for automotive fuel cell systems. ASME J. Dyn. Syst. Measur. Control **126**, 14–25 (2004)
2. Chang, Y.A., Moura, S.J.: Air flow control in fuel cell systems: an extremum seeking approach. In: American Control Conference, pp. 1052–1059 (2009)
3. Kim, B.M., Choi, Y.H., Yoo, S.J.: Adaptive control of proton exchange membrane fuel cell air supply systems with asymmetric oxygen excess ratio constraints. IEEE Access **8**, 5537–5549 (2020)
4. Garcia-Gabin, W., Dorado, F., Bordons, C.: Real-time implementation of a sliding mode controller for air supply on a PEM fuel cell. J. Process Control **20**(3), 325–336 (2010)
5. Williams, J.G., Liu, G., Chai, S., Rees, D.: Intelligent control for improvements in pem fuel cell flow performance. Int. J. Autom. Comput. **5**(2), 145–151 (2008)
6. Liu, Z., Chen, H., Peng, L., Ye, X., Xu, S., Zhang, T.: Feedforward-decoupled closed-loop fuzzy proportion-integral-derivative control of air supply system of proton exchange membrane fuel cell. Energy **240**, 122490 (2022)
7. Polak, A.: Simulation of fuzzy control of oxygen flow in PEM fuel cells. Energies **13**(9), 2372 (2020). https://doi.org/10.3390/en13092372
8. Beke, A., Kumbasar, T.: More than accuracy: a composite learning framework for interval type-2 fuzzy logic systems. IEEE Trans. Fuzzy Syst. **31**(3), 734–744 (2023)
9. Chandesris, M., Robin, C., Gerard, M., Bultel, Y.: Investigation of the difference between the low frequency limit of the impedance spectrum and the slope of the polarization curve. Electrochimica Acta **180**, 581–590 (2015)
10. Liu, M., Wang, C., Xie, F., Mao, Z.: A polymer electrolyte fuel cell life test using accelerating degradation technique. Int. J. Hydrogen Energy **38**(25), 11011–11016 (2013)

An Application of Cooperative Game Theory in Oil Refining Sites: Case Study of Dora Refinery in Iraq

Eman H. Ouda, Shaymaa H. Khazaal, and Jabbar Abbas(✉)

Department of Applied Sciences, University of Technology, Baghdad, Iraq
jabbar.a.ghafil@uotechnology.edu.iq

Abstract. Oil refineries are among the strategic projects that require long-term planning and which depend on them to meet the local and international needs of oil derivatives due to economic growth and thus have become one of the main economic pillars of the states. Cooperative games have been used in all cases in which producer actions maximize utility, and have shown to be a powerful gadget in several areas. In this contribution, we use an optimization problem to model cooperation among petroleum producers in oil refineries. Then, the methodology is applied to the processing of crude oil in the Dora refinery in Iraq. The results of this study will robustly support specialists to plan well-structured strategies for the oil management system of the future.

Keywords: Optimization Problem · Game Theory · Shapley Value · Oil Refining Sites

1 Introduction

The various industrial developments based on different types of petroleum products give great importance to the oil sector and for this reason the demand for oil of all kinds' increases. Therefore, there is increasing interest in using scientific methods and interest in productive capabilities in all vital facilities and in analyzing problems in order to make decisions to formulate well-structured strategies in the future. Cooperative game theory was applied in several areas and can be used successfully tool in evaluating participating players (e.g., see [1–4]). There are various forms of power and interaction indices that can be utilized as a solution concept and to assess the level of collaboration among parties. One of them is the Shapley power (or Shapley value) index proposed by Shapley in [5] for game theory. There are many application of the Shapley power such as: to Sharing congestion management costs among system operators has been given by Simon Voswinkel et al. [6], Eryganov I et al. [7] to calculate possible development of waste producers' costs in a case of cooperation in waste management, and Van Campen et al. [8] applied the Shapley power to the terrorist attack in WTC 9/11. In the scope of aggregation by the Choquet integral, the notion of interactions happening among a group of players which are a generalizations of the notion of Shapley indices, have been applied to several problems of decision-making [9–16].

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 592–599, 2023.
https://doi.org/10.1007/978-3-031-39777-6_69

The purpose of this contribution is to use an optimization problem to model cooperation among petroleum producers in oil refineries. First, we used the Shapley power index to caption the potential cost to oil producers in the situation of cooperation. Then, we apply the methodology to the processing of crude oil in the Dora[1] refinery in Iraq.

The structure of this paper is as follows. In Sect. 2, we give the preliminaries of the cooperative game theory. The optimization problem for oil producers is presented in Sect. 3. In Sect. 4, we introduce a study case for the processing of crude oil in the Dora refinery. Finally, some conclusions presented.

2 Preliminaries

Let $N = \{1, ..., n\}$ be a universal set of players and the power set of N denoted by $G(N)$. For all $\emptyset \neq S \subseteq N$is called a coalition. For each coalition $S \subseteq N$, a function $v(S)$ grants the profits than any group of players can get, which is called the characteristic function of the game. Note that the characteristic function of the set the empty $\emptyset$ set is $v(\emptyset) = 0$. Let N be a set of players, cooperative game is generally specified by a pair (N, v), where v is a coalition value function.

For any two players i and $j \in N$, $v(i)$ is one of representations of importance of $i \in N$. Hence, $v(ij) > v(i) + v(j)$(respectively, $<$) appears to model a positive (respectively, negative) interaction (respectively, substitutive) effect between players i and j. However, as discussed in [17], for a player i and a coalition $S \not\ni$i,

$$\alpha_{\,i} v(S) := v(S \cup i) - v(S) \tag{1}$$

shows to explain an value of importance of i in $S \cup i$. The Eq. (1) is called the *peripheral contribution* of a player i to a set S.

In cooperative game problems, the definition of Shapley value (which is usually called the power index) can be utilized as a solution concept of the game as follows.

Definition 1: the Shapley value $\boldsymbol{\varphi_i(N, v)}$, for the player $\boldsymbol{i} \in \boldsymbol{N}$, is a value on cooperative game $(\boldsymbol{N}, \boldsymbol{v})$ defined by.

$$\beta_i(N, v) = \sum_{S \subseteq N \setminus i} \frac{|S|!(|N| - 1 - |S|)!}{|N|!} [v(S \cup \{i\}) - v(S)]. \tag{2}$$

Consider three-person players game with $\boldsymbol{N} = \{1, 2, 3\}$, to calculate the Shapley value for this example of these players we need characteristic value as in the following $\boldsymbol{v}(\emptyset) = 0, \boldsymbol{v}(\{1\}) = 8, \boldsymbol{v}(\{2\}) = 7, \boldsymbol{v}(\{3\}) = 8, \boldsymbol{v}(\{1, 2\}) = 14, \boldsymbol{v}(\{1, 3\}) = 15, \boldsymbol{v}(\{2, 3\}) = 13, \boldsymbol{v}(\{1, 2, 3\}) = 20$

Now, for example calculate the shapely value for the first player, $\beta_1(N, v)$ by using Eq. (2): $\beta_1(N, v) = \frac{0!2!}{3!}(v(\{1\}) - v(\emptyset)) + \frac{1!1!}{3!}(v(\{1, 2\}) - v(\{2\})) + \frac{1!1!}{3!}(v(\{1, 3\}) - v(\{3\})) + \frac{2!0!}{3!}(v(\{1, 2, 3\}) - v(\{2, 3\})),$

$$\beta_1(N, v) = \frac{2}{6}(8 - 0) + \frac{1}{6}(14 - 7) + \frac{1}{6}(15 - 8) + \frac{2}{6}(20 - 13),$$

[1] The Dora Refinery is located in the south-eastern suburb of Baghdad, near the Tigris River.

$$\beta_1(N, v) = \frac{44}{6} \cong 7.33.$$

In the same technique we can find the Shapley value for the other players (2 and 3). Therefore, the Shapley vector $\beta_i(N, v) = (\beta_1, \beta_2, \beta_3) = (7.33, 5.83, 6.83)$.

3 The Optimization Problem for Modeling the Conflict Among Petroleum Producers in Oil Refineries

This section describes an optimization problem for modeling the general conflict among oil producers' in oil refineries by using cooperative game theory. In the oil management, the common dispute of petroleum producers can be formalized as a game with the characteristic function $v(S)$ given as follows. Let $M = \{1, \ldots, m\}$ be a set of oil refineries, $R_1^c, \ldots, R_m^c$ denote their capacities, and $c_1^p, \ldots, c_m^p$ denote their pipelines transfer fees. The group of petroleum fields is $N = \{1, \ldots, n\}$, their production of petroleum are R_1^p, $\ldots, R_n^p$. The transportation costs are given by matrix $\left[c_{i,j}^t\right]$, where $c_{i,j}^t$ represents the cost of transportation petroleum from the field $i \in N$ to the refinery $j \in M$. The quantity of petroleum arriving from oil field (in barrels) $i \in N$ to the refinery $j \in M$ is denoted by $x_{i,j}$. The function $v(S)$, $S \subseteq N$ is defined by Eqs. (3)–(10) as an optimization problem.

$$v(S) = \min_{x_{i,j}: i \in S, j \in M} \sum_{j \in M} \sum_{i \in S} \left(c_{i,j}^t + c_j^p\right) x_{i,j} \tag{3}$$

Subject to

$$\sum_{i \in S} x_{i,j} \le R_j^c - \sum_{i \in N \setminus S} x'_{i,j} \forall j \in M \tag{4}$$

$$\sum_{j \in M} x_{i,j} = R_i^p, \forall i \in S \tag{5}$$

$$x_{i,j} \ge 0, \forall i \in S, \forall j \in M \tag{6}$$

where

$$\begin{aligned} &\left\{x'_{i,j} : i \in N \setminus S, \forall j \in M\right\} \\ &= \min_{x_{i,j}: i \in N \setminus S, j \in M} \sum_{j \in M} \sum_{i \in N \setminus S} \left(c_{i,j}^t + c_j^p\right) x_{i,j} \end{aligned} \tag{7}$$

Subject to

$$\sum_{i \in N \setminus S} x_{i,j} \le R_j^c, \forall j \in M \tag{8}$$

$$\sum_{j \in M} x_{i,j} = R_i^p, \forall i \in N \setminus S \tag{9}$$

$$x_{i,j} \geq 0, \quad \forall i \in N \backslash S, \ \forall \ j \in M \tag{10}$$

The focal assumption of the entire model is that all petroleum generated can be treated by refineries. As a result, the resulting costs are compared by the Shapley values for fair allocation for the cooperative game problem for the oil producers using an algorithm and computer program (MatLab program) as a working procedure. The following algorithm is the path to calculate the values of $v(S)$, and the Shapley value $\beta_i(N, v)$.

ALGORITHM:

Step1: Let $\boldsymbol{N} = \{\mathbf{1}, \dots, \boldsymbol{n}\}$, $\boldsymbol{M} = \{\mathbf{1}, \dots, \boldsymbol{m}\}$,

Input: Pipelines transfer fees $\left(\boldsymbol{c}_{\mathbf{1}}^{\boldsymbol{P}}, \dots, \boldsymbol{c}_{\boldsymbol{m}}^{\boldsymbol{P}}\right)$, and the costs of transportation petroleum$(\boldsymbol{c}_{\mathbf{1},\mathbf{1}}^{\boldsymbol{t}}, \dots, \boldsymbol{c}_{\boldsymbol{n},\boldsymbol{m}}^{\boldsymbol{t}})$,

Step2:
$for \quad i = 1$ to number of players n , $i \in S$,
$for\ all \ j = 1 \ to \ m$ do
Input: $x_{i,j}$;
Step3: if
$$x_{i,j}^{'} = \underset{x_{i,j}:\, i \in N \backslash S, j \in M}{\text{Min}} \sum_{j \in M} \sum_{i \in N \backslash S} \left(c_{i,j}^{t} + c_{j}^{p}\right) x_{i,j} \ ;$$
end if
$end \ i$
$end \ j$
Step4: $for \quad i \in S$,
Find $v(S)$ using Eq. (2);
end for
Step5: Return to Step2
Step6: $for \ i \in \boldsymbol{s}$,
Calculate the Shapley value $\varphi_i(N, v)$using
$$\varphi_i(N, v) = \sum_{S \subseteq N \backslash i} \frac{|S|! \, (|N| - 1 - |S|)!}{|N|!} \ [v(S \cup \{i\}) - v(S)]$$
end for
Step7: Return to Step 6**.**

4 Case Study

The Arab region is classified as one of the most oil reserves in the world, and Iraq is the second Arab country that produces and exports oil, so the government pays great attention to this sector. Crude oil undergoes a refining process to obtain various types of petroleum products. The oil is transported from the producing fields to the refinery either by road transport or through oil pipelines. Figure 1 shows the petroleum fields (producers) and oil refineries in Iraq.

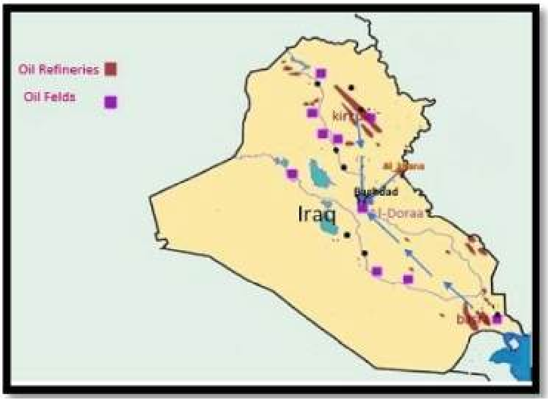


Fig. 1. The petroleum fields (producers) and oil refineries in Iraq

In this section, we deal with a case study of the Dora refinery in Iraq, which has taken the data for the data producers of crude oil (in $m^3 = 6.2898$ barrels) of the Dora refinery (for the year 2020) as shown in the Table 1.

Table 1. The data of producer's petroleum in the Dora refinery- Iraq

R_i^P	c_{ij}^t	c_j^p	R_j^c
$R_1^P = 5milionm^3$	$c_{11}^t = 4.02milion$	$c_1^p = 0.5milion$	$R_1^c = 7.5milionm^3$
$R_2^P = 0.05milionm^3$	$c_{21}^t = 0.04milion$	$c_1^p = 0.5milion$	$R_1^c = 7.5milionm^3$
$R_3^P = 1.3milionm^3$	$c_{31}^t = 0.82milion$	$c_1^p = 0.5milion$	$R_1^c = 7.5milionm^3$

To study this case, we apply the optimization problem described in Sect. 3 for modeling the conflict among petroleum producers in the oil refinery management process in Baghdad, and we calculate Shapley values for this study with the aim of displaying the potential cost to petroleum producers in case of cooperation. The Dora refinery is the only refinery in Baghdad, where crude oil undergoes a refining process to obtain various types of petroleum products. Dora refinery is funded by three oil fields: Al- Basra field, Kirkuk field, and Khana oil field. Therefore, the set of producers is $N = \{1, 2, 3\}$, and there is only one oil refinery 1 (i.e. $M = \{1\}$). The data were collected from the Dora refinery, with respect to the capacity of the Dora refinery, pipelines transfer fees, the fields producing crude oil and their quantities of supply for the Dora refinery, and transportation costs, as shown in Table 1 and Table 2, as illustrated in Fig. 2.

Table 2. The fields producing crude oil and their quantities of supply for Dora refinery

The field	field 1(*Basra*)	field 2 (*Khana*)	field 3(*Kirkuk*)
Quantity in barrels	5 Mil	0.05 Mil	1.3 Mil

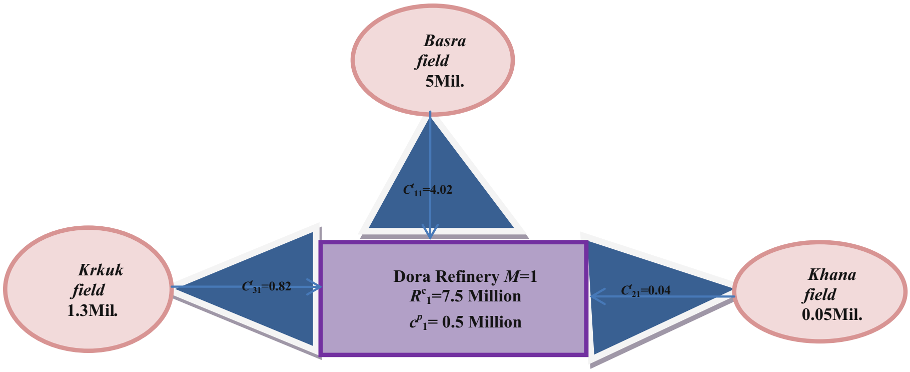


Fig. 2. The overview of the data of Dora refinery and petroleum fields

The characteristic functions for all coalitions of suppliers in the oil fields $v(S)$, $S \subseteq N = \{1, 2, 3\}$ are computed by the optimization problem as mentioned above through the algorithm presented in Sect. 3, as shown in Table 3.

Table 3. The characteristic value for the fields of Al-Dora refinery in Iraq

set of the fields	$\{1\}$	$\{2\}$	$\{3\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$	$\{1, 2, 3\}$
(N,δ)	22.6	0.027	1.716	22.627	23.864	1.743	24.34

Using Eq. (2), for each producer the Shapley value is presented in Table 4. For example, to calculate the Shapley value of the petroleum producer 1 (i. e. $\beta_1(N, v)$):

$$\beta_1(N, v) = \tfrac{0!2!}{3!}(v(\{1\}) - v(\emptyset)) + \tfrac{1!1!}{3!}(v(\{1, 2\}) - v(\{2\})) \\ + \tfrac{1!1!}{3!}(v(\{1, 3\}) - v(\{3\})) + \tfrac{2!0!}{3!}(v(\{1, 2, 3\}) - v(\{2, 3\}))$$

$$\beta_1(N, v) = \frac{2}{6}(22.6 - 0) + \frac{1}{6}(22.627 - 0.027) + \frac{1}{6}(23.864 - 1.716) + \frac{2}{6}(24.34 - 1.743)$$

$$\beta_1(N, v) \cong 22.523.$$

5 Results and Discussion

The Decision Maker (DM) of a set of alternatives (petroleum producers), adopts the preference relation that we denote it hereby $\prec$. For any two petroleum producers $i_1, i_2 \in N$, $i_1 \prec i_2$ means that the DM prefers alternative i_2 toi_1.

Table 4. The Shapley values of Dora refinery by fields

Game/petroleum production	f_1(Basra)	f_2(Khana)	f_3(Kirkuk)
$\beta_i(N,\delta)$	22.523	0. 176	1.63

Table 4 shows each of the oil suppliers to the Dora refinery using the Shapley value of each oil producer (22.523, 1.63, 0.176). According to the preference relation, the three producing fields will be ranked according to their contribution to the refinery is $2 \prec 3 \prec 1$. Evidently, the highest value is in the oil producer1 and its value is (22.523). We deduce that the first oil producer has the greatest potential to influence the outcome of the game and its contribution to the refinery will be greater than the rest oil producers. On the other hand, the oil producer 2 and its value (0.176) have the least value, so it will not have the ability to influence the outcome of the game.

Thus, this study shows that all producers finance the refinery with certain percentages of oil which mean that each large financier of the refinery's crude oil there will contribute to it the costs of industrial treatment and waste disposal and conversion into useful materials. Furthermore, it will be reducing the cost of transporting oil in cooperation with oil suppliers. As a result, this study will help those interested in formulating new strategies in the future in this field.

References

1. Skovsgaard, L., Jensen, I.G.: Recent trends in biogas value chains explained using cooperative game theory. Energy Econ. **74**, 503–522 (2018)
2. Hadi, H., Abbas, J.: An application of the Banzhaf values for cooperating among producers of waste processing in the Al-Mahmudiya factory. Baghdad Sci. J. (2023). https://doi.org/10.21123/bsj.2023.6788
3. Da Silva Rocha, A.B., Salomão, G.M.: Environmental policy regulation and corporate compliance in evolutionary game models with well-mixed and structured populations. Eur. J. Oper. Res. **279**(2), 486–501 (2019)
4. Abbas, J.: The Banzhaf interaction index for bi-cooperative games. Int. J. General Syst. **50**(5), 486–500 (2021). https://doi.org/10.1080/03081079.2021.1924166
5. Shapley, L.S.: Contributions to the Theory of Games, Chapter a Value for n-person Games (1953)
6. Voswinkel, S., Höckner, J., Khalid, A., Weber, C.: Sharing congestion management costs among system operators using the Shapley value. Appl. Energy **317**, 119039 (2022)
7. Ivan, E., Radovan, S., et al.: Application of cooperative game theory in waste management. Chem. Eng. Trans. **81**, 877–882 (2020)
8. Van Campen, T., Hamers, H., Husslage, B., Lindelauf, R.: A new approximation method for the Shapley value applied to the WTC 9/11 terrorist attack. Soc. Netw. Anal. Min. **8**(1), 1–12 (2017). https://doi.org/10.1007/s13278-017-0480-z
9. Abbas, J.: The balancing bipolar Choquet integrals. Int. J. Innovative Comput. Inf. Control **17**(3), 949–957 (2021). https://doi.org/10.24507/ijicic.17.03.949
10. Abbas, J.: The bipolar Choquet integrals based on ternary-element sets. J. Artif. Intell. Soft Comput. Res. **6**(1), 13–21 (2016). https://doi.org/10.1515/jaiscr-2016-0002

11. Hesham, M., Abbas, J.: Multi-criteria decision making on the optimal drug for rheumatoid arthritis disease. Iraqi J. Sci. **62**(5), 1659–1665 (2021). https://doi.org/10.24996/ijs.2021.62.5.28
12. Abbas, J.: The 2-additive choquet integral of bi-capacities. In: Rutkowski, L., Scherer, R., Korytkowski, M., Pedrycz, W., Tadeusiewicz, R., Zurada, J.M. (eds.) ICAISC 2019. LNCS (LNAI), vol. 11508, pp. 287–295. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20912-4_27
13. Naeem, A., Abbas, J.: A computational model for multi-criteria decision making in traffic jam problem. J. Autom. Mob. Rob. Intell. Syst. **15**(2), 39–43 (2022). https://doi.org/10.14313/JAMRIS/2-2021/12
14. Abbas, J.: Shilkret integral based on binary-element sets and its application in the area of synthetic evaluation. Eng. Technol. J. **33**(3), 571–577 (2015)
15. Abbas, J., Jafari, H.: The alternative representation of bipolar Sugeno integral. Eng. Technol. J. **40**(10), 1318–1324 (2022). https://doi.org/10.30684/etj.2022.135012.1256
16. Kareem, F., Abbas, J.: A generalization of the concave integral in terms of decomposition of the integrated function for bipolar scales. J. Appl. Sci. Nanotechnol. **1**(4), 81–90 (2021)
17. Grabisch, M., Roubens, M.: An axiomatic approach to the concept of interaction among players in cooperative games. Internat. J. Game Theory **28**, 547–565 (1999)

Formation of the Academic Index of a University Teacher Based on Weighted Evaluation Criteria

Misir Mardanov[1], Elchin Aliyev[2], Ramin Rzayev[2](✉), and Khanmurad Abdullayev[2]

[1] Institute of Mathematics and Mechanics, Vahabzadeh Str. 9, AZ1141 Baku, Azerbaijan
[2] Institute of Control Systems, Vahabzadeh Str. 9, AZ1141 Baku, Azerbaijan
raminrza@yahoo.com

Abstract. The assessment of the university teacher largely creates an idea of the quality of the university and its reputation. Nevertheless, in most universities in Azerbaijan, the assessment of the quality of the teachers' activities is not carried out on a systematic basis. For example, the basis for the ongoing procedure for replacing vacant posts at the department is not a clearly formulated concept on assessing the teacher activities. This complicates practical work and negatively affects the reputation of the university. Currently, various methods for evaluating the teaching staff are being developed and tested, but there is still no generally accepted methodology. Analyzing the reasons for the complexity of developing models and methods for evaluating the teachers' activities at Azerbaijan universities, the following main factors can be distinguished: 1) the difficulty of assessing the contribution of an individual teacher who performs many functional duties in the training of a specialist, which is inherently collective work; 2) the absence of reliable information about the qualitative results of the university as a whole and teachers, in particular; 3) the absence of some managers and teachers in conducting the objective assessment of the results of scientific and pedagogical activity based on complete verified information. Based on these considerations, the article proposes the method of expert evaluation of the teachers' activities according to the final result, systematized at the initial stage in the quality committee and university leadership.

Keywords: Academic Index · Expert Evaluation · Weighted Criterion

1 Introduction

In modern university rankings compiled by various international agencies, one of the key components is an integral indicator of the activities of teachers in the areas of scientific and methodological research and publications, training of scientific personnel, teaching, etc. [1]. It is not surprising that the scientific potential and methodological professionalism of the teaching staff is perceived as a reputation that guarantees the quality of education at the university [2]. In the context of informatization of society and increased competition in the market for the rendering of educational services, the acquisition and

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 600–608, 2023.
https://doi.org/10.1007/978-3-031-39777-6_70

maintenance of academic reputation become a priority for any university. Evaluation of the professional activity of university teachers is one of the most important components of the educational business process in terms of providing the conditions necessary for its full implementation under modern requirements. However, the establishment of academic indexes for teachers is an extremely difficult task, because. Scientific and pedagogical activity refers to a creative and innovative type of activity, where various evaluation options are possible. There are a number of approaches to assessing the activities of teachers: competence-based, systemic, multi-level, optimization and synergetic [1], which have been successfully implemented in numerous methods for calculating the academic ratings of teachers. On the assumption of this premise, the importance and relevance of creating new approaches to the formation of information profiles of teachers, taking into account both formal and weakly formalized (or non-metrizable) results of their activities at the university, becomes obvious.

2 Criteria for the Quality of Teachers' Activities

Most of the existing methods for evaluating the activity of university teachers are most actively used in US universities, which operate under the highly competitive market for the rendering of educational services. It is appropriate to note that the tradition of assessing the quality of a university teacher's work originated in the US academic environment, where private universities have always been the absolute majority. Therefore, in the process of our study, the American experience in the field of forming criteria for evaluating teachers is used. In particular, the developers of the Tennessee Technological University, analyzing the structural components of the educational business process at US universities, proposed a system of criteria for evaluating the quality of teachers' activities according to the following indicators that have a decisive impact on the university's reputation:

- *Pedagogical effectiveness* (x_1). It is assessed by independent and regular surveys of students on the quality of lectures, practical, laboratory and individual lessons, consultations, etc.
- *Research work* (x_2). It is evaluated by a special commission for the availability of scientific publications: monographs, articles and other scientific works.
- *Participation in research financed from abroad* (x_3). It is assessed by the quality committee for the participation of the teacher in projects financed from abroad.
- *Technical reports on research projects and disciplinary studies* (x_4). It is evaluated by the quality committee for participation of the teacher in grant projects. The teacher who wins the grant has to spend a lot of time and effort to generate the necessary reports that create a reputation not only for him, but for the university.
- *Performances in representative bodies* (x_5). It is evaluated by the quality committee for participation of the teacher in seminars, scientific forums and conferences, professional communities, etc. It is obvious that any performance of the teacher as an expert in public events creates a good advertisement for the university.
- *Representation and performances in public organizations not related to scientific activities* (x_6). It is assessed by the quality committee and the university administration for participation of the teacher in various boards of trustees, interest clubs, sports

sections or events, social services, foundations, etc. The involvement of the teacher in the public sphere creates additional advertising for the university.

- *Representation in bodies influencing public and state policy* (x_7). It is evaluated by the quality commission and the university administration for the performance of the teacher in Parliament and other state institutions, his work as an expert in legal bodies, etc.
- *Service in municipal, governmental and advisory bodies, various committees* (x_8). It is evaluated by the quality committee and the university administration, which take into account the amount of time spent by the teacher on this type of activity.
- *Citations of works in scientific publications* (x_9). It is evaluated by the quality committee using the Hirsch index.
- *Work in committees, commissions and working groups* (x_{10}). It is evaluated by colleagues and the university administration for the participation of the teacher in the internal and external life of the university, which indicates his loyalty to the ongoing educational business process.
- *Advising on current technological issues* (x_{11}). It is assessed by the university administration by calculating the number of days that the teacher spent on consultations with students, colleagues, administration.
- *Direct contact with production, employers* (x_{12}). It is evaluated by the university administration, taking into account the time spent by the teacher, and the effectiveness of such contacts depending on the area.
- *Technological licensing, certification, patents* (x_{13}). It is assessed by the quality committee to determine whether the teacher has proven inventions.

There is no unified approach to determining the integral assessment of the teacher's activities, since the compilation of an information profile is more of a scientific study than an administrative standard. The approach to assessing the quality of a teacher's activity, certainly, should be flexible, since each of the assessed types of teacher activity has its own characteristics. Moreover, in each specific case, the conclusion of the integral assessment of the teacher activity should be carried out by the university independently, depending on its profile, availability of resources, goals and objectives.

3 Problem Definition

To form information profiles of teachers based on integral assessments of their activities, it is first necessary to carry out a statistical analysis of the information compiled on the basis of decisions of expert commissions (quality committee and the university administration) for the teacher's compliance with the criteria x_i ($i = 1 \div 13$). This is a rather time-consuming and labor-intensive process, which should ensure maximum objectivity of the final information.

Suppose that the initial information from the expert commissions relative to the academic and other activities of the teacher for its compliance with the criteria x_i ($i = 1 \div 13$) is unified on a ten-point scale and averaged over each criterion. As a result, based on the collection and primary processing of data for all teachers using expert analysis methods, it is necessary to calculate the academic indices of teachers in the form of integral assessments of their activities. Thus, the main objective of this study is to develop a methodology for compiling the evaluation results of expert commissions regarding the activities of university teachers. Based on the qualitative characteristics of the factors x_i, which have a relative impact on the integral assessment of teachers' activities, it is necessary to calculate their academic indexes.

4 Identification of Weights of Evaluation Criteria Based on Expert Decisions

So, the teacher's index, as a result of an integral assessment of his activity, is a multifactorial category, which is characterized by the system of indicators x_i ($i = 1 \div 13$). From the point of view of the impact on the final result, each of these indicators has its own specific weight, demonstrating the degree of importance of the corresponding criterion in determining the teacher's index. Therefore, before initiate the multi-criteria assessment procedure, it is necessary to determine the weighting coefficients of indicators that each university establishes taking into account its own specifics.

Identification of the weights of indicators x_i as criteria for evaluating the teachers' activities implies: 1) expert ranking of indicators x_i for their priority, based on their relative impact on the final teacher index; 2) group expert evaluation of the normalized values of the generalized weights of indicators x_i, in accordance with their ranks.

Based on this scheme, we continue the reasoning as follows. Assume that by independent questioning of 15 experts, ranking estimates of the priority of indicators x_i are obtained. At the same time, each expert was asked to arrange x_i according to the following principle: designate the most important indicator with the number "1", the next less important one – with the number "2", and then in descending order of the expert's preference. The ranking estimates of indicators x_i obtained in this way are summarized in Table 1.

To proceed to the next stage, it is necessary to determine the degree of consistency of expert opinions. For this, the Kendall concordance coefficient [3] is usually used, which demonstrates the multiple rank correlation of expert opinions according to the following formula:

$$W = 12 \cdot S/ \left(m^2\right) n^3 - n\{\}, \tag{1}$$

where m is the number of invited experts, n is the number of evaluation criteria, and S is the deviation of expert opinions from the average ranking of indicators x_i, which is calculated by the formula [3]:

$$S = \sum\nolimits_{i=1}^{n} [\sum\nolimits_{j=1}^{m} r_{ij} - m(n+1)/2]^2, \tag{2}$$

where $r_{ij} \in \{1, 2, \ldots, n\}$ is the rank of the i-th indicator ($i = 1 \div n$) established by the j-th expert ($j = 1 \div m$).

In our case, with the value of $S = 37856$ calculated on the basis of (2) and the data from Table 1, the corresponding Kendall concordance coefficient is determined by formula (1) as $W = 12 \cdot 37856/[15^2(13^3 - 13)] = 0.9244$. The condition $W = 0.9244 > 0.6$ indicates a fairly acceptable agreement between expert opinions relative to the

Table 1. Expert ranking of teacher activities indicators.

Expert	Teacher performance indicators and their rankings (r_{ij})												
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}
1	1	2	3	5	7	10	12	13	4	11	9	8	6
2	2	1	4	5	7	10	12	13	3	9	11	8	6
3	1	2	4	6	8	10	12	13	3	11	9	7	5
4	2	1	3	6	7	12	10	11	5	13	8	9	4
5	1	2	3	5	7	12	10	13	4	9	11	8	6
6	1	4	3	5	7	13	10	12	2	9	8	11	6
7	3	1	2	4	7	13	10	12	5	9	8	11	6
8	1	3	4	6	8	10	13	12	2	11	9	7	5
9	1	2	3	4	7	13	12	10	5	9	11	8	6
10	1	2	4	6	7	10	12	13	3	8	11	9	5
11	2	1	3	6	8	11	12	13	4	10	9	7	5
12	1	2	3	6	7	13	10	11	4	12	8	9	5
13	1	2	4	5	7	13	11	12	3	9	10	8	6
14	2	1	3	5	7	13	10	12	4	9	8	11	6
15	3	1	2	4	7	12	10	11	5	9	8	13	6
$\sum$	23	27	48	78	108	175	166	181	56	148	138	134	83

priority rating of indicators x_i. At the next stage of the independent survey, the experts were also asked to establish the values of the normalized estimates of the generalized weights indicators x_i. As a result of such survey, the corresponding expert assessments were obtained and summarized in Table 2.

Identification of the generalized weights of indicators x_i is carried out in an iterative way according to the formula [3]

$$\alpha_i(t+1) = \sum_{j=1}^{m} w_j(t)\alpha_{ij}, \tag{3}$$

where $w_j(t)$ are indicators of expert competence at step t, determined from the following equalities [3]:

$$\begin{cases} w_j(t) = \frac{1}{\eta(t)} \sum\limits_{i=1}^{n} \alpha_i(t) \cdot \alpha_{ij} \; (j = \overline{1,\ m-1}), \\ w_m(t) = 1 - \sum\limits_{j=1}^{m-1} w_j(t),\ \sum\limits_{j=1}^{m} w_j(t) = 1, \end{cases} \tag{4}$$

Table 2. Expert estimates of normalized values of generalized weights of indicators x_i.

Expert	Indicators of teacher activity and normalized estimates of their weights (α_{ij})												
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}
1	0.140	0.130	0.115	0.095	0.075	0.045	0.030	0.020	0.105	0.040	0.055	0.065	0.085
2	0.130	0.135	0.110	0.095	0.075	0.040	0.035	0.020	0.115	0.055	0.040	0.065	0.085
3	0.140	0.130	0.105	0.085	0.065	0.045	0.030	0.020	0.115	0.040	0.055	0.075	0.095
4	0.130	0.140	0.115	0.085	0.075	0.030	0.045	0.040	0.095	0.020	0.065	0.055	0.105
5	0.140	0.130	0.115	0.095	0.075	0.030	0.045	0.020	0.105	0.055	0.040	0.065	0.085
6	0.140	0.105	0.115	0.095	0.075	0.020	0.045	0.030	0.130	0.055	0.065	0.040	0.085
7	0.110	0.145	0.130	0.105	0.075	0.020	0.045	0.030	0.095	0.055	0.065	0.040	0.085
8	0.140	0.115	0.105	0.085	0.065	0.045	0.020	0.030	0.130	0.040	0.055	0.075	0.095
9	0.140	0.130	0.115	0.105	0.075	0.020	0.030	0.045	0.095	0.055	0.040	0.065	0.085
10	0.140	0.130	0.105	0.085	0.075	0.045	0.030	0.020	0.115	0.065	0.040	0.055	0.095
11	0.130	0.140	0.115	0.085	0.065	0.040	0.030	0.020	0.100	0.050	0.055	0.075	0.095
12	0.140	0.130	0.115	0.085	0.075	0.020	0.045	0.040	0.105	0.030	0.065	0.055	0.095
13	0.140	0.130	0.105	0.095	0.075	0.020	0.040	0.030	0.115	0.055	0.045	0.065	0.085
14	0.130	0.140	0.115	0.095	0.075	0.020	0.045	0.030	0.105	0.055	0.065	0.040	0.085
15	0.115	0.140	0.130	0.105	0.075	0.030	0.045	0.040	0.095	0.055	0.065	0.020	0.085
$\sum$	2.005	1.970	1.710	1.395	1.095	0.470	0.560	0.435	1.620	0.725	0.815	0.855	1.345

where $\eta(t)$ is a normalizing factor that ensures the transition to the next iteration step and it is calculated by the formula [3]:

$$\eta(t) = \sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_i(t)\alpha_{ij}. \tag{5}$$

The process of identifying the weights of indicators x_i finishes under

$$\max_i\{|\alpha_i(t+1) - \alpha_i(t)|\} \le \varepsilon. \tag{6}$$

where ε is the calculation accuracy selected by the user.

Assuming the accuracy of calculations to be an arbitrarily small value, i.e. taking it, for example, as $\varepsilon = 10^{-5}$, and also assuming that at the initial stage ($t = 0$) the competence of experts is equally characterized by the value $w_j(0) = 1/m$, the average values for groups of normalized estimates of the generalized weights of indicators x_i in the 1st approximation are obtained from the particular form of formula (3): $\alpha_i(1) = \sum_{j=1}^{15} w_j(0)\alpha_{ij} = \sum_{j=1}^{15}\alpha_{ij}/15$ ($i = 1 \div 13$), as follows: $\alpha_1(1) = 0.13367$, $\alpha_2(1) = 0.13133$, $\alpha_3(1) = 0.11400$, $\alpha_4(1) = 0.09300$, $\alpha_5(1) = 0.07300$, $\alpha_6(1) = 0.03133$, $\alpha_7(1) = 0.03733$, $\alpha_8(1) = 0.02900$, $\alpha_9(1) = 0.10800$, $\alpha_{10}(1) = 0.04833$, $\alpha_{11}(1) = 0.05433$, $\alpha_{12}(1) = 0.05700$, $\alpha_{13}(1) = 0.08967$.

It is not difficult to see that condition (6) for the 1st approximation of the normalized estimates of the generalized weights of the indicators x_i is ruled out. Therefore, the transition to the next iteration step is carried out by the normalizing factor $\eta(1)$. According to (5), $\eta(1) = 1.40321$. Then, according to (4), the expert competence indicators adjusting in the form of

$$\begin{cases} w_j(1) = \frac{1}{\eta(1)} \sum_{i=1}^{14} \alpha_i(1) \cdot \alpha_{ij} \ (j = \overline{1,\ 14}), \\ w_{15}(1) = 1 - \sum_{j=1}^{14} w_j(1),\ \sum_{j=1}^{15} w_j(1) = 1, \end{cases}$$

in the 1st approximation take the following values: $w_1(1) = 0.06692$, $w_2(1) = 0.07171$, $w_3(1) = 0.06674$, $w_4(1) = 0.06652$, $w_5(1) = 0.06692$, $w_6(1) = 0.06644$, $w_7(1) = 0.06652$, $w_8(1) = 0.06643$, $w_9(1) = 0.06671$, $w_{10}(1) = 0.06673$, $w_{11}(1) = 0.06662$, $w_{12}(1) = 0.06679$, $w_{13}(1) = 0.06692$, $w_{14}(1) = 0.06684$, $w_{15}(1) = 0.06119$.

According to (3) or $\alpha_i(2) = \sum_{j=1}^{15} w_j(1)\alpha_{ij}$ the average values for groups of normalized estimates of the weights of indicators x_i in the 2nd approximation will be the following corresponding numbers: $\alpha_1(2) = 0.13376$, $\alpha_2(2) = 0.13131$, $\alpha_3(2) = 0.11389$, $\alpha_4(2) = 0.09295$, $\alpha_5(2) = 0.07300$, $\alpha_6(2) = 0.03138$, $\alpha_7(2) = 0.03728$, $\alpha_8(2) = 0.02889$, $\alpha_9(2) = 0.10810$, $\alpha_{10}(2) = 0.04833$, $\alpha_{11}(2) = 0.05419$, $\alpha_{12}(2) = 0.05725$, $\alpha_{13}(2) = 0.08966$.

Checking these values for the fulfillment of condition (6) and making sure that it is not fulfilled again:

$\max\{|\alpha_i(2)-\alpha_i(1)|\} = \max\{|0.13376–0.13367|, |0.13131–0.13133|, |0.11389–0.11400|, |0.09295–0.09300|, |0.07300–0.07300|, |0.03138–0.03133|, |0.03728–0.03733|, |0.02889–0.02900|, |0.10810–0.10800|, |0.04833–0.04833|, |0.05419–0.05433|, |0.05725–0.05700|, |0.08966–0.08967|\} = 0.00025 > \varepsilon$,

the normalizing factor corresponding to the next step is calculated as $\eta(2) = 1.40329$. Then the expert competence indicators in the 2nd approximation are the follows: $w_1(2) = 0.06692$; $w_2(2) = 0.07172$; $w_3(2) = 0.06675$; $w_4(2) = 0.06652$; $w_5(2) = 0.06693$; $w_6(2) = 0.06644$; $w_7(2) = 0.06651$; $w_8(2) = 0.06644$; $w_9(2) = 0.06671$; $w_{10}(2) = 0.06673$; $w_{11}(2) = 0.06662$; $w_{12}(2) = 0.06679$; $w_{13}(2) = 0.06693$; $w_{14}(2) = 0.06684$; $w_{15}(2) = 0.06117$.

Taking into account the updated degrees of expert competence, according to (3) in the 3rd approximation the group average values of normalized estimates of the weights of indicators x_i are the follows: $\alpha_1(3) = 0.13376$; $\alpha_2(3) = 0.13131$; $\alpha_3(3) = 0.11389$; $\alpha_4(3) = 0.09295$; $\alpha_5(3) = 0.07300$; $\alpha_6(3) = 0.03138$; $\alpha_7(3) = 0.03728$; $\alpha_8(3) = 0.02889$; $\alpha_9(3) = 0.10810$; $\alpha_{10}(3) = 0.04833$; $\alpha_{11}(3) = 0.05419$; $\alpha_{12}(3) = 0.05725$; $\alpha_{13}(3) = 0.08966$.

In this case, condition (6) is already satisfied, so the last obtained values $\alpha_1(3)$, $\alpha_2(3)$, ..., $\alpha_{13}(3)$ can be considered as the generalized weights of indicators x_i.

The calculation of the academic index of the university teacher, theoretically ranging from 0 to 100, is carried out using the formula:

$$R = \frac{\sum_{i=1}^{n} \alpha_i e_i}{\max\{\sum_{i=1}^{n} \alpha_i e_i\}} \times 100, \tag{7}$$

where n is the number of evaluation criteria; α_i is the weight of the evaluation criterion x_i ($i = 1 \div n$); e_i is consolidated expert assessment of the academic and other activities of the teacher on a ten-point scale. The higher the index, the higher the teacher's rating in the hierarchy of the educational process.

As an example, five teachers a_k ($k = 1 \div 5$) were selected, for whom the initial information from expert commissions regarding their academic and other activities for compliance with the x_i ($i = 1 \div 13$) was unified on a ten-point scale and averaged for each criterion. After such processing, for each teacher, the average data for each criterion are systematized in the form of Table 3, which also contains the weights of the criteria and academic indexes calculated according to the formula (7).

Table 3. Average estimates of unified expert decisions obtained on a ten-point scale.

Quality criteria (x_i)	Criterion weight (α_i)	University teachers (a_k)					max $\{e_{ik}\}$
		a_1	a_2	a_3	a_4	a_5	
x_1	0.13376	7.47	8.46	2.37	4.18	2.54	8.46
x_2	0.13131	7.98	3.11	6.12	8.32	2.97	8.32
x_3	0.11389	5.33	3.51	6.50	1.84	8.90	8.90
x_4	0.09295	1.67	6.15	1.04	9.89	1.79	9.89
x_5	0.07300	1.62	9.87	9.72	0.33	3.96	9.87
x_6	0.03138	9.18	9.17	4.45	9.04	8.10	9.18
x_7	0.03728	3.68	2.05	4.20	3.05	1.15	4.20
x_8	0.02889	2.48	0.01	3.86	2.63	5.69	5.69
x_9	0.10810	5.29	6.52	5.97	7.39	3.20	7.39
x_{10}	0.04833	3.27	5.07	8.73	6.24	4.51	8.73
x_{11}	0.05419	6.17	2.09	5.30	9.79	5.58	9.79
x_{12}	0.05725	2.98	8.00	4.91	2.46	2.93	8.00
x_{13}	0.08966	7.81	6.43	3.35	9.06	3.51	9.06
Academic index		63.00	66.92	58.88	68.91	47.12	100.00

5 Conclusion

The article proposes an approach to calculating the academic indices of teachers by weighted summation of consolidated expert assessments of key indicators of their scientific and pedagogical activities. The system of teaching quality indicators, which are systematically used in US universities for multi-criteria assessment of teachers' activities, was chosen as the basis. To establish the priority of evaluation criteria, an analytical method was used to identify the weights of quality indicators. The proposed approach was tested on the example of five teachers, who are characterized by averaged estimates of unified expert decisions obtained at the initial stage on the ten-point scale.

References

1. Altbach, P.G.: The Global Future of Higher Education and the Academic Profession: The BRICs and the United States. Springer, Heidelberg (2013)

2. Kehm, B.M.: University Rankings, Diversity and the New Landscape of Higher Education. Sense Publishers, Rotterdam (2009)
3. Lin, A., Hedayat, S., Wu, W.: Statistical Tools for Measuring Agreement. Springer, New York, NY (2012). https://doi.org/10.1007/978-1-4614-0562-7
4. Mardanov, M., Rzayev, R.: One approach to multi-criteria evaluation of alternatives in the logical basis of neural networks. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Sadikoglu, F.M. (eds.) ICAFS 2018. AISC, vol. 896, pp. 279–287. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04164-9_38

Classification of XTEKS Companies During COVID-19 Pandemic using Fuzzy-Analytic Hierarchy Process and Fuzzy-C-Means

Fatih Yiğit(✉)

Altinbas University, 34217 İstanbul, Turkey
fatih.yigit@altinbas.edu.tr

Abstract. Assessment of companies is vital for an accurate investment decision. Financial ratios are essential performance indicators. However, there is no consensus in their comparison among all financial ratios. Expert opinions are an indispensable resource for such assessment. This study uses an integrated approach to benefit from expert opinions. Clustering is an important area of unsupervised learning. Clustering, when assigned to classes, can also be used for classification. It is vital to classify data to apply for decision-making. This study applies the Fuzzy Analytic Hierarchy Process (FAHP) and Fuzzy C-Means (FCM) for clustering and classification as a part of machine learning. Financial ratios are widely used to compare different companies. This study focuses on companies under the "Textile Leather Index" registered in the Istanbul Stock Exchange (BIST) for applying the proposed model.

The study employed current financial results and the positive and negative trends of the last year for classification. The results allow the decision-maker to choose the right company to invest in. Among 17 companies, 2 are classified as A class.

To the best of our research, using trend values and integrating FAHP and FCM for classification is new in the literature.

Keywords: Fuzzy Analytic Hierarchy Process (FAHP) · Multi-Criteria Decision Making · COVID-19 · Company Evaluation

1 Introduction

The selection of the right company to invest in is a critical decision. Money is scarce, and the alternatives have no limits. Therefore, it is a hard decision. This study aims to use the Fuzzy Analytic Hierarchy Process (FAHP) and Fuzzy C-Means (FCM) to assess and classify potentials correctly. Companies in the textile industry registered in Istanbul Stock Exchange are in XTEKS. The study focuses on 17 companies registered in XTEKS. The mentioned companies are deeply affected by the recent pandemic. Due to lockdowns and relevant curfews, many shopping malls are closed. Most people stayed at home and preferred to work at the home office. This change also reduced the demand for apparel.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 609–616, 2023.
https://doi.org/10.1007/978-3-031-39777-6_71

This study allows the decision-maker to classify the companies based on financial ratios and changes in these ratios. The proposed methodology shows not only current ratios calculated based on expert opinion but also based on trends. The criterion called trend represents the changes in the performance of companies per quarter based on financial ratios. A positive trend represents an increase in financial ratios, such as profit and inventory turnover, and a negative trend represents a degeneration in the same financial ratios. Expert opinion is efficient data to solve problems. FAHP is used to convert these opinions into weights. Those weights are used to compare alternative companies using financial ratios.

Gaining insight into the right company to invest in is a complicated task. This study incorporates the FAHP with the FCM to solve this multiple-criteria decision-making (MCDM) problem. To our best knowledge, these two methodologies are combined to classify companies based on financial ratios for the first time in literature. The proposed method uses calculated financial ratios for the last quarter and their progress. This way, decision-makers can still choose a low-performance company due to improved financial ratios.

The remainder of this study is as follows. In Sect. 2, the literature review is given. In Sect. 3, details of the proposed methodology are given; in Sect. 4, the evaluation and classification of the companies registered in the textile sector; and in Sect. 5, the conclusion is given.

2 Literature Review

Financial ratios have been used for a long time. They are used not only for assessment but also for prediction. Beaver [1] used financial ratios for the prediction of failing companies. Since then, many other applications related to financial ratios have been published. Ertuğrul and Karakaşoğlu [2] applied FAHP with TOPSIS. Turkish cement firms are evaluated using these combined methods based on financial ratios. The rankings of the firms are concluded based on the calculated results. Tseng et al. [3] used similar performance indicators for measuring the business performance of TFT-LCD panel companies based in Taiwan. The goal of the study is to help high-tech manufacturing executives. Based on the outcomes, they may act according to strengths and weaknesses. Similar results can be acquired from the proposed study. Executives or investors may benefit from the proposed approach.

Bulgurcu [4] used the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method to compare 13 technology companies using ratio analysis. The ratio analysis investigates the correlation between market value and rankings. It has been found that companies should improve general efficiency. Fenyves et al. [5] implemented a benchmarking method to evaluate the performance of companies based on financial analysis. The study employed Data Envelopment Analysis (DEA). Kliestik et al. [6] assessed the financial ratios used in the models of transition countries. The study found that the most used financial ratios are the current ratio, total liabilities to total asset ratio, and total sales to total asset ratio. These ratios are different from the suggested ratio. As a result, it shows the area for improvement. According to our research, studies focused mainly on static evaluation. The studies in the literature do not cover the performance changes between periods. In other words, they use the results of a period for classification. Also, DEA and TOPSIS are MCDM approaches. Since the proposed study was MCDM in the first step, there is a similarity in the methods preferred between the mentioned studies and proposed approaches.

The analytic hierarchy process (AHP) is a widely used method for MCDM. According to Saaty [7], AHP is a general theory of measurement used to conclude discrete and continuous pair comparisons. FAHP is a widely used extension of AHP methodology with the addition of fuzzy logic. This improvement enabled the integration of fuzzy behavior inherent in human decisions. Among MCDM, FAHP is ranked the second most used methodology, following AHP, according to Kubler et al. [8]. Both AHP and FAHP are efficient methodologies for solving MCDM problems. As a result, they are widely used for academic studies. According to Kubler et al. [8], 190 studies are published using FAHP methodology between 2004–2016. This common usage is one of the main reasons for choosing the FAHP method in this study. FAHP is used in many different MCDM problems. Gündoğdu and Kahraman [9] used a novel approach of FAHP and applied the proposed model to the renewable energy location selection problems.

Clustering is performed to analyze data and group members of the population based on their similarities. The proposed study performs clustering using FCM. FCM has distinct advantages as the outcomes are membership values of each class member. FCM was first proposed by Dunn [10]. FCM has been widely used in many applications since then. Esnaf and Küçükdeniz [11] used FCM for the multi-facility location problem. The results show that the proposed methodology can improve the logistics network's transportation cost performance. Ma et al. [12] applied the Pythagorean Fuzzy C-Means (PFCM) approach in the image segmentation area. The application is made on a dataset available. In this study, the proposed model performed better than other alternative models.

The literature research showed that financial ratio assessment, FAHP, and FCM are vast areas of interest. Financial ratios are easy to calculate and widely accepted performance indicators. FAHP model is easy to understand and apply the model for to use decision makers' experience. FCM is a flexible model due to membership functions gained as output. The outputs as membership values can change the number of members in each cluster. The flexibility allows the change in the number of members in each cell.

3 Method

The study uses the financial ratios of the XTEKS companies for valuation. The proposed method uses FAHP to integrate the expert's view on the assessment and valuation. FCM is used to classify companies into three classes based on weights and financial ratios. The stages of the proposed method are given in Fig. 1. As presented in Fig. 1, Stage 1 covers the assessment of financial ratios. In stage 2, FAHP is executed. In stage 2, fuzzy pair-wise comparisons performed by decision-makers are used to assess the weights of each financial ratio. In the last stage, stage 3, FCM is performed. The goal of stage 3 is to classify companies based on their performance.

These three classes are also called A, B, and C. The proposed methodology aims to assess alternative companies based on their performance. Financial ratios of the companies are helpful values to evaluate their performance. They are widely used by practitioners, academicians, and other companies such as banks, creditors, and investors. The study used the financial ratios used by Ertuğrul and Karakaşoğlu [2]. Details of the ratios can be found in the relevant research.

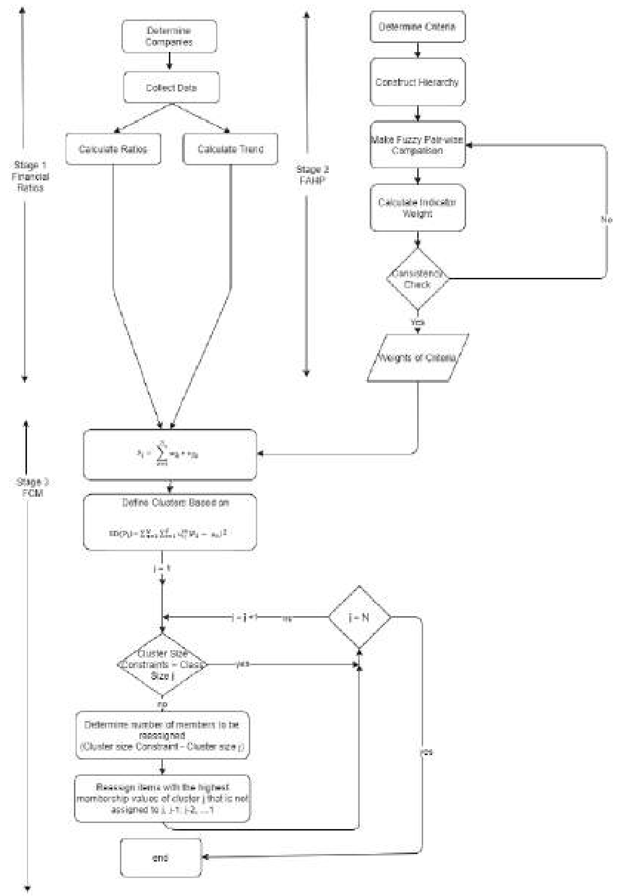


Fig. 1. Flowchart of the Proposed Method

Trend values mean the change in the financial ratios of the companies based on quarterly values. A positive trend implies the company performs better than in previous quarters and vice versa. To our best research, using trends for assessment is new in the literature. As an extension, in this study, the author used the trend with current values of financial ratios. Financial values are categorized into four main groups, and the following main ratios are used for this study. This study focuses on the financial ratios of the XTEKS companies. Current ratios represent the financial ratios of the companies reported in Q3

2020. Trend value is calculated based on the companies' progress between Q4 2019 and Q3 2020. The study uses the slope of a regression line between four financial periods to calculate the trend values.

FAHP is proposed to convert experts' opinions to scalar values by Van Laarhoven and Pedrycz [12]. FAHP incorporates AHP with fuzzy theory. The linguistic scale of the traditional AHP method is used to express the fuzzy uncertainty inherent in a decision-making process [13]. FCM uses membership values for clustering. Details of the technique can be found in the study of Bezdek [16].

4 Numerical Study

In this section, the proposed model is applied to the XTEKS companies. There are 19 companies registered in XTEKS. The study covers 17 of these companies. 2 companies do not share the optional quarterly financial ratios. The proposed research uses XTEKS due to the COVID-19 pandemic and since the textile industry is traditionally essential for Turkey. Some of these companies are either producing textile products. DAGI and DESA are among such companies. Other companies supply products to retail stores, and HATEK and LUKSK are among such companies. These companies are affected indirectly by COVID-19 due to lost sales to their customers. As a result, the study focused on XTEKS, as many people prefer to invest in such companies on the stock exchange.

4.1 Stage 1 -Calculation of Financial Ratios

Financial ratios between Q4 2019 and Q3 2020 of the companies are collected. For some financial ratios, higher values represent a better status and vice versa. Debt Ratio is an example of such ratios representing a lower value preferred. The numerator and denominator places are changed to represent the ratios similarly. This way, higher values represent a better position for all financial ratios.

4.2 Stage 2 -Application of FAHP

The proposed method's first stage is applied to 17 companies. In this FAHP stage, the expert opinions of three financial professionals are used. The financial ratios used for comparison are taken from the study of Ertuğrul and Karakaşoğlu [2]. The decision-maker also accepts these financial ratios. Experts from different business functions have different views on financial ratios. The author chose various experts from other areas of finance for this study.

There are multiple methods for the calculation of FAHP. The author used Buckley's [14] model in this study. However, the method proposed by Chang [15] is widely used. It has the main drawback. It may assign weights that do not represent the priorities. Final weights calculated based on multiplying the weight of main criteria and sub-criteria weights are given in Fig. 2.

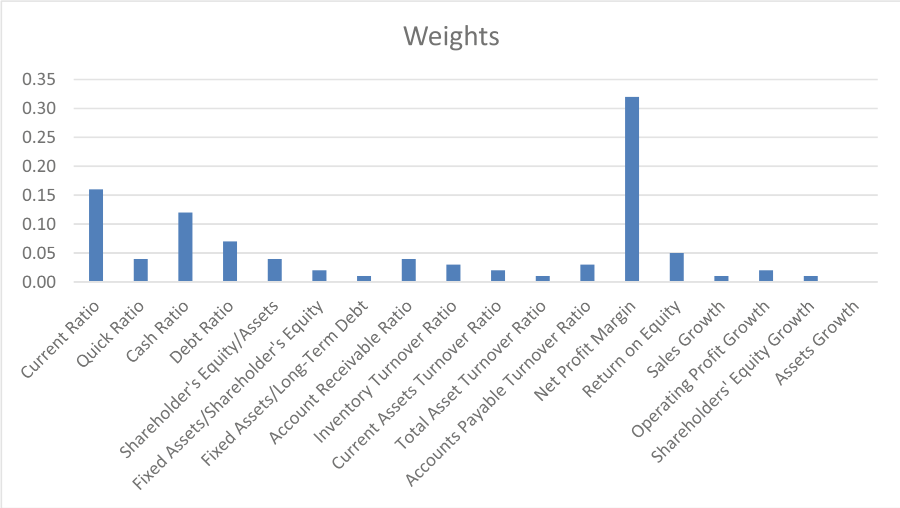


Fig. 2. Weights of Criteria

4.3 Stage 3 -Application of FCM

Classification is performed for 3 classes using an embedded function. A class covers the highest-performing company, and the C class covers the lowest-performing company. B class covers the mediocre performing companies. The classification of each class is concluded based on membership function. The membership function allows the decision-maker to classify different classes when needed.

5 Results and Conclusion

Based on the calculations, the classification among XTEKS companies is prepared. In Fig. 3., the classification and membership values are given. In this classification, FCM proved to be a successful application. Two companies are classified under the A ranking. The membership values for each company also represent flexibility. The proposed model assigns the companies to relevant classes. A decision-maker may find the number of companies not suitable. In the case study, only two companies are assigned to top-class A. Membership values allow the user to assign other companies to the desired class A. The use of trends in the financial ratios of the companies has a significant advantage. It allows the decision makers to assess the companies not only due to their latest performance. Also, classification enables the decision-maker to choose different companies assigned to the same class. In the application, class A represents the highest-performing companies. Only two companies are assigned.

The study proved that the proposed method quickly classifies companies based on their financial ratios. Other areas of applications can easily be applied with the proposed hybrid method. Future studies will focus on alternative FAHP techniques for comparison. In future studies, more decision-makers will be integrated into the process. Other classification methods will be used for future studies. A software application can be beneficial, and visualization may benefit further applications.

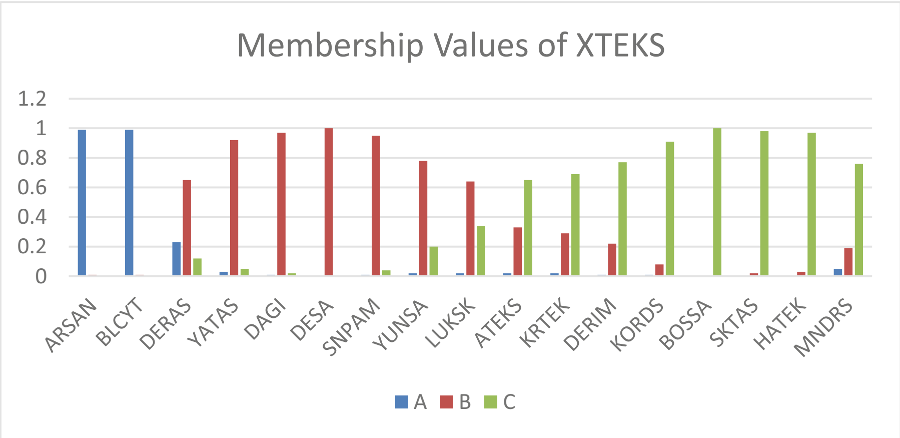


Fig. 3. Membership Values of XTEKS Companies

References

1. Beaver, W.H.: Financial ratios as predictors of failure. J. Account. Res. **4**, 71–111 (1966). http://www.jstor.org/stable/2490171
2. Ertuğrul, I., Karakaşoğlu, N.: Performance evaluation of Turkish cement firms with fuzzy analytic hierarchy process and TOPSIS methods. Expert Syst Appl **3**(1), 702–715 (2009). https://doi.org/10.1016/j.eswa.2007.10.014
3. Tseng, F.M., Chiu, Y.J., Chen, J.S.: Measuring business performance in the high-tech manufacturing industry: a case study of Taiwan's large-sized TFT-LCD panel companies. Omega (Westport) **37**(3), 686–697 (2009)
4. Bulgurcu, B.K.: Application of TOPSIS technique for financial performance evaluation of technology firms in istanbul stock exchange market. Procedia Soc. Behav. Sci. **62**, 1033–1040 (2012)
5. Fenyves, V., Tarnóczi, T., Zsidó, K.: Financial performance evaluation of agricultural enterprises with DEA method. Procedia Econ. Finan. **32**(15), 423–431 (2015)
6. Kliestik, T., Valaskova, K., Lazaroiu, G., Kovacova, M.: Remaining financially healthy and competitive: the role of financial predictors. J. Compet. **12**, 74–92 (2020)
7. Saaty, R.W.: The analytic hierarchy process—what it is and how it is used. Math. Model. **9**(3), 161–176 (1987)
8. Kubler, S., Robert, J., Derigent, W., Voisin, A., Le Traon, Y.: A state-of the-art survey & testbed of fuzzy AHP (FAHP) applications. Expert Syst. Appl. **65**, 398–422 (2016)
9. Kutlu Gündoğdu, F., Kahraman, C.: A novel spherical fuzzy analytic hierarchy process and its renewable energy application. Soft. Comput. **24**(6), 4607–4621 (2019). https://doi.org/10.1007/s00500-019-04222-w
10. Dunn, J.C.: A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters (1973)

11. Esnaf, Ş, Küçükdeniz, T.: A fuzzy clustering-based hybrid method for a multi-facility location problem. J. Intell. Manuf. **20**(2), 259–265 (2009)
12. Van Laarhoven, P.J.M., Pedrycz, W.: A fuzzy extension of Saaty's priority theory. Fuzzy Sets Syst. **11**(1–3), 229–241 (1983)
13. Hsu, Y.L., Lee, C.H., Kreng, V.B.: The application of fuzzy delphi method and Fuzzy AHP in lubricant regenerative technology selection. Expert Syst. Appl. **37**(1), 419–425 (2010)
14. Buckley, J.J.: Fuzzy hierarchical analysis. Fuzzy Sets Syst. **17**(3), 233–247 (1985)
15. Chang, D.Y.: Applications of the extent analysis method on fuzzy AHP. Eur. J. Oper. Res. **95**(3), 649–655 (1996)
16. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2–3), 191–203 (1984)

An Intelligent System for Determination of Stop – Loss and Take – Profit Limits: A Dynamic Decision Learning Approach

Mahmut Sami Sivri[1], Ahmet Berkay Gultekin[1,2], Alp Ustundag[1], Omer Faruk Beyca[1], Emre Ari[1](✉), and Omer Faruk Gurcan[2]

[1] Istanbul Technical University, İstanbul, Turkey
{sivri,gultekinah,ustundaga,beyca,ari18}@itu.edu.tr
[2] Sivas Cumhuriyet University, Sivas, Turkey
ofgurcan@cumhuriyet.edu.tr

Abstract. Stock market prices are notoriously difficult to predict for traders and investors alike. However, accurate stock market price predictions can result in high returns and substantial percentages of returns for investors and traders. Unfortunately, stock price data is inherently complex, noisy, and nonlinear, making it a challenging task. As technology continues to advance, trading strategies are beginning to adapt to automated systems instead of relying on manual analysis. Dynamically determining buying and selling levels in automated systems has become increasingly important. Many traders and investors seek to minimize losses and maximize profits by utilizing technical analysis methods and implementing stop-loss and take-profit strategies. Technical analysis methods are commonly used by traders and investors to determine and set predetermined thresholds for existing positions, as well as enter positions with stop-loss and take-profit orders. In this study, the main objective is to determine stop-loss and take-profit levels dynamically by analysing historical data using standard deviation and Sharp Ratios. To decide on the selling (short) or buying (long) position, TP\SL levels have been divided into two separate parts with different approaches. The approaches in this study aim to compare the end-of-day Open to Close returns with TP\SL level returns to determine the best course of action. Overall, this study aims to develop effective trading strategies that can minimize losses and maximize profits in the volatile world of stock market trading.

Keywords: Stop – Loss (SL) · Take – Profit (TP) · Artificial Intelligence (AI) · Stock Prediction · Open to Close (OTC) · Sharp Ratio (SR)

1 Introduction

The stock market is a challenging and constantly changing environment, making it difficult for traders and investors to predict stock prices accurately. However, accurate price predictions can yield high returns for investors and traders. With the advancement of technology, automated trading strategies have gained popularity among traders and

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 617–624, 2023.
https://doi.org/10.1007/978-3-031-39777-6_72

investors. One of the major obstacles in automated trading is dynamically determining the buying and selling levels to optimize profits and minimize losses. Traders and investors often use technical analysis methods to establish thresholds for existing positions, as well as to enter positions with stop-loss and take-profit strategies.

The main objective of this study is to develop an approach for dynamically determining the stop-loss and take-profit levels in automated trading systems using historical data and standard deviation analysis. The approach proposed in this study utilizes artificial intelligence techniques, particularly machine learning algorithms, to analyse historical stock market data and determine the TP/SL levels for buy and sell positions separately.

The study evaluates the effectiveness of the proposed approach by comparing the end-of-day returns with the TP/SL levels returns using the Sharpe ratio metric. The study's results suggest that the proposed approach can potentially improve the accuracy of automated stock trading strategies with TP & SL strategies via Sharpe ratio.

In essence, this study aims to create a method that can dynamically determine the optimal buying and selling levels in automated trading systems through the use of standard deviation analysis and artificial intelligence techniques. The findings of this research have the potential to enhance the accuracy and profitability of automated trading strategies for both traders and investors.

The paper is structured into five sections. The Sect. 1 provides an introduction to the study. The Sect. 2 will be a literature review discussing related works. The Sect. 3 will present the experiments and results, with the algorithm's steps provided in detail. In the Sect. 4, the discussion and conclusion will summarize the results and explore possible directions for future related research. Finally, the fifth and last section will contain the references for the study's related works.

2 Related Works

Predicting stock market prices accurately is a challenging task for traders and investors due to the complex, noisy, and nonlinear nature of stock price data. However, accurate predictions can result in high returns for traders and investors. As technology has advanced, trading strategies have increasingly been adapted to automated systems rather than manual processes. One key challenge in automated trading is dynamically determining the buying and selling levels to optimize profits and minimize losses. To achieve this, traders and investors often use technical analysis methods and strategies such as stop-loss and take-profit. Several studies have highlighted the benefits of algorithmic trading for improving liquidity and generating profits [1–4].

In recent times, there has been a surge in the use of artificial intelligence (AI) to predict stock prices. Machine learning algorithms like artificial neural networks, decision trees, and support vector machines have been employed for this purpose. These algorithms can capture intricate patterns and relationships in the data, leading to improved accuracy of stock price predictions.

Technical analysis indicators, which are based on past price and volume data, can be used to predict stock prices and identify trends, support and resistance levels, and momentum in stock prices. Bollinger Band approach is a commonly used method that utilizes mean and standard deviation to determine buying and selling levels. However, this method may not be suitable for low volatile data. Other advanced regression analysis techniques have also been applied for trading and investment, such as in [6, 7]. Additionally, machine learning algorithms, like artificial neural networks and support vector machines, have been used to predict stock prices and improve trading strategies, as discussed in [5]. These approaches have been shown to capture complex patterns and relationships in the data, leading to more accurate predictions and higher profits.

Traders and investors often employ stop-loss and take-profit strategies to minimize losses and maximize profits, as discussed in the literature [8–11]. A stop-loss order involves placing an order with a broker to sell a security (or short) when it reaches a certain price level, while a take-profit order involves placing an order with a broker to buy a security (long) when it reaches a certain price level, with the aim of making a profit. Technical analysis methods such as support and resistance levels are often used to determine the appropriate stop-loss and take-profit levels [8–11]. These methods are discussed extensively in the literature, including Algorithmic Trading & DMA [8], Trading Systems and Methods [9], The Evaluation and Optimization of Trading Strategies [10], and Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market [11].

In their study, the authors utilized historical data to dynamically determine stop-loss and take-profit levels based on the standard deviation and sharp ratio. They compared two different methods to determine these levels and evaluated the end-of-day returns in comparison to the TP/SL level returns. The study's results suggest that the proposed method can be effective in dynamically determining stop-loss and take-profit levels and potentially improve the accuracy of stock price predictions when compared to Open to Close returns [12–14].

3 Application

The approach proposed in this study employs artificial intelligence techniques, particularly ensemble learning algorithms such as Random Forest (RF), Extreme Gradient Boosting (XGB), and Light Gradient Boosting (LGB), which are derived from Decision Tree (DT) Algorithms in the field of machine learning. The aim is to analyze historical stock market data and dynamically determine the stop-loss and take-profit levels. To select relevant features, various methods such as Correlation, Chi-Square, Recursive, Shapley, and Lime are utilized. The study employs the Python programming language and several libraries, including Pandas, NumPy, and Scikit-Learn, to conduct data analysis and machine learning.

The proposed approach involves using artificial intelligence techniques, specifically ensemble learning algorithms such as Random Forest (RF), Extreme Gradient Boosting (XGB), and Light Gradient Boosting (LGB), which are based on Decision Tree (DT) Algorithms. These algorithms are used to analyse historical stock market data and dynamically determine the stop-loss and take-profit levels. In the first step, a pool of models is created using the selected ensemble learning algorithms and features selected through methods such as Correlation, Chi-Square, Recursive, Shapley, etc. The 30 models with the highest Sharp Ratio and based on 6 months of historical data for a given stock are selected from this pool. In the second step, these selected models and features are implemented on 30 selected stock market equities. In the final step, the TP & SL levels are determined based on Standard Deviation (SD) and Sharp Ratio (SR), and their performance is compared with Open to Close returns. This entire process can be considered pre-processing before applying the proposed approach, which is outlined in more detail in the following steps.

In the first step of our approach, we selected the most appropriate features and models from a pool of options. Ensemble learning models, including Random Forest, Extreme Gradient Boosting, and Light Gradient Boosting, were trained on the historical data of a given stock over the last six months. Before selecting the models, we used feature selection methods such as Correlation, Chi-Square, Shapley, Recursive, and Lime to identify the most important features to include.

Step 2 involved the application of the selected models and features to all equities, except for the equity data used in building the models. This process involved the training of equities based on six months data of the stock market, and ordering them based on their Sharp Ratio results using Open to Close, TP & SL. The best result equities were selected by comparing their Open to Close and TP & SL performance.

Step 3 involved examining the previous day's returns in two directions: the first 5 negative returns and the first 5 positive returns were grouped separately for analysis.

In Step 4, the TP and SL positions were determined based on the information obtained in the previous steps. For the sell (short) position, the TP/SL levels were determined using a specific approach.

To evaluate the effectiveness of the proposed approach, the study compared the end-of-day returns with the TP/SL levels returns against the Open to Close returns. The Sharpe Ratio was used as a metric to assess the performance of the proposed approach, with the highest SR score between TP & SL and Open to Close returns being the desired outcome. The definition and formula of the SR were provided in the study.

3.1 Take – Profit and Stop Lose Method (TP & SL)

The TP & SL method employs a specific approach to determine the TP & SL levels, which involves the following steps:

1. The algorithm looks back to find the first five positive returns.
2. The algorithm also looks back to find the first five negative returns.
3. The algorithm calculates the average of the positive and negative returns separately.
4. If the following day's returns reach these average values, the algorithm triggers an order to sell (short) or buy (long).

This method helps the algorithm to dynamically determine the stop-loss and take-profit levels based on the stock's historical data. The effectiveness of this approach is evaluated by comparing the end-of-day returns with the TP & SL levels returns and Open to Close returns. The Sharpe Ratio (SR) is used as a metric to assess the performance of this approach. The highest score of the SR between TP & SL and Open to Close returns indicates the effectiveness of the proposed approach. The formula and definition of the SR are provided to assess the performance of the approach.

Fig. 1. FROTO Candle Graph Daily Price

3.2 Sharpe Ratio

In this application, the Sharpe Ratio serves as the primary indicator for dynamically determining the TP & SL of equities in the Stock Market using daily data from a six-month period. The formula for the SR is as follows:

$$\text{Sharpe Ratio (SR)} = \frac{R_p - R_f}{\nu_p} \tag{1}$$

R_p : *Return of portfolio*
R_f : *Risk − Free Rate*
σ_p : *standart deviation of the portfolio*$'$*s excess return*

where R_p is the portfolio return, R_f is the risk-free rate (assumed to be zero in this study), and ν_p is the standard deviation of the portfolio's excess return. The SR is calculated using data from the previous six months and the resulting scores are compared between TP & SL and Open-to-Close prices. The highest SR scores are then selected (Table 1).

Based on Fig. 1, the calculated Sharpe Ratio (SR) values for TP & SL and Open to Close returns are 2.39 and 0.95, respectively. The bold values in the figure indicate an improvement in the SR values for TP & SL. Overall, the TP & SL values have a higher SR than the Open to Close return values.

Table 1. Sharpe Ratios Open to Close and TP & SL

Models	6 Months TP & SL- SR	6 Months Open to Close- SR
AKBNK	**3,4**	0,3
ARCLK	**1,5**	0,8
ASELS	**1,3**	0,9
BIMAS	**2,4**	0,8
DOHOL	**2,8**	1,6
EKGYO	**4,3**	0,8
EREGL	**2,3**	–0,1
FROTO	**1,7**	1
GARAN	**2,6**	0,8
HALKB	**4,7**	–0,6
ISCTR	1,5	1,5
KCHOL	**1,8**	1,2
KOZAA	**2,7**	1,6
KOZAL	2,1	**3,4**
KRDMD	**4,4**	1,1
MGROS	0,6	0,6
PETKM	**4,1**	0,8
PGSUS	1	**1,2**
SAHOL	**2,7**	1,1
SISE	**4,6**	1,8
TAVHL	**2**	0,5
TCELL	**1**	0,5
THYAO	**1,2**	0,7
TKFEN	**2,8**	1,5
TOASO	**1,6**	1,1
TSKB	**2,6**	1,9
TTKOM	**3,1**	0,4
TUPRS	**1,9**	–0,1
VAKBN	**1,5**	0,1
YKBNK	1,5	1,5
Average	**2,39**	0,95

4 Discussions and Concluding Remarks

Accurately predicting stock market prices is essential for investors and traders to achieve high returns. However, the complexity, noise, and nonlinearity of stock price data make it challenging to predict. Thus, the development of automated trading strategies has become crucial in dynamically determining buying and selling levels. In this context, technical analysis methods, such as stop-loss and take-profit strategies, have gained popularity among traders and investors. This study aimed to determine stop-loss and take-profit levels dynamically by analysing standard deviation in historical data compared to Open to Close returns. By dividing TP/SL levels into two parts using different approaches, the study aimed to decide which level to take for a sell (short) position or a buy (long) position. The approaches in this study were based on comparing end-of-day returns with TP/SL levels returns. Overall, this study provides a useful framework for traders and investors to improve their decision-making process while minimizing losses and maximizing profits.

Upon analysing the results of comparing SRs of the TP & SL and Open to Close returns for 30 equity shares over a 6-month period, the average SR value of TP & SL was found to be approximately 2.39, while the average SR value of Open to Close returns was found to be 0.95. Overall, the TP & SL returns approach proved to be more effective than Open to Close returns, as indicated by the higher average SR value. By using the TP & SL method, we were able to achieve higher returns while minimizing risk, as compared to the Open to Close approach. In the future studies, the aim of this study to make dynamic approaches developing this static approach.

References

1. Treleaven, P., Galas, M., Lalchand, V.: Algorithmic trading review. Commun. ACM **56**(11), 76–85 (2013)
2. Hendershott, T., Riordan, R.: Algorithmic Trading and Information. University of California, Berkeley (2009)
3. Nuti, G., Mirghaemi, M., Treleaven, P., Yingsaeree, C.: Algorithmic trading. Computer **44**(11), 61–69 (2011)
4. Hendershott, T., Jones, C.M., Menkveld, A.J.: Does algorithmic trading improve liquidity? J. Finan. **66**(1), 1–33 (2011)
5. Liang, Y., Thavaneswaran, A., Yu, N., Hoque, M. E., Thulasiram, R.K.: Dynamic data science applications in optimal profit algorithmic trading. In: 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 1314–1319. IEEE (2020)
6. Dunis, C.L., Williams, M.: Applications of advanced regression analysis for trading and investment. Appl. Quant. Methods Trad. Invest. (2003)
7. Jegadeesh, N., Titman, S.: Returns to buying winners and selling losers: implications for stock market efficiency. J. Finan. **48**(1), 65–91 (1993)
8. Johnson, B.: Algorithmic Trading & DMA: An Introduction to Direct Access Trading Strategies, vol. 1. Myeloma Press, London (2010)
9. Kaufman, P.J.: Trading Systems and Methods, + Website, vol. 591. John Wiley & Sons, Hoboken (2013)
10. Pardo, R.: The Evaluation and Optimization of Trading Strategies. John Wiley & Sons, Hoboken (2011)

11. Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., Vega, C.: Rise of the machines: algorithmic trading in the foreign exchange market. J. Finan. **69**(5), 2045–2084 (2014)
12. Narang, R.K.: Inside the Black Box: The Simple Truth about Quantitative Trading. -Description Based on Print Version Record. J. Wiley et Sons, Hoboken (2009)
13. Pesaran, M.H.: Predictability of asset returns and the efficient market hypothesis (2010)
14. Kissell, R., Malamut, R.: Algorithm decision making framework. J. Trading **1**(1), 10 (2006)

A Real-Time Assessment Method Based on the Detection of Human Facial Emotions

Bhupinder Singh(✉), Divyansh Tak, and Swapnil Verma

Lovely Professional University, Phagwara, Punjab 144411, India
bhupinder.28636@lpu.co.in, bhupinder0741@gmail.com

Abstract. The abstract summarizes a research paper on Facial Emotion Recognition (FER), which aims to develop an effective system for recognizing emotions from facial images and videos. The study collected a diverse dataset of facial images and videos, pre-processing it using various techniques so that the FER model could be trained on consistent data.

The authors used a Convolutional Neural Network (CNN) to perform foreign language acoustic recognition. They optimized its performance through hyperparameter tuning, and evaluated the FER model's accuracy, precision, recall and F1-score on a separate test set.

The real-time FER system was tested in the application and enabled users to receive emotion feedback from their peers during a training session. The abstract provides a summary of a study that creates a system for identifying facial expressions from a variety of photos and videos. For optimization, the article uses a convolutional neural network and hyperparameter tweaking. The abstract provides a summary of a study that creates a system for identifying facial expressions from a variety of photos and videos. For optimization, the article uses a convolutional neural network and hyperparameter tweaking. This study investigates the most recent developments in emotion recognition systems, exploring the methodology used, the datasets used, and the difficulties encountered. The report also looks at how research should be improved to provide better results. Before concluding, the study's findings will be thoroughly examined and contrasted with those of recent publications.

Keywords: SVM Classifier · Deep Intuitive Learning · CNN recognition · Emotion Detection · Image Processing · FER Model

1 Introduction

Due to recent developments in deep learning-based models and the growing accessibility of face image and video data, Facial Emotion Recognition (FER) has emerged as an important study area in computer vision. Due to its useful uses in a number of industries, such as marketing, healthcare, and human-computer interface, it has attracted a lot of interest. This research paper's major objective is to present a study on the creation of a Facial Emotion Recognition (FER) system that is accurate and effective in swiftly identifying emotions from facial photos and videos. In order to accomplish this, a

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 625–633, 2023.
https://doi.org/10.1007/978-3-031-39777-6_73

deep learning model—more specifically, a Convolutional Neural Network (CNN)—was trained using a large dataset of videos and photos of faces displaying a range of emotions. The accuracy, precision, recall, and F1-score measures were used to assess the efficacy of the FER model and to enhance its performance. The following is the structure of the research paper. Initially, we give a general summary of the history and relevant research on FER. The methods employed in the study's dataset collecting and preprocessing are then described. The CNN-based FER model is then shown, and the hyperparameter tuning procedure is described. The findings and analysis of the performance of the FER model, including its accuracy, clarity, recall, and F1-score, are then covered. Lastly, we wrap up the investigation and go over some potential real-world uses for the suggested FER system. Overall, by outlining a thorough strategy for creating a precise and effective FER system, this research study adds to the body of knowledge on FER. The results highlight the significance of preprocessing and hyperparameter adjustment in achieving high performance and show the efficacy of deep learning-based models for FER. Applications for the proposed FER system include marketing, human-computer interaction, and the monitoring of mental health.

Objective
The inquiry into the efficacy of merging a CNN algorithm with Tensor Flow for face emotion identification is what distinguishes this research article from others. The report also addresses implementation- related problems and suggests alternative remedies to enhance the performance of the technology. This work extends face emotion identification technology and its potential applications in several fields by filling up these gaps in the existing literature.

Originality
This research paper's introduction highlights how it differs from earlier studies by concentrating on merging a CNN algorithm and Tensor Flow for facial emotion recognition. This work fills gaps in the existing literature by addressing implementation issues and providing alternate methods to enhance performance. It also advances face emotion recognition technology and its potential applications across numerous industries.

2 Literature Survey

According to the study [1], there are five primary emotions: happiness, sadness, fear, anger, and neutrality. These emotions were determined by watching body movements, including head and joint movements, upper and lower body and arm movements, notwithstanding the possibility for additional accuracy improvements in emotion recognition systems. Using video recordings taken while moving, sitting, and in situations where there was no action, the researchers were able to extract motion or kinetic parameters from various body parts, such as speed, space, and symmetry. ANOVA and MANOVA were used to evaluate the applicability of the features acquired and normalize them. In order to merge functions, score and grade level fusion methods were also used.

2.1 Problem Formulation

With applications in many disciplines, including psychology, marketing, and human-computer interaction, Facial Emotion Recognition (FER) is a significant research challenge in computer vision. The main objective of FER is to automatically identify the emotion that a person is showing in a certain facial image or video sequence. This entails recognizing the face and facial features, gathering pertinent data, and categorizing the emotion.

Another significant obstacle in FER is the dearth of labeled datasets. Large and varied datasets of facial expressions and emotions are needed for the development of precise FER systems. It can be costly and time-consuming to gather and annotate such datasets, though. Moreover, inconsistent annotation quality and consistency may result in inconsistent data. Hence, selecting and preparing the available datasets with care is necessary for creating effective FER systems.

Another problem with FER is real-time processing. FER systems must operate in real-time in applications like marketing and human-computer interaction in order to deliver timely and pertinent feedback. Nonetheless, FER processing time can be substantial, particularly for deep learning-based methods. Consequently, it is crucial for the practical implementation of FER to create effective algorithms that can handle real-time processing [14].

In short, the goal of FER is to accurately categorize a person's emotion in a given face image or video sequence. The complexity and variety of facial expressions, the dearth of labeled datasets, and the requirement for real-time processing make this problem difficult. Robust algorithms that can manage these difficulties and function in a range of applications are needed to create efficient FER systems [4].

2.2 Discrepancies

There isn't much cross-cultural research on emotion recognition algorithms because they are often trained on data from just one culture or language. Yet, the effectiveness of emotion recognition systems might be impacted by cultural variations in how emotions are presented and understood. Future research might focus on creating emotion recognition systems that can recognize emotions despite cultural variances. Moreover, little attention has been paid to the effects of data imbalances. Several emotion recognition algorithms are trained on unbalanced datasets, where some emotions are represented by more samples than others. Due to data imbalances, emotion identification algorithms may make mistakes when detecting underrepresented emotions. So, additional study may be necessary to look into this problem and offer solutions. Also, there hasn't been much research on multimodal emotion identification, which recognizes emotions expressed across a variety of modalities, including speech, body language, and facial expressions [10].

Future research may therefore concentrate on creating emotion detection systems that can recognize emotions through a variety of modalities. Furthermore, since the majority of emotion recognition algorithms are regarded as "black boxes," it is challenging to comprehend how they produce predictions. Consequently, additional study is required to enhance these systems' interpretability. There may be a need for greater research into how to make such systems more visible and interpretable so that consumers can understand how emotion detection algorithms arrive at their predictions. Few studies have been done on the generalizability of emotion detection systems. Many emotion recognition algorithms have been developed and assessed using controlled datasets, which might not correctly represent the variety of emotions that exist in real life. So, a research gap can be filled by investigating the generalizability of emotion recognition systems to real-world situations and diverse populations [11].

Carlos Busso, Zhigang Deng [1] used three different approaches to distinguish between four emotions (sadness, happiness, anger, and neutral) based on audio, facial expression, and bimodal information. The goal was to compare the performance of unimodal systems, evaluate the pros and cons of these approaches, and compare various fusion methods to improve the system's overall recognition rate. The database used in the trials consisted of an actress reading 258 phrases with her facial motion data recorded using a VICON motion capture system with three cameras, sampling at a rate of 120Hz and concluded that the effectiveness and limitations of facial expression and aural emotion classifiers were assessed in the study. For most misclassifications, the bimodal emotion classifier performed better than the unimodal systems [13]. Although the two fusion techniques performed similarly, there were notable differences in the recognition of particular emotions [12]. The study showed that using audio and visual modalities, it is possible to precisely distinguish human affective states, which may one day enhance the usability and engagement of human-computer interfaces [9].

Paweł Tarnowski and Marcin Kołodziej [2] involved five distinct age groups: 26 children aged between 5 and 6 (12 girls and 14 boys, meanD Age range: 5 years, 9 months 24 7 and 8-year-olds (10 girls and 14 boys, meanD; ages ranged from 5 months to 6 years and 2 months) Age range: 7 years, 10 months 7 months to 8 years and 4 months), 24 9 and 10 year olds (meanD: 13 girls, 11 males Age range: 9 years, 9 months 26 children between the ages of 11 and 12 (meanD: 17 females and 9 boys; range: 9 months to 10 years 10 months) Age range: 11 years, 10 months 26 adults (14 women and 12 men, meanD; ages range: 11 months to 12 months) Age range: 22 years, 5 months 18 months and as much as 29 years and five months and concluded that Age was used as a between-subject factor, and emotion and orientation were used as within-subject variables, respectively [5].

Krumhuber et al. [3] conducted a systematic search for pertinent research, evaluated the articles in accordance with predetermined criteria, retrieved pertinent data from the studies that were included, and then synthesized the results to detect recurring themes and patterns. The study offered a thorough analysis of how dynamic features of facial expressions affect the ability to recognize emotions and concluded that dynamic facial expressions have a major impact on how well emotions are detected, with dynamic expressions being more accurate than static ones. Depending on the emotion being communicated, dynamic variables have different consequences, and individual differences like age, gender, and culture can also have an impact. The authors stressed the need of taking these aspects into account while creating automated emotion recognition systems and comprehending how people perceive and express emotions in natural contexts [15].

3 Research Methodology

A computer vision and artificial intelligence (AI) tool called Facial Emotion Recognition (FER) tries to identify and categorize human emotions based on facial expressions. To guarantee the system's correctness, reliability, and validity, the development of a FER system requires a clearly defined research approach [17]. This is an illustration of a FER project methodology: Defining the research design that will be used for the project is the first stage [16]. Figure 1 describes the process for facial recognition procedure Choosing the study's design, such as experimental or observational, and its sample method, such as random or purposeful sampling, are all parts of this process.

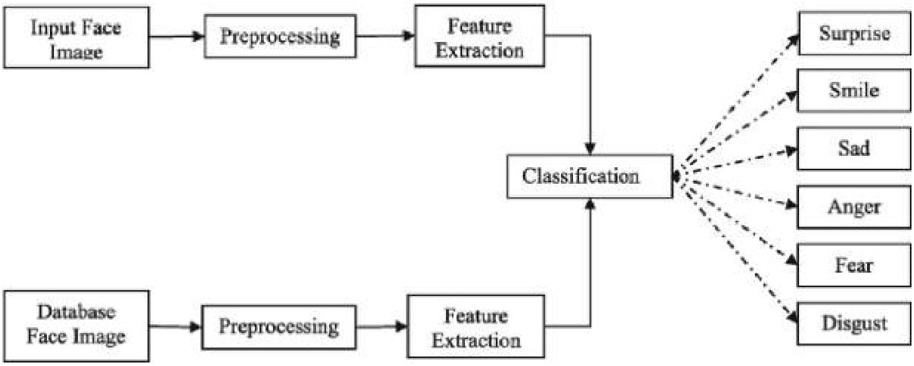


Fig. 1. Describes the process for facial recognition procedure

Data Collection: Acquiring and preparing the dataset are steps in the data collection process. The dataset in this instance ought to be made up of emoticon-labeled pictures of faces [8]. The dataset ought to be varied and representative of the intended audience [7]. The amount of the dataset can also be increased by using data augmentation methods like flipping and rotating the photographs.

Feature Extraction: The necessary features are derived from the facial photos in this particular step. The photos then can be easily used to extract features using methods like Principal Component Analysis (PCA), Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG).

Model Selection and Training: Based on the retrieved features, a suitable machine learning model is chosen. Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Random Forests are a few examples of models that can be applied. The preprocessed dataset is then used to train the chosen model [5].

Model Evaluation: A variety of metrics are used to evaluate the trained model, including accuracy, precision, recall, and F1 score. A different test dataset that wasn't used during the training phase is used to evaluate the model. The model's performance is contrasted with those of other models that have been published in the past [18].

Results Interpretation: To make judgements on the effectiveness of the FER system, the findings are examined and analyzed. Also, the system's shortcomings are highlighted, and suggestions for further investigation are given [19].

Implementation: After that, the FER system is put into practice in a real-world situation. In order to do this, the system must be integrated with additional hardware or software, such as mobile applications or security systems. As a result, creating a FER system necessitates a clear research approach that covers data gathering, feature extraction, model choice, training, evaluation, and result interpretation [4]. A methodical approach to FER can increase the system's precision, dependability, and validity, making it applicable to a variety of real-world applications. Using methods like CNNs or other deep learning models, which can learn features directly from the photos, this can be accomplished. Creating a hybrid supervised learning model by fusing various models and methods will increase the precision of cancer detection. Figure 2 represents Block Diagram represents Research Methodology [6].

4 Discussion

In response to the increased demand for robust and trustworthy automatic identification systems for artificial intelligence and Internet of Things applications, recent research publications have concentrated on emotion recognition and sentiment analysis. These results imply that in addition to facial expressions and textual information, bodily movements must also be considered in order to increase the accuracy of emotion identification systems. Naturalist databases have been found to be helpful for identifying compound and subtle emotions, including CIU handwriting database, AMFED+, and iCV-MEFED, among others [20]. In comparison to other techniques, Convolutional Neural Networks (CNNs) have also been proven to have higher accuracy for feature extraction and classification. It has been demonstrated that the Multimodal Automated Emotion Recognition System has application potential in a number of fields, including healthcare and military operations in far-flung locales. Figure 2 Shows different human emotions as a result, it is essential to incorporate emotion recognition systems with modern technology, especially for end users like teenagers who live in a tech-savvy society [21]. Research in emotion recognition analysis and sentiment analysis has been sparked by the numerous uses of emotion recognition systems [22].

5 Conclusion

The current study reviews new findings in sentiment analysis, machine learning techniques, emotion detection systems, and their applications. The findings show that self-learning CNN algorithms are effective for minimizing overfitting and imbalance in naturalistic databases. CNN identifies a number of applications in the realms of robotics, virtual reality, and healthcare.

The system's security can also be increased by leveraging cloud storage and cancelable biometrics. The essay emphasizes the significance of creating trustworthy and dependable automatic identification systems to suit the expanding demand for AI and

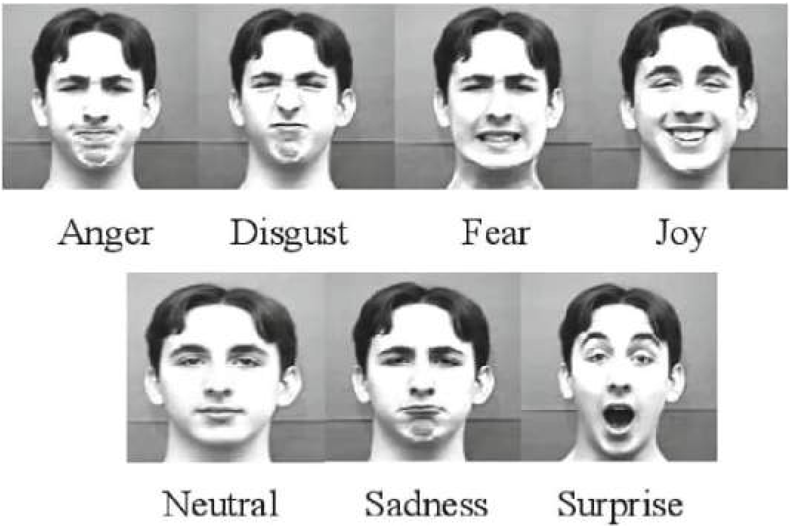


Fig. 2. Shows different human emotions

IoT applications. As a result, in order for emotion detection systems to be effective, bodily motions must be taken into account in addition to face expressions and textual data.

CNN has been proven to be more accurate in feature extraction and classification than other techniques in the research publications that examined naturalistic databases that can distinguish compound and nuanced emotions. It is crucial to heavily integrate emotion recognition systems with current technology to serve the general public, especially teenagers who play a crucial role in the tech-savvy world. Emotion recognition systems have numerous applications in various fields, including healthcare and military operations in remote locations.

6 Future Scope of Work

Future study should leverage RGB datasets created under uncontrolled conditions, deep neural networks as emotion classifiers, and multimodal behavioral systems that encompass body movements, facial expressions, and voice to improve the accuracy of emotion identification systems. According to the report, there is still room for advancement in how the technology is used. The training dataset may be increased, the model architecture can be improved, and various deep learning methods can be investigated in future study. Further research can be done on the potential uses of this technology in human-robot interaction and mental health diagnosis. Overall, this research lays the groundwork for further research to improve the precision and efficiency of face emotion identification systems.

References

1. Busso, C., Deng, Z., Yildirim, S.: Analysis of emotion recognition using facial expressions, speech and multimodal information (2007)
2. Tarnowski, P., Kołodziej, M., Majkowski, A., Rak, R.J.: Emotion recognition using facial expressions. J. Phys: Conf. Ser. **896**(1), 012142 (2017)
3. Krumhuber, E.G., Kappas, A., Manstead, A.S.: Effects of dynamic aspects of facial expressions: a review. Emot. Rev. **5**(1), 41–46 (2013)
4. Hulliyah, K., Bakar, N.S.A.A., Ismail, A.R.: Emotion recognition and brain mapping for sentiment analysis: a review. In: 2017 Second International Conference on Informatics and Computing (ICIC), Jayapura, Indonesia (2017). https://doi.org/10.1109/IAC.2017.8280568
5. Yeh, S.C., Hwang, W.Y.: A study for the application of body sensing in assisted rehabilitation training. In: 2012 International Symposium on Computer, Consumer and Control, p. 92 (2012). https://doi.org/10.1109/IS3C.2012.240
6. Gong, C., Lin, F., Zhou, X., Lü, X.: Amygdala-inspired affective computing: To realize personalize intracranial emotions with accurately observed external emotions. China Commun. **16**, 115 (2019). https://doi.org/10.23919/JCC.2019.08.011
7. Wei, W., Jia, Q., Feng, Y., Chen, G., Chu, M.: Multi-modal facial expression feature based on deep-neural networks. J. Multimodal User Interf. **14**(1), 17–23 (2019). https://doi.org/10.1007/s12193-019-00308-9
8. Hossain, M.S., Muhammad, G.: Emotion recognition using deep learning approach from audio-visual emotional big data. Inf. Fusion **49**, 69–78 (2018). https://doi.org/10.1016/j.inffus.2018.09.008
9. Rahman, S.A., AlOtaibi, F.A., AlShehri, W.A.: Sentiment analysis of twitter data. In: 2019 International Conference on Computer and Information Sciences (ICCIS), Saudi Arabia (2019). https://doi.org/10.1109/ICCISci.2019.8716464
10. Deng, D., Jing, L., Yu, J., Sun, S., Ng, M.K.: Sentiment lexicon construction with hierarchical supervision topic model. IEEE/ACM Trans. Audio Speech Lang. Process. **27**, 704–718 (2019). https://doi.org/10.1109/TASLP.2019.2892232
11. Hong, S., Kim, Y.: Dynamic pose estimation using multiple RGB-D cameras. Sensors **18**, 3865 (2018). https://doi.org/10.3390/s18113865
12. Vryzas, N., Vrysis, L., Kotsakis, R., Dimoulas, C.: Speech emotion recognition adapted to multimodal semantic repositories. In: 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), pp. 31–35 (2018). https://doi.org/10.1109/SMAP.2018.8501881
13. Torres, A.D., Yan, H., Aboutalebi, A.H., Das, A., Duan, L., Rad, P.: Patient facial emotion recognition and sentiment analysis using secure cloud with hardware acceleration. In: Computational Intelligence for Multimedia (2018)
14. Vij, A., Pruthi, J.: An automated psychometric analyzer based on sentiment analysis and emotion recognition for healthcare. Procedia Comput. Sci. **132**, 1184−1191 (2018). https://doi.org/10.1016/j.procs.2018.05.033
15. Nandakumar, K., Jain, A.K.: Soft biometrics. In: Encyclopedia of Biometrics, pp. 1425–1429 (2009). https://doi.org/10.1007/9784899-7488-4
16. Girard, J.M., Cohn, J.F., Rosenwald, P.: Nonverbal social withdrawal in depression: evidence from manual and automatic analysis. Image Vis Comput. **32**(10), 641–647 (2014). https://doi.org/10.1016/j.imavis.2013.12.007
17. Gong, C., Lin, F., Zhou, X., Lü, X.: Amygdala-inspired affective computing: To realize personalize intracranial emotions with accurately observed external emotions. China Commun **16**, 115 (2019). https://doi.org/10.23919/JCC.2019.08.011

18. Faiz, M., Daniel, A.K.: A multi-criteria cloud selection model based on fuzzy logic technique for QoS. Int. J. Syst. Assur. Eng. Manag. (2022). https://doi.org/10.1007/s13198-022-01723-0
19. Faiz, M., Daniel, A.K.: FCSM: fuzzy cloud selection model using QoS parameters. In: 2021 First International Conference on Advances in Computing and Future Communication Technologies (ICACFCT), Meerut, India, pp. 42–47 (2021). https://doi.org/10.1109/ICACFCT53978.2021.9837347
20. Singh, B., Henge, S.K.: Prediction of stock price movements using statistical and hybrid regression techniques to reduce diversify risk. In: Kahraman, C., Tolga, A.C., Cevik Onar, S., Cebi, S., Oztaysi, B., Sari, I.U. (eds.) Intelligent and Fuzzy Systems. INFUS 2022. Lecture Notes in Networks and Systems, vol. 505, pp. 456–462. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09176-6_52
21. Singh, B., Henge, S.K.: The computation trend of fuzzy rules for effective decision support mechanism on basis of supervised learning for multiple periods. In: Kahraman, C., Tolga, A.C., Cevik Onar, S., Cebi, S., Oztaysi, B., Sari, I.U. (eds.) Intelligent and Fuzzy Systems. INFUS 2022. Lecture Notes in Networks and Systems, vol. 505, pp. 117–123. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09176-6_14
22. Singh, B., et al.: ML-based interconnected affecting factors with supporting matrices for assessment of risk in stock market. Wirel. Commun. Mob. Comput. (2022). https://doi.org/10.1155/2022/2432839

Data Driven Positioning Analysis of Music Streaming Platforms

Ayse Basak Incekas(✉) and Umut Asan

Istanbul Technical University, Istanbul, Turkey
{incekas15,asanu}@itu.edu.tr

Abstract. This study investigates the market position of music streaming platforms by analyzing user sentiment and topics expressed in customer reviews. In contrast to traditional methods, this study employs machine learning techniques to extract less biased and more authentic comments from user review data. Sentiment and topic analysis are utilized to identify the emotional tone of the language used and distinct topics discussed within customer reviews of the four most popular music streaming platforms, namely Spotify, Amazon Music, Apple Music, and YouTube Music. The study comprises four main steps, including data collection, cleaning and pre-processing, sentiment analysis, and topic modeling. The results reveal that Amazon Music is prominent in functionality aspects, while Spotify ranks highest across all topics. Apple and YouTube Music have the highest scores in reviews related to customization. The proposed approach provides valuable insights into user perceptions and preferences, which can assist brands in improving their market position. The paper concludes with a summary of the findings, marketing implications, and suggestions for future research.

Keywords: BERTopic · Sentiment analysis · Customer reviews · Music streaming platforms

1 Introduction

People's interactions with one another have been transformed by social media and other online communication tools. Music streaming platforms have grown in popularity over time, with users spending more than 1 h and 37 min per day on these platforms in the second quarter of 2022 [8]. The rise of music streaming platforms has transformed the way people consume music, making it more accessible than ever before. As a result, the music streaming industry has become increasingly competitive, with major players such as Spotify, Apple Music, Amazon Music and Youtube Music vying for market dominance. In such a competitive market, brand positioning is crucial for determining customers' perceptions and preferences. Only a limited number of studies have explored the current state of music streaming platforms despite their increasing popularity [9].

Traditional methods of assessing brand positioning involve asking customers questions, which often lead to biased responses and limited insight. In this study,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 634–641, 2023.
https://doi.org/10.1007/978-3-031-39777-6_74

an alternative approach utilizing Natural Language Processing (NLP) techniques is suggested to analyze user reviews of music streaming platforms and determine the relative position of each brand in the market. Specifically, applying state-of-the-art NLP algorithms provide a better understanding of the sentiment and underlying topics in the reviews. Extracting meaning and emotions from large volumes of text data, allows a more efficient and cost-effective way of assessing brand positioning, which is less biased and faster than traditional methods.

Past studies have suggested employing diverse methods ((e.g. Latent Dirichlet Allocation [3], Contextual Valence Shifters [6]) for detecting and classifying sentiments in reviews [7], to understand emotions and recognize the individuals expressing them.

This study contributes to the field by providing a detailed understanding of the emotional aspects of customer reviews on music streaming platforms. It distinguishes itself from prior research by constructing separate topic models for positive, neutral, and negative reviews. Overall, the work provides better insight into customer perceptions and can guide brand positioning decisions for music streaming platforms.

The rest of this paper is organised as follows. Section 2 outlines the proposed methodology. Section 3 explains the application, which involves an overview of the music streaming platforms, step by step construction of the sentiment and BERTopic model, analysis of the results, and an interpretation of the identified topics and their significance. In Sect. 4, a summary of the findings and their marketing implications are provided. Finally, the limitations of the study and future research directions are discussed.

2 Methodology

As mentioned above, the purpose of this study is to analyse social media data to gain insight into user opinions towards top music streaming platforms using sentiment analysis and BERTopic. To achieve this, a methodology consisting of four steps is performed. First, data is obtained from various social media platforms. Second, the data is cleaned and pre-processed to remove irrelevant information. Third, sentiment analysis is performed to understand the emotions expressed in the data. Finally, topic modelling is conducted using the BERTopic algorithm to identify key themes and topics discussed in the data. The details of these steps are given below.

Step 1: Data can be collected from social media platforms such as Twitter and/or from online user platforms such as App Store. The decision for the relevant platform(s) should be based on the quality and volume of the data.

Step 2: The data is cleaned and pre-processed by removing irrelevant or duplicated data, correcting spelling and grammatical errors, converting all text to lowercase, removing hashtags, hyperlinks, and Twitter handles, tokenizing the text into individual words, removing common stop words and special characters, removing numbers, and removing any extra white space to ensure the accuracy and reliability of the analysis. The pre-processed data is then transformed into a structured format suitable for text mining analysis.

Step 3: Sentiment analysis is performed using machine learning techniques, specifically a supervised learning algorithm. The pre-processed data is labeled with the sentiment of the sentence such as positive, neutral and negative.

Step 4: Topic modelling is applied for positive, neutral, and negative reviews/texts separately. Topic modelling is a method for learning the basic topics (in terms of word distributions) for a set of documents as well as the commonalities of each document to these topics [10]. In the literature, there are several approaches to topic modelling such as Latent Semantic Analysis (LSA) [3], Probabilistic Latent Analysis (PLSA) [5], and Latent Dirichlet Allocation (LDA) [2]. Labelling is not required in these approaches but the number of categories must be determined beforehand. A more recent approach, BERTopic, offers several advantages over traditional topic modelling techniques, including, no need for pre-processing, ease of use, high consistency and stability, even when changing the number of topics. It is an unsupervised statistical text mining technique used to find underlying topics for a set of documents [10]. It is specifically designed to work well with short texts, such as social media posts or reviews, as it uses a pre-trained language model that is trained on large amounts of data, including short texts. This enables BERTopic to generate high-quality topic representations for short texts, which can be challenging for traditional topic modeling techniques that are designed for longer documents [1].

BERTopic works in three steps. First, it converts each document to a numerical representation (embedding) using a pre-trained language model. Second, it groups similar documents together using UMAP to reduce the dimensionality of the embeddings before clustering (HDBSCAN) them. Finally, it extracts topics from the clusters using a modified version of TF-IDF called c-TF-IDF, which models the importance of words in sets of documents rather than individual documents. The method repeatedly merges the representations of the least common topic with the most similar one, until the number of topics reaches a user-specified value. BERTopic obtains keyword-based descriptions for each topic, allowing us to infer meaning from the groupings produced by unsupervised clustering [4].

Step 5: The resulting tables can be analyzed using various quantitative or qualitative methods, such as perceptual mapping or thematic analysis, to gain insights into the topics of interest across different platforms.

3 Application

This section demonstrates how the suggested methodology was employed to analyze review data about competing music streaming platforms.

Many different music streaming platforms are available to users in order to meet the rapidly increasing demand for listeners. Spotify has the largest market share with 32% of the market, followed by Apple Music with 16% and Amazon and Tencent Music with 13% [11]. The data used in this study is obtained from English-written user reviews on the App Store. The study includes data from the four leading music streaming platforms, excluding Tencent Music, since this platform primarily provides music streaming services to users in China, which means that the available English data on Tencent will be restricted. App Store provides valuable feedback on apps to both researchers and potential users. User reviews help understand how users interact with the app, what features they like or dislike, and what issues they might be experiencing. User reviews can help potential users make informed decisions about whether or not to download a specific app. Overall, user reviews are an important part of the App Store ecosystem because they allow users to provide feedback and researchers to understand the user experience.

Step 1: The data is collected from the App Store using the app_store_scraper package in Python. 500 reviews were collected randomly for each platform, based on reviews from the United States. The filtered data includes the user's review, review date, and the platform name. The time-span of the data covers the reviews from 2017-09-25 to 2023-04-06. The data from the different platforms was combined to create a single table with 2000 rows.

Step 2: User review data in text format is pre-processed and cleaned using the steps explained in the methodology section.

Step 3: The sentiment analysis is performed using pysentimiento library on the pre-processed data. The results provided a probability distribution of positive, negative, and neutral sentiment. For example,

SentimentOutput (output = POS, probas = POS: 0.994, NEU: 0.005, NEG: 0.000)

Out of 2000 reviews, 672 reviews are classified as positive, 759 as neutral, and 569 as negative. According to the sentiment of the sentence, the data is divided into three data frames as follows: df_POS, df_NEU, and df_NEG.

Step 4: Topic modeling was performed on the three data frames using the BERTopic algorithm ('bertopic' package in Python). Since the topics of each sentiment were expected to be different, three distinct BERTopic models were constructed - one for each data frame, representing each sentiment. The resulting models contained 10 topics for positive reviews, 10 topics for neutral reviews, and 2 topics for negative reviews. To achieve parsimony in the analysis, only the top 5 topics with the highest frequency were considered for further analysis. Table 1 displays the most significant topic words and their scores for the top topics for each sentiment, accompanied by a brief description of each topic. The table presents that the topics identified in the positive reviews cover various aspects of the music application, such as its features, functionality, playlists, customization, and emotional associations with music. On the other hand, the neutral reviews include opinions regarding music consumption, app-related challenges, music subscription services, advertising, monetization, and offline use. Lastly, the negative reviews discuss general music activities and technical difficulties.

Step 5: A new table is generated for each data frame, containing the review sentence, the platform name, and columns representing the topics. To populate the topic columns, the weights of each sentence are aggregated by summing up the weight of the words in the sentence that are associated with the topic.

Table 1. The first 10 words for each topic.

Topic id	Words	Frequency	Topic description
Positive_0	app', 'music', 'like', 'great', 'songs', 'free', 'listen', 'get', 'good', 'ads'	125	App features
Positive_1	'song', 'app', 'play', 'music', 'work', 'listen', 'sometimes', 'phone', 'fix', 'songs'	49	Functionality
Positive_2	'playlist', 'would', 'add', 'songs', 'song', 'playlists', 'love', 'could', 'app', 'like'	49	Playlists
Positive_3	'songs', 'song', 'love', 'app', 'playlist', 'like', 'liked', 'change', 'music', 'add'	23	Customization
Positive_4	'emoji', 'face', 'with', 'smiling', 'music', 'app', 'musical', 'songs', 'really', 'even'	20	Emotions related to music
Neutral_0	'artist', 'songs', 'album', 'artists', 'music', 'like', 'albums', 'song', 'playlist', 'library'	99	Music consumption
Neutral_1	'app', 'music', 'song', 'play', 'playing', 'problem', 'back', 'go', 'phone', 'pause'	60	App problems
Neutral_2	'prime', 'unlimited', 'music', 'pay', 'service', 'app', 'use', 'like', 'one', 'subscription'	51	Music subscription services
Neutral_3	'ads', 'ad', 'premium', 'free', 'listen', 'get', 'app', 'songs', 'minutes', 'music'	38	Advertising and monetization
Neutral_4	'playlist', 'playlists', 'offline', 'songs', 'song', 'one', 'downloaded', 'app', 'music', 'play'	31	Offline use
Negative_0	'music', 'songs', 'app', 'song', 'like', 'playlist', 'listen', 'want', 'play', 'get'	276	General music related activities
Negative_1	'app', 'music', 'song', 'play', 'songs', 'playlist', 'playing', 'downloaded', 'like', 'fix'	200	Technical issues

Subsequently, the data frame is grouped by platform name and topic. The median value is computed for each topic in every platform.

Creating the BERTopic model for each data frame, Tables 2, 3 and 4 were obtained for positive, neutral and negative reviews, respectively.

For the positive reviews, the results indicate that Amazon Music is highly associated with functionality aspects (positive_1) with a score of 0.48. Following this, reviews related to customization (positive_3) have a high score of 0.43. Both Apple Music and Youtube Music have the highest scores in reviews related to app customization (positive_3). Finally, Spotify has the highest score in all topics, with the highest score of 0.6 in the reviews related to app customization (positive_3).

Table 2. Results of the positive reviews.

Platform	Positive_0	Positive_1	Positive_2	Positive_3	Positive_4
Amazon	0.30	0.48	0.31	0.43	0.27
Apple	0.23	0.26	0.33	0.37	0.18
Spotify	0.41	0.53	0.58	0.60	0.30
YouTube	0.23	0.27	0.23	0.31	0.19

Table 3. Results of the neutral reviews.

Platform	Neutral_0	Neutral_1	Neutral_2	Neutral_3	Neutral_4
Amazon	0.24	0.37	0.19	0.18	0.31
Apple	0.17	0.15	0.13	0.14	0.18
Spotify	0.31	0.42	0.18	0.36	0.43
YouTube	0.21	0.26	0.12	0.12	0.21

Table 4. Results of the negative reviews.

Platform	Negative_0	Negative_1
Amazon	0.17	0.18
Apple	0.32	0.30
Spotify	1.29	1.09
YouTube	0.18	0.14

The results of the analysis on neutral reviews indicate a strong association between Spotify and music consumption (neutral_0) and app problems (neutral_1), followed by Amazon Music. Among music subscription services (neutral_2), Amazon Music scored the highest with a score of 0.19, followed by Spotify with a score of 0.18. In terms of advertising and monetization (neutral_3) and offline use (neutral_4), Spotify took the lead with a score of 0.36 and 0.43, respectively.

The results from the negative reviews indicate that Spotify is most closely associated with general music-related activities (negative_0), followed by Apple, Youtube, and Amazon Music, respectively. Technical problems (negative_1) are strongly associated with Spotify, while Apple, Amazon, and YouTube Music show comparatively lower levels of such issues.

4 Conclusion

In summary, this study addresses the problem of obtaining an accurate understanding of customer perceptions and brand positioning in the music streaming platform market. The traditional approach of asking customers questions has been shown to be often biased, time-consuming, and costly. To overcome these limitations, an alternative approach that utilizes customer reviews in App Store and NLP techniques is proposed.

The study found that the sentiment expressed in user reviews varied across platforms, with neutral sentiments dominating the reviews, as well as that the topics discussed also differed across platforms. The application of BERT-based algorithms provided a more accurate understanding of the sentiment and topics expressed in the user reviews, which distinguishes this study from previous research in the field.

The benefits of the proposed methodology include the ability to obtain unbiased and rich data in a short period of time. The findings highlight important managerial implications, including the potential to save money and time, provide a more unbiased evaluation, and enable continuous monitoring of customer perceptions and brand positioning in the highly competitive music streaming platform market. The brands can use this information to improve their customer satisfaction and retention by addressing common concerns, identifying areas for improvement, and tailoring their products to meet customer needs.

Although the study has limitations, such as the use of a single data source and the focus on English language reviews, future research could expand the analysis to include other languages, demographics, and user ratings to further enhance the understanding of user perceptions and experiences. Overall, this study contributes valuable insights into data-driven positioning analysis and highlights the importance of brand positioning for music streaming platforms in establishing a unique and favorable perception among their target customers.

References

1. Albanese, N.C.: Topic Modeling with LSA, PLSA, LDA, NMF, BERTopic & Top2Vec: a comparison. Towards Data Sci. (2021). https://towardsdatascience.com/topic-modeling-with-lsa-plsa-lda-nmf-bertopic-top2vec-a-comparison-5e6ce4b1e4a5
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
3. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by latent semantic analysis. J. Am. Soc. Inform. Sci. Res. **41**(6), 391–407 (1990). https://doi.org/10.1002/(SICI)1097-4571(199009)41:6⟨391::AID-ASI1⟩3.0.CO;2-9
4. Grootendorst, M.: BERTopic: Neural topic modeling with a class-based TF-IDF procedure (2022)
5. Hofmann, T.: Probabilistic latent semantic analysis. In: Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI1999), pp. 289–296 (1999). https://doi.org/10.48550/arXiv.1301.6705
6. Kennedy, A., Inkpen, D.: Sentiment classification of movie reviews using contextual valence shifters. Comput. Intell. **22**(2), 110–125 (2006)
7. Kim, S.M., Hovy, E.: Determining the sentiment of opinions. In: COLING 2004: Proceedings of the 20th international conference on Computational Linguistics, p. 1367. Association for Computational Linguistics (2004)
8. Marketing Charts. Gen Z Streams Music for Almost 2 Hours A Day. Marketing Charts (20 December 2022). https://www.marketingcharts.com/digital/audio-and-radio-119052

9. Mena, A., Heidemann, J.: An empirical study of real audio traffic. IEEE INFOCOM (2000). https://doi.org/10.1109/INFCOM.2000.832178
10. Nikolenko, S.I., Koltcov, S., Koltsova, O.: Topic modelling for qualitative studies. J. Inf. Sci. **43**(1), 88–102 (2017). https://doi.org/10.1177/0165551515617393
11. Statista Search Department, Music Streaming Worldwide - Statistics & Facts. Statista (2022). https://www.statista.com/topics/6408/music-streaming/ (Retrieved 23 April 2023)

Application of Extrapolation Methods for Predicting Traffic Flows Using Elements of Fuzzy Logic

Ramin Rzayev[1(✉)] and Emil Ahmadov[1,2]

[1] Institute of Control Systems, Vahabzadeh Street 68, AZ1141 Baku, Azerbaijan
raminrza@yahoo.com

[2] Intelligent Transport Management Centre of the Ministry of Internal Affairs of the Republic of Azerbaijan, Heydar Aliyev Avenue 137, Baku, Azerbaijan

Abstract. The article analyses the problem of practical modelling of traffic flows within the existing road infrastructure of an urban agglomeration. As a methodological basis for predictive analytics, well-studied extrapolation methods for constructing matrices of interborough movements and making forecasts based on them are considered. The features of the practical application of extrapolation methods by a single growth coefficient, average growth coefficients, Detroit and Fratara approaches are analyzed. As a result of the analysis of methods and models for constructing correspondence matrices in relation to the transport system of the urban agglomeration, the disadvantages and advantages of the above extrapolation methods are revealed. To predict the growth rate, which is a key indicator of the study, it is proposed to use extrapolation methods in terms of fuzzy logic. The main paradigm is that observations of interborough movements within the urban agglomeration are carried out at the level of "Soft Measurements". Therefore, the data of such observations should be interpreted as weakly structured, that is, for example, in the notation of fuzzy sets.

Keywords: Transport System · Traffic Flow · Correspondence Matrix · Extrapolation Method · Fuzzy Set

1 Introduction

Absolutely all megacities of the world permanently suffer from traffic queues, which negatively affects the business environment and the economy as a whole. One of the reasons for the formation of congestion is the lack of adequate transport infrastructure or when it works at the limit of its capabilities. However, in most cases, ineffective traffic management remains the main cause. Based on this premise, the importance and relevance of the tasks of strategic planning for the development of transport infrastructure and traffic management with the widespread use of adaptive control elements become obvious. To solve these problems, it is advisable to use traffic patterns that are able to provide management structures with useful information in the form of long-term forecasts for strategic planning of transport infrastructure development, as well as in the

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 642–650, 2023.
https://doi.org/10.1007/978-3-031-39777-6_75

form of short-term forecasts to support operational decision-making on traffic management in real time. These forecasts are built on the basis of quantitative characteristics of the state of traffic flows (speed and intensity of traffic in each section of the transport network, travel time, volumes of interborough movements, etc.), which are generated by appropriate mathematical models at the preliminary stage of the study.

The current level of development of information and computer technologies allows developers to quickly build adequate transport models for cities, regions and countries on the whole. At the moment, a huge number of approaches to building models of traffic flows are described in the specialized scientific literature. This thematic review considers predictive models that answer questions about future volumes of interborough movements (correspondence between individual borough) and their distribution (routing) along the transport network graph. In most cases, the results of calculations obtained using predictive models regarding future volumes of demand for movement are input characteristics of dynamic models of traffic flows.

2 Transport Demand and Correspondence Matrix

By analogy with the consumer market, in the theory of traffic modelling, the main concepts are also supply and demand. A transport offer is understood as a set of means of delivery that are available in a particular territory. Transport demand, which is of particular interest in the context of this review, is considered as an aggregate assessment of the need for the movement of people or goods along a specific transport network using various modes of transport – public, personal (including various types of individual mobility) and freight. In particular, the satisfaction of the transport demand for passenger transportation in cities is quantified by the so-called indicators of the transport mobility of the urban population.

Transport demand modeling is a paramount task when analyzing the current transport situation and forecasting the load on the transport network and/or its individual sections. So, to determine the volumes of transport flows and loading of elements of the transport network, various mathematical models are applied, the inputs of which are data on the number of people and goods moving around the city. This initial information, as a rule, is systematized in the form of the so-called matrix of interborough movements or, simply, correspondence matrix, and the dimension of the matrix is determined by the number of boroughs in the transport network. For the urban agglomeration consisting of n transport boroughs, the square correspondence matrix is generally presented in the form of Table 1, in which r_{ij} symbolizes the volume of movements (correspondences) from the i-th transport area to the j-th area, that is, the volume of the flow of people and goods from the i-th source-area to the j-th sink-area.

Based on the elements of the correspondence matrix, the load on individual sections of the transport network is determined. This allows, firstly, to identify the most intensive nodes and, as a result, redistribute traffic flows for their unloading, and, secondly, to evaluate the intensity of transport and passenger flows to optimize the mode of public transport. The final calculation of the load – the excess of freight and passenger traffic in a separate section of the transport network is carried out step by step using a model to calculate the volume of transport demand. The hierarchical structure of such model for the common case is shown in Fig. 1 and is described in detail in [2, 3]. In particular, in [1] it is noted that this model is widely used in the preparation of transport projects covering medium and long-term planning horizons.

Table 1. Transport correspondence matrix.

	1	2	…	j	…	n	$\sum_{j=1}^{n} r_{ij}$
1	$r11$	r_{12}	…	$r1j$	…	r_{1n}	
2	$r21$	r_{22}	…	$r2j$	…	r_{2n}	
…	…	…	…	…	…	…	
i	$ri1$	r_{i2}	…	rij	…	r_{in}	
…	…	…	…	…	…	…	
n	$rn1$	r_{n2}	…	rnj	…	r_{nn}	
$\sum_{i=1}^{n} r_{ij}$							

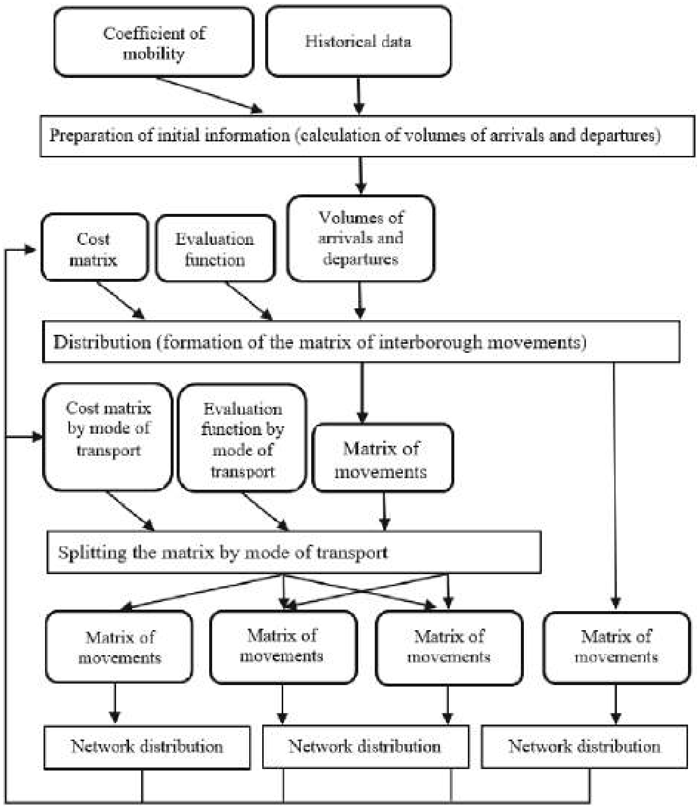


Fig. 1. Hierarchical structure of the model for calculating transport demand

At the top, tentatively speaking, at the first level of the model, based on statistical data and mobility coefficients that reflect the correlation between the socio-economic indicators of the given urban agglomeration and mobility, identified from surveys of respondents, common information is generated on the volumes of arrivals and departures for each transport area. Based on this information, at the next second level of the model, the correspondence matrix is formed in the form of Table 1. Further, at the third level, the correspondence matrix is "splitting" into separate matrices for various types of transport. As a result, at the last fourth level of the model, the desired solution is determined relative to the distribution of freight and passenger flows over the network, which makes it possible to calculate the load on each section of the transport network.

3 Classic Approaches to the Formation of the Correspondence Matrix

Predictive analytics is primarily a set of methods of statistics, data mining and game theory that are used to study regularities between the historical data of a particular indicator in order to predict its values in the future: short-term, medium- and long-term periods. For successful predictive analysis it is recommended to strictly adhere to the following stages: goal-setting, obtaining data from various sources, preparing data, creating a predictive model, evaluating the model, implementing the model, monitoring the effectiveness of the model. Without taking into account the last three points, the step-by-step procedure for modeling transport demand described above is fully integrated into the scheme for implementing predictive analytics.

The key link in the transport demand model is the correspondence matrix, the elements of which are the quantities of movements from source-areas to sink-areas. To determine the quantitative characteristics of movements, various predictive methods can be applied, using actual data on interborough movements and forecasts of their growth as initial information. The following is an overview of some of the extrapolation methods that are described in so many specialized scientific publications that it is impossible to identify the original sources.

3.1 The Method of the Single Coefficient of Transport Demand Growth

In a series of predicative approaches, this method is the most trivial. It is implemented through the establishment of the so-called growth coefficient, which is determined by the following ratio [1, 4]

$$C = F/A \tag{1}$$

where F is the predicted volume of interborough movements; A is its actual value.

In general, the growth coefficient is an economic multifactorial category, that is, it depends on changes in many factors, for example, the population in the urban agglomeration (A_p), the average per capita income (A_{inc}), the level of motorization (A_{ml}), etc. Then, taking into account the above factors, by analogy with (1), the growth coefficient can be established from the following ratio [4]:

$$C = \frac{f(F_p, F_{inc}, F_{ml})}{f(A_p, A_{inc}, A_{ml})} \tag{2}$$

where $f(\cdot)$ is a certain function that converts the values of socio-economic indicators into the value of the total volume of traffic flows within the given urban agglomeration. Obviously, in order to calculate the growth coefficient for the agglomeration under consideration using formula (2), it is necessary, firstly, to identify the function f: $R^3 \to R^1$ and, secondly, to use the apparatus of predictive analytics to analyze and predict the dynamics of changes in the indicated socio-economic indicators.

Thus, assuming the value of the growth coefficient to be known, it is possible to determine all elements of the predicted correspondence matrix by following equalities

$$r_{ij}^* = C \cdot r_{ij}^0 \tag{3}$$

where hereinafter in the text, '*' is the index of the predicted volume of movements from the *i*-th region to the *j*-th region of the agglomeration, that is, the predicted volume of the flow of people and goods from the *i*-th source-area to the *j*-th sink-area; '0' is the index of the current value of the corresponding amount of movement.

Building a predictable correspondence matrix using the method of a single growth coefficient according to formulas (1) and (3) is very rough operation, since this approach does not take into account the dynamics of development between individual socio-economic indicators of the urban agglomeration and has low reliability. However, in practice, this method may be required for approximate estimates of the volume of traffic flows in the short term when designing a separate part of the city.

3.2 Method of Average Growth Coefficients

As in the previous case, the method of average growth coefficients is based on the predictive analysis of the volumes of interborough movements. Here, as the basic parameters of the predictive model, the growth coefficients C_i and C_j are used for each *i*-th and *j*-th transport area, respectively, which, similar to (1), are calculated by the following ratios

$$C_i = \frac{r_i^*}{r_i^0},\ C_j = \frac{r_j^*}{r_j^0}, \tag{4}$$

where $i, j = 1 \div n$; n is the number of transport areas in the urban agglomeration; r_i^0 and r_j^0 are actual (current) volumes of traffic flows in the *i*-th and *j*-th areas, respectively; r_i^* and r_j^* are their predicted volumes. Then the predicted distribution of traffic flows in the cells of the correspondence matrix according to the method of average growth coefficients is carried out on the basis of the formula

$$r_{ij} = C(0) \cdot r_{ij}^0 \tag{5}$$

where $C(0) = (C_i + C_j)/2$. Since the values of traffic flows between n transport areas, determined from ratios (5), as a rule, do not satisfy the equalities

$$r_i^* = C(0) \cdot \sum\nolimits_{i=1}^{n} r_{ij}(1),\ \ r_j^* = C(0) \cdot \sum\nolimits_{j=1}^{n} r_{ij}(1), \tag{6}$$

the method of average growth coefficients for predicting the volume of interborough movements is implemented iteratively in several steps according to the formula

$$r_{ij}(t+1) = C(t) \cdot r_{ij}(t), \tag{7}$$

where $C(t) = [C_i(t) + C_j(t)]/2$ are average values of the growth coefficients for the calculated moments $t = 0, 1, 2\ldots$

The calculation results at each iteration step form the initial data for the computational procedure of the next step. The process will continue until equality is reached between the predetermined value of the transport turnover of the area and the amount of movements obtained as a result of the calculation for this area using (6).

It is not difficult to see that the method of average growth coefficients takes into account the different dynamics of the development of certain areas of the urban agglomeration. However, with a significant increase in the mobility of the population due to the emergence of new residential areas, large shopping centers and/or large industrial enterprises, the application of this method leads to unacceptable errors, which does not allow it to be used for practical modeling of traffic flows.

3.3 Detroit Method

Unlike the previous methods, the Detroit method, in addition to the growth coefficients of individual areas of the urban agglomeration, takes into account the growth coefficient of the entire urban agglomeration. In order to maintain the correspondence between the predicted and calculated volumes of interborough movements, the predicted volumes of interborough movements are also determined iteratively in several steps. At the first step, these values are established as $r_{ij}(1) = \frac{C_i(0) \cdot C_j(0)}{C} \cdot r_{ij}(0)$, where $i, j = 1 \div n$; n is the number of areas; $C_i(0)$ and $C_j(0)$ are the growth coefficients for each i-th and j-th transport area, respectively, which are calculated in the form of (4); C is the growth coefficient of the entire urban agglomeration, determined by (1).

At the next steps, the calculation of the volumes of predicted interborough movements is carry out by the following equalities

$$r_{ij}(t) = \frac{C_i(t-1) \cdot C_j(t-1)}{C} \cdot r_{ij}(t-1), \quad t = 1, 2 \ldots \tag{8}$$

The Detroit method of predicting the volume of interborough movements is also trivial in terms of calculations. In practice, this method provides greater reliability of forecasts compared to previous methods of predictive analytics. However, the accuracy of forecasts deteriorates markedly in cases where the growth rate of the individual transport area differs significantly from the growth rate of the urban agglomeration.

3.4 Fratar Method

In practical modeling of traffic flows, the Fratar method has become most widespread as the most effective extrapolation method for forecasting correspondence matrix. According to the Fratar method, the predicted volume of movements from the i-th transport area to the j-th one for some future point in time is proportional to the actual volume of movements from the i-th transport area, multiplied by the growth coefficient of transport demand in the j-th area:

$$r_{ij}(1) = C_i(0) \cdot C_j(0) \frac{K_i(0) + K_j(0)}{2} \cdot r_{ij}(0),$$

where

$$K_i(0) = \frac{\sum_{k=1}^{n} r_{ik}}{\sum_{k=1}^{n} C_k r_{ik}} \text{ and } K_j(0) = \frac{\sum_{k=1}^{n} r_{jk}}{\sum_{k=1}^{n} C_k r_{jk}}$$

are the coefficients of growth in movements in the k-th area, due to the development of demand in the i-th and j-th transport areas, respectively; C_k is the demand growth factor for transport area k. The approximate desired solution by the Fratar method is also determined iteratively based on the following expression

$$r_{ij}(t) = C_i(t-1) \cdot C_j(t-1) \frac{K_i(t-1) + K_j(t-1)}{2} \cdot r_{ij}(t-1), \quad t = 1, 2 \ldots \tag{9}$$

3.5 Fuzzy Approaches to the Formation of the Correspondence Matrix

In practice, traffic flow management occurs under uncertainty, one of the factors of which is the incompleteness and/or inaccuracy of information relative to the actual volumes of interborough movements of human and freight resources, which is a "stumbling block" in the formation of correspondence matrices. Within the framework of predictive analytics, the possibility of overcoming this type of uncertainty is increasingly being considered in the application of new approaches. One such approach bears on a more adequate representation of the available information.

Regardless of the level of formalization, all observations of the volume of interborough movements within the transport network, in fact, are carried out at the level of so-called "Soft Measurements", the results of which are reflected in the form of weakly structured estimates [5], for example, in the form of intervals. At the moment, the most adequate formalisms of such assessments are fuzzy sets (fuzzy numbers) [6], which also allow overcoming semantic uncertainty in the value evaluative judgments of specialized experts. Thus, the fuzzy representation of relevant data makes it possible to apply the methodology of contextual fuzzy relations for the formation of actual and forecast correspondence.

When working with fuzzy numbers (sets), various membership functions can be chosen. For example, to describe the actual and predicted values of the elements of the correspondence matrices, as well as the growth and development coefficients, triangular fuzzy numbers are used in [7]. However, from the point of view of the realism and adequacy of the fuzzy description in the study of transport systems, the most convenient formalisms are still trapezoidal functions, or even better, membership functions of the Gaussian or "bell-shaped" type. In particular, if a weakly structured estimate of the volume of the traffic flow between the i-th area-source and the j-th area-sink is expressed as the interval $[a_{ij}, d_{ij}]$, then it can be represented as the fuzzy set $\tilde{A}_{ij} = \tilde{A}(r_{ij})$ with the following trapezoidal membership function (see Fig. 2).

$$\mu_{\tilde{A}}(r_{ij}) = \begin{cases} \frac{r_{ij} - a_{ij}}{b_{ij} - a_{ij}}, & \text{if } a_{ij} \le r_{ij} \le b_{ij}, \\ 1, & \text{if } b_{ij} \le r_{ij} \le c_{ij}, \\ \frac{d_{ij} - r_{ij}}{d_{ij} - c_{ij}}, & \text{if } c_{ij} \le r_{ij} \le d_{ij}, \\ 0, & \text{in other cases} \end{cases}$$

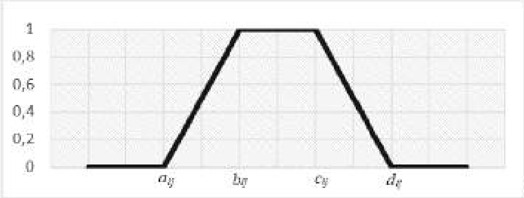


Fig. 2. Trapezoidal membership function

By analogy with [7], in each cell of the correspondence matrix the actual volumes of interborough movements can be reflected by the corresponding fuzzy sets $\tilde{X}_{ij} =< a_{ij}, b_{ij}, c_{ij}, d_{ij} >$ and the volumes of predicted correspondence can be found in the form of the fuzzy set $\tilde{X}^*_{ij} =< a^*_{ij}, b^*_{ij}, c^*_{ij}, d^*_{ij} >$, respectively. Of course, based on the rules of elementary operations on fuzzy sets [6], in accordance with (1) it is possible to represent the growth coefficients in the form of corresponding fuzzy sets, for example, as $\tilde{C} = \tilde{F}/\tilde{A} =< c_1, c_2, c_3, c_4 >$ [7], where c_k are the parameters of the fuzzy growth coefficient obtained from the results of operations on the corresponding fuzzy sets, and, further, carry out calculations according to the schemes of extrapolation methods in the usual manner with crisp α-cuts of contextual fuzzy formalisms.

It seems to us that the formation of the predictable correspondence matrix within the fuzzy information environment can be implemented in a more interesting manner, i.e. not by simply "immersing" the formulas of extrapolation methods (1)–(8) into the fuzzy environment, but using, for example, a Fuzzy Inference System, both for case (1) and for a more objective analysis of the growth in transport demand according to formula (2). To begin with, it is necessary to correctly select discrete universes, identify appropriate membership functions, and fuzzify the available weakly structured data relative to interborough movements of the transport network.

Another approach to the formation of the predictable correspondence matrix in a fuzzy paradigm can be the analysis of the dynamics of historical data of interborough movements by modeling the corresponding fuzzy time series with their subsequent forecasting in nominal numbers after defuzzification [5, 8, 9]. In particular, assuming historical data on transport demand in previous periods of observation to be weakly structured, it is possible through modeling the corresponding fuzzy time series to obtain forecasts in nominal units for short-term and even medium-term periods, and then calculate the value of the growth coefficient, and with it all the rest parameters of extrapolation methods for forming the correspondence matrix. But this is the subject of our next research.

4 Conclusion

In practical modeling of traffic flows, extrapolation methods are not widely used, primarily due to the irregular dynamics of the structural development of the urban agglomeration. Therefore, the extrapolation methods discussed above are usually used to form correspondence matrix with the forecasting horizon of no more than 5 years. However, the main reason for the inefficiency of extrapolation methods is still the lack of verified (adequate) information relative to the actual volumes of interborough movements, the values of which, as noted above, depend on numerous factors of very different nature. This problem requires a separate consideration, and one of the ways to overcome it lies in the plane of using fuzzy logic elements.

References

1. Thematic Review of the Association: Mathematical Methods for Modeling Traffic Flows. Release 5. KOSTA, St. Petersburg (2019). (in Russian)
2. Ortizar, I.D., Williumsen, L.G.: Modelling Transport, 4th edn. Wiley, Hoboken (2011)
3. Gorev, A.E., Boettger, K., Prokhorov, A.V., Gizatullin, R.R.: Fundamentals of Transport Modeling. KOSTA, St. Petersburg (2015). (in Russian)
4. Seliverstov, Ya.A., Seliverstov, S.A.: Methods and models of the construction of transport correspondence matrix. Sci. Tech. Bul. St. Petersburg State Polytechnical Univ. Inform. Telecommun. Control **217–222**(2–3), 49–70 (2015). (in Russian)
5. Rzayev, R.R.: Data Mining in Decision Support Systems. Palmerium Academic Publishing, Saarbruchen (2016). (in Russian)
6. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning. Inf. Sci. **8**(3), 199–249 (1965)
7. Koshunyaeva, N.V.: Application of fuzzy characteristics in the construction of correspondence matrices by extrapolation methods. Mod. High Technol. **5**, 209–213 (2022). (in Russian)
8. Rzayev, R.R., Abdullayev, K., Alizada, P.E., Salmanov, F.M.: Fuzzy time series forecasting by the example of the Dow Jones Index dynamics. In: Arai, K. (eds.) Advances in Information and Communication. FICC 2022. LNNS, vol. 438, pp. 216–228. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98012-2_17
9. Rzayev, R.R., Alizada, P.E., Mehtiyev, T.Z.: Volatile time series forecasting on the example of the dynamics of the Dow Jones index. Probl. Inf. Soc. **14**(1), 14–25 (2023)

Geoinformation Model for Smart Grid Sustainability Management

Stanislav Belyakov, Alexander Bozhenyuk(✉), Leontiy Samoylov, and Polina Nikashina

Southern Federal University, Nekrasovskiy Street, 44, 347928 Taganrog, Russia
avb002@yandex.ru

Abstract. The article deals with the problem of managing the distribution of energy capacity in the territory covered by an intelligent energy grid. The goal of control is to stabilize the power flow in the presence of external influences caused by a change in the state of the environment. Geoinformation models are described that use the representation of knowledge by images. Conceptually, the image includes the center and the allowed transformations of the center in some context. The case is considered when the threat of disruption of the smart grid is assessed by transforming the image into a given area of space in which the smart grid is located. The formation of a list of threats for a given state of the external environment is the essence of sustainability management. The features of the algorithmization of the procedure for transforming images are analyzed. A method for assessing the reliability of transformation is given.

Keywords: Geoinformation modeling · Smart Grid · Intelligent GIS · Representation of knowledge

1 Introduction

The sustainable functioning of the smart grid is of interest both in practical and theoretical terms. Really functioning equipment in energy supply systems is only partially able to withstand external negative influences. These include the impact of the natural environment, the technosphere, failures of equipment and programs of the smart grid itself. The problem of sustainability should be addressed by the development of the management system in relation to its intellectualization. Smart things, appliances, transport, and cities are now seen as an undeniable prospect. From a theoretical point of view, achieving sustainability requires adequate models of control objects. Of particular interest are geoinformation models that reflect the phenomena and processes of the real world in cartographic form.

Intelligent geoinformation models are a special class of figurative-sign models that include maps, plans, diagrams of real-world systems and knowledge about their behaviour. Part of the knowledge is embedded in geographic electronic maps already at the stage of their production. The rich content of the cards increases the efficiency of

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 651–658, 2023.
https://doi.org/10.1007/978-3-031-39777-6_76

decisions made. However, this knowledge is not enough to work in conditions of uncertainty. Ensuring sustainability requires an assessment of the current state of the network, its classification, and the development of a control action. Such a control paradigm requires a special conceptual model for describing and evaluating spatial situations. The specificity is to effectively use the rich experience of decision-making in the conditions of incompleteness, uncertainty, and ambiguity of data on the state of the smart grid. Geographic information systems (GIS) software tools allow the user to record a large amount of detailed data on decision-making precedents. However, the data does not provide an idea of which path decision making should take. This issue is investigated in the present work.

Particular attention when using geoinformation models, from our point of view, should be given to the transfer of knowledge contained in them (transfer learning). Unlike machine learning models, the pretraining of GIS models is not associated with the preprocessing of data sets and the establishment of some reasonable initial values for the model parameters. The behavior of geo-information models is highly dependent on the topology of space. For example, it seems impossible to transfer the experience of eliminating emergency situations of power supply systems from a swampy area to a sandy desert zone. However, expert analysts can "imagine" what a known situation might look like in a completely different area. This feature is extremely useful and attracts to the study of how experience should be presented, what is the logic of the transfer of experience and how one or another decision can be argued.

The purpose of this work is to describe a geoinformation model for the transfer of knowledge about anomalous situations that affect the quality of the smart grid. The originality of the study lies in the use of a special form of knowledge representation and an algorithm for its use.

Section 2 sets out the sustainability management problem. Section 3 contains a literature review. Section 4 describes the concept of knowledge representation about anomalous situations. Section 5 presents the knowledge transfer algorithm. Section 6 describes an example of applying a geoinformation model. Section 7 includes conclusions from the results of the study.

2 Problem Statement

Under the sustainability property for a smart grid, we will understand its ability to withstand threats of disruption of normal functioning. The nature of threats is different. Threats have a wide range of variable parameters. As a result, objectively there is uncertainty about which of the threats are relevant at a given time. The spatial localization of threats is also essential since smart grids are geographically dispersed objects. In the task of predicting threats and developing protective measures, the experience of observing events and phenomena that previously led to smart grid disruptions will always be important. The ability to transfer the experience of applying protective measures from one situation to another is of particular importance for spatial objects. The ability to use GIS to solve the problem of "how THIS will look THERE" from this point of view is a promising tool for ensuring sustainability.

We will represent the geoinformation model for maintaining sustainability for the smart grid as a set of objects:

$$M = \{C, B, G, P\}, \tag{1}$$

where C is the set of contexts in which problems of spatial analysis can be solved; B s a set of spatial areas in which the smart grid is located; G is a set of procedures for transferring experience in studying precedents of dangerous situations; P is the set of precedents observed earlier in the smart grid.

The following task should be solved using a geoinformation model:

$$\begin{cases} W(G(B_w, P_a, C_w) \approx \ max, \\ B_w \int \ B, P_a \int \ P, C_w \int \ C. \end{cases} \tag{2}$$

Here W is a criterion for the reliability of experience transfer; B_w is the work area of the space in which threat analysis is performed; P_a list of current threats; C_w is the context of the analysis.

To provide a solution to problem (2), it is necessary:

- to determine the form of presentation of knowledge about the precedent, which will allow transferring them to a given spatial area;
- to present the procedure of knowledge transfer;
- to substantiate the criterion for evaluating the transfer result.

If the listed tasks are solved, the smart grid control system does the following:

1. It receives information in cartographic form about an anomalous event or process. The necessary cartographic objects are applied to the corresponding layers of the electronic map;
2. It forms an analysis area on the map, which is updated with live data about the external environment;
3. It selects precedents of known situations that could potentially occur in the analyzed area;
4. Precedents, the possibility of repetition of which is high enough, are considered threats in the near future;
5. It generates control signals aimed at reducing the likelihood of threats and minimizing damage in the event of an attack.

3 Literature Review

The construction and use of geoinformation models for smart grid management has been used for many years [1]. This approach is considered as the development of control actions based on a multi-criteria assessment of the state of an object [2]. A feature of this mechanism is a comprehensive assessment of territories and the accumulation of precedents of emergency situations. Management is based on the results of statistical generalizations of continuously accumulated data on the state of the environment in which the smart grid operates [3]. Attention is drawn to the structuring of precedents, which minimally distinguishes them from the general classification of GIS objects [4]. On the one hand, this reduces the cost of maintaining the GIS information base, but on the other hand, it reduces the reliability of the spatial analysis of impacts on the smart grid. There is an urgent need to study more complex case information structures.

Case based reasoning (CBR) is one of the long-established methods of experience reuse [5]. It is based on the use of precedent proximity metrics and the compactness hypothesis. In accordance with this, decisions in similar situations should be close in meaning. This principle is effective for the analysis of spatial situations, however, the difficulty of choosing an expressive and adequate measure of similarity forces one to turn to other principles [6]. In particular, the question of what CBR modifications should be used to transfer experience from one area of space to another requires research.

Considering that GIS are focused on the continuous accumulation of spatial data and belong to big data systems, many publications are devoted to the application of data analysis methods in big data systems [7]. Considerable interest is shown in machine learning models [8, 9]. This approach is appropriate when large amounts of data directly characterize the decision-making process. The practice of ensuring sustainability for the smart grid does not have such data. In terms of content, this is data on the state of the external environment and single logically justified decisions. Here it is necessary to explore a special way of learning an intelligent system, using extensive data about the topology of the space in which the smart grid is located.

The transfer of knowledge is an element of research on the global problem of the extraction and use of knowledge. Works in this area explore models for presenting existing knowledge for its transfer to the interested party. An example is the work [10]. Here, the Throughput Model is used as a basis, which includes perception, information, judgment and decision making. The listed elements are considered as intelligent operations based on knowledge. Knowledge transfer is seen as supporting the algorithmic process of the Throughput Model. The level of conceptual representation of knowledge in the works of this direction corresponds to the organizational one, when the smart grid management cycle is implemented by experts. Thus, research in this direction concentrates on a system of concepts that partially cover the intelligence of the smart grid automated control system.

Smart grid research from the point of view of cybersecurity analyzes threats and vulnerabilities that can generate incidents of failures and failures [11]. This helps to systematize knowledge about the dangers and develop protective measures. At the same time, the use of geoinformation models is not considered in the works on this issue.

4 Representing Precedents of Spatial Situations

A well-known method of applying experience is CBR [5]. Given the spatial position of two precedents p_a and p_b, their similarity is estimated by the accepted distance metric $d(p_b, p_a)$. The minimum distance corresponds to the maximum similarity. The CBR procedure in a geoinformation model is difficult to apply for two reasons:

- distance $d(p_b, p_a)$ requires two precedents on the map. In the case where the exact location of one of them is not known, complicates the situation;

- for a cartographic space, the metric $d(p_b, p_a)$ is non-linear. For any map object, a set of topological relationships must be satisfied, determined by the purpose of creating a map.

Analyzing possible ways to solve problem (2), we will take into account that any precedent is a subset of spatial objects (K) and relations (R), represented in the geoinformation model:

$$p_i \nabla P \parallel \quad p_i \int \; K \perp R, p_i \neq \emptyset. \tag{3}$$

Each precedent is placed in an area on the map B_{p_a}. Transferring a precedent means defining a new area $\tilde{B}_{p_a}$ in which cartographic objects and relations of the new precedent $\widetilde{p_i}$ will be built. In the general case, all elements of the pair

$$(\tilde{p}_i, \; K \perp R) \tag{4}$$

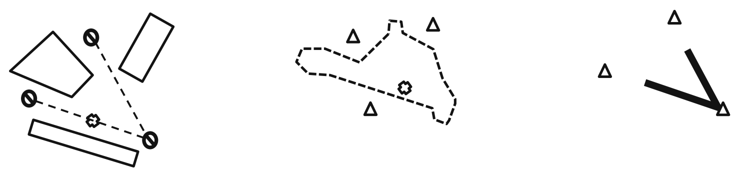
must be topologically compatible. Substantially, this means that the new position of any element of the precedent $\widetilde{p_i}$ must exactly repeat the spatial position of a similar element of the original precedent p_i. Because the areas B_{p_a} and $\tilde{B}_{p_a}$ cover different sets of cartographic features, exact topological consistency becomes extremely unlikely. At the same time, the practice of spatial analysis shows that the location of the elements of the precedent $\widetilde{p_i}$ is not rigidly fixed relative to each other and the surrounding objects. Possible position variations are determined by the meaning of the used context $C_w \int \; C$. However, for this, the boundaries of such changes in position must be known.

The need to use knowledge about admissible transformations leads to the following conceptual model [12]. Any precedent is described as an image:

$$<c, h_1(c), h_2(c), ..., h_M(c)>, \tag{5}$$

in which c is the center of the pattern and $h_i(c)$ are the allowed transformations of the pattern.

The center includes a cartographic representation of the use case in a given context. Any valid transformation is a set of cartographic objects that reflect the change in the values of the properties of the center but retain the meaning of the precedent. Figure 1 shows two instances of valid transformations for an electrical cable break use case defined by a point object.



(a) center of the image (b) transformation h_1 (c) transformation h_2

Fig. 1. Example of valid transformations.

Figure 1a shows the precedent image as a point object on the map (conditionally marked with a cross) in its area B_{p_a}. All significant objects characterizing the emergency situation and the subsequent decision are shown (not shown here). A valid transformation h_1 is the position of the open space wire break location relative to nearby buildings (Fig. 1b). This factor significantly influenced the subsequent decision. A possible transformation of h_2 s considered to be the possible position of the break on two adjacent lines (Fig. 1c). The mapping image of each valid transformation includes feature references that define their geometry (shown as triangles). Such references are used to explain the geometric characteristics of a valid transformation.

The conceptual model (5) defines the knowledge representation required for the subsequent transfer of precedents. It is essential that the topological consistency, in contrast to (4), must be controlled for the pair ($\tilde{p}_i$, $\tilde{B}_{p_i}$). This not only reduces the complexity of setting and checking constraints, but also makes the transfer of the use case more intelligent, i.e. adapted to the previously unknown features of the terrain.

We emphasize that model (5) is represented by cartographic objects, and does not exclude the binding of knowledge of known models to them: production, semantic, frame, and others.

5 Transferring a Precedent

The precedent transfer procedure in the form (5) consists in placing the corresponding elements of the precedent $\tilde{p}_a$ in the given area $\tilde{B}_{p_a}$ if there are areas of admissible transformations. The algorithmic features of this procedure are as follows:

- each valid transformation is placed on a separate layer as a set of cartographic objects. Their types and attribute values must be arguments to a function to map each use case element to a given location;
- each layer is assigned a program function that interprets the objects of this layer and returns as a result the area for placing the precedent element. If the area is empty, the use case cannot be moved;
- each layer is assigned a program function that constructs an element of a new precedent in a non-empty placement area;
- the final step is to call the function for evaluating the reliability of the result.

The key to the transformation function is the assessment of reliability. The result may include only a subset of the elements that match the original use case. For example, transformation h_2 in Fig. 1 may not transform if the two cables are too close together. However, this feature, according to the expert, may retain the possibility of re-applying the known solution, be a valid transformation of the result of the transformation. Therefore, an estimate of the transformation result can also be obtained according to model (5). The image can be represented as:

$$<F_{TR}\left(p_a, \tilde{B}_{p_a}\right), h_1\left(F_{TR}\left(p_a, \tilde{B}_{p_a}\right)\right), \ldots, h_M\left(F_{TR}\left(p_a, \tilde{B}_{p_a}\right)\right)>.$$

Here $F_{TR}\left(p_a, \tilde{B}_{p_a}\right)$ is the function of transforming the precedent into the given area. The expert will be required to indicate the value of the confidence score for each valid transformation:

$$\left(v_1, h_1\left(F_{TR}\left(p_a, \tilde{B}_{p_a}\right)\right), \ldots, \left(v_M, h_M\left(F_{TR}\left(p_a, \tilde{B}_{p_a}\right)\right)\right)\right),$$
$$v_r \nabla [0, v_{max}], r = \overline{1, M}.$$

The reliability estimation algorithm consists in successive enumeration of layers with h_r images until the closest acceptable transformation is found.

6 Application

As an option for applying the proposed approach, a geoinformation model was developed to protect solar panels from the consequences of a mass flight of mosquitoes of the Chironomidae family in the city of Taganrog [13]. Insects completely cover the solar panels, disabling them. Four precedents of a sharp increase in the population of Chironomidae in the steppe, coastal, suburban and urban areas were described and analyzed. Here, the relief, composition of the vegetation cover, two-week temperature regime, humidity and a number of other indicators were taken into account. The area of knowledge transformation was about 97 km^2. The number of sites with solar panels is 8. The highest level of forecast reliability was 0.92 for autumn.

7 Conclusion

The study made it possible to draw the following conclusions. Firstly, the sustainability of smart grid depends significantly on environmental factors. Their dynamics and unpredictability make intelligent geoinformation models useful. They include procedures and knowledge for reproducing processes and phenomena in different locations. Secondly, the knowledge representation model has specific features due to the need to transform precedents. Conceptually, this is a description of the permissible transformations of objects, events or phenomena that retain their semantic integrity. We consider the search for new forms of presentation of precedents for the transformation of experience to be a further direction of research on this topic.

Acknowledgments. The research was funded by the Russian Science Foundation Project No. 23-21-00206, https://rscf.ru/en/project/23-21-00206/ implemented by the Southern Federal University.

References

1. Alhamwi, A., Medjroubi, W., Medjroubi, T., Agert, C.: GIS-based urban energy systems models and tools: Introducing a model for the optimisation of flexibilisation technologies in urban areas. Appl. Energy **191**, 1–9 (2017)
2. Mardania, A., Zavadskas, E., Zakuana, K.: A review of multi-criteria decision-making applications to solve energy management problems: two decades from 1995 to 2015. Renew. Sustain. Energy Rev. **71**, 216–256 (2017)

3. Bhatt, J., Shah, V., Jan, O.: An instrument at engineer's review on smart grid: Critical applications and parameters. Renew. Sustain. Energy Rev. **40**(C), 2017–2039 (2014)
4. Mrówczyńska, M., Skiba, M., Sztubecka, M., Bazan-Krzywoszańska, A., Kazak, J.K., Gajownik, P.: Scenarios as a tool supporting decisions in urban energy policy: the analysis using fuzzy logic, multi-criteria analysis and GIS tools. Renew. Sustain. Energy Rev. **137**, 110598 (2021)
5. Kolodner, J.: Case-Based Reasoning. Morgan Kaufmann, San Mateo (1993)
6. Zhuoa, H.H., Kambhampatib, Z.: Model-lite planning: case-based vs. model-based approaches. Artif. Intell. **246**, 1–21 (2017)
7. Mostafa, N., Mohamed Ramadan, H.S., Elfarouk, O.: Renewable energy management in smart grids by using big data analytics and machine learning. Mach. Learn. Appl. **9**, 100363 (2022)
8. Jing, Q., Yan, J., Wang, Y., He, R., Lu, L.: A novel differentiable neural network architecture automatic search method for GIS partial discharge pattern recognition. Measurement **195**, 111154 (2022)
9. Li, Y., Yabuki, N., Fukuda, T.: Integrating GIS, deep learning, and environmental sensors for multicriteria evaluation of urban street walkability. Landsc. Urban Plan. **230**, 104603 (2023)
10. Rodgers, W., Cardenas, J.A., Gemoets, L.A., Sarfi, R.J.: A smart grids knowledge transfer paradigm supported by experts' throughput modeling artificial intelligence algorithmic processes. Technol. Forecast. Soc. Chang. **190**, 122373 (2023)
11. Yaacoub, J.P.A., Fernandez, J.H., Noura, H.N., Chehab, A.: Security of power line communication systems: issues, limitations and existing solutions. Comput. Sci. Rev. **39**, 100331 (2021)
12. Belyakov, S., Belyakova, M., Bozhenyuk, A., Rozenberg, I.: The features of generations of solutions by intellectual information systems. Adv. Intell. Syst. Comput. **451**, 221–229 (2016)
13. Bulysheva1, N.I., Nabozhenko, M.V., Savikin, A.I., Shokhin, I.V.: Phenology of flight of Cironomid flies (Diptera: Chironomidae) in Taganrog Bay of the Sea of Azov. Caucasian Entomol. Bull. **9**(2), 329–332 (2013)

Intelligent Benders' Decomposition Algorithm for Sequencing Multi-model Assembly Line with Sequence Dependent Setup Times Problem: A Case Study in a Garment Industry

Elvin Sarı(✉), Mert Paldrak, Yaren Can, Tunahan Kuzu, Sude Dila Ceylan, Devin Duran, Mustafa Arslan Örnek, and Ozan Can Yıldız

Industrial Engineering Department, Yaşar University, Bornova, İzmir, Turkey
elvinsari00@gmail.com, {mert.paldrak,arslan.ornek}@yasar.edu.tr

Abstract. In recent years, clothing manufactures aim at producing various of products with low stock in order to meet customer demand. Besides this fact, the ready-made clothing industry needs to pursue science, technology, and innovation policies to keep up with the rapid change in the fashion industry. One of the most commonly used production systems in garment industry is assembly lines where parts are subsequently added until the end product is obtained. In garment industry, on time delivery plays a vital role in increasing customer satisfaction while ensuring demand. Consequently, setup times in assembly lines are of paramount importance to track the performance of production system. In this study, the minimization problem of long setup times due to the wide variety of models produced in the garment industry is handled. A real-life production management problem is defined, formulated as an MIP model, and solved to improve customer delivery rate and to increase efficiency by minimizing setup time. To solve this problem to optimality, two exact solution techniques namely Branch and Bound and Benders' Decomposition Techniques are taken into consideration. The proposed mathematical model is solved with ILOG CPLEX OPTIMIZATION STUDIO version 20.1 and the solutions obtained using each technique are compared with respect to solution quality and computational time.

Keywords: Garment Industry · Multi-Model Assembly Line · Sequence Dependent Setup times · MIP · Multi-Objective

1 Introduction

The garment industry is a significant sector of the global economy that involves the production, design, and marketing of clothing and textiles, and has a long history of evolution. The industry comprises diverse manufacturing processes, technologies, and materials that aim to meet the changing demand of customers. This sector is crucial to global trade and provides employment opportunities to millions of people worldwide. However, due to the constantly changing customer preferences and emerging trends,

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 659–668, 2023.
https://doi.org/10.1007/978-3-031-39777-6_77

garment manufacturers must stay current with the latest industry trends to remain competitive. Among the various production processes, the sewing process is crucial as it employs different tools and techniques to join fabric pieces together. Although computerized sewing machines are widely used, manual techniques are still prevalent in most garment companies. As the garment industry is labor-intensive, efficient utilization of machines and workers is vital. Therefore, balancing the line and assigning tasks equitably is critical for maximizing efficiency in both workers and machines, leading to the emergence of the line balancing problem in the garment industry.

Line balancing is utilized to optimize production processes by reducing idle time and maximizing machine efficiency through workload distribution across different workstations. The importance of line balancing is particularly heightened in the garment industry due to the presence of multiple stages involved in the manufacturing process. In the context of sewing lines, line balancing involves the allocation of sewing tasks across workstations to minimize idle time and increase efficiency. The line balancing process requires analysis of the production process to identify tasks and group them into workstations that can perform tasks in parallel, resulting in even workload distribution across the workstations.

The focus of this study is on the sewing line balancing problem in a garment company. The efficiency of the sewing line is primarily influenced by the completion time of each product, which is affected by the set-up time required for machine or equipment preparation between two consecutive tasks. The set-up time has a significant impact on the overall efficiency and productivity of the sewing line. Hence, a mathematical model is proposed in this study with the objective of minimizing the total set-up time of the sewing line and the total lateness, considering the variety of products being produced.

The remaining sections of the study are as follows. In Sect. 3, multi-objective MIP model for our problem is formulated. In Sect. 4, our proposed solution techniques, namely B&B and Benders' Decomposition algorithm along with master and sub problem of decomposition model are given in detail. In Sect. 5, computational results and our findings are summarized. Lastly, overall conclusion related to our study and future directions are provided in Sect. 6.

2 Literature Review

The garment industry is a complex and tremendously competitive market that necessitates garment manufacturers to be efficient and innovative in their operations to make profit. In this section, a literature review on the garment industry and sewing line is conducted to provide a general overview of the key issues and challenges that garment manufacturers and companies face regularly considering Benders Decomposition algorithm.

The assembly line is the processing of materials by taking advantage of the labor force along the line and the sorting of these processes along the line. Assembly lines were first used by the Ford Company in the United States. The assembly line in the garment industry comprises a number of workstations where a certain operation is carried out in a predetermined order. Production lines might be single-model, multi-model, or mixed-model [2]. Frequently used Multi-Model Assembly lines (MuMAL) are used to produce different types of models of a basic product in split lots. In MuMAL problems, lot size is more than one. Despite some shared jobs, there are some notable different tasks amongst models [4].

Setup contains the machine, process, or bench for parts or the cycle. The setup comprises getting tools, placing WIP, returning tooling, cleaning, installing jigs and fixtures, adjusting tools, and examining material [1]. It is common for the lengths of setup times to be influenced by the degree to which the technological needs of two successive processes are comparable. Several researchers have given recognition to the sequence dependence of setup times [8]. Also, [6] has pointed out that the assumption of sequence-independent setup times does not hold true in many real-world situations. They came up with a Branch-and-Bound algorithm for scheduling jobs with setup times that depend on the order in which they are run on a single processor so that the number of tool changes is kept to a minimum. [5] have developed a multi-objective mathematical model. The mathematical model contains three different objective functions. The first one is minimizing setup time to increase efficiency. The Second objective function is decreasing work in process by minimizing earliness and backlog level. The Last objective function is minimizing the lateness of orders to improve service level. [4] developed a multi-objective model to improve the current assignment of tasks to several workstations and achieve an optimized cycle time for each model.

More than 50 years have passed since Bender's (1962) created the Bender's Decomposition (BD) technique, whose primary goal is to address issues with complicated variables that, when temporarily resolved, give a situation that is much simpler to manage [7]. Bender's Decomposition is a technique for addressing mixed integer programming (MIP) problems with a specific structure in the constraint set. Infanger [3] used and developed Bender's decomposition techniques and Monte Carlo Sampling for solving two-stage stochastic linear programs to solve problems.

3 Mathematical Modeling

Manufacturing companies generally use a production policy to be able to meet customer demand while minimizing respond time and increasing customer satisfaction. Availability of necessary raw materials quality of and products and priced are of paramount important of customer satisfaction. Considering this fact, they track the performance of production policies in terms of some indicators related to production control and management. Among all manufacturing industries garment industry is notorious for be in highly labour-intensive product variability and non-environmentally friendly. Most of garment industries focus on optimization of their processes to become lucrative in competitive market which is getting tougher every day. On the other hands cost incurrent related to production are perpetually increasing due to an increase in labour and raw material cost for limited resources and variety in customer trends. Manufacturers like Sun Textile aim at reducing cost increasing flexibility and quality while maintaining product variability in the competitive market.

In this study a real-life production management problem in a garment industry is described and modelled using operation research techniques where the objectives are to minimize setup times improve services quality by minimizing lateness of orders.

This to objectives are pivotal the since Sun Textile take care of maximizing customer satisfaction while making profit. These goals constrict each other in other words it is impossible to optimize the one without deteriorating the other. Consequently, a multi-objective mathematical formulation is proposed with a setup carryover strategy. Firstly, we proposed a *MIP* model where two objective functions are considered separately and then a Goal Programming model will be formulated based on these *MIP* models.

The assumptions made while creating this mathematical model are as follows:

- Initially, all models are assumed to be ready to go into the production line.
- Setup times between two consecutive models are known in advance.
- No pre-emption between models is allowed. In other words, a model is not allowed to be interrupted in the middle of execution until its process terminates.
- The due date of each model is determined according to the customer orders.

Mathematical Model.

Indices:

$i, k \in I = \{0, \ldots, n\}$ –index of models produced on the line.

$j, a \in J = \{0, \ldots, n\}$ –index of position (sequence) of a model on the line.

Parameters:

Pt_i = Processing time of model i.

DD_i = Due date of model i.

ST_{ik} = Setup time between model i and model k.

Decision Variables:

CT_j = Completion time of a model in the position j.

X_{ij} = 1; if model i is processed in position j, otherwise 0.

Y_{ij} = 1; if model i is processed before its due date in position j, otherwise 0.

Z_{jik} = 1; if model k is processed immediately after model i in position j, otherwise 0.

L_j = Lateness of a model in position j.

E_j = Earliness of a model in position j.

Objective Functions:

$$Z_1 = \min \sum_{j=0}^{n} \sum_{i=0}^{n} \sum_{k=0}^{n} Z_{jik} * ST_{ik} \tag{1}$$

$$Z_2 = \min \sum_{j=0}^{n} L_j \tag{2}$$

Constraints:

$$\sum_{j=0}^{n} X_{ij} = 1, \qquad \forall i \tag{3}$$

$$\sum_{i=0}^{n} X_{ij} = 1, \qquad \forall j \tag{4}$$

$$X_{00} = 1 \tag{5}$$

$$\sum_{i=0}^{n}\sum_{k=0}^{n} Z_{jik} = 1, \qquad \forall j; j \neq 0 (j = 1, \ldots) \tag{6}$$

$$X_{ij-1} + X_{kj} - Z_{jik} \leq 1, \qquad \forall j, \forall i \; and \; \forall k; k \neq i; j \neq 0 (j = 1, \ldots., n) \tag{7}$$

$$Z_{0ik} = 0, \qquad \forall i, \forall k \tag{8}$$

$$CT_j = \sum_{a=1}^{j}\sum_{i=0}^{n} Pt_{i*}x_{ia} + \sum_{i=1}^{n}\sum_{k\neq i}^{n} Z_{jik} * ST_{ik}, \forall j \tag{9}$$

$$CT_j - DD_i * X_{ij} + \mathrm{M} * Y_{ij} \geq 0, \qquad \forall j and \forall i \tag{10}$$

$$\sum_{i=0}^{n} DD_i * Y_{ij} - CT_j \leq E_j, \qquad \forall j \tag{11}$$

$$\sum_{i=0}^{n} DD_i * X_{ij} - CT_j + L_j \geq 0, \qquad \forall j \tag{12}$$

$$L_j \geq 0, \qquad \forall j \tag{13}$$

$$E_j \geq 0, \qquad \forall j \tag{14}$$

$$X_{ij}, Y_{ij}, Z_{jik} \; binary, \quad \forall j, i, k, a \tag{15}$$

Our objectives are to minimize total setup times (1) and total lateness (2) on the assembly line while changing the model type. Constraints (3) and (4) ensure that only one model can be assigned to a position and only one position to a model. Since the last position of the previous model type is the first position of the next model, the assignment of model 0 to the 0th position is provided by constraint (5). For all positions except the starting position, the process of only one model immediately after a model is provided by the equation in constraint (6). As seen in constraint (7), if two models are processed one after the other, the first model should be processed right after the other model. In the equation in Constraint (8), a model at position 0 indicates that any model cannot be found in its previous position. Constraint (9) calculates the completion time of the model at each position. Constraint (10), (11), and (12) determine earliness and lateness of a position according to the assigned model. As seen in constraint (13) and (14) the lateness and earliness are integer variables. Constraint (15) gives the domain of decision variables.

4 Data Generation and Solution Methodologies

4.1 Data Generation

The data set for the problem is collected after face-to-face meetings with sewing line production managers. However, due to the confidential agreements, we are not allowed to use these data as they are. Consequently, we had to generate our data with the help of MATLAB R2016. Processing time, setup times between models and due dates are generated based on a Uniform distribution with relevant parameters. The instances are generated starting from 5 models in each case and then increasing the number of models to 8,10 and 20. Our instances are classified as small, medium, and large according to their size where number of models determines the size of a problem. Each instance is solved through IBM ILOG CPLEX OPTIMIZATION STUDIO version 20.1 on a personal computer and elapsed time to solve each instance is recorded. The average statistics of solutions regarding each size of models are shown in Table 1 both objective functions. Moreover, objective function values of each instance are provided in below Table 2.

Table 1. Optimal Solution Values of Each Instance for each Objective Function

	SETUP			LATENESS		
OBJ. FUN	5	8	10	5	8	10
Solution Time(sec.)	0,19	0,31	0,64	0,11	0,14	0,39
Constraints	258	792	1408	258	792	1408
Variables	306	918	1606	306	918	1606
Binary	288	891	1573	288	891	1573
Integer	18	27	33	18	27	33
Non-zero Coefficient	1156	3907	7151	1156	3907	7151
Nodes	0	0	3043	0	0	0
Iteration	334	10619	27293	435	486	749

Table 2. Optimal Solution Values of Each Objective for Small and Medium Instances

		Data1	Data2	Data3	Data4	Data5	Data6	Data7	Data8	Data9	Data10	Data11	Data12	Data13	Data14	Data15
type		5	5	5	5	5	8	8	8	8	8	10	10	10	10	10
Setup	Z1	84	66	79	80	63	99	83	84	110	77	93	96	75	123	84
Lateness	Z2	1046	1442	2315	1290	2265	5904	7342	9155	7814	5790	11495	13815	11296	7904	8154

4.1.1 Solution Methodologies for Mathematical Model

As explained in the first section, our problem is formulated as an *MIP* model in which both continuous and binary variables are used. To solve this problem to optimality, two exact solution techniques, namely Branch & Bound and Bender's Decomposition are employed. Since the problem complexity increases exponentially with increasing problem size, Branch & Bound fails to find optimal solutions to big-sized problems in a reasonable computational time. The exact solution techniques are summarized below.

4.1.2 Branch and Bound Technique

The Branch & Bound Technique is a method used for solving optimization problems in discrete and combinatorial domains. The algorithm performs a systematic search of the state-space by exhaustively generating and evaluating candidate solutions. It achieves this by exploring branches of candidate solutions and finding the optimal solution through comparison against upper and lower estimated bounds. To solve the proposed mixed-integer programming (MIP) model, we first attempt to solve its linear relaxation version, and then proceed to branch on fractional values of the variables X_{ik}, Y_{ij}, Z_{jik} until the optimal solution is found. Due to the integration of IBM CPLEX OPTIMIZER with the Branch & Bound algorithm, it becomes convenient to solve our MIP model using the CPLEX solver.

4.1.3 BENDER's Decomposition Method

In the process of modelling optimization problems, a major issue is the problem size, which is determined by the number of decision variables and constraints involved. The size of the problem increases the complexity of the solution procedure and affects the computational time required to obtain a solution. Traditional techniques such as Branch & Bound may not be efficient in finding optimal solutions within a reasonable time as the number of decision variables and constraints increases. To address these challenges, partitioning methods such as Bender's decomposition can be a suitable approach to adopt.

Partitioning methods, unlike traditional techniques, divide the solution process into two parts. Bender's Decomposition algorithm is a well-known approach to easily solve NP-hard problems such as integer, mixed integer, and nonlinear programming models. This technique is designed to solve mixed integer problems by dividing the problem into a master problem and a sub-problem. The master problem is the relaxed version of the original problem, where some variables are extracted, and only binary variables are included. In contrast, the sub-problem consists of only continuous or other complex decision variables. This partitioning procedure significantly simplifies the problem, enabling the algorithm to solve it efficiently. The division of decision variables according to the problem type is illustrated in Fig. 1.

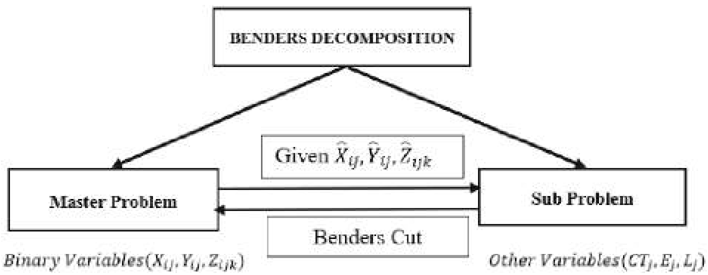


Fig. 1. Benders' Decomposition Model for Proposed MIP Model

Master Problem

Master problem includes constraints that merely have complicating variables. As all variables in constraints (3) to (8) are binary (complicating variables), master problem is formulated as follows:

Objective Function:

$$MinZ_{lower} \tag{16}$$

Constraints:

$$X_{ij}, Y_{ij}, Z_{jik}\ binary, \ \ \forall j, i, k \tag{17}$$

Sub Problem.

To formulate sub-problem, complicating variables are assumed as fixed values. For simplicity, X_{ij}, Y_{ij}, Z_{ijk} are fixed and shown by $\widehat{X}_{ij}$, $\widehat{Y}_{ij}$, $\widehat{Z}_{ijk}$, respectively.

Objective Function:

$$MinZ = \sum_{j=0}^{n} L_j \tag{18}$$

Constraints:

$$CT_j = \sum_{a=1}^{j}\sum_{i=0}^{n} Pt_{i*}x_{ia} + \sum_{i=1}^{n}\sum_{k\neq i}^{n} Z_{jik} * ST_{ik}, \qquad \forall j \tag{19}$$

$$CT_j - DD_i{*}X_{ij} + \mathrm{M}{*}Y_{ij} \geq 0, \quad \forall j\ and\ \forall i \tag{20}$$

$$\sum_{i=0}^{n} DD_i{*}Y_{ij} - CT_j \leq E_j, \qquad \forall j \tag{21}$$

$$\sum_{i=0}^{n} DD_i{*}X_{ij} - CT_j + L_j \geq 0, \qquad \forall j \tag{22}$$

$$CT_j, L_j, E_j \geq 0, \qquad \forall j \tag{23}$$

To formulate dual sub-problem, new dual variables should be defined. To do this, U_j, V_{ij}, W_j and Q_j are associated to constraints (19) to (22), respectively. Finally, dual sub-problem is formulated as follows:

Objective Function:

$$\begin{aligned} MaxW = {} & \sum_{j=0}^{n}\left(\sum_{a=1}^{j}\sum_{i=0}^{n} Pt_{i*}\hat{X}_{ia} + \sum_{i=1}^{n}\sum_{k\neq i}^{n} \hat{Z}_{ijk}{*}ST_{ik}\right) * U_j + \sum_{j=0}^{n}\sum_{i=0}^{n}(DD_i{*}\hat{X}_{ij} \\ & - \mathrm{M}{*}\hat{Y}_{ij}) * V_{ij} + \sum_{j=0}^{n}\left(\sum_{i=0}^{n} DD_i{*}\hat{Y}_{ij}\right) * W_j \\ & - \sum_{j=0}^{n}\left(\sum_{i=0}^{n} DD_i{*}\hat{X}_{ij}\right) * Q_j \end{aligned} \tag{24}$$

Constraints:

$$U_j + \sum_{i=0}^{n} V_{ij} + 2W_j \leq 1, \qquad \forall j \tag{25}$$

$$U_j\, urs,\; V_{ij} \geq 0, W_j \geq 0, Q_j \geq 0 \tag{26}$$

5 Computational Results

To compare the efficiency of two exact solution techniques, large instances with size 15 and 20 are randomly generated and solved using Branch & Bound and Bender's Decomposition separately. Both algorithms are run in CLPEX platform where Branch & Bound and Bender's Decomposition are coupled. The obtained results are tabulated as in Table 3. Based on the below table, it is concluded that both algorithms can solve instances to optimality, however Bender's Decomposition outperforms Branch & Bound with respect to computational time and convergence speed.

Table 3. Optimal Solution Values of Each Objective for Large Instances

	BRANCH & BOUND		BENDERS DECOMPOSITION	
	LATENESS		LATENESS	
OBJ. FUN	15	20	15	20
Objective	47,95	82,045	47,95	81,645
Solution Time(sec.)	33,3	2617	22	2088
Constraints	4208	9408	4208	9408
Variables	4656	10206	4656	10206
Binary	4608	10143	4608	10143
Integer	48	63	48	63
Non-zero Coefficient	22121	50191	22121	50191
Nodes	4993	103881	6386	104033
Iteration	113819	3457114	226682	3793973

6 Conclusions and Future Works

This study focuses on a multi-objective optimization problem in a garment manufacturer in Turkey, IZMIR. The aim of the study is to minimize both total set-up times and total lateness in order to increase the efficiency of sewing line and meet customers' demands on time. In order to solve the problem to optimality, we considered each objective function separately and applied two exact solution techniques, namely Branch & Bound and Benders' Decomposition Technique. To better analyze the performance of proposed solution techniques, we generated three levels of test instances with different model numbers. Based on our results, it is concluded that although Branch & Bound technique can find promising solutions for small and medium instances, Benders' Decomposition technique is a better alternative to solve large instances thanks to its fast convergence to the optimal solution.

As a future direction, try to propose new algorithms and methods with nice properties to deal with such a multi-objective problem could be considered as a general topic for the future study.

References

1. Eren, T., Güner, E.: A bicriteria scheduling with sequence-dependent setup times. Appl. Math. Comput. **179**(1), 378–385 (2006)
2. Fortuny-Santos, J., Ruiz-de-Arbulo-López, P., Cuatrecasas-Arbós, L., Fortuny-Profitós, J.: Balancing workload and workforce capacity in lean management: application to multi-model assembly lines. Appl. Sci. **10**(24), 8829 (2020)
3. Infanger, G.: Monte Carlo (importance) sampling within a benders decomposition algorithm for stochastic linear programs. Ann. Oper. Res. **39**(1), 69–95 (1992)
4. Jafari Asl, A., Solimanpur, M., Shankar, R.: Multi-objective multi-model assembly line balancing problem: a quantitative study in engine manufacturing industry. Opsearch **56**(3), 603–627 (2019). https://doi.org/10.1007/s12597-019-00387-y
5. Karakutuk, S.S., Örnek, M.A.: A goal programming approach to lean production system implementation. J. Oper. Res. Soc., 1–14 (2022)
6. Lockett, A.G., Muhlemann, A.P.: A scheduling problem involving sequence dependent changeover times. Oper. Res. **20**(4), 895–902 (1992)
7. Rahmaniani, R., Crainic, T.G., Gendreau, M., Rei, W.: The Benders decomposition algorithm: a literature review. Eur. J. Oper. Res. **259**(3), 801–817 (2017)
8. Tan, K.C., Narasimhan, R., Rubin, P.A., Ragatz, G.L.: A comparison of four methods for minimizing total tardiness on a single processor with sequence dependent setup times. Omega **28**(3), 313–326 (2000)

Electroencephalograms Processing Methods for Research of Verbal Intelligence

Oksana Kharchenko[1] and Zlatinka Kovacheva[1,2](✉)

[1] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Acad. Georgi Bonchev Str., Block 8, Sofia, Bulgaria
{kharchenko,zkovacheva}@math.bas.bg
[2] University of Mining and Geology, Sofia, Bulgaria

Abstract. In most cases, electroencephalography is used for diagnostic purposes to detect diseases of the nervous system. However, the information contained in the electroencephalogram can be used much more widely, in particular, for the analysis of verbal intelligence. In this paper, the possibility of using various methods of processing electroencephalograms (EEG) for assessing cognitive functions is considered. Currently, the use of numerical-analytical methods for estimation of verbal intelligence is discussed.

One of the most frequent disorders of neurological symptoms is a violation of cognitive functions. Cognitive functions are related to the activity of the brain as a whole. Changes in verbal intelligence seem to relate to the changes in brain functioning. A technique to assess brain functioning is electroencephalography. The methods, used to study cognitive functions using electroencephalograms, have to show the dynamics of the changes.

Fuzzy set theory plays an important role in dealing with uncertainty for decision making in medical applications. Electroencephalograms represent a broadband oscillatory process that is inherently random and non-stationary. Therefore, it is difficult to determine its parameters precisely. In this article, the methods for estimation of the EEG frequencies are presented. Windowed Fourier Transform and Wavelet Transform are applied to electroencephalograms. A comparative analysis of these methods is carried out.

The considered methods make it possible to observe the change in frequency over time as opposed to the Fourier transform. This is important for the evaluation of cognitive functions.

Keywords: Verbal intelligence · Electroencephalography · Windowed Fourier Transform · Wavelet Transform

1 Introduction

Verbal intelligence is the ability to analyze information and solve problems using language-based reasoning. Language-based reasoning may involve reading or listening to words, conversing, writing, or even thinking. Modern world is built around listening to or reading words for meaning and expressing knowledge through spoken language.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 669–675, 2023.
https://doi.org/10.1007/978-3-031-39777-6_78

One of the most frequent disorders of neurological symptoms is a violation of cognitive functions. Cognitive functions are related to the activity of the brain as a whole. Especially often cognitive disorders occur in old age. The increase in life expectancy and, accordingly, the increase in the number of older people makes the problem of cognitive impairment very relevant for neurologists and doctors of other specialties. Cognitive functions are usually understood as the most complex functions of the brain, with which the process of studying the world takes place. Cognitive functions include memory, gnosis, speech, praxis [1].

Obviously, the assessment of verbal intelligence is an important task. Analysis of electroencephalograms (EEG) allows to get quantitative estimates of verbal intelligence [2, 3].

Electroencelography is a method of studying the brain based on the registration of its electrical potentials. An EEG is a complex random process that can be represented as a non-stationary stochastic process. [4, 5]. The spectrum of EEG is quite complicated: The classification of EEG rhythms by some basic ranges is introduced. The concept of "rhythm" on the EEG refers to a certain frequency band matching to a certain brain condition. In addition, the rhythms are unstable, the parameters of instability are also informative for diagnosis. The EEG rhythms of an awake healthy adult includes alpha, beta, mu, gamma - rhythms. Alpha rhythm is characterized by a frequency in the range of $8 \div 13$ Hz, and amplitude of up to 100 μV. It is registered in $85 \div 95\%$ of healthy adult patients. The largest amplitude alpha-rhythm has in a calm and relaxed waking state.

Beta rhythm is characterized by a frequency in the range of $14 \div 40$ Hz, and amplitude of up to 15 μV. Normally, it is very weakly expressed and usually has amplitude of $3 \div 7$ μV.

Types of activity what are pathological of an awake patient include the appearance of slow rhythms: theta and delta. The lower its frequency is and higher its amplitude is, the more pronounced is the pathological process. Theta rhythm is characterized by a frequency of $4 \div 6$ Hz, the amplitude exceeds 40 μV, reaching 300 μV or more in some pathological conditions. The delta rhythm is characterized by a frequency of $0.5 \div 3$ Hz, the amplitude is the same as that of the theta-rhythm.

The classification of EEG rhythms is given in Table 1.

The use of numerical methods in electroencealography is usually limited to the Fourier transform and correlation functions. These traditional numerical methods do not take into account the non-stationarity of the EEG process, i.e. its change over time. This drawback significantly hinders the development of diagnostics. Therefore, the aim of this article is to develop methods that illustrate the change in EEG characteristics over time. This allows determining the in time reaction to any stimulus.

The paper is organized as follows: Section 2 presents Windowed Fourier Transform of the EEG of an adult healthy person. In Sect. 3 we consider Wavelet Transform of the EEG of an adult healthy person. The advantages of the considered methods and suggestions for the future researches are briefly mentioned in the Conclusion.

Table 1. EEG rhythms

EEG rhythms of an awake adult		
Rhythm	Frequency, Hz	Amplitude, μV
α	8 ÷ 13	up to100
β	14 ÷ 40	up to 15 normally 3 ÷ 7
Types of activity that are pathological for an awake adult		
δ	0,5 ÷ 3	exceeds 40 μV, reaching 300 μV or more in some pathological conditions
θ	4 ÷ 6	–«–

2 Estimation of the EEG Frequencies Based on Windowed Fourier Transform

Figure 1 shows the EEG of an adult healthy person. Since the frequency band of 0.5 ÷ 40 Hz is informative, the signal is filtered by a band-pass filter with cut-off frequencies of 0.4 and 45 Hz. This figure clearly shows the non-stationarity of the EEG.

The spectrum of this implementation is shown in Fig. 2. The spectral method in the case of a non-stationary process is incorrect to use [6–8], but, nevertheless, it has found wide application in electroencephalography [4, 5], because it allows to get approximate rhythms of the EEG.

In Fig. 2, a high level of alpha rhythm is observed, which is natural in a state of calm wakefulness. The rest of the rhythms are weakly expressed.

In [2, 3] the calculation of power on six resting-state EEG components (beta, alpha1, alpha2, gamma, theta, and delta) to estimate verbal intelligence is performed. But these articles did not take into account the non-stationarity of EEG.

Windowed Fourier Transform (WFT) is often used to study non-stationary signals, since the EEG frequency varies over time. In fact, the simultaneous localization in the time and frequency domains that any function can achieve is constrained by a lower bound, a statement usually known as the Heisenberg uncertainty principle. In the present situation, such a function is a product of a window function and a complex exponential function.

A WFT is a variation of the Fourier transform defined as follows [9, 10]:

$S(\alpha, t) = \int_{-\infty}^{\infty} W(\beta - t)s(\beta)e^{-j\alpha\beta}d\beta,$

where $s(\beta)$ is the original signal,

$W(\beta - t)$ is the window function.

There are several different types of window functions that can be applied depending on the signal. We apply WFT to implementations of EEG.

Figure 3 illustrated WFT of the EEG. The window size is 0.68 s., the window is rectangular.

Figure 1 shows a slow oscillation between 4 and 5 s. Comparing Fig. 1 and Fig. 3, it can be seen that in the interval between 4 and 5 s. the frequencies are low. Figure 3

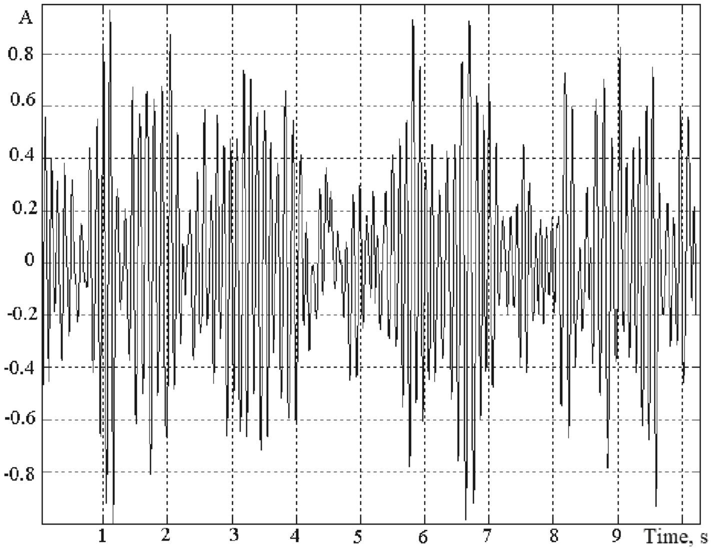


Fig. 1. EEG of an adult healthy person.

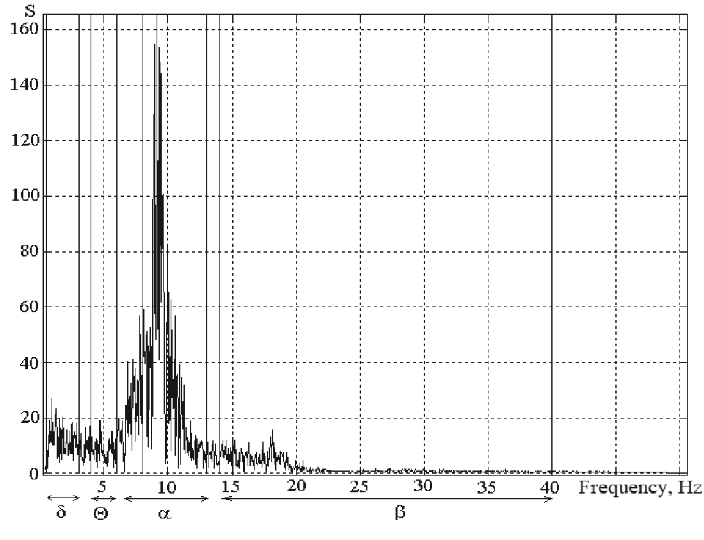


Fig. 2. Spectral density of the EEG shown in Fig. 1.

also shows that the basic energy of the signal is concentrated in the alpha range. But the signal energy in this range is not constant.

The WFT makes it possible to trace the change in frequency over time, but based on Heisenberg's uncertainty principle, it is impossible to accurately determine frequency and time at the same time. Therefore, as the size of the window decreases, uncertainty in frequency increases and vice versa.

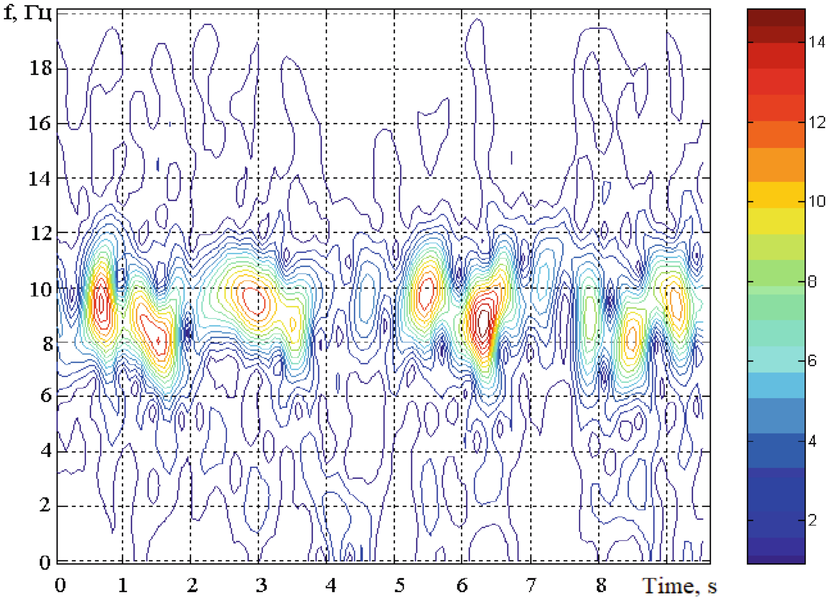


Fig. 3. Windowed Fourier Transform of the EEG shown in Fig. 1.

3 Estimation of the EEG Frequencies Based on Wavelet Transform

The Wavelet Transform (WT) is an alternative to the classical Windowed Fourier Transform. The Windowed Fourier Transform serves as a tool to describe or compare the fine structure of a function at different resolutions. Although quite successful, Fourier analysis is not able to describe highly localized functions.

Currently, Wavelet transform is widely used in the tasks of representing non-stationary processes. WT is a type of spectral analysis in which the simple oscillations are presented by functions, called wavelets. WT provides a time-frequency signal representation and it is applied to overcome some of the WFT problems of the poor resolution.

Wavelets have big advantages over Fourier series both in the general and accurate representation of functions and their various local features. They are represented by a much more diverse set of types than the only sinusoidal function in Fourier Transform [11].

Wavelet analysis is performed similarly to WFT, signal is multiplied with a function (wavelet), as well as with a window function in WFT. However, there are two significant differences between WFT and WT:

- the FT of the wavelet-weighted signal is not performed;
- the width of the window varies.

The signal is represented as a sum of the basic functions $\theta\ (t)$ multiplied by coefficients C_k

$$s(t) = \sum_k C_k \theta_{\,k}(t),$$

where $\theta\left(\frac{t-b}{a}\right)$ is the basis of wavelets;

b is the shift,

a is the scale.

Figure 4 illustrates WT of the EEG 1. Simlets wavelets are used. These wavelets belong to orthogonal wavelets with a compact carrier; they are close to symmetrical ones [12].

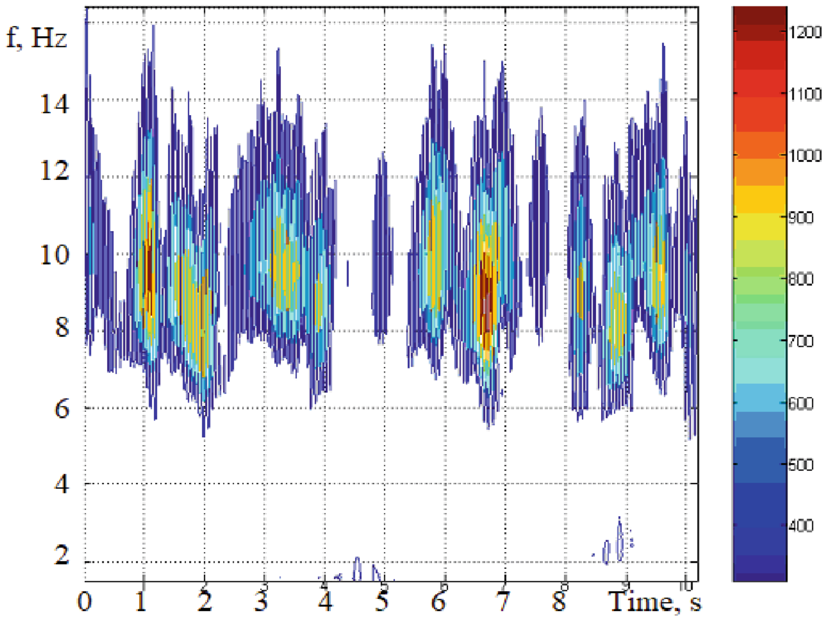


Fig. 4. Wavelet Transform of the EEG shown in Fig. 1.

Comparing Fig. 1 and Fig. 4, it can be seen that in the interval between 4 and 5 s. the frequencies are almost the same. Figure 4 also shows that the basic energy of the signal is concentrated in the alpha range. But the signal energy in this range is not constant. Comparing Fig. 3 and Fig. 4, it can be seen that in Fig. 4 low frequencies are expressed more clearly.

4 Conclusion

Cognitive functions characterize verbal intelligence and they can be investigated using the methods of analysis of EEG used in this work. It is shown that in contrast to traditional medical electroencephalography, aimed at identifying pathologies, these methods can be used much more broadly to determine the cognitive functions, which makes it possible to obtain quantitative assessments of verbal intelligence.

Windowed Fourier Transform and Wavelet Transform are considered. The presented methods make it possible to trace the dynamics of changes in the EEG. They make it possible to numerically characterize the duration and change of the basic physiological rhythms, as well as to monitor the change in frequency over time within each rhythm. That is important to study cognitive functions.

Analysis of a healthy person based on these methods showed that the basic energy of the signal is concentrated in the alpha range. But the signal energy in this range is not constant. There is also a short band of low frequencies.

Low frequencies are expressed more clearly with WT. This is due to the better resolution of WT compared to WFT.

For the further researches, applying the signal processing methods used in radio engineering to EEG signal processing has a perspective. For example, the Hilbert Transform can be applied, because frequency and amplitude are functions of time in this method.

Acknowledgment. This work was supported by a grant from The Simons Foundation International (grant no. 992227, IMI-BAS).

References

1. Zakharov, V.: Cognitive impairment in neurological practice. Difficult Patient **3**(5), 4–9 (2005)
2. Boyle, R., et al.: Verbal intelligence is a more robust cross-sectional measure of cognitive reserve than level of education in healthy older adults. Alzheimers Res. **13**(1), 128 (2021)
3. Ferrari-Díaz, M., Bravo-Chávez, R., Silva-Pereyra, J., Fernández, T., García-Peña, C., Rodríguez-Camacho, M.: Verbal intelligence and leisure activities are associated with cognitive performance and resting-state electroencephalogram. Front. Aging Neurosci. Sec. Neurocognitive Aging Behav. **14**, 1–13 (2022)
4. Zenkov, L.P., Ronkin, M.A.: Functional Diagnostics of Nervous Diseases: Guide for the Doctor. Medichina, Moscow (1991). (in Russian)
5. Zenkov, L.P.: Clinical Electroencephalography (with Elements of Epileptology). A Guide for Doctors. Medpressinform, Moscow (2004). (in Russian)
6. Baskakov, S.I.: Radiotechnical Circuits and Signals. Vyssh. School, Moscow (1983). (in Russian)
7. Voloshchuk, Yu.I.: Signals and Processes in Radio-Technics. Xarkhiv OO Company smith (2005). (in Ukranian)
8. Siebert, W.: Circuits, Signals, and Systems. MIT Press, Computers (1986)
9. Torre, A.: The Wigner Distribution Function: Optical Production, Linear Ray and Wave Optics in Phase Space (2005)
10. Jonas Gomes, J., Velho, L.: From Fourier Analysis to Wavelets, IMPA, vol. 3 (2015)
11. Bassam, N.A., Ramachanrad, V., Parameswaram, S.E.: Wavelet theory and application in communication and signal processing, pp. 145–256 (2021)
12. Merry, R.J.E.: Wavelet theory and applications a literature study, Eindhoven University of Technology Department of Mechanical Engineering Control Systems Technology Group Eindhoven, pp. 5–41 (2005)

Site Selection of Solar Powered Electric Vehicle Smart Charging Station: Istanbul Case

Serra Mandaci and A. Çağrı Tolga(✉)

Industrial Engineering Department, Galatasaray University, Ortakoy, 34349 Istanbul, Turkey
ctolga@gsu.edu.tr

Abstract. In parallel with the goal of reducing the use of fossil fuels, adopted by many authorities on a global scale within the scope of creating a sustainable future, the use of electric vehicles has started to increase, and it is expected to become more widespread in Türkiye in the upcoming years. However, the prevalence of electric vehicle charging stations (EVCSs) is a factor that directly affects consumer appetite. From this perspective, to make the transition to electrification in transportation in the smoothest way, the city of Istanbul, which is the most populous city of Türkiye, should be prepared in terms of infrastructure. In this study, when the available opportunities are evaluated, İSPARK A.Ş., the subsidiary of Istanbul Metropolitan Municipality that provides parking services, is seen as a great opportunity for this transition. Among the existing car parks in Istanbul, a site selection and ranking are made with weighted Euclidean distance-based approach (WEDBA) to install EVCSs. The feasibility of installing an EVCS that generates electricity from solar energy, i.e., photovoltaic charging station (PVCS), with micro wind turbines as an intelligent backup system, is examined among 150 car parks. Considering the potential solar energy productivity and the mean density of the wind, car parks are ranked for each of the 3 DC charging station power scenarios.

Keywords: EVCS · PVCS · MCDM · WEDBA

1 Introduction

Reducing fossil fuel dependence is recognized as one of the major concerns in the world today for creating a sustainable future and therefore dissemination of electric vehicles is supported by different authorities around the world. At earlier times, electric vehicles were not developed enough to be competitive compared to the gas-fueled cars. But in today's world, most of the well-known automobile brands have developed electric and/or hybrid models with intelligent systems to offer environment-friendly products worldwide. With the positive developments in battery technologies, production of electric vehicles has been greatly affected. But lack of availability of recharging infrastructure influences the customer appetite in electric vehicle market and this situation is mostly observed in the developing countries such as Türkiye. According to the Sabanci University - Istanbul International Center for Energy and Climate's (IICEC) report on Türkiye

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 676–683, 2023.
https://doi.org/10.1007/978-3-031-39777-6_79

Electric Vehicle Outlook 2021 [12], even if the electric vehicle sales increased in 2021, the market share is still below 1% in Türkiye. It is aimed to achieve a more rapid increase in the market share with the expected significant impact of the introduction of TOGG (Türkiye's Automobile Enterprise Group, introduced from the beginning as an electric vehicle producer). There are also other ongoing projects of other brands within Türkiye for production of electric vehicles to support the electrification of transportation sector but there is still a significant gap in the electric vehicles charging stations. As stated in IICEC's report, total number of charging stations in Türkiye was above 3,000 due to the private sector investments.

On city basis, it can be stated that densely populated cities require more investments than others. Istanbul is a mega city in Türkiye with almost 16 million inhabitants and Istanbul Metropolitan Municipality (IMM) is the local authority which offers local services. When the transition to electrification of the transportation sector accelerates, Istanbul's infrastructure should be ready for a smooth transition to respond to the needs of the residents along with intelligent systems. With this objective in mind, İSPARK A.Ş. Which is one of the subsidiaries of IMM and that provides car parking service in many spots of the city, could be supportive for expanding the electric vehicles charging stations infrastructure. İSPARK operates in 37 districts out of 39 districts in Istanbul and has 745 car parks with different sizes and capacity. 212 of the car parks are in the Asian side and 533 are in the European side. [13] With this information in hand, these car parks can be considered as an opportunity to provide EVCSs. This study makes a scientific site selection with multi-criteria decision-making among İSPARK car parks and assign rankings in terms of their efficiency to install an EVCS, that generates electricity from solar energy for contributing to yield a sustainable future. Also, micro wind turbines are added to the power generation system as a backup.

Over the past years, since electric vehicles have become a hot topic, various methods have been applied by scholars to help decision makers effectively determine the location and/or size of the proposed EVCSs. A great portion of these studies include multi criteria decision making. Guo & Zhao [3] studied optimal site selection problem of EVCS with fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS); Feng et al. [2] developed an integrated multi-criteria decision-making (MCDM) method through the linguistic entropy weight method and fuzzy axiomatic design for selecting a suitable site for EVCS; Erbaş et al. [1] applied a geographic information system (GIS)-based multi-criteria decision analysis approach to address the EVCS site selection in Istanbul. Then, this study used fuzzy analytical hierarchy process (AHP) and TOPSIS methods for choosing the optimal EVCS sites. Liu et al. [8] introduced multiple interrelated criteria for EVCS site selection problems and proposed an integrated MCDM approach by a grey decision-making trial and evaluation laboratory (DEMATEL) with uncertain linguistic multi-objective optimization by ratio analysis and full multiplicative form (UL-MULTIMOORA) for determining the most suitable EVCS site. There are also some studies where EVCSs are powered by solar energy, i.e., PVCS. But they are not very common compared to the others. Khan et al. [7] conducted a comprehensive review on solar powered EVCS, and their study includes various technical details on solar powered EVCS. Ji et al. [4] studied optimizing the locations and sizes of solar assisted EVCSs in an urban area. Zhou et al. [11] proposed a practical model with GIS and TODIM (an acronym in Portuguese of interactive MCDM, Tomada de Decisao Interativa Multicriterio) for location selection of PVCS and conducted empirical research for the city of Beijing. There are two more EVCS site selection research in the current literature to the best of authors' knowledge, that include analysis on Istanbul. Kaya et al. [5] identified the most suitable charging location alternatives based on GIS and they evaluated the current and proposed EVCS location alternatives with multi-criteria decision approach. Kaya et. al[6] focused on the exact locations of charging stations for five large EVCS companies in Istanbul and developed a three-step methodology for analyzing those locations.

Overall, there are many studies carried out on optimal EVCS site selection and ranking with different methods, but few studies exist on PVCS site selection. It is obvious that MCDM methods are widely found useful in selection problems since they help combining qualitative and quantitative characteristics.

In the upcoming section, materials and methods included in this study are defined, application is presented, with minimum required slot calculation and appropriate multi-criteria that are introduced for selection among car parks. Then, results are presented in the next section and suggestions for further analysis are given in conclusion.

2 Materials, Method, and Application

As mentioned earlier, İSPARK has 745 car parks with different sizes and capacity. Among these car parks, 469 of them are available for public use while others are only open for subscriber use or are completely rented by rent-a-car companies or different firms. There are basically three types of car parks: open, multistorey and on-street car parks. Open car parks are subject to this study due to the photovoltaic nature of planned charging stations; therefore 469 car parks are narrowed down to 152. Also, when dropping the parking data with incomplete information, 150 car parks are left in the main data. A location-allocation problem model is built based on the criteria that is specific to the MCDM problem. A ranking is assigned to the car parks based on their suitability for installing a photovoltaic charging station with micro wind turbines and the application is evaluated. Materials and methods that are used in this study can be found in this section in detail along with application.

2.1 WEDBA

As Rao & Singh [9] defined, The Weighted Euclidean Distance Based Approach calculates weighted distances of various alternatives from the most and least desired situations. The steps of WEDBA are defined below:

Step 1: Construct the decision matrix. The i^{th} alternative with n attributes is denoted as $Y_i = \{y_{i1}, \ldots, y_{ij}, \ldots, y_{in}\}$ where y_{ij} is the performance value of attribute j for alternative i. ($for\ i = 1, \ldots, m; j = 1, \ldots, n$)

$$D = \begin{bmatrix} y_{11} & \cdots & y_{1j} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{i1} & \cdots & y_{ij} & \cdots & y_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mj} & \cdots & y_{mn} \end{bmatrix} \tag{1}$$

Step 2: Standardize the decision matrix. An important property of a standard score is that it has a mean of 0 and a variance of 1 (i.e., the standard deviation is 1).

$$D' = \begin{bmatrix} Z_{11} & \cdots & Z_{1j} & \cdots & Z_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ Z_{i1} & \cdots & Z_{ij} & \cdots & Z_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ Z_{m1} & \cdots & Z_{mj} & \cdots & Z_{mn} \end{bmatrix} \quad (2)$$

where,

$$Z_{ij} = \frac{x_{ij} - \mu_j}{\alpha_j} \quad (3)$$

$$x_{ij} = \frac{y_{ij}}{max_j \sum y_{ij}}; \text{ if } j^{th} \text{ attribute is considered beneficial} \quad (4)$$

$x_{ij} = \frac{min_j(y_{ij})}{y_{ij}}$; if j^{th} attribute is considered non - beneficial(5)

$$\mu_j = \frac{1}{m} \sum_{i=1}^{m} x_{ij} \quad (6)$$

$$\alpha_j = \sqrt{\frac{\sum_{i=1}^{m} (x_{ij} - \mu_j)^2}{m}} \quad (7)$$

Step 3: Define the ideal and non-ideal points. The ideal points that are denoted by 'a*' and anti-ideal points, denoted by 'b*' are extracted from standardized decision matrix.

$$a^* = \left\{a_j^*\right\} \text{ and } b^* = \left\{b_j^*\right\} \text{ where } j = 1, \ldots, n \quad (8)$$

Step 4: Decide the attribute weights. The weights can be determined as objective weights from data in a decision matrix or the subjective preferences of the decision maker, or they can be determined by combining objective weights and subjective preferences. Rao [10] proposed the entropy method to calculate the objective weight of attribute importance. Entropy is using probability theory to calculate the uncertainty in information. If an attribute has very similar values amongst all the alternatives, it is considered less important than other attributes, therefore a small weight is assigned to that attribute since it does not help to differentiate alternatives. The entropy value E_j can be measured with the amount of information that is included in the decision matrix and associated with each attribute as in the following equation:

$$E_j = \frac{-\sum_{i=1}^{m} p_{ij} \ln p_{ij}}{\ln m}; \text{ j} = (1, \ldots, \text{n}) \quad (9)$$

where performance rating p_{ij} is,

$$p_{ij} = \frac{y_{ij}}{\sum_{k=1}^{m} y_{kj}} \quad (10)$$

The value of attribute j for alternative i is y_{ij} and m is the number of alternatives. The calculation of the degree of divergence (d_j) of the average intrinsic information contained by each attribute is given as:

$$d_j = 1 - E_j;\ (1 \leq j \leq n) \tag{11}$$

For each attribute j, the objective weight w_j is calculated as:

$$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j} \tag{12}$$

Step 5: Calculate weighted Euclidean distance, index score and ranking. The concept of WEDBA basically requires that the chosen alternative should have the shortest distance to the ideal solution whereas it should have the farthest distance to the non-ideal solution. The distance is calculated in the Euclidean form that measures the distance between two points. For each alternative, the overall performance index score can be determined by its Euclidean distance to ideal and non-ideal solutions. The alternatives are not directly compared among themselves but with the ideal and non-ideal solutions, therefore the distance is interrelated with weights of the attributes and should be incorporated in the distance measurement. For this reason, weighted Euclidean distances are considered in WEDBA. As known in the mathematics literature, the Euclidean distance between two points, say P and Q, in n dimensional space is basically the length of the line segment, PQ. If points are to be represented in a Cartesian coordinate in Euclidean n-space, say $P = (p_1, \ldots, p_n)$ and $Q = (q_1, \ldots, q_n)$, the distance is calculated by $d(PQ) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}$.

WED_i^+ stands for the weighted Euclidean distance between an alternative i and ideal point $a*$ and WED_i^- for the distance between an alternative i and non-ideal point $b*$.

$$WED_i^+ = \sqrt{\sum_{j=1}^{n} \left\{w_j.(Z_{ij} - a_j^*)\right\}^2};\ i = (1, \ldots, m) \tag{13}$$

$$WED_i^- = \sqrt{\sum_{j=1}^{n} \left\{w_j.(Z_{ij} - b_j^*)\right\}^2};\ i = (1, \ldots, m) \tag{14}$$

Therefore, the index score of an alternative i is,

$$Index\ Score_i = \frac{WED_i^-}{WED_i^+ + WED_i^-} \tag{15}$$

The desirability of an alternative increases when the index score is higher because the index score shows the relative closeness of an alternative to the ideal solution. The alternatives are ranked according to their index scores.

2.2 Application

WEDBA method is chosen for this study because the carparks that are closer to the optimum points are to be determined and the weights of the criteria can either be objective,

Table 1. Details of evaluation criteria for the WEDBA problem.

Criteria	Explanation	Unit	Data Source
Photovoltaic Power Output (PVOUT)	The power produced per installed PV capacity unit per year. (On a coordinate basis)	kWh/kWp	Global Solar Atlas [14]
Mean Power Density of the Wind	It is the mean annual power available per square meter of swept area of a turbine and is calculated for 10 m height above ground. (On a district basis)	W/m^2	Global Wind Atlas [15]
Capacity	Number of vehicles that can be accommodated in the car park	Quantity	İSPARK
Area	The car park should have a certain area to install the PVCS and the possibility of expanding the capacity	m^2	İSPARK
Total Number of Parking in 2022	Indicative measure for possible yearly demand	Quantity	İSPARK
Average Number of Parking Per Day in 2022	Indicative measure for an average daily power demand for the PVCS	Quantity	İSPARK
Average Vehicle Parking Time	Indicative measure for the possible length of use of charging units	Minute	İSPARK
Parking Tariff (Full Day)	Supportive in terms of pricing and demand	Turkish Lira	İSPARK

subjective, or integrated. All the criteria used in the application are given in Table 1, which are all objective and treated as benefit type.

Three scenarios are analyzed for the power of the charging units: 120 kW, 180 kW and 360 kW, and 500-W solar panels are chosen for the power generation. When the 500-W solar panels on the market are examined, it is seen that their dimensions are 1.1 m x 2.2 m on average. The dimension of each parking slot in Istanbul must be 1.5 m x 2.5 m for cars [16]. Table 2 shows the minimum required slot calculation for each of the scenario to eliminate the carparks that do not have enough slots to install EVCS.

As a preliminary requirement of this study, car parks that do not have enough slots to install solar panels are avoided in accordance with the calculations shown in Table 2.

Table 2. Minimum required slot calculation.

DC charging station power	Required #of Panels	Required Area of the Panels (m^2)	Construction Flexibility (%10) Added (m^2)	Corresponding Area for Parking Slots (m^2)	Minimum Required Slots
360 kW	720	1,742.4	1,916.64	153.33	154
180 kW	360	8,71.2	958.32	76.67	77
120 kW	240	580.8	638.88	51.11	52

3 Results

Considering the potential solar energy productivity and the mean density of the wind, 5 car parks are suggested for each of the 3 scenarios: 120 kW, 180 kW and 360 kW, and 500-W solar panels are chosen for the power generation. After applying above methodology to the alternative set, below results in Table 3 are obtained:

Table 3. Ranking of the car parks obtained from WEDBA.

RANKING		
For 120 kW	For 180 kW	For 360 kW
Eski Salı Pazarı	Eski Salı Pazarı	Eski Salı Pazarı
Büyükçekmece Sosyal Tesisleri Arkası	Avcılar Meydan Açık Otoparkı	Avcılar Meydan Açık Otoparkı
Bostancı İDO Karşısı Açık Otopark	Sultanahmet Açık Otoparkı	Avcılar Denizköşkler Toplantı ve Davet Salonu Açık Otoparkı
Kartal Savarona Sahil Cep Otoparkı	Avcılar Denizköşkler Toplantı ve Davet Salonu Açık Otoparkı	Ortaköy Açık
Avcılar Meydan Açık Otoparkı	İspark West İstanbul Marina Açık Otoparkı	Kadıköy Açık Otoparkı

The results presented in Table 3 indicate that Eski Salı Pazarı comes as first to install an EVCS for various power alternatives. Since the values of PVOUT and mean power density of the wind data were very close to each other among the alternatives, it was observed that the criterion weights were low and did not significantly affect the result. The most outstanding attribute of the closest alternative to the optimum is that circulation is quite high in the sense of average number of parking per day along with average parking duration. Practically considering that the population density of the district Kadıköy (where this car park is located) is high during the day and it is in a central location, it seems quite consistent that this car park stands out in the results.

4 Conclusion

The purpose of this study was to investigate the existing carparks of the IMM's subsidiary İSPARK A.Ş. in terms of their feasibility to install photovoltaic charging stations that have a micro wind turbine as a smart backup system. Considering the potential solar energy productivity and the mean density of the wind, 5 car parks are suggested for each of the 3 scenarios. Further analysis can be done by adding more attributes, using different methodologies, and comparing the results of those can be supportive to reach a more detailed result. Also, economic feasibility, costs and risks associated with this project of installing these charging stations can be analyzed.

References

1. Erbaş, M., Kabak, M., Özceylan, E., Çetinkaya, C.: Optimal siting of electric vehicle charging stations: a GIS-based fuzzy multi-criteria decision analysis. Energy **163**, 1017–1031 (2018)
2. Feng, J., Xu, S.X., Li, M.: A novel multi-criteria decision-making method for selecting the site of an electric vehicle charging station from a sustainable perspective. Sustain. Cities Soc. **65**, 102623 (2021)
3. Guo, S., Zhao, H.: Optimal site selection of electric vehicle charging station by using fuzzy TOPSIS based on sustainability perspective. Appl. Energy **158**, 390–402 (2015)
4. Ji, D., Lv, M., Yang, J., Yi, W.: Optimizing the locations and sizes of solar assisted electric vehicle charging stations in an urban area. IEEE Access **8**, 112772–112782 (2020)
5. Kaya, Ö., Tortum, A., Alemdar, K.D., Çodur, M.Y.: Site selection for EVCS in Istanbul by GIS and multi-criteria decision-making. Transp. Res. Part D Transp. Environ. **80**, 102271 (2020). ISSN 1361-9209. https://doi.org/10.1016/j.trd.2020.102271
6. Kaya, Ö., Alemdar, K., Campisi, T., Tortum, A., Kayaci Çodur, M.: The development of decarbonisation strategies: a three-step methodology for the suitable analysis of current EVCS locations applied to Istanbul. Türkiye Energies. **14**, 2756 (2021). https://doi.org/10.3390/en14102756
7. Khan, S., Ahmad, A., Ahmad, F., Shafaati Shemami, M., Saad Alam, M., Khateeb, S.: A comprehensive review on solar powered electric vehicle charging system. Smart Sci. **6**(1), 54–79 (2018)
8. Liu, H.C., Yang, M., Zhou, M., Tian, G.: An integrated multi-criteria decision-making approach to location planning of electric vehicle charging stations. IEEE Trans. Intell. Transp. Syst. **20**(1), 362–373 (2019). https://doi.org/10.1109/TITS.2018.2815680
9. Rao, R.V., Singh, D.: Evaluating flexible manufacturing systems using Euclidean distance-based integrated approach. Int. J. Decision Sci. Risk Manag. **3**(1/2), 32–53 (2011)
10. Rao, R.V.: Decision Making in the Manufacturing Environment Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods. Springer-Verlag, London (2007)
11. Zhou, J., Wu, Y., Wu, C., He, F., Zhang, B., Liu, F.: A geographical information system based multi-criteria decision-making approach for location analysis and evaluation of urban photovoltaic charging station: a case study in Beijing. Energy Convers. Manage. **205**, 112340 (2020)
12. Türkiye Electric Vehicle Outlook. https://iicec.sabanciuniv.edu/tevo
13. İSPARK carparks. https://ispark.istanbul/otoparklarimiz/
14. Global Solar Atlas. https://globalsolaratlas.info/map
15. Global Wind Atlas: https://globalwindatlas.info/en
16. Istanbul Parking Lot Regulation. https://www.ibb.istanbul/BBImages/Slider/Image/otopark-yonetmeligi.pdf

Evaluation of Interior Spaces in Educational Areas in Terms of Prospect and Refuge Theory Using Fuzzy Logic

Kübra Yılmaz Şenel(✉) and Aslı Çekmiş

Istanbul Technical University, Istanbul, Turkey
{yilmazkub,cekmis}@itu.edu.tr

Abstract. This study aims that how the physical conditions of the educational areas are perceived psychologically by the people in that environment. Also, It reports examination of educational areas using fuzzy logic in the context of prospect and refuge theory. The theory of "prospect and refuge" seeks to describe why certain environments feel secure and thereby meet basic human psychological needs. Environments that meet such needs will often provide people with the capacity to observe (prospect) without being seen (refuge) (Dosen and Michael 2013). Prospect and refuge theory has usually been applied to landscapes but recent works suggest that it could also be applied to interior spaces. In order to do this, the effects of the study area on human psychology are examined using fuzzy logic based on past studies. This study shows that; in terms of prospect and refugee theory, the way users perceive educational areas and the outputs of the system created using fuzzy logic are enormously similar. It is possible to create productive systems in future research that will increase this perceptual experience by developing sample educational area models that are produced as a result.

Keywords: prospect and refuge theory · fuzzy logic · educational areas

1 Introduction

Today, it is undeniable fact that student's behaviors are affected from environment in which they live and work. In fact, learning and teaching don't happen in the space. So, the existence of a physical, social and psychological environment is necessary. Firstly, the environment which consist of various buildings, equipments, models, arrangements, temperature, color and various objects is shaped by humans, then humans are shaped by the same environment. According to the educators, learning may occur in a suitable environment in terms of physical, social and psychological aspects. For effective education, this environment needs to be organized in a way that is appropriate to the teaching and learning activities. This necessitates organizing and directing the various interaction dimensions between the environment and the individual according to the educational objectives (Küçükoğlu and Özerbaş 2010).

As a result, many recent studies show that the physical quality of the building is correlated with student achievement, learning performance and motivation (Al 2014).

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 684–691, 2023.
https://doi.org/10.1007/978-3-031-39777-6_80

The education and training process is usually carried out in classrooms, which are important components of educational structures. The realization of the educational philosophy and training program is directly related to the design of the classrooms. Designing the classrooms in accordance with the right architectural properties plays an important role in the success of individuals in the process of learning and teaching (Diker 2019).

In this study, especially the psychological effect of the study area on students is examined through the prospect and refuge theory with fuzzy logic method.

2 Prospect and Refuge Theory

Fig. 1. (1) Strong refuge, (2) Prospect-refuge balanced, (3) Prospect (original image).

Many researches show that preferences for environments depends on both prospect (farthest visible point) and refuge (hiding against possible enemies) (see Fig. 1). The reason for this situation is the human instinct to survive. If the one is in a position where one can see the whole environment and detect possible enemies, then one feels prospect because the more one can perceive the environment, the more one feels safe.

Prospect and Refuge theory has usually been applied to landscapes but recent works suggest that it could also be applied to interior spaces. Interior refuge is created by darkness, solid walls rather than windows, closure on three sides, small horizontal dimensions, narrow horizontal aspect ratio, and, especially, low ceilings (Hildebrand 1999, p. 32). Interior prospect is created, conversely, by bright light, large horizontal, by transparent surfaces, and especially by large vertical dimensions. A connection is proposed between gradient of light (dark to light vs light to dark). In particular, it is suggested that, if one is moving from darker to lighter, one would be able to see without being seen, and so it will be possible to explore with safety and hence produce pleasure (Hildebrand 1999, p. 54). "Refuge and prospect are opposites: refuge is small and dark; prospect is expansive and bright. It follows that they cannot exist in the same space. They can occur contiguously, however, and must, because we need them both, and we need them together. From the refuge we must be able to survey the prospect; from the prospect we must be able to retreat to the refuge" (Hildebrand 1999, p. 22).

People's perception of the environment and their psychological reactions to it are subjective. Likewise, researchers acknowledge that studies on this subject are contradictory. For instance; According to Wohlwill and Heerwagen, the way people perceive environments is clearly different depending on their physical stature and size or the personal experiences (Wohlwill 1976; Heerwagen 1993). Although this, architectural and design texts tend to selectively emphasize the universal aspects of spatial experience (Kellert 2005; Lippmann 2010). This cannot be totally inconsequent since people do experience aspects of space, including its organization, in similar ways (Ellard 2009). For example, scientific studies have shown that a close visual connection between habitable space and nature is beneficial for psychological wellbeing, recovery and stress relief (Kaplan and Kaplan 1989; Heerwagen and Orians 1993). Studies have also observed that restricted views may cause negative reactions (Heerwagen 2008) while visual connections might encourage movement and evoke pleasure through the exploration of space (Kaplan 1987).

3 Methodology

In this study, fuzzy logic inference system is created by using MatLab program. In the study area model, which was designed with fuzzy logic, the inputs of the system; ratio of window area to wall area, depth and view, while the output of the system creates the prospect and refuge value. The sample study area models created based on this study can be seen in Fig. 2. The designed sample study area models vary between 0 and 1 in terms of prospect and refuge values through the system outputs.

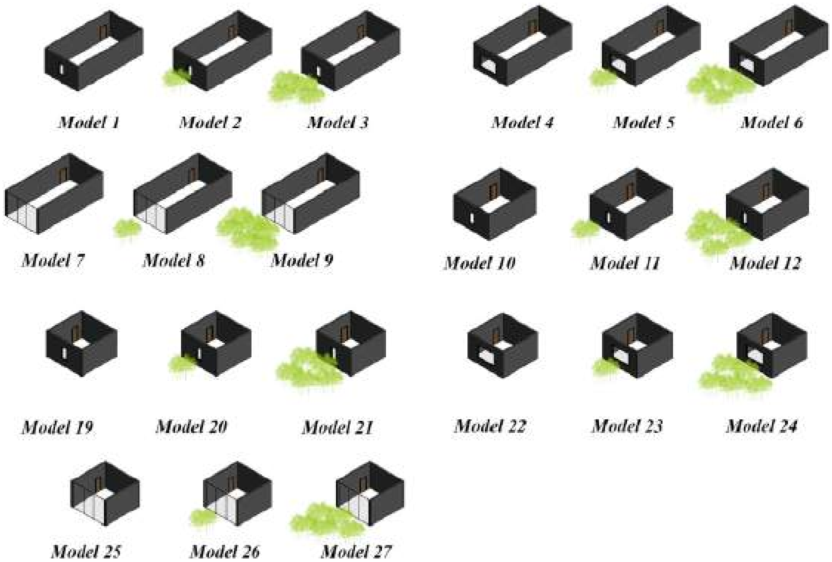


Fig. 2. Study area models

In the model, fuzzy model as the Sugeno type and "weighted average" method as the clarification method were used. In membership functions, Values are expressed for WWA "high-middle-low", for depth "high-middle-low" and for view "high-middle-low". For system output which represents prospect and refuge values, three functions are defined as "refuge, balanced and prospect". For the fuzzy logic model, 27 fuzzy rules have been formed in the type of IF, which are associated with the ratio of the window area to wall area (WWA), the depth and the view to the defined prospect and refuge value of area. The rules defined in the system are generally based on the principle that the prospect and refugee value increases when the WWA ratio increases, the depth decreases and the view increases.

3.1 Inputs Selected in Fuzzy Logic Model

In order to evaluate the prospect and refuge value of education areas in MATLAB2022a SDU-student- version, a sugeno-type fuzzy inference systems was used. Three inputs (parameters) were selected for this application:

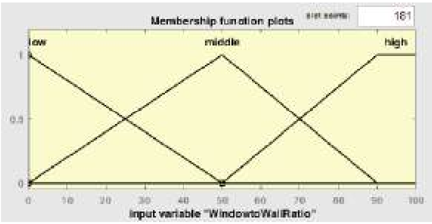


Fig. 3. Membership functions of WWR

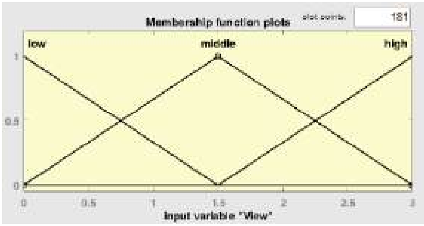


Fig. 4. Membership functions of the view

The Ratio of Window to Wall Area: Windows are designed for ventilating of the space, illuminating of the space and creating a visual connection between building and environment. For users of the educational structures, The daylight, which will be taken from the windows during the day, is crucial in terms of physical and mental health.

As part of an effort to understand the psychological impact of environments, prospect-refuge has been proposed to explain the preferred environment as one in which people cannot be seen by others, but can see others. Within the space that defines this concept, there is provision for restoration, concentration and safety, ultimately improving comfort. Thus, the spatial elements that provide good prospect and refuge conditions have been extensively studied in research on visual preference. For example, some researchers have mentioned that, for interior spaces, fenestration helps optimize visual access to indoor or outdoor areas, enhancing prospect-refuge conditions. In particular, the design of windows provides a sense of openness, which has a relaxing psychological effect and characterizes visual comfort in buildings. Moreover, window design, which is difficult to change later, should be considered in the early stages of the architectural design process, considering multiple aspects of design at the same time (Hwang and Lee 2018).

According to the Educational Structures Minimum Design Standards Guide (2015), regulation on designing window and the taking in daylight are explained as follows:

1. A minimum of 60% of the learning and teaching area floor area, with a target of 90%, must receive a minimum daylight factor of 2.5 as measured at the floor level under a uniform design sky.
2. A minimum of 60% of the learning and teaching area floor area, with a target of 90%, must have a direct line of sight externally.

3. A minimum of 60% of the learning and teaching area floor area, with a target of 90%, must have a clear line of sight to a high quality internal or external view.

According to these data, Three membership functions were created for the ratio of the window to Wall Area. These was evaluated as low, medium and high. The sub-sets (MF's) for the Window to Wall ratio: "low" trimf [0 0 50], "middle" trimf [0 50 90] and "high" trapmf [50 90] is given in Fig. 3. The situation between this input and the output is as follows: the closer is the window to wall ratio to the best, the more advantageous it is.

The View: Although there are contradictions in the interior prospect and refugee studies, some results are obtained. In several of the studies, people rated views from the interior to nature or to adjacent spaces as being preferred over a lack of natural views or of being totally enclosed in place. According to Appleton (1975) proposed that environments which offer a balance of both natural outlook and enclosure of space would be most preferred. After then, he advised that in many cases a "balance […] can be achieved from serial vision, involving the successive experiences of exposure to strongly contrasting landscape types, strong prospect and then strong refuge, is more potent than that which comes from trying to achieve a balance all at once" (Appleton 1984, p 102). However, there is proof that people have a preference for spaces which feature access to nature and daylight (Ulrich 1993) and for environment that support social interaction and thereby create safer environments (Kuo and Sullivan 2001; Ellard 2015). Besides, a preference for water views, which are perceived as calming, is often taked in studies (Ulrich 1984; Kaplan and Kaplan 1989; Heerwagen and Orians 1993).

According to these data, three membership functions were created for the view. These was evaluated as low, middle and high. Some assumptions are made for inputs. 0 represents the situation where there is lack of natural views or being totaly enclosed. 1 represents the situation where there is some natural view and view built environment. 2 represents the situation where there is spectacular natural view (forest, garden, waterfalls etc.). The sub-sets (MF's) for the View: "low" trimf [0 0 1], "middle" trimf [0 1 2] and "high" trimf [1 2 2] (see Fig. 4).

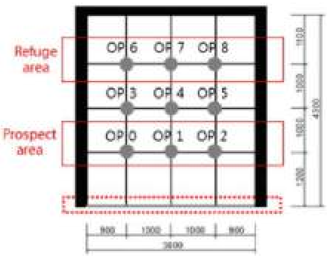


Fig. 5. Prospect and refuge areas (Hwang and Lee 2018)

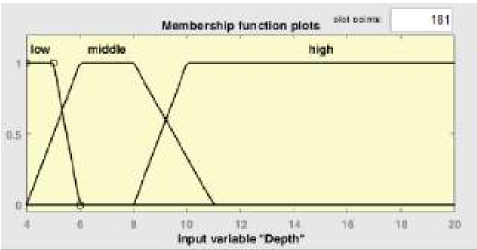


Fig. 6. Membership functions of the depth

The Depth: According to Hwang and Lee's study which is called "Parametric Model for Window Design Based on Prospect-Refuge Measurement in Residential Environment", the distance to the windows increases, the feeling of shelter will increase (refuge) (see Fig. 5). The minimum depth in the study areas is determined as 4 m. The depth between 4–6 m is defined as a less deep field. The depth between 4–6 m is defined as a less deep field. The depth between 6–11 m is defined as middle deep field. The depth between 10–20 m is defined as very deep field.

Based on these, three membership functions were created for the depth of place. These was evaluated as low, medium and high. The sub-sets (MF's) for the Window to

Fig. 7. Fuzzy outputs (P&R) and survey results on created models

Wall ratio: "low" trapmf [4 5 6], "middle" trapmf [4 6 8 11] and "high" trapmf [8 10 20] (see Fig. 6).

3.2 Outputs Selected in Fuzzy Logic Model

In Sugeno type Fis, three membership functions were created for prospect and refuge values. The prospect and refuge functions was evaluated as prospect, balanced and refuge. The fuzzy rule base includes the IF, THEN type rules written between logical input and output variables (Şen 1999; Tayfur 2003; Kazanasmaz 2010). In this paper, 27 fuzzy rules have been created, which are related to the prospect and refuge, the depth, the view, and the window to wall ratio.

The part that starts with (if) and goes until (then) is based on the Depth, Window to Wall Area and the View. The part after (then) forms a situation on the prospect and refuge value. The fuzzy inference engine generates fuzzy output from a given input set, taking all the rules in the fuzzy rule base into account. Clarification converts the fuzzy outputs generated by the fuzzy inference engine into a single number (Şen 1999; Tayfur 2003; Kazanasmaz 2010). There are clarification methods such as Centroid, bisector, mom, lom, som etc. In this paper, centroid was used as the clarification method.

4 Conclusion

In this paper, a fuzzy logic algorithm was formed in order to grade in terms of prospect and refuge theory of the working spaces which are the main spaces of the educational areas. The model was constructed by using the parameters of the depth, the Window to Wall Ratio and the View. As a result of evaluations and comparisons of study area models, a prospect and refuge value was established.

In order to understand the psychological impact of sample study areas on people, a survey was conducted with 50 people. The models of these areas were shown and their properties were specified. They were asked how they felt when they thought they were in these areas (prospect-balanced - refuge). The results of the survey and the results obtained from the fuzzy model were compared. The Superpose of results are shown in figure above. (see Fig. 7) According to this data, when the survey results are compared with the fuzzy logic model, there are mostly similarities in the outputs. When the sample models recommended for the study areas are examined, model 18 and model 27 are defined as the most suitable models since they have the highest prospect value.

From the findings and explanations, it is suggested that future studies can reveal a generative design by connecting a parametric tool such as Grasshopper or Dynamo. A productive model chain can be created by connecting the fuzzy input system to the output of the generated workspace models. When parameters such as View, Depth, WWR are changed, its 3-dimensional representation will also change with creating a flexible design. In this way, it will be possible to produce and analyze numerous exemplary models specific to Prospect-Refuge theory.

References

Dosen, A.S., Michael, J.O.: Prospect and refuge theory: constructing a critical definition for architecture and design. Int. J. Des. Soc. **6**(1), 9–24 (2013)

Küçükoğlu, A., Özerbaş, M.A.: Educational ergonomy and organization of in-class physical variables. J. Atatürk Univ. Social Sci. Inst. **4**(2) (2010)

Hildebrand, G.: Origins of Architectural Pleasure. University of California Press, Berkeley (1999)

Heerwagen, J., Orians, G.: Humans, habitats and aesthetics. In: Kellert, S.R., Wilson, E.O. (eds.) The Biophilia Hypothesis, pp. 138–172. Island Press, Shearwater Books, Washington (1993)

Kaplan, R., Kaplan, S.: The Experience of Nature: A Psychological Perspective. Cambridge University Press, Cambridge (1989)

Wohlwill, J.F.: Environmental aesthetics: the environment as a source of affect. In: Altman, I., Wohlwill, J.F. (eds.) Human Behavior and Environment, pp. 37–86. Plenum, New York (1976)

Kellert, S.R.: Building for Life: Designing and Understanding the Human-Nature Connection. Island Press, Washington (2005)

Ellard, C.: You are Here: Why We can Find our Way to the Moon, But Get Lost in the Mall. Anchor Books, New York (2009)

Diker, F.: Evaluation of the efficiency of university classrooms with the fuzzy logic. In: Architecture and Civil Engineering, pp. 341–357(2019)

Kaplan, S.: Aesthetics, affect, and cognition: environmental preference from an evolutionary perspective. Environ. Behav. **19**(1), 3–32 (1987)

Heerwagen, J.: Psychological value of space. National institute of building sciences (2008)

Hwang H., Lee, H.: Parametric model for window design based on prospect-refuge measurement in residential environment (2008)

Al, S.: Eğitim Yapılarının Fiziksel Konfor Koşullarının Öğrenci Başarısına Etkisi. Karadeniz Technical University, Institute of Science, Doctorate Thesis, 375, Trabzon, (2014)

Kazanasmaz, T., Tayfur, G.: Hasta Bakım Ünitelerinin Tasarım Verimliliklerinin Bulanık Mantık Modeli Bağlamında Değerlendirilmesi. Megaron **5**(1), 11–22 (2010)

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/905692/BB103_Area_Guidelines_for_Mainstream_Schools.pdf

https://www.education.sa.gov.au/sites/default/files/decd-design-standards.pdf?acsf_files_redirect

Design of a Fuzzy Linear Quadratic Regulator for the Speed Control of SPMSM

Kursad M. Gul(✉), Ayhan Kural, and Tufan Kumbasar

Istanbul Technical University, Istanbul 34437, Turkey
{gulk21,kuralayh,kumbasart}@itu.edu.tr

Abstract. Surface-Mounted Permanent Magnet Synchronous Machines (SPMSM) operate in wide torque-speed envelopes which impose the necessity of designing an effective control system to prevent performance degradation. However, the conventional Cascade PI (CPI) control structure is still utilized in the vast majority of SPMSM applications even though it becomes inadequate against emerging nonlinearities in different operating regions. In this paper, we propose a Fuzzy Linear Quadratic Regulator (F-LQR) for the speed control of SPMSMs to maintain robustness while improving the speed control performance. We designed a Single input Fuzzy Logic Controller (SFLC) to manipulate the error compensation dynamics of the LQR to improve reference tracking and disturbance rejection performances. We employed a geometrical analysis approach when designing the SFLC which enables us to explicitly shape the fuzzy mapping with just two parameters. The practical design goals considered in the design phase of LQR and SFLC are determined by taking into account the physical limitations of the inverter board as well as the system dynamics of the SPMSM. We experimentally analyzed and compared the performances of the proposed F-LQR, LQR, and CPI on the real-world SPMSM test setup with two inclusive test scenarios. The comparative test results revealed that the proposed F-LQR improves the speed control performance while maintaining the robustness of the control system when compared to LQR and CPI.

Keywords: Fuzzy control · LQR · SPMSM · vector control

1 Introduction

Designing a controller to efficiently drive a Surface-Mounted Permanent Magnet Synchronous Machine (SPMSM) in its full torque-speed envelope is a challenge due to emerging nonlinearities such as inverter nonlinearities, rotor flux harmonics, cogging torque, structural eccentricities, magnetic saturation, etc. [11]. The conventionally employed Cascade PI (CPI) control structure in which the outer PI control loop regulates the speed whereas the inner individual PI control

This work is supported by FIGES A.S.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 692–699, 2023.
https://doi.org/10.1007/978-3-031-39777-6_81

loops separately control the torque and flux experience performance degradation as the nonlinearities emerge [3,8,10], thus the necessity of developing more effective control structures arise. In [9,10], performance degradation with the CPI is analyzed and a Fuzzy Logic Controller (FLC) is proposed for the speed control of SPMSMs while a Linear Quadratic Regulator (LQR) is designed considering the periodic disturbances in [12]. In [8], a sensitivity analysis is conducted and the results imply that the state-feedback controllers are more robust against disturbances than CPI. A Fuzzy-LQR (F-LQR) in which the sum of positions and velocities of the state variables multiplied by LQR gains used as inputs to the FLC is designed for a double link pendulum in [4]. In [1], an F-LQR is designed for a motorcycle which is composed of several LQRs each one is optimal for different regions and their weight in the final controller output is determined from an FLC. In [6], an F-LQR is designed for generic aircraft in which the state variables are transformed into error and rate of change of error using a linear fusion function and processed within the FLC.

In this paper, we propose the design of an F-LQR for the speed control of SPMSMs to improve the reference tracking and disturbance rejection performances while maintaining optimality and robustness. The performance improvement with the proposed F-LQR is verified by the experimental tests conducted on the real-world SPMSM test setup shown in Fig 1a. We designed a Single input FLC (SFLC) that manipulates the error compensation dynamics of the F-LQR to improve speed control performance, unlike the F-LQR approaches in the literature. The Fuzzy Mapping (FM) of SFLC is explicitly shaped by adjusting two parameters via a geometrical approach. The practical design goals which are used to design F-LQR are determined by taking into account the SPMSM dynamics and physical limitations of the inverter board. We experimentally analyzed and compared the performances of the proposed F-LQR, LQR, and CPI on the real-world experimental test setup for SPMSMs with two inclusive test scenarios. The comparative test results revealed that the F-LQR shows superior speed control performance in terms of voltage ripple reduction, robustness, and external disturbance rejection when compared to LQR and CPI.

The rest of this paper is organized as follows. Section 2 introduces LQR design for the real-world SPMSM. Section 3 presents the proposed F-LQR alongside its design. Section 4 provides the comparative experimental test results while Sect. 5 gives the driven conclusions and future work.

2 LQR Design for Real-World SPMSM

In this paper, we handle the experimental SPMSM test setup illustrated in Fig. 1a. For the design of the LQR that is the baseline of the proposed F-LQR, we derive an augmented state space model from the nonlinear model of an SPMSM

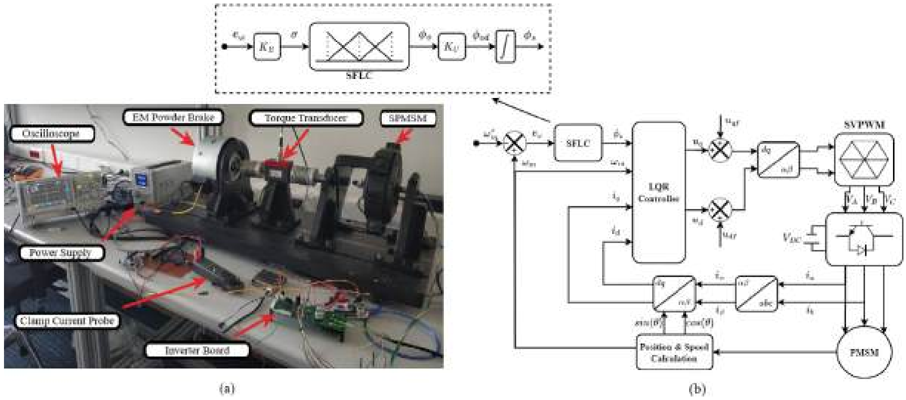


Fig. 1. Experimental test setup (a), F-LQR control structure (b).

in the d-q reference frame presented in [7] as follows:

$$\dot{\mathbf{x}}_a = \mathbf{A}_a\mathbf{x}_a(t) + \mathbf{B}_a\mathbf{u}_a(t) + \mathbf{F}_a r_a(t)$$
$$r_a(t) = \omega_m^*(t), \quad \mathbf{u}_a(t) = \begin{bmatrix} u_d(t) + u_{df}(t) \\ u_q(t) + u_{qf}(t) \end{bmatrix}, \quad \begin{cases} u_{df} = \omega_m P L_s i_q(t) \\ u_{qf} = -\omega_m P(L_s i_d(t) + \psi_m) \end{cases} \tag{1}$$

where

$$\mathbf{A}_a = \begin{bmatrix} \frac{-R_s}{L_s} & 0 & 0 & 0 \\ 0 & \frac{-R_s}{L_s} & 0 & 0 \\ 0 & \frac{K_T}{J_m} & \frac{-B_m}{J_m} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{B}_a = \begin{bmatrix} \frac{1}{L_s} & 0 \\ 0 & \frac{1}{L_s} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{x}_a(t) = \begin{bmatrix} i_d(t) \\ i_q(t) \\ \omega_m(t) \\ x_\nu(t) \end{bmatrix}, \quad \mathbf{F}_a = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

Here, ω_m^* is the desired angular speed whereas ω_m is the measured angular speed. $u_{d,q}$ are the voltage control components while $i_{d,q}$ are the current components in the d-q reference frame, respectively. The state variable x_ν correspond to the integral of angular speed error i.e., $x_\nu(t) = \int_0^t e_\nu(\tau)\, d\tau$ where $e_\nu = \omega_m - \omega_m^*$. For the setup in Fig.1a, the SPMSM parameters are: stator resistance $R_s = 0.225$ffl, synchronous inductance $L_s = 0.66\,\mathrm{mH}$, permanent magnet flux linkage $\psi_m = 0.0281\mathrm{W}$, the number of pole pairs $P = 20$, moment of inertia $J_m = 0.018\,\mathrm{kg/m^2}$, static friction coefficient $B_m = 0.025\,\mathrm{Nm/rad/s}$, DC bus voltage $V_{dc} = 48\,\mathrm{V}$, and rated torque $T_{rated} = 7\,\mathrm{Nm}$.

A state-feedback control law can be designed for the system in Eq. (1) as $\mathbf{u}(t) = [u_d \;\; u_q]^T = -\mathbf{K}\mathbf{x}_a(t)$ where $\mathbf{K}$ is the state-feedback gain matrix which is obtained through the LQR approach. The optimal $\mathbf{K}$ is obtained by solving the Riccati equation to minimize the performance index V [7]:

$$V = \frac{1}{2}\int_0^\infty (\mathbf{x}_a^T(t)\mathbf{Q}\mathbf{x}_a(t) + \mathbf{u}_a^T(t)\mathbf{R}\mathbf{u}_a(t))\, dt \tag{2}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are the state and control vector weight matrices, respectively. We set $\mathbf{Q}$ and $\mathbf{R}$ such that to achieve satisfactory system response, fast disturbance rejection, low voltage ripples, and zero d-axis currents to achieve maximum torque per ampere [10]. The settings of $\mathbf{Q}$ and $\mathbf{R}$ that achieve the control objectives alongside the resulting $\mathbf{K}$ are as follows:

$$\mathbf{Q} = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 2.9453 & 0 & 0 & 0 \\ 0 & 0.8090 & 0.3006 & 2.2361 \end{bmatrix} \tag{3}$$

3 Design and Tuning of F-LQR

As shown in Fig. 1b, the proposed F-LQR utilizes an SFLC to manipulate the speed control dynamics of the LQR through the state $x_\nu(t) = \int_0^t e_\nu(\tau)\, d\tau$. The SFLC uses and processes e_ν as its input $\sigma(t) = K_E e_\nu(t)$ where K_E is the input Scaling Factor (SF) used to normalize the e_ν into the universe of discourse of the Membership Functions (MF) and is defined as $K_E = 1/(e_\nu(t_f))$, and t_f is the speed reference variation time instance [5]. The output of SFLC is generated via the FM $\phi_o(\sigma(t))$ which is designed by using the geometrical analysis approach introduced in [2]. Then, the ϕ_o is denormalized by the output SF K_U which defined as $K_U = K_E^{-1}$. The integral of the denormalized output of SFLC i.e., $\phi_s(t) = \int_0^t \phi_{od}(\tau)\, d\tau$ is processed by the LQR, thus the state vector $\mathbf{x}_a(t)$ in F-LQR is redefined as $\mathbf{x}_a(t) = [i_d(t) \;\; i_q(t) \;\; \omega_m(t) \;\; \phi_s(t)]^T$.

The rule base of SFLC is composed of 5 rules in the following form.

$$R^i: \text{ IF } \sigma \text{ is } E_i, \text{ THEN } \phi_o \text{ is } \phi_i = B_i \tag{4}$$

where $i = \{-2, -1, 0, 1, 2\}$, B_i are crisp MFs satisfying $B_{-p} = -B_p$, $p = 1, 2$ and $B_0 = 0 < B_1 < B_2 = 1$. The antecedent part is defined with 50% overlapping triangular MFs (E_i) with the cores c_i satisfying $c_{-p} = -c_p$, and $c_0 = 0 < c_1 < c_2 = 1$ as shown in Fig. 2a. The SFLC uses the product implication and the center of sets defuzzification method. The FM of SFLC i.e., ϕ_o for an input $\sigma(t) \in [c_1, c_2]$ can be defined by the following equations [5]:

$$\phi_o(\sigma(t)) = k^i \sigma(t) + \eta^i, \quad k^i = \frac{B_{i+1} - B_i}{c_{i+1} - c_i}, \quad \eta^i = \frac{B_i c_{i+1} - B_{i+1} c_i}{c_{i+1} - c_i} \tag{5}$$

For the handled SFLC, the only design parameters to be defined are B_1 and c_1. A geometrical analysis approach is proposed in [2] to analyze the effect of parameters B_1 and c_1 on the FM of SFLC. In this approach, the FM is explicitly shaped by introducing two parameters i.e., radius l and slope γ defined as:

$$l = \sqrt{B_1^2 + c_1^2}, \quad \gamma = \tan^{-1}\left(\frac{B_1}{c_1}\right) \tag{6}$$

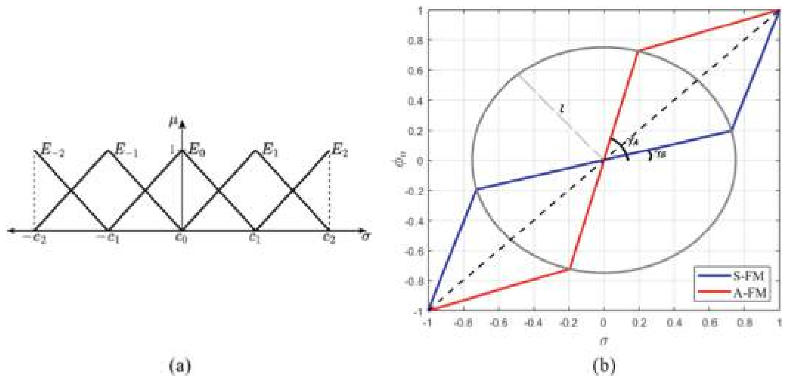


Fig. 2. Antecedent MFs of the SFLC (a), FMs of SFLC (b).

The parameter γ determines whether the FM has either Aggressive (A) or Smooth (S) characteristics while the l is defining the region of the characteristics. The FM of SFLC is an A-FM if $\gamma > 45°$ while an S-FM for $\gamma < 45°$. Note that if $\gamma = 45°$ then SFLC reduces to unit mapping $\phi_o(\sigma(t)) = \sigma(t)$, thus the F-LQR reduces to its conventional LQR counterpart. To explicitly analyze the resulting FM characteristics, A-FM and S-FM are constructed as shown in Fig. 2b by setting $\gamma_A = 75°$ for A-FM while $\gamma_S = 15°$ for S-FM when $l = 0.75$ in both cases.

Obviously, the A-FM will improve disturbance rejection and tracking error compensation performance of the F-LQR whereas the S-FM will reduce due to their mapping characteristics as it shapes the integral term of the F-LQR. Moreover, since the F-LQR is composed of a single loop structure rather than a cascade control one, adopting an SFLC with A-FM will increase the response speed. Thus, we design an A-FM to improve the speed control performance of F-LQR to achieve the control objectives explained in the LQR design phase. Accordingly, we set $\gamma = 75°$, and $l = 0.75$. Once l and γ are set, the design parameters B_1 and c_1 are then calculated as $B_1 = 0.7244$ and $c_1 = 0.1941$ by numerically solving Eq. (6).

4 Experimental Test Results and Discussion

In order to comparatively analyze the speed control performance of the proposed controller, we conducted reference tracking and disturbance rejection tests on the experimental test setup shown in Fig. 1a. To compare the speed control performance of the proposed design and highlight the performance improvements, a CPI as explained in [8] is designed with $8.3 rad/s$ and $3024 rad/s$ bandwidths for speed and current control loops, respectively. The speed control performances were quantified by using the Sum of Squared Error (SSE) defined as $SSE = \sum_{l=1}^{L}(e_\nu(l))^2$, and Total Variation (TV) of the control signal defined as $TV = \sum_{l=1}^{L} ||\mathbf{u}_a(l)||_2$ where L is the total number of samples. The SSE is a measure for reference tracking performance whereas TV provides information on the level of aggressiveness/ smoothness of the inverter control voltage.

4.1 Reference Tracking Performance

In this test, the controllers were subjected to consecutive step reference inputs without a torque load. The resulting controller performances are shown in Fig. 3 while the corresponding performance measures are tabulated in Table 1.

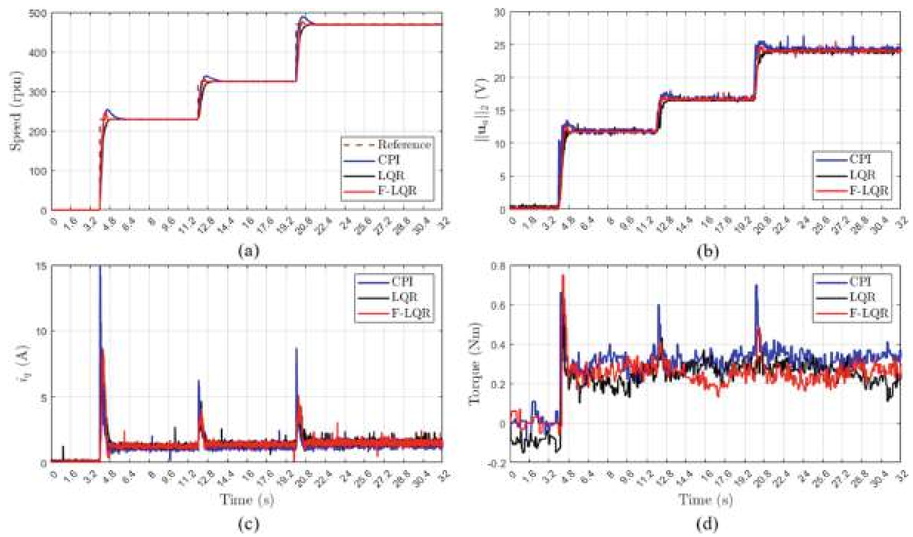


Fig. 3. Reference tracking test results. (a) Angular speeds, (b) voltage vector magnitudes, (c) q-axis currents, (d) measured shaft torques.

For the reference tracking test, the proposed F-LQR showed the best performance in terms of steady-state voltage ripple reduction as can be seen in Fig. 3b which coincides with the resulting TV measures in Table 1. Also, even though the CPI ended up with the least SSE, the F-LQR provided the smallest settling times and fewer overshoots than CPI as can be seen in Fig. 3a. Moreover, the performance improvement with the F-LQR is obvious compared to the LQR controller in terms of both SSE and TV measures.

4.2 Disturbance Rejection Performance

In this test, after the control loops converged to 375 rpm speed reference, we applied 15 Nm torque load disturbance which was even higher than two times the rated torque of the SPMSM at $t = 12$ s and released the disturbance at

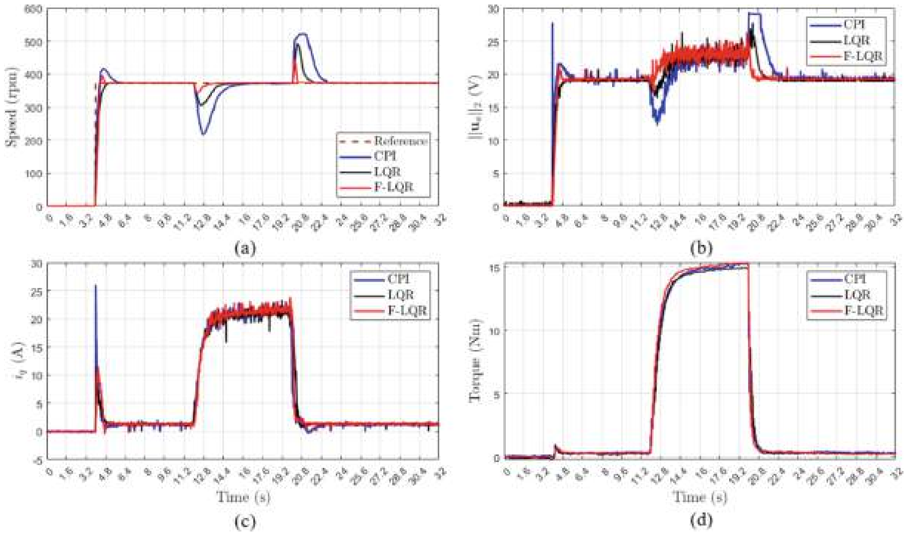


Fig. 4. Disturbance rejection test results. (a) Angular speeds, (b) voltage vector magnitudes, (c) q-axis currents, (d) measured shaft torques.

$t = 20$ s. The resulting controller performances are shown in Fig. 4 while the corresponding performance measures are presented in Table 1.

For the disturbance rejection test, the proposed F-LQR shows superior performance compared to the CPI and LQR in terms of disturbance rejection as can be seen in Fig. 4a which coincides with the resulting SSE measures in Table 1. Moreover, the F-LQR ends up with less TV than CPI, thus producing smoother control voltages. Bringing up the inferences made along with the SSE and TV measures of the two experimental tests together, it can be stated that the proposed F-LQR outperformed the LQR and CPI regarding both disturbance rejection and reference tracking performance.

Table 1. Experimental Test Results

	Reference Tracking Test			Disturbance Rejection Test		
	CPI	LQR	F-LQR	CPI	LQR	F-LQR
SSE($x10^5$)	622.3	1281	974.9	7336	2512	1973
TV	808.9	509.3	488.0	1092	977	1040

5 Conclusion and Future Work

This paper presents the design and experimental validation of a novel F-LQR for the speed control loop of real-world SPMSMs to prevent performance degradation along its wide operating range. To improve the speed control performance, an SFLC is designed which manipulates the error compensation dynamics of the LQR. We conducted experimental tests on the real-world SPMSM test setup to evaluate and compare the control performance of the proposed design with conventionally employed CPI and LQR. The comparative test results revealed that the proposed F-LQR provides fewer output voltage ripples and superior external disturbance rejections compared to the CPI and LQR structures.

As for our future work, we plan to design type-2 fuzzy nonlinear optimal speed controllers for SPMSMs.

References

1. Goodarzi, A., Armion, A.: Integrated fuzzy-optimal motorcycle dynamic control. Veh. Syst. Dyn. **48**(sup1), 505–524 (2010)
2. Guzay, C., Kumbasar, T.: Aggressive maneuvering of a quadcopter via differential flatness-based fuzzy controllers: from tuning to experiments. Appl. Soft Comput. **126**, 109223 (2022)
3. Harnefors, L., Saarakkala, S.E., Hinkkanen, M.: Speed control of electrical drives using classical control methods. IEEE Trans. Ind. Appl. **49**(2), 889–898 (2013)
4. Hazem, Z.B., Fotuhi, M.J., Bingül, Z.: A study of anti-swing fuzzy lqr control of a double serial link rotary pendulum. IETE J. Res., 1–12 (2021)
5. Kumbasar, T.: Robust stability analysis and systematic design of single-input interval type-2 fuzzy logic controllers. IEEE Trans. Fuzzy Syst. **24**(3), 675–694 (2016)
6. Singh, D.J., Verma, N.K., Ghosh, A.K., Malagaudanavar, A.: An application of interval type-2 fuzzy model based control system for generic aircraft. Appl. Soft Comput. **121**, 108721 (2022)
7. Tarczewski, T., Grzesiak, L.M.: Constrained state feedback speed control of PMSM based on model predictive approach. IEEE Trans. Ind. Electron. **63**(6), 3867–3875 (2016)
8. Tarczewski, T., Szczepanski, R., Erwinski, K., Hu, X., Grzesiak, L.M.: A novel sensitivity analysis to moment of inertia and load variations for PMSM drives. IEEE Trans. Power Electron. **37**(11), 13299–13309 (2022)
9. Uddin, M.N., Rahman, M.A.: High-speed control of IPMSM drives using improved fuzzy logic algorithms. IEEE Trans. Ind. Electron. **54**(1), 190–199 (2007)
10. Wang, C., Zhu, Z.Q.: Fuzzy logic speed control of permanent magnet synchronous machine and feedback voltage ripple reduction in flux-weakening operation region. IEEE Trans. Ind. Appl. **56**(2), 1505–1517 (2020)
11. Xia, C., Ji, B., Yan, Y.: Smooth speed control for low-speed high-torque permanent-magnet synchronous motor using proportional-integral-resonant controller. IEEE Trans. Ind. Electron. **62**(4), 2123–2134 (2015)
12. Xia, C., Liu, N., Zhou, Z., Yan, Y., Shi, T.: Steady-state performance improvement for LQR-based pmsm drives. IEEE Trans. Power Electron. **33**(12), 10622–10632 (2018)

Intelligent Flight Scheduling for Transit Passengers by Introducing New Flights in Hub-and-Spoke Network

Melis Tan Taçoğlu[1,2] and Mustafa Arslan Örnek[1,2](✉)

[1] Logistics Management Department, Yaşar University, Bornova, İzmir, Turkey
{melis.tacoglu,arslan.ornek}@yasar.edu.tr

[2] Industrial Engineering Department, Yaşar University, Bornova, İzmir, Turkey

Abstract. Passenger demand in the aviation industry is growing exponentially, flight schedules are of paramount importance for airlines and airports, and resources must be used efficiently. Flight scheduling is more complex in the hub-and-spoke (HS) network in which transit passengers visit more than one airport to reach their final destination An airline's most important asset is its aircraft and the decision that most directly concerns aircraft utilization is its schedule. Mathematical models are proposed with the aim of minimizing passenger waiting time and fulfilling available flight seat capacity for scheduled flights. The first mathematical model assigns transit passengers to scheduled flights to ensure the minimum waiting time. In the case of excess demand, a new flight has introduced the schedule without disrupting the scheduled flights so as to provide the shortest waiting time for the remaining transit passengers in the second mathematical model. This article contributes to the flight scheduling literature by providing an objective of considering minimizing transit passenger waiting time and launching a minimum number of new flights at the same time in the HS network. The computational experiment is conducted to represent the verification of mathematical models.

Keywords: Hub-and-Spoke Network · Flight Scheduling · Transit Passengers

1 Introduction

In recent years, air transportation volume has been significantly growing with the help of recent technological development related to the airline industry. Airline companies utilize their scarce resources such as aircraft, crew, and slot time efficiently to meet passenger demand. Due to the high demand for scarce resources, strategic planning needs to be done critically regarding each decision to allocate scarce resources efficiently. The most crucial capital for an airline is aircraft, and the decision that most directly concerns aircraft utilization in the aviation sector is their schedule. Therefore, inbound and outbound planning is a vital problem in the high passenger demand environment. Overcoming this problem can be sustained by changing the business model. Major airlines' business models have been shifting from point-to-point networks to hub and spoke networks in recent years. An HS network model contains two or more airports or

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 700–708, 2023.
https://doi.org/10.1007/978-3-031-39777-6_82

airport hubs attached to the flight network to deliver passengers from the origin to the final destination, on the other hand, airports are connected directly without a hub airport in the point-to-point business model [1].

One of the essential tasks in the aviation industry is the planning of flights, especially in the hub and spoke network, where route planning is more complicated compared to a point-to-point network high number of connecting passengers in the flight network. Airline companies must construct a flight schedule for as many arriving and departing flights as they can in a short, reasonable short time if they want to attract transit passengers by minimizing their waiting times and offering them various flight options. Generally, the flight schedule is made ahead of six months with the aim of profit maximization [2]. This problem is known as the flight scheduling problem (FSP) in schedule design literature. The vein that this problem apart from the literature is that it generates a plan that will meet the excess demand without disturbing the already existing flight schedule. A novel mathematical model is proposed with the integration of proposed new flights for the selected hub. Most of the studies focus on analyzing factors for a newly introduced line or network optimization of all connections, including many different hubs; however, an integrated model of new flight introduction and flight schedule for the selected hub has been relatively scant. This paper is organized as a literature review, mathematical models, computational experiment, and conclusion, respectively.

2 Literature Review

The air transportation planning process for an airline is complex, and scheduling is a vital step in the planning process. Flight scheduling planning has been examined broadly in the air transportation literature. [2] and [3] conducted a detailed literature review survey on flight scheduling planning.

HS networks have been a crucial part of many airline networks worldwide since the change in the regulation of the domestic air transport industry in the United States in 1978. Optimal usage of the HS network is when the airline's air transportation network is expanding vastly and has many destination points in its network [4]. Although large-scale airlines' business models are HS networks, such as Delta, Lufthansa, Turkish Airlines, and Emirates, some airline companies manage their network by a point-to-point business model [5]. Point-to-point network is used for direct transportation from origin to destination at a low cost. In addition, recently, there has been a rapid change in airline network planning literature to hybridization of network models that remove distinct borders between pure HS and point-to-point networks [6].

Numerous studies incorporating the scheduling problem with other planning process components in aviation have been conducted to support airline schedule improvement by developing optimization models such as fleet assignment, crew routing, and aircraft routing [7, 8]. Some studies consider more than one integration with flight schedules. Fleet assignment, aircraft routing, and crew pairing problems are integrated with airline scheduling and solved optimally by the column generation approach [9].

In the FSP literature, most of the articles have examined planning a base/master schedule. [10] developed a mathematical and heuristic model for new flight schedule generation while maximizing aircraft fleet utilization and minimizing the waiting time for transfer passengers. [11] proposed a model that determines new flight schedules regarding planned crew transfers, rest periods, passenger connections, and aircraft maintenance.

In the literature, few studies examine flight schedules without changing the master flight schedule. [5] proposed a unique mathematical model and heuristic approach for the airline bank optimization problem that generates new flight schedules with improved flight connection times intending to minimize transfer passenger waiting time for a specified cluster. [12] proposed two unique approaches for reconstructing flight schedules instantaneously by launching new flights to planned flight schedules considering airline profit maximization. New flights are added to the schedule without significantly changing scheduled flights. Flight scheduling articles are given in Table 1.

Table 1. Summary of Literature Review

Article	Year	Scheduling Type	Model Formulation	Objective Function	Solution Procedure
[11]	2002	Introducing a new flight to the master plan	Linear Programming Model	Minimize total waiting time of aircraft, crew, and passenger	Heuristic
[8]	2013	Master Plan Scheduling	Mixed Integer Linear model	Maximize the expected profits	Optimization Solver Xpress 20.00.11
[13]	2014	Master Plan Scheduling	Multi Commodity Network Flow Model	Maximize utilization of remaining flying time	Branch-and-bound and Compressed annealing heuristic
[9]	2016	Master Plan Scheduling	Mixed Integer Linear Programming	Minimize the sum of the number of aircraft routes, the number of crew pairings, and the waiting times of crews between consecutive flights	LP Relaxation with column generation and Heuristic
[10]	2016	Master Plan Scheduling	Mixed Integer Linear Programming	Maximize average fleet utilization	Heuristic

(*continued*)

Table 1. *(continued)*

Article	Year	Scheduling Type	Model Formulation	Objective Function	Solution Procedure
[12]	2019	Introducing a new flight to the base plan	Mixed Integer Conic Quadratic Model	Maximize airline profit	Mixed-integer conic quadratic optimization with McCormick inequalities
[5]	2020	Introducing a new flight to the base plan	Binary Integer Programming	Minimize transfer passenger waiting time	Simulated Annealing and Tabu Search
[7]	2021	Introducing a new flight to the base plan	Multi Commodity Network Flow Model	Maximize the expected operating profits	Differential Evolution algorithm

3 Mathematical Models

Two mathematical models have been developed and explained, respectively; a basic mathematical model and a basic mathematical model with new flights. Both models assume that transit passengers coming from different locations reach their destination point only via the selected specific hub in the HS network.

3.1 Basic Mathematical Model

The basic mathematical model is constructed to evaluate whether the current flight capacity is adequate for the transit passenger demand. In the model, B number of transit passengers and M number of scheduled flights are available. In addition, M + 1'st flight is added to satisfy feasibility. If the current flight's seat capacity exceeds the number of transit passengers, transit passengers will be assigned to M + 1'st flight. The indices, sets, parameters, and models are given below.

Indices

t: Transit passenger index, t = 1, …, B

f: Scheduled flight index, f = 1, …, M, M + 1

Sets

F_t: Set of eligible scheduled flights for transit passenger t.

A scheduled flight may be ineligible for assignment due to its departure time (The ready time of the transit passenger may be at a later time than the departure time of the scheduled flight).

T_f: Set of eligible transit passengers for scheduled flight f.

A transit passenger may be ineligible for assignment due to its ready time (The departure time of the scheduled flight may be at an earlier time than the ready time of the transit passenger).

Parameters

M_t: Maximum allowable time for passenger t in the airport.

The maximum allowable time for a transit passenger is the maximum allowed for that passenger to stay at the airport without any cost incurred. The cost is applied when the maximum allowable time is exceeded, and it is assumed that it is the same for each passenger.

(if waiting time > maximum allowable time, the cost is incurred.).

r_t: Ready time of the transit passenger t.

e_f: Expected departure time of the scheduled flight f.

w_{tf}: Waiting time of transit passenger t if the transit passenger t is assigned to scheduled flight f,

$$wtf = e_f - r_t,\ f \in F_t,\ t \in T_f$$

$$h_{tf} = \begin{cases} 1, & \text{if } w_{tf} \geq m_t \\ h, & \text{otherwise where } h > 1 \end{cases}$$

h_{tf} parameter is the cost coefficient based on the waiting time of transit passenger t for scheduled flight f

c_{tf}: The weighted cost of assigning transit passenger t to scheduled flight f ctf = htf wtf

Weighted cost is obtained by multiplication of h_{tf} and w_{tf}. The waiting time of the transit passenger t for scheduled flight f is multiplied by cost h if its waiting time exceeds the maximum allowable time; otherwise, the transit passenger is multiplied by 1.

P_f: Capacity of scheduled flight f

Decision Variables

$$x_{tf} = \begin{cases} 1, & \text{if transit passeneger t is assigned flight f} \\ 0, & \text{otherwise} \end{cases}$$

Model

$$\text{Min} \sum_t \sum_{f \in F_t} c_{tf} x_{tf} \tag{1}$$

s.t.

$$\sum_{f \in F_t} x_{tf} = 1 \quad *t \tag{2}$$

$$\sum_{t \in T_f} x_{tf} \geq P_f \quad *f \tag{3}$$

$$x_{tf} \in \{0, 1\} \quad *t, f \tag{4}$$

The objective function minimizes the total weighted cost of assigning transit passenger t to the scheduled flight f. Constraint (2) ensures that each transit passenger must be solely assigned to a scheduled flight or flight M + 1. Constraint (3) provides that the total number of passengers assigned to a scheduled flight or M + 1 cannot exceed each scheduled flight capacity. M + 1 flight seat capacity is at least the number of available transit passengers, and the value of the $c_{t(M+1)}$ for all t is tremendous compared to all c_{tf} values.

3.2 Basic Mathematical Model with New Flights

New flights are introduced based on passengers and flight availabilities to arrival time, departure time, and seat capacities. In the previous mathematical model, passengers who are not assigned to a scheduled flight due to the limited seat capacity are assigned to the specified flight $M + 1$. In this model, the scheduled flight index takes values from 1 to M. Indices, sets, and parameters are given in addition to the first model. The mathematical model is also shown below.

Indices

n: New flight index, $n = 1, \ldots, N$

Sets

N_t: Set of eligible new flights for transit passenger t.

A new flight may be ineligible for assignment due to its departure time (The ready time of the transit passenger may be at a later time than the departure time of the new flight).

T_n: Set of eligible transit passengers for new flight n.

A transit passenger may be ineligible for assignment due to its ready time (The departure time of the new flight may be at an earlier time than the ready time of the transit passenger).

Parameters

e_n: Expected departure time of the new flight n.

w_{tn}: Waiting time of transit passenger t if the transit passenger t is assigned to the scheduled flight f,

$$w_{tn} = e_n - r_t, \ n \in N_t, \ t \in T_f$$

$$h_{tn} = \sqrt{\begin{matrix} 1, \ \text{if} \, w_{tn} \geq m_t \\ h, \ \text{otherwise where} \, h > 1 \end{matrix}}$$

htn parameter is the cost coefficient based on the waiting time of transit passenger t for new flight n

k_{tn} = The weighted cost of assigning transit passenger t to the new flight n, $k_{tn} = h_{tn} * w_{tn}$.

The multiplication of h_{tn} and w_{tn} obtains weighted cost. The waiting time of the transit passenger t for new flight n is multiplied by cost h if its waiting time exceeds the maximum allowable time; otherwise, the transit passenger is multiplied by 1.

Z_n: Capacity of new flight f

O_n: Cost of opening a new flight f

Decision Variables

$$s_{tn} = \sqrt{\begin{matrix} 1, \ \text{if transit passeneger t is assigned new flight n} \\ 0, \ \text{otherwise} \end{matrix}}$$

$$y_n = \sqrt{\begin{matrix} 1, \text{if a new flight n is opened} \\ 0, \ \text{otherwise} \end{matrix}}$$

Model

$$\text{Min} \sum_t \sum_{f \in F_t} c_{tf} x_{dtf} + \sum_t \sum_{f \in F_t} k_{tn} s_{tn} + \sum_n O_n y_n \quad (5)$$

s.t.

$$\sum\nolimits_{f \in F_t} x_{tf} + \sum\nolimits_{n \in N_t} s_{tn} = 1 \quad *t \tag{6}$$

$$\sum_{t \in T_f} x_{tf} \geq P_f \quad *f \tag{7}$$

$$\sum_{t \in T_n} s_{tn} \geq Z_n \quad *n \tag{8}$$

$$s_{tn} \geq y_n \quad *t, n \tag{9}$$

$$x_{tf} \in \{0, 1\} \quad *t, f \tag{10}$$

$$s_{tn} \in \{0, 1\} \quad *t, n \tag{11}$$

$$y_n \in \{0, 1\} \quad *n \tag{12}$$

The objective function minimizes the total weighted cost of assigning transit passenger t to the scheduled flight f, the total weighted cost of assigning transit passenger t to the new flight n, and the total cost of initializing the new flight to the current flight schedule. Constraint (6) ensures that each transit passenger must solely assign a scheduled flight or a new flight. Constraint (7) provides that the total number of passengers assigned to each scheduled flight cannot exceed each scheduled flight capacity. Constraint (8) ensures that the number of passengers assigned to new flights cannot exceed each new flight capacity. Constraint (9) indicates the logical relationship between transit passenger assignment to a new flight and the introduction of a new flight to the current schedule. If any transit passenger is assigned to a new flight, this constraint ensures that a new flight will be introduced.

4 Computational Experiment

The computational experiment is conducted to assess the mathematical models. These models are conducted for 100 passengers. Passenger index and arrival time are given in the following respectively: 1 to 10, 10 to 22, 22 to 36, 36 to 42, 42 to 50, 50 to 59, 59 to 70, 70 to 79, 79 to 91, 91 to 100; 3 am, 5 am, 6 am, 8 am, 9 am, 10 am, 11 am, 12 am, 13 am, and 14 am. Maximum allowable time (mt) is set to 3 h, and the h value is set to 3. Suppose a transit passenger waits more than 3 h. In that case, the weighted cost of assigning the transit passenger t to scheduled flight f (c_{tf}) will be found by multiplication of 3 and the transit passenger waiting time for that specified flight regardless of flight type.

Basic mathematical model is used for 100 passengers with 7 scheduled flights shown in Table 2. As a result, 20 passengers are assigned to flight M + 1, since total flight seat capacity (80) is inadequate for transit passenger demand (100). Therefore, new flights are needed to compensate the demand.

Table 2. Scheduled and New Flight Information Data

Flight Type	Flight Number	Departure Time	Flight Capacity	Opening Cost
Scheduled	1	06:00	10	–
Scheduled	2	08:00	10	–
Scheduled	3	10:00	10	–
Scheduled	4	11:00	10	–
Scheduled	5	12:00	10	–
Scheduled	6	13:00	10	–
Scheduled	7	15:00	20	–
New	1	17:00	20	$10000
New	2	19:00	20	$10000
New	3	21:00	20	$10000

Basic model with new flight is used with 100 transit passengers, 7 scheduled flights, and 3 new flights shown in Table 2. According to the optimal solution, solely new flight 1 is opened, and 20 transit passengers are assigned to the new flight, and these passengers are 23, 24, 25, 27, 30, 34, 67, 68, 69, 70, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100. The remaining passengers are assigned to the scheduled flights, and total capacity is used for the scheduled flights. Both problems are solved in IBM CPLEX.

5 Conclusion

Flight scheduling is a crucial task for airlines to efficiently utilize their most valuable asset which is aircraft. Airlines need to make intelligent decisions to meet the excessive inbound demand for transit passengers by adding a connecting flight to the master plan in advance of six months of the flight time. Therefore, mathematical models become the guide for airlines in the decision-making process. Models are developed in such a way as to meet its purpose on both sides at the same time by minimizing the waiting time of the passengers and minimizing the number of launcing new flights without disturbing the master plan. This study differs from other studies by considering aforementioned objectives at the same time. Future direction could investigate multi destinations points as an extension of the proposed model.

References

1. Morlotti, C., Birolini, S., Cattaneo, M., Redondi, R.: Introducing connecting flights in LCCs' business model: Ryanair's network strategy. J. Air Transp. Manag. **87**, 101849 (2020)
2. Eltoukhy, A.E., Chan, F.T., Chung, S.H.: Airline schedule planning: a review and future directions. Ind. Manag. Data Syst. **117**, 1201–1243 (2017)
3. Lohatepanont, M., Barnhart, C.: Airline schedule planning: integrated models and algorithms for schedule design and fleet assignment. Transp. Sci. **38**(1), 19–32 (2004)

4. Gillen, D., Morrison, W.G.: Regulation, competition and network evolution in aviation. J. Air Transp. Manag. **11**(3), 161–174 (2005)
5. Çiftçi, M.E., Özkır, V.: Optimising flight connection times in airline bank structure through Simulated Annealing and Tabu Search algorithms. J. Air Transp. Manag. **87**, 101858 (2020)
6. Birolini, S., Besana, E., Cattaneo, M., Redondi, R., Sallan, J.M.: An integrated connection planning and passenger allocation model for low-cost carriers. J. Air Transp. Manag. **99**, 102160 (2022)
7. Birolini, S., Antunes, A.P., Cattaneo, M., Malighetti, P., Paleari, S.: Integrated flight scheduling and fleet assignment with improved supply-demand interactions. Transp. Res. Part B: Methodol. **149**, 162–180 (2021)
8. Pita, J.P., Barnhart, C., Antunes, A.P.: Integrated flight scheduling and fleet assignment under airport congestion. Transp. Sci. **47**(4), 477–492 (2013)
9. Cacchiani, V., Salazar-González, J.J.: Optimal solutions to a real-world integrated airline scheduling problem. Transp. Sci. **51**(1), 250–268 (2017)
10. Kepir, B., Koçyiğit, Ç., Koyuncu, I., Özer, M.B., Kara, B.Y., Gürbüz, M.A.: Flight-scheduling optimization and automation for AnadoluJet. Interfaces **46**(4), 315–325 (2016)
11. Stojković, G., Soumis, F., Desrosiers, J., Solomon, M.M.: An optimization model for a real-time flight scheduling problem. Transp. Res. Part A: Policy Pract. **36**(9), 779–788 (2002)
12. Şafak, Ö., Atamtürk, A., Aktürk, M.S.: Accommodating new flights into an existing airline flight schedule. Transp. Res. Part C: Emerg. Technol. **104**, 265–286 (2019)
13. Başdere, M., Bilge, Ü.: Operational aircraft maintenance routing problem with remaining time consideration. Eur. J. Oper. Res. **235**(1), 315–328 (2014)

Epic Poetry Summarization and Visualization by Using Sentiment Analysis, Topic Modeling and DALL-E

Derya Köşnek, Sinem Bozdağ(✉), and Tolga Kaya

Department of Management Engineering, Istanbul Technical University, 34367 Istanbul, Turkey
{kosnek18,bozdags18,kayatolga}@itu.edu.tr

Abstract. Epics are the narratives of generally exceptional illustrations of the nation's cultural heritage and reflect the cultural elements of the era that epic occurred. Epics are also studied by social scientists and linguists and they are about to vanish from the history. Since these are works from centuries ago, there are no images or visual records of the events. However, the detailed descriptions of characters and environments are accessible through the chapters of the epics. This study aims to suggest a framework for epic poetry summary extraction based on text mining and image generation techniques. As a result of combining the results of the used algorithms, a final output was obtained that will reveal the themes of the chapters of the epics. The outputs of the model were given as input to the text-to-image application DALL-E, and the main ideas were visualized. Results show that obtained summaries are in line with the epic literature.

Keywords: Epic Poetry · Latent Dirichlet Allocation · Sentiment Analysis · Topic Modelling · Dall-E · LexRank · Machine Learning · Text Mining · Text to Image Generation

1 Introduction

Throughout the paper, it is aimed to find the main themes and dominant emotions of epic poetries. The reason for conducting the study is to capture the presence of various sentiments that are shown by the nations from the epics both verbally and visually. The purpose of this study is to provide a model that outputs the importance and sentiments of epic poetries by combining different text analysis algorithms such as Sentiment Analysis, Latent Dirichlet Allocation, and LexRank. There are very limited studies in the literature analyzing epic poems with machine learning algorithms. The scope of the study is limited by the German epic poetry, Nibelungenlied [15], the Anglo-Saxon epic poetry, Beowulf [7], and the Mesopotamia epic poetry, Gilgamesh [5]. From the poetry of Nibelungenlied, chapters named "How Brunhild was Received at Worms" and "How Kriemhild Fared to the Huns", from Beowulf chapter named "Brave Though Aged", and from Gilgamesh tablet 6 were used to train and explain the outputs of the model in this report. All chapters from Nibelungenlied and almost half of the chapters of Beowulf and Gilgamesh were examined. Although in this report three epic poetry were taken into account due to their

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 709–716, 2023.
https://doi.org/10.1007/978-3-031-39777-6_83

place in the literature, the study is designed to be suitable for every chapter of each epic poetry.

The rest of the paper is organized as follows: A literature review was conducted in the next section. The third section is devoted to the methodology of used algorithms. The overall information about data was provided in the fourth chapter. Lastly, findings were discussed and the report ended with a conclusion.

2 Literature Review

Epic reflects the myths, emergence of a nation's heroes, history, and beliefs and is a tool for transferring these cultural elements from generation to generation and to other nations [21]. Because poetry is a form that evokes emotions with its semantic content, prosody, and aesthetics [10], epic poetry evokes a common national emotion.

According to Ashliman's review, the 10th chapter of the Nibelungenlied epic describes the weddings of Kriemhild and Siegfried, and Brunhild and Gunther, and the 21st chapter depicts Kriemhild's journey to Etzel's palace [18].

According to Hooker's work, in the 6th tablet of the Gilgamesh epic, when Ishtar's love proposal was rejected by Gilgamesh, Ishtar makes the Bull of Heaven attack the city of Uruk and Gilgamesh for revenge, and Enkidu and Gilgamesh slayed the Bull of Heaven [6].

In his work, Ashliman says that in the 33rd chapter of the Beowulf saga, the furious dragon, whose golden cup was stolen, burst into flames and set fire to everything, including Beowulf's house [1].

2.1 Sentiment Analysis

Sentiment analysis is the analysis of the author's attitude, emotion, or emotional state of the content of the text from the text [14]. Due to the fact that texts were classified according to their sentiment, sentiment analysis procedures can be thought of as classification [12]. Cultural differences, usage of creative and non-standard language, and lack of para-linguistic information are challenges that can affect the accuracy of sentiment analysis outputs [14].

2.2 Summarization by Topic Modeling

Summarizing is an NLP technique that condenses a text into a smaller version with the same meaning, without omitting important information [16]. There are different approaches used in summarizing. One of them is the extractive approach. In this approach, the sentences in the text are ordered according to their importance, and a summary is generated by the most important sentences [3]. In the abstractive method, a summary is made with new sentences and words that are not in the original text. Two general approaches to abstractive summarization are linguistic and semantic [20]. According to the study of Widyassari et al., the most used text summarization model between 2008–2019 is fuzzy logic [20]. In topic modeling, various parts of the input that consist of a basket of words text given are classified based on the topics. This process ends with the distribution of parts to the most likely topic outputs [9].

Latent Dirichlet Allocation. Word, subject, and document levels make up the LDA approach's three levels of data structuring. Using probability distributions over words, LDA portrays documents as having been created by combining hidden subjects [19]

2.3 DALL-E

DALL-E is a transformer language model that is created by OpenAI that allows converting written texts into images [17].

In a study conducted by Marcus, Davis, and Aaronson, DALL-E's capacity to understand and extract meaning from the texts given as input was measured by giving specific examples. In general, it turned out that the quality of the pictures obtained is high, it can easily apply various artistic styles and is successful in understanding the input given. On the other hand, there are some disadvantages: it can output unfinished pictures, it can be difficult to decipher and interpret the relationships between items, and it has small multimodality in combining meanings between sentence elements. Fourteen different inputs are given in the study and DALL-E was successful in generating synthetic images that satisfy the input's meaning in most of them [11].

3 Methodology

3.1 Sentiment Analysis

Determining and emotionally categorizing written expressions in a text is the process of sentiment analysis [8]. NRC was used to group words according to predetermined word emotion association by NRC Emotion Lexicon. By NRC, words were allocated to emotions such as anger, fear, joy, sadness, disgust, surprise, trust, and anticipation, and the overall positive and negative scores were calculated [13].

3.2 Latent Dirichlet Allocation

In LDA, each topic is generated through a Dirichlet prior distribution over words [19]. LDA classifies words into different topics and the goal is maximizing Marginal Log Likelihood Method [19]. The mathematical expression of Dirichlet Prior Distribution is as follows:

$$p(\lambda|\pi) = \frac{\upsilon\left(\sum_{i=1}^{k} \pi_i\right)}{\prod_{i=1}^{k} \upsilon(\pi_i)} \lambda_1^{z_1-1} \ldots \lambda_k^{z_k-1} \quad (1)$$

3.3 LexRank

It is an unsupervised approach used in text summarization and it scores sentences using the graph method. Based on eigenvector centralities [2], each sentence is scored according to similarity with many others in the text and if a sentence has high similarity with others, it will be more important [4]. The similarity between two sentences is calculated by the formula below:

$$idf - modified - cosine(x, y) = \frac{\sum_{w\in x,\, y} tf_{w,x} tf_{w,y} (idf_w)^2}{\sqrt{\sum_{x_i \in x} \left(tf_{x_i,x} idf_{x_i}\right)^2} \times \sqrt{\sum_{y_i \in y} \left(tf_{y_i,y} idf_{y_i}\right)^2}} \quad (2)$$

$tf_{w,s}$: number of occurrences of the word w in the sentence s.

3.4 DALL-E

It trains the model by receiving the text and image at the same time and developing all the tokens with the maximum likelihood method [17].

4 Data

One input sentence from each chapter used in the analysis can be seen in Table 1.

Table 1. Input sentences of epic poetries.

	Input	Type
Nibelungenlied Chapter 10	On yonder side Rhine river/they saw a stately band	Character
Nibelungenlied Chapter 21	The messengers leave we riding./Now shall ye understand	Character
Gilgamesh Tablet 6	He washed out his marred hair and cleaned up his equipment(?),	Character
Beowulf Chapter 33	Soon as the valiant Dankwart/stood beneath the door	Character

By using R, the data from three different epics are retrieved. Data cleaning and preparation methods were applied to the raw text. All letters were converted into lowercase for each word to be the same as its other counterparts that have some uppercase letters. Stop words were omitted from the text by the stopwords function in the tm package and blank items in the list were removed. Wordclouds were generated for chapters (see Fig. 1).

Sentiment analyses were conducted to get the dominant emotions in each chapter. By LDA, words were assigned to the defined number of main topics. The number of main topics was determined by FindTopicNumbers_plot function (see Fig. 2).

Each sentence was scored based on similarity by the LexRank algorithm.

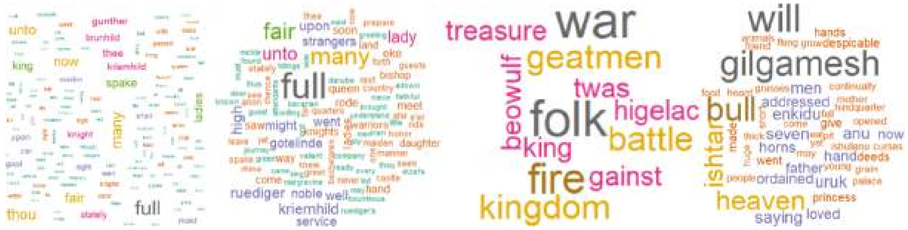


Fig. 1. Wordclouds of chapters 10 and 21 from Nibelungenlied, chapter 33 from Beowulf, and tablet 6 from Gilgamesh.

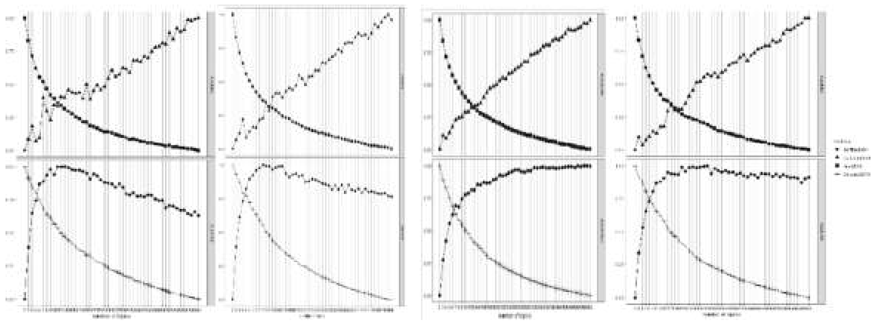


Fig. 2. The number of main topics of chapter 10 and 21 from Nibelungenlied, chapter 33 from Beowulf and tablet 6 from Gilgamesh.

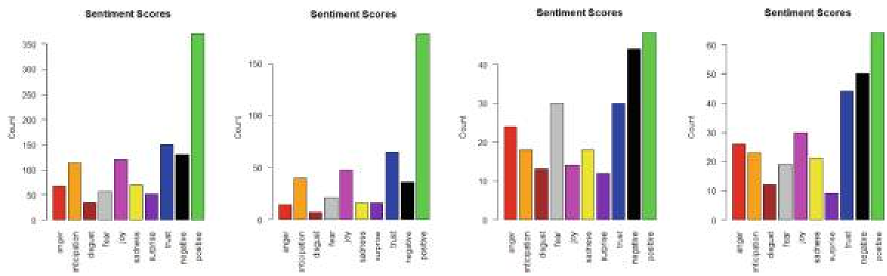


Fig. 3. Sentiment classifications of each word in chapters 10 and 21 from Nibelungenlied, chapter 33 from Beowulf, and tablet 6 from Gilgamesh.

5 Findings

As an output of sentiment analysis, dominant emotions, and overall positivity and negativity scores can be seen from the graphs for each chapter (see Fig. 3).

For Nibelungenlied, in the 10th chapter, “joy”, “trust”, and “anticipation” sentiments are dominant. It also tunes with the words “maid”, “bride”, “wife”, and “ladies” generated in the wordcloud. The dominant emotions in the 21st chapter are “trust”, “joy”, and “anticipation”. For Beowulf chapter 33, the sentiments “fear”, “anger”, and “trust” are common. This negative atmosphere of the chapter aligns with the summary of Prof. D. L.

Ashliman. Sentiment analysis of Gilgamesh tablet 6 shows that emotions “trust”, “joy”, and “anger” are dominant. This finding is compatible with the analysis of Hooker.

In LDA, for chapter 10 from Nibelungen, 8 main topics were selected with the help of Fig. 2. The same procedure was applied to all chapters that are analyzed. Table 2 was generated based on the output of Gilgamesh tablet 6.

Table 2. Table of topics and most common words for Gilgamesh tablet 6.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
gilgamesh	saying	father	enkidu	bull	horns	heaven	will
ishtar	grow	bearer	seven	addressed	deeds	give	uruk

Then, by using LexRankr function, each sentence was scored and the top 3 important sentences were chosen with respect to the order of appearance in the text. Sentences are displayed in Table 3.

Table 3. Top 3 important sentences from epics.

Epic Chapter	Sentences
Nibelungenlied Chapter 10	Then came to know each other/full many a knight and fair lady
	And many a knight full gallant/was stately lady’s guide
	But clasped in arms full fondly/and oft the lady fair,
Nibelungenlied Chapter 21	That thither was come Kriemhild,/the bishop’s niece full fair,
	Full fair was the reception:/whereat was Ruediger right glad
	Who with full ready service/upon the ladies then did wait
Beowulf Chapter 33	The wise one weened then that wielding his kingdom
	Lone-goer lorn to his land-folk returning,
	Where Hygd to him tendered treasure and kingdom,
Gilgamesh Tablet 6	a princess Ishtar raised her eyes to the beauty of Gilgamesh
	Ishtar led the Bull of Heaven down to the earth.
	Woe unto Gilgamesh who slandered me and killed the Bull of

Outputs from LexRankr function and sentiment analysis were combined for visualization. For visualization, topic modeling outputs were not used due to the limitations of DALL-E. They were visualized by using DALL-E with the style of Rembrandt in baroque oil painting (see Fig. 4). The sentences that are given as input to DALL-E are as follows.

Chapter 10 from Nibelungenlied. “Then came to know each other/full many a knight and fair lady., And many a knight full gallant/was stately lady’s guide., But clasped in arms full fondly/and oft the lady fair,, trust, joy, anticipation, in rembrandt style, baroque oil painting”

Chapter 21 from Nibelungenlied. “That thither was come Kriemhild,/the bishop’s niece full fair,, Full fair was the reception:/whereat was Ruediger right glad., Who with full ready service/upon the ladies then did wait., trust, joy, anticipation, in rembrandt style, baroque oil painting”

Chapter 33 from Beowulf. “The wise one weened then that wielding his kingdom, Lone-goer lorn to his land-folk returning,, Where Hygd to him tendered treasure and kingdom,, fear, trust, anger, in rembrandt style, baroque oil painting”

Tablet 6 from Gilgamesh. “a princess Ishtar raised her eyes to the beauty of Gilgamesh., Ishtar led the Bull of Heaven down to the earth., \”Woe unto Gilgamesh who slandered me and killed the Bull of, trust, joy, anger, in rembrandt style, baroque oil painting”

Fig. 4. Visual reflections of summarized chapters by DALL-E.

DALL-E outputs of Nibelungenlied chapters 10 and 21, and tablet 6 from Gilgamesh represent the atmosphere and events described in the chapters. However, the visual output of Beowulf’s chapter 33 does not symbolize the general ambiance of the chapter.

6 Conclusion

The methods mentioned in this report have been applied to all 39 chapters of the Nibelungenlied saga. In 35 of 39 chapters, the sentiments align with the summary of Prof. Ashliman. Therefore, the accuracy of sentiment analysis is 89.7%. When the topic modeling outputs were analyzed together with the summary, it was seen that the success rate was 61.5%. The performance of summarizing the chapter of an epic poem by the first 3 sentences handled through importance analysis was found to be low with a rate of 56.4%. It has been seen that the importance analysis method gave successful results in the tablets that were analyzed from the Gilgamesh epic. The success rate of DALL-E visualization outputs remained at 46.1% for Nibelungenlied. The reason for this low rate was evaluated as summary performance. Sentiment analysis has been outstanding in all 3 epics. Topic modeling also was successful in the epics of Beowulf and Gilgamesh. Future research on synthesizing sentences for meaningful summaries through machine learning algorithms may improve the clarity of the summary. This improvement on summaries might be addressed on future studies.

References

1. Ashliman, D.L.: Beowulf: A Summary in English Prose (2010). https://sites.pitt.edu/%7Edash/beowulf.html
2. Crabcamp/lexrank: LexRank algorithm for text summarization, GitHub (n.d.). https://github.com/crabcamp/lexrank
3. El-Kassas, W.S., Salama, C.R., Rafea, A.A., Mohamed, H.K.: Automatic text summarization: a comprehensive survey. Expert Syst. Appl. **165**, 1–26 (2021)
4. Erkan, G., Radev, D.R.: J. Artif. Intell. Res. **22**, 457–479 (2004)
5. Gallery Kovacs, M.: The Epic of Gilgamesh (1998). https://uruk-warka.dk/Gilgamish/The%20Epic%20of%20Gilgamesh.pdf
6. Gilgamesh Summary (n.d.). https://www.faculty.umb.edu/gary_zabel/Courses/Phil%20281b/Philosophy%20of%20Magic/World%20Soul/GILG.HTM
7. Hall, L.: Beowulf: An Anglo-Saxon Epic Poem (2005). https://www.gutenberg.org/cache/epub/16328/pg16328-images.html
8. Holster, J.D.: Introduction to R for Data Science: A LISA (2020)
9. Jelodar, H., et al.: Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools Appl. **78**(11), 15169–15211 (2019). https://doi.org/10.1007/s11042-018-6894-4
10. Johnson-Laird, P.N., Oatley, K.: How poetry evokes emotions. Acta Physiol. (Oxf) **224**, 1–12 (2022)
11. Marcus, G., Davis, E., Aaronson, S.: A very preliminary analysis of DALL-E 2 (2022)
12. Mejova, Y.: Sentiment Analysis: An Overview [Comprehensive Exam Paper] (2009)
13. Mohammad, S.M., Turney, P.D.: NRC Emotion Lexicon. National Research Council Canada (2013)
14. Mohammad, S.M.: Sentiment Analysis. Emotion Measurement, pp. 201–237 (2016)
15. Needler, G.H.: The Nibelungenlied (2005). https://www.gutenberg.org/ebooks/7321
16. Onah, D.F.O.: A Data-driven Latent Semantic Analysis for Automatic Text (2022). https://arxiv.org/abs/2207.14687
17. OpenAI. Dall·E: Creating images from text. OpenAI (2022). https://openai.com/blog/dall-e/
18. The Nibelungenlied: A Summary in English Prose (n.d.). https://sites.pitt.edu/%7Edash/nibelungenlied.html
19. Vayansky, I., Kumar, S.A.: A review of topic modeling methods. Inf. Syst. **94** (2020)
20. Widyassari, A.P., et al.: Review of automatic text summarization techniques & methods. J. King Saud Univ. Comput. Inf. Sci. **34**(4), 1029–1046 (2022)
21. Yoshida, A.: Epic | Definition, Characteristics, Examples, Development, & Facts (1999). https://www.britannica.com/art/epic

A Fuzzy System for Forecasting Mild Cognitive Impairment

Cheng-Li Liu[1](✉), Che-Jen Chuang[1], and Kuo-Wei Su[2]

[1] Vanung University, 320614 Taoyuan, Taiwan
johnny@vnu.edu.tw
[2] National Kaohsiung University of Science and Technology, 824005 Kaohsiung, Taiwan

Abstract. Mild cognitive impairment (MCI) is a condition in which brain function declines. It is a transition stage between normal aging and Alzheimer's disease (AD). Improper use of test methods in clinical practice often leads to failed tests, resulting in variations in test results. The objective of this study is to use the 3D virtual environment to create an image testing scenario that integrates Mini-Cog, Short Portable Mental Status Questionnaire (SPMSQ), Mini-mental Status Examination (MMSE), Saint Louis University Mental Status Examination (SLUMS), Clinical Dementia Rating (CDR), and Cognitive Abilities Screening Instrument (CASI), in addition, a fuzzy logic control (FLC) was used to develop an MCI forecasting system. The following findings were observed: The MCI forecasting system is highly correlated with the traditional hardcopy tests, including Mini-Cog, SPMSQ, MMSE, SLUMS, and CASI. The forecasting system can provide test scores for testers to make judgments. In addition, the distribution of the System Usability Scale (SUS) evaluation scores for the MCI forecasting system revealed that 87.5% were grade C (suitable to use) or above, and 29.2% were grade B (extremely good to use) or above. There is good positive feedback for the system from the subjects.

Keywords: Mild Cognitive Impairment · Alzheimer's Disease · 3D Virtual Environment · Fuzzy Logic Control · Screening Tool

1 Introduction

With advances in cognitive impairment research, the term "Mild Cognitive Impairment" first appeared in tertiary symptoms in the Global Deterioration Scale developed by Reisberg et al. (1982) [1]. MCI is a decline in brain function that includes cognitive impairment, memory defects, and learning or attention difficulties but does not meet the criteria for dementia [2]. Although the manifestations of patients with MCI do not interfere with activities of daily living, some patients with MCI may exhibit dementia symptoms, such as difficulty in the calculation, persistent amnesia, and organizational difficulties. Therefore, MCI is a transition stage between normal aging and AD, and it may be an underlying symptom of AD [3], a gray area between normal cognition and dementia. How to evaluate the MCI effectively is an important issue? In Sect. 2, a

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 717–724, 2023.
https://doi.org/10.1007/978-3-031-39777-6_84

literature review on MCI is discussed. In Sect. 3, we describe our research methods and the system scenario we designed. In Sect. 4, we present the results and discussion about the effectiveness of the MCI forecasting system.

2 Literature Review

Epidemiological studies have shown that MCI may be a prodromal stage of dementia [2, 4]. Many neuropsychological tests are often used to assess cognitive impairment and make subsequent diagnoses. Comprehensive tests can be performed when abnormalities occur. These tests include computed tomography, magnetic resonance imaging, single-photon emission computed tomography, and positron emission tomography to rule out meningioma, subdural hemorrhage, and other brain diseases or dementia subtypes [5]. Among these methods, the diagnostic value of hippocampal volume try is more accurate, but it is difficult to find these patients for consultation in many hospitals, particularly those in the early stages of the disease [6]. Hence, the majority of early MCI diagnoses are made through neuropsychological assessment. Most of these screening tools have threshold values for determining MCI, and there are reasonably suspected intervals for determining AD, e.g., an MMSE score of 25 points indicates a sign of MCI, and attention should be paid to the possibility of mild AD if the score decreases to 23–24 points. Therefore, combining the test results of multiple tools will increase the accuracy of diagnosing MCI. However, every screening tool requires time for testing and explanation. Hence, it is not easy to simultaneously implement multiple screening tests. The primary objective of this study is to use the 3D virtual environment to create an image testing scenario that integrates Mini-Cog, Short Portable Mental Status Questionnaire (SPMSQ), Mini-mental Status Examination (MMSE), Saint Louis University Mental Status Examination (SLUMS), Clinical Dementia Rating (CDR), and Cognitive Abilities Screening Instrument (CASI), in addition, combine fuzzy logic control (FLC) technology to develop an MCI forecasting system.

3 Material and Methods

3.1 Material and Methods

System Framework. This study uses a 3D virtual environment to construct a scenario-based image test environment and employs fuzzy logic control (FLC) technology to predict the possibility of MCI. Firstly, in terms of testing, MCI occurs mainly due to cerebral information processing dysfunction. Therefore, this study uses the five primary information processing functions in humans (perception, cognition, attention, memory, and action) as a basis, and according to the six tools mentioned above to classify the test items into eight major dimensions (orientation, Identification, construction, judgment, calculation, attention, memory, and control ability).

System Assessment Method. Based on the commonly used MCI and AD screening tools (Mini-Cog, SPMSQ, MMSE, SLUMS, CDR, and CASI), the system compiles and generates eight main dimensions and 11 assessments sub-items (Table 1). These 11 assessment sub-items are formulated based on scenario concepts, in addition to ensuring the assessment method and content of the original assessment tools. According to the scoring characteristics of various screening tools, different scores are given to each assessment sub-item, which allows scores for multiple screening tools to be obtained on a single test and reduces the subject's frustration in undergoing multiple tests.

MCI Fuzzy Forecasting Module. After the establishment of the assessment items in Table 1, the scores of each assessment sub-item are added to obtain the corresponding scores of the original six screening tools. The fuzzy analysis is a very suitable method to combine the results of the six screening tools to obtain the best value for MCI assessment. The fuzzy theory was proposed by Zadeh in 1965 [8], while fuzzy decision research began with "Decision-Making in a Fuzzy Environment," published by Bellman and Zadeh in 1979 [8]. Fuzzy logic control (FLC) is an actual application of fuzzy theory, and its core consists of a fuzzy rule base and fuzzy inference engine (FIE).

The fuzzy forecasting mechanism designed in this study first uses the fuzzy inference semantic functions for variable inputs, memberships, fuzzy inference, and output. Each input variable defined by six traditional screening tools includes three membership functions (Normal, Mild, and MS (Moderate & Severe)). The output variable is defined as a cognitive impairment estimate (CI), including three membership functions (Normal, Doubted, and Confirmed). Therefore, the rule base is established to have a total of 729 rules, of which seven rules belong to "Normal," 57 rules belong to "MCI (Doubted), and the other 665 rules belong to "MCI (Confirmed)."

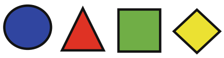
System Scenario Design. This study used the following 4-step systematic development method: development of theoretical basis, system analysis, system design, and system test. Since this study used a 3D virtual environment to simulate a scenario image test environment, this environment can encourage middle-aged to older subjects to undergo testing and reduce the anxiety and boredom associated with hardcopy tests. The navigation guides the subject to naturally integrate into the test scenario, allowing a more precise assessment of whether the subject has MCI and the possibility of transferring to AD in the future. There are night test regions based on eight dimensions to be designed in an MCI testing center.

Orientation Test. The orientation test included orientation in three areas (people, space, and time), and generally, ordinary people should be able to observe and identify the person, place, and time, i.e., who? where? and when? If this ability was absent and orientation lost, the subject would be considered disorientated. Figure 1(a) shows one of the tests: Chinese zodiac testing.

Identification Test. The left temporal lobe is mainly responsible for language expression and language comprehension. This study used language expression, image recognition, and language comprehension to assess the ability of identification. Figure 1(b) shows the language expression test: use the mouse to click _cannot be eaten.

Memory Test. The most significant early symptom of patients with AD is memory loss. Studies showed that in cases of short-term memory loss, there is a significant decline in visual connection, e.g., Object's color and image connection. Therefore, the following two tests are designed in this study: the object's shape and image connection and the

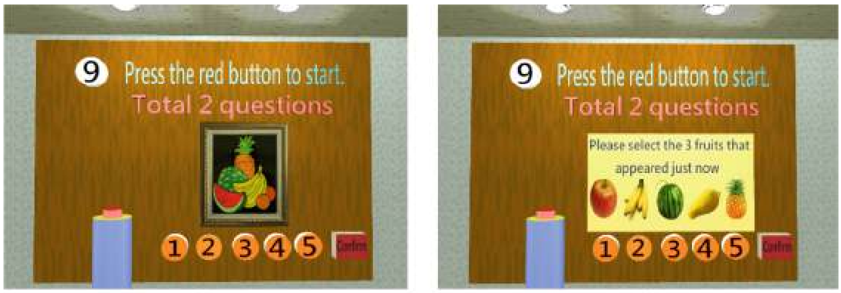
Table 1. Assignment table of assessment items conversion to various screening tools

Stage	Dimension	Sub-item	Content
1. Perception	1. Orientation	1. Orientation	1. Temporal orientation: (1) Zodiac sign? (2) What season is it? (3) What month is it? (4) What day of the week is it? 2. Spatial orientation: where is the Mona Lisa painting?
	2. Identification	2. Language expression	1. ___cannot be consumed (images: apple, bottle, pear) 2. Closed_(images: hands, eyes, pants) 3. I want to ride in a/an (images: mobile phone, airplane, camera) 4. I ___ to eat (words: want, see, drink) 5. I walk__(words: home, recollection, family)
		3. Image recognition	1. Images (toys, knives, tools); select the knife 2. Images (television, fan, refrigerator); select the television
		4. Language comprehension	1. Clapping; select an image (clapping, applause, bringing the palms together) 2. Walking; select an image (jumping, walking, squatting)
2. Memory	3. Memory	5. Short-term memory	1. Images of banana, pineapple, watermelon, etc., will appear, and you are required to select the banana, pineapple and watermelon from the 5 objects for a total of 2 times. 2.  After memorizing, shape and color matching are conducted twice.
		6. Long-term memory	1. Who is the current president in Taiwan? Photographs of Chen Shui-Bian, Ma Ying-Jeou, and Tsai Ing-Wen. 2. Place of residence? Maps of Taiwan, Japan, and Hainan Island. 3. How many minutes are there in 1 hour? 4. How many months are there in 1 year? 5. In which direction does the sun set?
3. Attention	4. Attention	7. Attention	1. Random shapes: determine the shape shown on the screen as △, □, and ○; 5 consecutive times.
4. Cognition	5. Construction	8. Construction	1. How do you draw Mickey Mouse using △, □, and ○? 2. How do you construct the right shape using △, □, and ○?  3. How do you construct the right shape using △, □, and ○?  4. Label the 2 o'clock and 8 o'clock positions in the image (2 positions).
	6. Judgment	9. Judgment	1. Select the right clock if the time is 10:10. 2. What do physicians use? (Images: scalpel, bottle, and wooden rod). 3. What do teachers use? (Images: pot, books, and hammer).
	7. Calculation	10. Calculations	1. How many fishes are there in total in the tank? 2. What is 3 + the previous answer? 3. What is 3 + the previous answer? 4. What is the previous answer + 3 twice? 5. You used 50 NTD to buy a 3 NTD apple and a 20 NTD notebook. How much do you have left?
5. Action	8. Control ability	11. Control ability	1. Rotate the 3D cube to the front color is 2. Rotate the 3D cube to the front color is 3. Rotate the 3D cube to the front color is 

object's color and image connection are used for short-term memory assessment (Figs. 2 (a)(b)).

(a) (b)

Fig. 1. (a) Orientation test: Chinese zodiac test; (b) Identification test: language expression test.



(a) (b)

Fig. 2. Short-term memory test: (a) A fruit drawing appears for 5 s before disappearing; (b) A fruit list appears, and the subject must select the fruit that had just appeared.

Attention Test. In this study, several random temporary images are used to assess whether the subject could pay attention to the characteristics of events that occur within a short period, and select the correct answers.

Construction Test. In this study, an image formed from shapes is shown, and the subject is asked what shape composes of the image. The objective of this test is to test whether the subject understands shape composition.

Judgment Test. After early memory loss occurs in dementia patients, it will affect judgment. In this study, one of the tests is to select the right clock if the time is 10:10.

Calculation Test. In this study, subjects observe fish in a fish tank and are instructed to count the number of fishes in a relaxed manner to avoid sensitivity and fear toward numbers, which would result in misjudgment.

Control Ability Test. Control ability is a behavioral presentation of cognitive decision-making. When cognitive impairment occurs, it will affect control behavior to worse. It may be an auxiliary proof of cognitive impairment. In this study, mouse operation was used to assess the control ability.

Test Completion Screen. After the subject completes the eight major tests, the system will carry out the MCI assessment result. The finally result is shown in Fig. 3, the right image shows the corresponding values of the six traditional screening tools (blue rectangles show the corresponding values). The lower left image shows the fuzzy MCI forecasting value, which provided the forecasting score of the possibility of MCI to the subject.

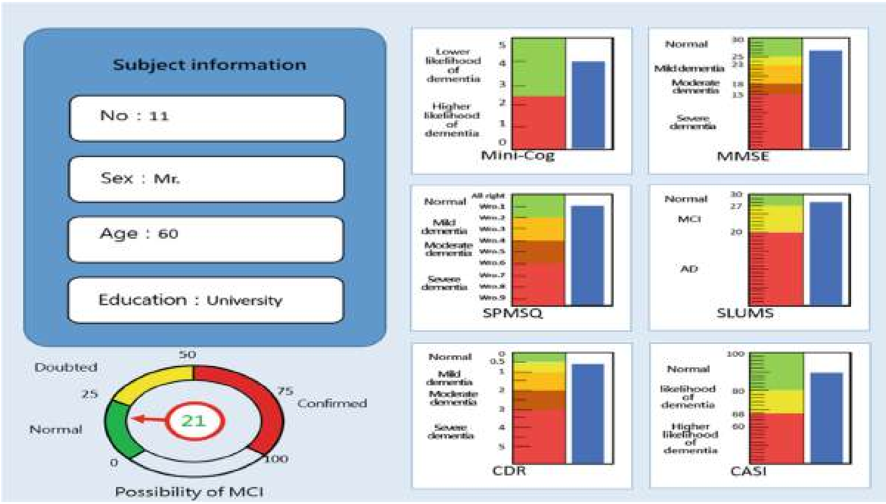


Fig. 3. Output display of testing result.

3.2 Subjects

In this study, 24 middle-aged to old people aged 50 to 65 years, which is the age group when early cognitive impairment tends to occur, were invited to participate in MCI forecasting system testing. All subjects received visual acuity and laptop-operated tests to decrease experimental variation due to inter-individual differences.

4 Results and Discussion

4.1 Correlation Analysis of MCI Forecasting System

The results showed that the MCI forecasting system developed in this study is highly correlated with traditional hardcopy Mini-Cog, SPMSQ, MMSE, SLUMS, and CASI tests, of which MMSE and CASI had extremely high correlation, but the correlation with CDR was lower. The reason may be that the CDR is an additive test, the scoring method differs slightly from other tests, and the total score is only 3. When the MCI system generates the corresponding value of CDR, minor errors might increase the score by 0.5 points, resulting in a large discrepancy. Overall, the forecasting system test can preliminarily obtain results consistent with most hardcopy tests.

4.2 Effectiveness Analysis of Fuzzy Prediction

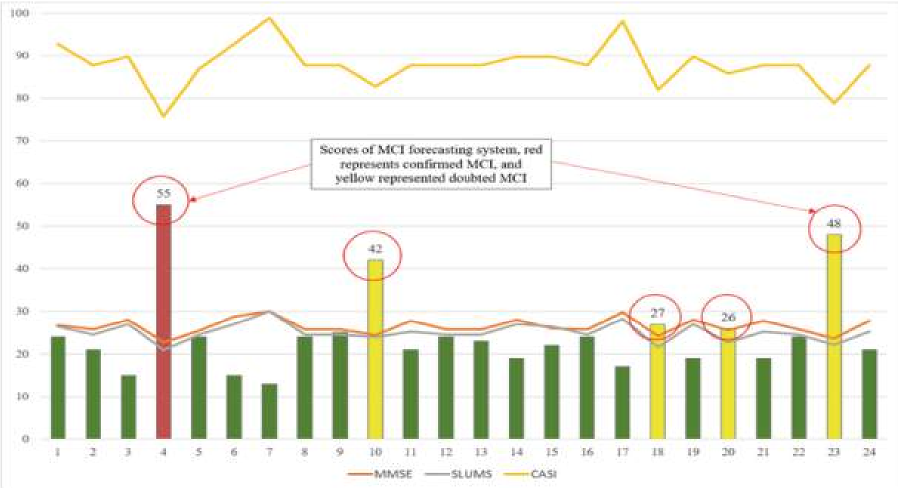


Fig. 4. Distribution plots of fuzzy prediction scores and hard copy test scores for MMSE, SLUMS, and CASI.

In Fig. 4, the bar chart shows the scores for 24 subjects resulting from the MCI fuzzy prediction module. As the objective of the output result is for warning, it can promptly provide a reference value for the tester. In particular, if a subject has a situation similar to subject No. 20 occurred in which the test scores between the MCI forecasting system and traditional hardcopy tests are inconsistent: traditional Mini-Cog and CDR demonstrate a result of suspected MCI. However, the corresponding Mini-Cog and SLUMS in the forecasting system demonstrate a result of suspected MCI (see Fig. 4). At that point, the fuzzy prediction score of the forecasting system was 26, which reached the warning threshold for suspected MCI. Therefore, the subject can be judged as a doubted MCI patient. In addition, subject No. 4 had a higher fuzzy prediction score, showing that it is highly likely that the subject has MCI. This result also demonstrated the particular function of the MCI forecasting system.

5 Conclusion

This study employed VE technology and fuzzy logic control (FLC) to integrate Mini-Cog, SPMSQ, MMSE, SLUMS, CDR, and CASI into an image "MCI forecasting system." The results showed that the MCI forecasting system is highly correlated with traditional hardcopy Mini-Cog, SPMSQ, MMSE, SLUMS, and CASI tests, of which MMSE and CASI had a significant correlation but insignificant with CDR. Furthermore, the fuzzy predicting marker can provide an MCI reference score to the assessor to assist the assessor in making judgments and achieving the objective of warning.

Because this study focused on the development and piloting of MCI forecasting by FLC, a small group of ages between 50 and 65 adults was selected. However, the age group at which MCI may occur may be less than 50 years old. The MCI occurs at a younger age, the possibility of conversion to AD will be higher. In the future, relevant research on the 40–50 age group will be considered to obtain complete evaluation information.

Acknowledgements. The authors gratefully acknowledge the financial support for this work from the National Science and Technology Council of Taiwan, R.O.C.

References

1. Reisberg, B., Ferris, S.H., DeLeon, M.J., Crook, T.: The global deterioration scale for assessment of primary degenerative dementia. Am. J. Psychiatry **139**(9), 1136–1139 (1982)
2. Rossetto, F., et al.: Cognitive and affective theory of mind in mild cognitive impairment and Parkinson's disease: preliminary evidence from the Italian version of the yoni task. Dev. Neuropsychol. **43**(8), 764–780 (2018)
3. Qi, H.H., Liu, H., Hu, H.M., He, H.J., Zhao, X.H.: Primary disruption of the memory-related subsystems of the default mode network in alzheimer's disease: resting-state functional connectivity MRI study. Front. Aging Neurosc. **10**, 344 (2018)
4. Gosztolya, G., Vincze, V., Toth, L., Pakaski, M., Kalman, J., Hoffmann, I.: Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. Comput. Speech Lang. **53**, 181–197 (2019)
5. National Institute for Health and Clinical Excellence: Dementia: Quick Reference Guide. UK, London (2006)
6. Flier, W.M., et al.: Alzheimer's disease. Lancet **2016**(388), 505–517 (2016)
7. LNCS Homepage. http://www.springer.com/lncs, (Accessed 21 Nov 2016)
8. Zadeh, L.A.: Fuzzy Sets. Inf. Control **8**, 338–353 (1965)
9. Bellman, R.E., Zadeh, L.A.: Decision-making in a fuzzy environment. Manage. Sci. **17**, 141–164 (1979)

Fuzzy Multi-criteria Sorting Models FMAVT-Sort and FTOPSIS-Sort: Features and Application Within the Case Studies

Alexander Radaev[1(✉)], Elif Haktanir[2,3], Boris Yatsalo[4], and Cengiz Kahraman[2]

[1] Institute of Atomic Energy of the National Research Nuclear University MEPHI (IATE MEPHI), Obninsk, Russian Federation
alexander.radaev.1994@gmail.com

[2] Department of Industrial Engineering, Istanbul Technical University, Istanbul 34367, Besiktas, Turkey
kahramanc@itu.edu.tr

[3] Department of Industrial Engineering, Altinbas University, Istanbul 34217, Bagcilar, Turkey
elif.haktanir@altinbas.edu.tr

[4] Foundation Progress and Business, Krakow, Poland

Abstract. The need to sort alternatives according to ordered categories, especially in problems consisting of a large number of alternatives, makes it effective to use of multi-criteria sorting methods. A family of fuzzy multi-criteria sorting models FMAVT-Sort and FTOPSIS-Sort based on a fuzzy extension of MAVT and TOPSIS methods respectively and the their features are presented. In the developed models, different approaches can be used both for calculating functions of fuzzy numbers and for ranking of fuzzy numbers. Case studies showing the application of the developed models to sort problems related to the Internet of Things and anti-COVID measures are presented.

Keywords: Fuzzy numbers · Fuzzy multi-criteria sorting · Fuzzy ranking methods · Fuzzy MAVT · Fuzzy TOPSIS · Fuzzy MCDA

1 Introduction

Multi-criteria decision analysis (MCDA) can be used in diverse classes of problems, and one of them are problems with a large number of alternatives. Multi-criteria sorting models can be used to assign an alternative from a given set to one of the predefined ordered categories [26]. The most popular multi-criteria sorting models are extensions of existing MCDA methods. The most common are the following multi-criteria sorting methods: ELECTRE TRI, UTADIS, Flow-Sort. For decision making under uncertainty, ordinary MCDA sorting methods can be extended to fuzzy MCDA sorting (FMCDA-Sort) models.

Calculating functions of Fuzzy Numbers (FNs) and ranking of FNs are the main stages within FMCDA. The ordinary fuzzy sets by Zadeh [24] are considered in this contribution.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 725–734, 2023.
https://doi.org/10.1007/978-3-031-39777-6_85

In this work, the following calculation methods are used: Approximate assessing functions based on the basic type of FNs: Triangular FNs (TrFNs) or/and Trapezoidal FNs (TpFNs) [12]; Standard Fuzzy Arithmetic (SFA) [6,8] and Reduced Transformation Method (RTM) [8].

In addition, the most used approach in FMCDA is based on approximate calculations. In this case, the overestimation problem may occur and the final result will differ from one obtained by the Zadehs extension principle [16,25].

As for the methods of ranking of FNs, in this contribution, we use two well-known and widely used defuzzification methods: Centroid Index (CI) [13], Integral of Means (IM) [17].

The novelty of the contribution is the implementation of families of original FMCDA sorting models: FMAVT-Sort and FTOPSIS-Sort. The main difference between these models and the existing ones is the possibility of using not only approximate calculations and defuzzification methods, but also SFA and RTM for calculating functions of FNs. FNs of the general form, as well as linguistic variables, can be used to evaluate alternatives by criteria, weights coefficients, and limiting profiles. The developed models can be used in different problems involving sorting alternatives based on the use of criteria values presented by FNs of the general form.

This paper is structured as follows. Section 2 briefly describes the families of models presented in the work. In Sect. 3 case studies on sorting of Internet of Things platforms and non-pharmaceutical measures against COVID-19 illustrate application of developed models. Finally, Sect. 4 concludes this paper.

2 Fuzzy Multi-criteria Sorting Models

The history of multi-criteria sorting methods begins with the ELECTRE-TRI method [14]. This method is the most used method in publications. Other ones that are sufficiently used in case studies include methods UTADIS, ELECTRE TRI-C and Flow-Sort. As for fuzzy models, most of them are fuzzy extensions of methods Flow-Sort [3], ELECTRE TRI [7] and AHPSort [11]. It is necessary to note that the above publications do not use SFA and transformation methods. Multi-criteria sorting methods are quite popular in supplier selection, risk assessment and performance analysis. A more detailed analysis of sorting methods is presented in [2].

2.1 FMAVT and FTOPSIS Models

There are many publications describing the fuzzy extension of the MCDA methods. E.g., FTOPSIS: [4,5,20], FMAVT: [19,21].

The MAVT method is widely used MCDA method; in this work, an additive model is used [10]:

$$V(\boldsymbol{a}_i) = \sum_{j=1}^{m} w_j V_j(c_{ij}), \tag{1}$$

here, c_{ij} is the value of alternative a_i for criterion j, w_j is the weight coefficient, and $V_j(x)$ is the partial value function (PVF) for criterion j, $j = 1, \ldots, m, i = 1, \ldots, n$; $V(a_i)$ is interpreted as the generalized value of the alternative a_i.

In this work, the following PVFs is used (for a benefit criterion j):

$$V_j(x) = (1 - exp(\alpha(x - c_{j,min})))/(1 - exp(\alpha(c_{j,max} - c_{j,min}))), \quad (2)$$

here $c_{j,min} = \min_i c_{ij},\ c_{j,max} = \max_i c_{ij},\ i = 1, ..., n.$

PVFs, $V_j(x)$, implement normalization of criteria values: $c_{ij} \rightarrow x_{ij} = V_j(c_{ij})$, $x_{ij} \in [0, 1]$. In this paper, we consider PVFs are *strictly monotonically* increasing for benefit criteria and decreasing for cost criteria.

A *fuzzy extension* of additive MAVT [18,19,21] is based on Eq. (1) with the use of fuzzy criteria values, c_{ij}, and fuzzy weight coefficients, w_j. In case of local FMAVT, FMAVT$_{loc}$, criteria values, $c_{ij} = \{[A_\alpha^{ij}, B_\alpha^{ij}]\}, \alpha \in [0, 1]$, consider marginal values (real numbers) for criterion j: $A_0^j = \min_i A_0^{ij},\ B_0^j = \max_i B_0^{ij}$, and (real) PVF, $V_j(x)$, is set in segment $[A_0^j, B_0^j]$ as for MAVT.

As for the TOPSIS method [9], the calculation of the generalized criterion is based on the following expression.

$$D_i = D(a_i) = \frac{D_i^-}{D_i^- + D_i^+} = \frac{(\sum_1^m w_k^2(x_{ik} - x_k^-)^2)^{1/2}}{(\sum_1^m w_k^2(x_{ik} - x_k^-)^2)^{1/2} + (\sum_1^m w_k^2(x_k^+ - x_{ik})^2)^{1/2}}, \quad (3)$$

here x_{ik} is a normalized criterion value of alternative a_i for criterion k, $i = 1, \ldots, n$, $k = 1, \ldots, m$, x_k^+ and x_k^- are, respectively, normalized coordinates of *ideal* and *anti-ideal* points/alternatives in $\mathbb{R}^m$, w_k is a weight coefficient, D_i^+ and D_i^- are, correspondingly, weighted distances from the alternative $a_i = (x_{i1}, \ldots, x_{im})$ to ideal, $I^+ = (x_1^+, \ldots, x_m^+)$, and anti-ideal, $I^- = (x_1^-, \ldots, x_m^-)$, alternatives.

Combining methods for calculating functions of FNs and methods of ranking FNs, a family of models (FMAVT/FTOPSIS) can be obtained which are a fuzzy extension of the classical method (MAVT/TOPSIS).

It should be noted that a more detailed description of the fuzzy extension of methods MAVT and TOPSIS (FMAVT and FTOPSIS accordingly), their features and application are presented in the works of the authors [19,23].

2.2 Fuzzy Multi-criteria Sorting Models FMAVT-Sort and FTOPSIS-Sort

MCDA/FMCDA sorting models allow alternatives to be sorted into ordered categories $Q_1, \ldots, Q_K$ (category Q_h is preferred to Q_{h+1}). To form the ordered categories within a multi-criteria sorting problem, limiting (used in this work) or central profiles are assigned [26].

Fuzzy criteria values, c_{ij}, of alternative a_i for criterion j, and fuzzy weight coefficients, $w_j,\ j = 1, ..., m,\ \ i = 1, ..., n$, are considered within an FMCDA problem along with *fuzzy limiting profiles*, $P_h = (P_{h1}^l, ..., P_{hm}^l),\ h = 1, ..., K+1$, where P_{hj}^l is a FN (a limiting profile h for criterion j), and vector P_{h+1} is dominated in Pareto by P_h: $P_{h+1} \prec_{P(M)} P_h$, here M is a used fuzzy ranking method, and FNs P^l_{hj}, $h = 1, ..., K+1$, form a set of distinguishable FNs [22].

The decision rule for assigning an alternative a_i to corresponding category $Q(a_i)$ within FMAVT-Sort/FTOPSIS-Sort model with ranking method M is as follows:

$$Q(a_i) = Q_1 \ if \ V(P_2) \prec_M V(a_i), Q(a_i) = Q_h \ if \ V(P_{h+1}) \prec_M V(a_i) \preceq_M V(P_h), h = 2, \ldots, K-1; Q(a_i) = Q_K \ if \ V(a_i) \preceq_M V(P_K), \quad (4)$$

here $V(a)$ (or $D(a)$ for FTOPSIS-Sort models) is the value of generalized criterion (1)/(3) for alternative a. According to [19], for the set of (strictly) distinguishable vectors of FNs, P_h, $h = 1, ..., K+1$, $V(P_{h+1}) \prec_M V(P_h)$, $h = 1, ..., K$, for ranking method $M = CI, IM$. In addition, a careful analysis of the Eq. (4) shows that for correct implementation of this decision rule, the marginal limiting profiles, P_1 and P_{K+1}, $h = 1, ..., K+1$, should be considered as singletons.

The following FM-Sort (FMAVT-Sort) and FT-Sort (FTOPSIS-Sort) models are considered: FMTrCI/IM-Sort, FMSCI/IM-Sort, FTTrCI/IM-Sort, FTSCI/IM-Sort, FTRCI/IM-Sort, where Tr means approximate calculations, S means SFA, R means RTM, CI and IM are methods for ranking of FNs.

3 The Case Studies

To illustrate the developed models, two case studies were considered. Functions of FNs within FMAVT-Sort and FTOPSIS-Sort models under consideration were determined with the number of α-cuts $N_\alpha = 20$ and then recalculated for $N_\alpha = 40$; according to assessments, no differences were found in the output results.

3.1 Sorting Internet of Things Platforms

The growth in the number of devices connected to the network, the growth in the amount of information transmitted between them, goes hand in hand with such a concept as the Internet of Things (IoT). Today, companies of different sizes from almost all sectors use various IoT applications to work more efficiently, improve business processes, make more effective decisions, and increase the value of their business. When evaluating different purchasing options, Industrial IoT Platforms users compare competencies (criteria) in categories such as evaluation and contracting, integration and deployment, service and support, and specific product capabilities [1].

Table 2 (performance table) presents 15 industrial IoT provider alternatives: A_1: Azure IoT, A_2: Oracle IoT Cloud, A_3: PTC ThingWorx, A_4: IBM Watson IoT, A_5: SAP Leonardo IoT, A_6: Software AG Cumulocity IoT, A_7: Altizon Datonis Industrial IoT Platform, A_8: Everyware IoT, A_9: Siemens MindSphere, A_{10}: Atos Codex IoT; A_{11}: iLens, A_{12}: PI System, A_{13}: Accenture Connected Platforms as a Service, A_{14}: Davra IoT Platform, A_{15}: Predix Platform, and 4 criteria, which are *benefit* ones.

Table 1. Linguistic terms and their TrFNs correspondences

Linguistic Term	Corresponding TrFN (for benefit criterion)
Low (L)	(0, 0, 0.25)
Between Low and Middle (BLM)	(0, 0.25, 0.5)
Middle (M)	(0.25, 0.5, 0.75)
Between Middle and High (BMH)	(0.5, 0.75, 1)
High (H)	(0.75, 1, 1)

Table 2. Performance Table, IoT Case Study

Alternatives/Criteria	C_1	C_2	C_3	C_4
A_1	BMH	BMH	H	H
A_2	BMH	BMH	H	BMH
A_3	BMH	M	H	H
A_4	BLM	M	M	M
A_5	L	BLM	M	M
A_6	M	BMH	M	BMH
A_7	H	M	H	BMH
A_8	L	BLM	M	BLM
A_9	BMH	H	H	H
A_{10}	L	BLM	L	BMH
A_{11}	H	H	H	H
A_{12}	BLM	BMH	BMH	M
A_{13}	L	L	BMH	L
A_{14}	H	H	H	H
A_{15}	L	L	L	M

Table 3. Limiting profiles, IoT Case Study

Criteria/Lim. Profiles	P_5	P_4	P_3	P_2	P_1
C_1, C_2, C_3, C_4	0	(0.2, 0.25, 0.3)	(0.45, 0.5, 0.55)	(0.7, 0.75, 0.8)	1

Table 4. Criteria weight coefficients (aggregated), IoT Case Study

	C_1	C_2	C_3	C_4
Weights	(0.4, 0.65, 0.85)	(0.4, 0.65, 0.85)	(0.2, 0.45, 0.7)	(0.65, 0.9, 1)

Within fuzzy multi-criteria sorting problem, four categories are considered: Q_1 – Absolutely capable alternatives; Q_2 – High level capable alternatives; Q_3 – Middle level capable alternatives; Q_4 – Low level capable alternatives.

Criteria weight coefficients are determined by aggregating the views of five experts. Experts evaluations with the use of linguistic representations, Table 1, and the aggregated fuzzy criteria weights are presented in Table 4.

Experts agreed to set linear partial value functions for criteria C_1 and C_2, and non-linear (in accordance with Eq. (2)) for criteria C_3 (over-linear PVF, $\alpha = -5$) and C_4 (sub-linear PVF, $\alpha = 5$).

In Table 5, the value $I(V(A_i))$, $I = CI, IM$, $i = 1, \ldots, 15$, for the generalized criterion $V(A_i)/D(A_i)$ (1)/(3) and corresponding model (FMAVT-Sort/FTOPSIS-Sort), is indicated in brackets.

Table 5. The results of assigning alternatives to categories by FMAVT-Sort and FTTrCI-Sort/FTRCI-Sort models

Models	Categories			
	Q_4	Q_3	Q_2	Q_1
FMTrCI-Sort	A_{15} (0.43); A_{13} (0.59)	A_8 (0.713); A_5 (0.803); A_{10} (0.869); A_4 (1.086)	A_{12} (1.271)	A_6 (1.717); A_7 (1.837); A_2 (1.908); A_3 (2.01); A_1 (2.168); A_9 (2.256); A_{11}, A_{14} (2.343)
FMTrIM-Sort	A_{15} (0.339); A_{13} (0.553)	A_8 (0.683); A_{10} (0.755); A_5 (0.764); A_4 (1.058)	A_{12} (1.243); A_6 (1.658)	A_7 (1.798); A_2 (1.851); A_3 (2.048); A_1 (2.207); A_9 (2.314); A_{11}, A_{14} (2.42)
FMSCI-Sort	A_{15} (0.414); A_{13} (0.582)	A_8 (0.695); A_5 (0.774); A_{10} (0.812); A_4 (1.052)	A_{12} (1.237); A_6 (1.646)	A_7 (1.77); A_2 (1.837); A_3 (1.952); A_1 (2.11); A_9 (2.204); A_{11}, A_{14} (2.297)
FMSIM-Sort	A_{15} (0.322); A_{13} (0.543); A_8 (0.665)	A_{10} (0.695); A_5 (0.735); A_4 (1.023)	A_{12} (1.209); A_6 (1.586)	A_7 (1.73); A_2 (1.779); A_3 (1.986); A_1 (2.145); A_9 (2.256); A_{11}, A_{14} (2.366)
FTTrCI-Sort		A_{13} (0.441); A_{15} (0.548); A_8 (0.627)	A_{10} (0.704); A_5 (0.746)	A_4 (1.006); A_{12} (1.082); A_3 (1.151); A_7 (1.243); A_6 (1.284); A_{11}, A_{14} (1.297); A_9 (1.302); A_1 (1.341); A_2 (1.568)
FTRCI-Sort		A_{13} (0.259); A_8 (0.272); A_{15} (0.321); A_5 (0.375); A_{10} (0.44); A_4 (0.443)	A_{12} (0.524); A_6 (0.646); A_7 (0.703); A_3 (0.727)	A_2 (0.76); A_1 (0.806); A_9 (0.84); A_{11}, A_{14} (0.917)

According to Table 5, alternatives A_6, A_8, and A_{10} are assigned to different categories for different FMAVT-Sort models. At the same time, for models differing in the ranking method of FNs, the above alternatives belong to categories with a lower index (i.e., they are more capable) for models ranked by CI (Q_1

Table 6. Limiting profiles, anti-COVID Case Study

Criteria	Limiting profiles			
	$P^l_{1,j}$	$P^l_{2,j}$	$P^l_{3,j}$	$P^l_{4,j}$
C_1: Common satisfaction of the society (Benefit)	1	(0.5, 0.6, 0.7)	(0.4, 0.5, 0.6)	0
C_2: Maximum required time to implement (Cost)	0	(0, 0.2, 0.4)	(0.2, 0.4, 0.6)	1
C_3: Financial burden to the economy (Cost)	0	(0, 0.2, 0.4)	(0.2, 0.4, 0.6)	1
C_4: Effectiveness on virus spread (Benefit)	1	(0.2, 0.4, 0.6)	(0, 0.2, 0.4)	0

for FMTrCI-Sort and Q_3 for FMSCI-Sort). This fact may indicate a greater influence of overestimation on models that use ranking by CI.

Additionally, FTTrCI-Sort and FTRCI-sort models were used, as belonging to another family, it can see that the results differ from the FMAVT-Sort models: no alternative is assigned to the category Q_4, and same time for the FTTrCI-Sort (approximate) model the category Q_1 contains more alternatives than the one for the FTRCI-Sort (proper) model.

3.2 Sorting Anti-COVID Non-pharmaceutical Interventions/Measures

Despite the decrease in the number of cases of Covid-19, the problems associated with the application of measures to control the spread of diseases remain important. In this study, the NPI alternatives defined in [15] are sorted by the proposed models. These NPI alternatives are: A_1: Quarantine of patients and those suspected of infection, A_2: Internal border restrictions (within a country), A_3: Social distancing, A_4: Health monitoring, A_5: Public awareness campaigns, A_6: Restriction of nonessential businesses, A_7: Restrictions of mass gatherings, A_8: External border restrictions, A_9: Closure of schools, A_{10}: Declaration of emergency, A_{11}: Formation of new task units/bureaus and government policies changing administrative capacity to respond to the crisis, A_{12}: Common health testing, A_{13}: Curfew, A_{14}: Restriction of nonessential government services. These alternatives are analyzed based on four criteria: C_1: Common satisfaction of the society (benefit), C_2: Maximum required time to implement (cost), C_3: Financial burden to the economy (cost), C_4: Effectiveness on virus spread (benefit). The performance table (criteria values for all alternatives) are presented in Table 7. Fuzzy weight coefficients have been evaluated as averaging expert judgments in linguistic scale, see Table 8.

The four limiting profiles for this fuzzy multi-criteria sorting problem, see Table 6, define three ordered categories, G_h, $h = 1, 2, 3$, with the following interpretation: G_1: the category of effective interventions; G_2: slightly; G_3: poorly.

The decision rule (4) is used for assigning each alternative A_i, $i = 1, ..., 14$, to categories Q_h, $h = 1, 2, 3$.

In accordance with Table 9, models with different ranking methods but the same approach to assess functions of FNs lead to the same sorting of alternatives. At the same time, models with different computational algorithms, Table 9, lead

Table 7. Performance Table, anti-COVID Case Study

Alternatives/ Criteria	C_1 (Benefit)	C_2 (Cost)	C_3 (Cost)	C_4 (Benefit)
A_1	BLM	LM	MH	H
A_2	L	BLM	H	H
A_3	MH	L	BLM	H
A_4	MH	M	BMH	MH
A_5	H	BMH	M	M
A_6	BLM	BMH	M	M
A_7	L	L	L	H
A_8	L	H	H	H
A_9	L	BLM	H	MH
A_{10}	L	BLM	BMH	H
A_{11}	H	BMH	MH	M
A_{12}	H	BMH	MH	MH
A_{13}	L	BLM	H	H
A_{14}	L	M	BLM	LM

Table 8. Criteria weight coefficients (aggregated), anti-COVID Case Study

	C_1	C_2	C_3	C_4
Weight	(0.42, 0.67, 0.83)	(0.5, 0.75, 0.92)	(0.5, 0.75, 0.92)	(0.67, 0.92, 1)

Table 9. The results of assigning alternatives to categories by FTOPSIS-Sort and FMTrCI-Sort/FMSCI-Sort models

Method	Categories		
	G_3	G_2	G_1
FTTrCI-Sort		A_8 (0.524); A_6 (0.531); A_9 (0.639); A_2, A_{13} (0.681)	A_{14} (0.735); A_{10} (0.752); A_{11} (0.78); A_1 (0.823); A_{12} (0.831); A_7 (0.877); A_5 (0.881); A_4 (0.977); A_3 (1.36)
FTTrIM-Sort		A_6 (0.451); A_8 (0.499); A_9 (0.59); A_2, A_{13} (0.639)	A_{14} (0.675); A_{10} (0.701); A_{11} (0.705); A_{12} (0.76); A_1 (0.764); A_5 (0.793); A_7 (0.818); A_4 (0.874); A_3 (1.217)
FTSCI-Sort		A_6 (0.446); A_8 (0.493); A_9 (0.586); A_2, A_{13} (0.635)	A_{14} (0.671); A_{10} (0.695); A_{11} (0.695); A_{12} (0.752); A_1 (0.755); A_5 (0.781); A_7 (0.814); A_4 (0.868); A_3 (1.198)
FTSIM-Sort		A_6 (0.351); A_8 (0.466); A_9 (0.532); A_2, A_{13} (0.589)	A_{14} (0.605); A_{11} (0.612); A_{10} (0.64); A_{12} (0.674); A_5 (0.682); A_1 (0.691); A_7 (0.75); A_4 (0.752); A_3 (1.036)
FTRCI-Sort	A_6 (0.227)	A_8 (0.427); A_9 (0.452); A_{11} (0.469); A_{14} (0.498)	A_5 (0.515); A_2, A_{13} (0.518); A_{12} (0.531); A_{10} (0.544); A_4 (0.558); A_1 (0.568); A_7 (0.634); A_3 (0.773)
FTRIM-Sort	A_6 (0.222)	A_8 (0.425); A_9 (0.449); A_{11} (0.472); A_{14} (0.498)	A_2, A_{13} (0.517); A_5 (0.518); A_{12} (0.535); A_{10} (0.545); A_4 (0.56); A_1 (0.573); A_7 (0.636); A_3 (0.778)
FMTrCI-Sort	A_6 (0.73)	A_8 (1.03); A_9 (1.287); A_{11} (1.496); A_2, A_{13} (1.531)	A_{14} (1.572); A_{10} (1.67); A_5 (1.677); A_{12} (1.712); A_1 (1.795); A_4 (1.802); A_7 (2.107); A_3 (2.448)
FMSCI-Sort	A_6 (0.702)	A_8 (1.015); A_9 (1.268); A_{11} (1.466); A_2, A_{13} (1.51)	A_{14} (1.547); A_{10} (1.644); A_5 (1.647); A_{12} (1.682); A_1 (1.764); A_4 (1.769); A_7 (2.087); A_3 (2.419)

to different sorting alternatives. For comparison, the results of the FMTrCI-Sort and FMSCI-Sort models are shown (in this case, linear VFs are used). The result on sorting alternatives by FMAVT-Sort is the same. At the same time, alternatives A_2, A_{13} belong to the second category, in contrast to the proper FTOPSIS-Sort models. These alternatives are close to the limiting profile ($CI(G(P_3)) = 1.525$, $CI(G(A_2)) = 1.51$ for FMSCI-Sort model), which indicates the possibility of assigning to different categories for different methods.

In accordance with the Tables 5 and 9, the results of applying models with proper assessing functions of FNs are the same regardless of the methods for ranking FNs. Thus, the authors expect that the use of models FMSCI-Sort and FMSIM-Sort (this models lead to proper assessing due to there is no dependent variables in the expression (1)), FTRCI-Sort and FTRIM-Sort, is reasonable.

4 Conclusions

In this contribution, families of Fuzzy MAVT and Fuzzy TOPSIS Sorting (FMAVT-Sort and FTOPSIS-Sort) models are implemented within the two case studies with the use of different fuzzy extensions of ordinary MAVT and TOPSIS methods accordingly. An important feature of these models is the use of proper assessing functions of FNs, in contrast to the existing fuzzy sorting models, which do not use either SFA or RTM. The other feature of this work is the comparison of models based on the extension of the same MCDA method using the same input values. These comparisons make it possible to analyze how the method of calculating functions of FNs and/or the fuzzy ranking method influence the application of FMCDA sorting models representing extensions of the same ordinary MCDA method. In addition, a comparison of fuzzy sorting models representing the fuzzy extensions of different MCDA methods is given.

The further directions of research and development will be include: development of new families of FMCDA-Sort models based on existing MCDA methods (eg. PROMETHEE); analysis the frequency and the level/significance of distinctions in sorting alternatives by FMCDA-Sort models, which extend the same ordinary MCDA method by fuzzy numbers.

References

1. AWS IoT Alternatives. https://www.gartner.com/reviews/market/industrial-iot-platforms/vendor/amazon-web-services/product/aws-iot/alternatives, (Accessed 28 March 2023)
2. Alvarez, P., Ishizaka, A., Martinez, L.: Multiple-criteria decision-making sorting methods: a survey. Expert Syst. Appli. **183**(3) (2021)
3. Campos, A.C., Mareschal, B., de Almeida, A.T.: Fuzzy FlowSort: an integration of the FlowSort method and fuzzy set theory for decision making on the basis of inaccurate quantitative data. Inf. Sci. **293**, 115–124 (2015)
4. Chen, C.T.: Extensions of the TOPSIS for group decision-making under fuzzy environment. Fuzzy Sets Syst. **114**(1), 1–9 (2000)
5. Chen, S., Hwang, C.: Fuzzy Multiple Attribute Decision Making: Methods and Applications, vol. 375. Springer-Verlag, Berlin (1992). https://doi.org/10.1007/978-3-642-46768-4

6. Dubois, D., Prade, H.: Operations on fuzzy numbers. Int. J. Syst. Sci. **9**(6), 613–626 (1978)
7. Galo, N.R., Calache, L.D., Carpinetti, L.C.: A group decision approach for supplier categorization based on hesitant fuzzy and ELECTRE TRI. Int. J. Prod. Econ. **202**, 182–196 (2018)
8. Hanss, M.: Applied Fuzzy Arithmetic. Springer-Verlag (2005). https://doi.org/10.1007/b138914
9. Hwang, C.L., Yoon, K.: Multiple Attribute Decision Making: Methods and Applications. LNEME, vol. 186. Springer-Verlag, Berlin (1981). https://doi.org/10.1007/978-3-642-48318-9
10. Keeney, R., Raiffa, H.: Decisions with multiple objectives: preferences and value tradeoffs. John Wiley & Sons, New York (1976)
11. Krejčí, J., Ishizaka, A.: FAHPSort: a fuzzy extension of the AHPSort method. Int. J. Inform. Technol. Decision Making **17**(04), 1119–1145 (2018)
12. Lee, K.: First Course on Fuzzy Theory and Applications. Advances in Soft Computing. Springer-Verlag, Berlin (2005). https://doi.org/10.1007/3-540-32366-X
13. Ostergaard, J.: Fuzzy logic control of a heat exchanger process. Tekniske Hojskole, Starkstromsafdelingen, Danmarks (1976)
14. Roy, B.: Multicriteria Methodology for Decision Aiding. Springer, US (1996). https://doi.org/10.1007/978-3-031-28465-6
15. Samanlioglu, F., Kaya, B.E.: Evaluation of the COVID-19 pandemic intervention strategies with hesitant f-AHP. J. Healthcare Eng. **2020**, 1–11 (2020)
16. Wang, X., Ruan, D., Kerre, E.: Mathematics of Fuzziness Basic Issues. Springer-Verlag (2009). https://doi.org/10.1007/978-3-540-78311-4
17. Yager, R.: On choosing between fuzzy subsets. Kybernetes **9**(2), 151–154 (1980)
18. Yatsalo, B., Korobov, A., Martínez, L.: Fuzzy multi-criteria acceptability analysis: a new approach to multi-criteria decision analysis under fuzzy environment. Expert Syst. Appl. **84**, 262–271 (2017)
19. Yatsalo, B., Korobov, A., Martínez, L.: From MCDA to Fuzzy MCDA: violation of basic axioms and how to fix it. Neural Comput. Appl. **33**(5), 1711–1732 (2021)
20. Yatsalo, B., Korobov, A., Oztaysi, B., Kahraman, C., Martínez, L.: A general approach to fuzzy TOPSIS based on the concept of Fuzzy multicriteria acceptability analysis. J. Intell. Fuzzy Syst. **38**, 979–995 (2020)
21. Yatsalo, B., Korobov, A., Radaev, A., Qin, J., Martínez, L.: Ranking of independent and dependent fuzzy numbers and intransitivity in fuzzy MCDA. IEEE Trans. Fuzzy Syst. **30**(5), 1382–1395 (2022)
22. Yatsalo, B., Martínez, L.: Fuzzy rank acceptability analysis: a confidence measure of ranking fuzzy numbers. IEEE Trans. Fuzzy Syst. **26**(6), 3579–3593 (2018)
23. Yatsalo, B., Radaev, A., Martínez, L.: From MCDA to Fuzzy MCDA: presumption of model adequacy or is every fuzzification of an MCDA method justified? Inf. Sci. **587**, 371–392 (2022)
24. Zadeh, L.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
25. Zadeh, L.: The concept of a linguistic variable and its applications to approximate reasoning. Inform. Sci. **Part I, II, III**(8,9), 199–249, 301–357, 43–80 (1975)
26. Zopounidis, C., Doumpos, M.: Multicriteria classification and sorting methods?: a literature overview. Eur. J. Oper. Res. **138**, 229–246 (2002)

Hybrid Fuzzy Approach for Bus Type Selection: IETT Case Study

Büşra Buran[1](✉) and Mehmet Erçek[2]

[1] Graduate School, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey
buran18@itu.edu.tr

[2] Management Engineering Department, Istanbul Technical University, 34367 Beşiktaş, Istanbul, Turkey
ercekme@itu.edu.tr

Abstract. Sustainability has been a hot topic in recent years in the literature. It focuses not only on environmental issues but also on economic and social topics. From the public transportation side, it contributes sustainable future taking into account bus type selection with a holistic perspective. In this study, the bus type selection problem for public transportation is modeled in a real case study. The hybrid fuzzy AHP method is applied to the proposed problem based on the SERVQUAL approach. Using Spherical Fuzzy Set, comparison matrices are constructed by both perception and expectation views. Also, gap analysis is conducted to discover differences between expectations and perceptions. Results show that when economic factors are the most important criterion for expected views, operational issues, business, and strategy come first for perceived views. For bus type selection, the diesel bus is chosen by customers with the perceived view. But from the expected perspectives, customers prefer e-bus regarding the proposed criteria. Regarding findings, there is a gap between perceived and expected views. At this point, this study will be a guide for decision-makers to discover the gap regarding customer needs.

Keywords: Fuzzy AHP · SERVQUAL approach · Gap analysis · Public transportation

1 Introduction

Sustainability is also has been an important topic for transportation. Nowadays, we have to handle climate change and global warming problems. From this perspective, transportation plays a critical role due to decreasing greenhouse gas emissions. Regarding the United Nations of sustainable transportation conference, approximately one-quarter of greenhouse gas emissions come from transportation [1]. When we look at the United States region, this ratio rises to twenty-nine percent [2]. At this point, public transportation will help to reduce the emissions regard with taking action. United Nations results showed that there is a reduction of about two percent in Europe between 2000–2019 [1].

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 735–743, 2023.
https://doi.org/10.1007/978-3-031-39777-6_86

To decrease greenhouse emissions, both transport operators and authorities take actions such as transforming their fleets into green transportation. For example, RATP which is the transportation authority for France released a plan that is called sustainable mobility [3]. Also, The New York State Department of Transportation (NYSDOT) presented a transportation environmental sustainability rating program that is called GreenLITES. The program includes protecting the environment, conserving energy and natural resources, and encouraging innovative approaches for sustainable design [4].

This study focuses on bus type selection for Istanbul city taking into account not only operational, and economic topics, but also business and strategy issues by customer views. Although there are lots of studies about vehicle type selection problems in the literature [5–8] there is a lack of hybrid method based on SERVQUAL approach for bus type selection cases. This study proposed a hybrid fuzzy AHP method using Spherical Fuzzy Set to fill the gap with a holistic perspective based on both customer expectations and perceptions.

The paper is organized into five main sections that start with the introduction. The second section presents literature studies about vehicle selection problems. And the methodology is proposed including criteria definition, model design, and proposed method in Sect. 3. In the fourth part, case study and findings are presented. Finally, the conclusion and discussion are given in five parts of the study.

2 Literature Review

There are different studies on vehicle type selection problems in the literature that are presented as follows. Deveci and Torkayesh [5] presented a study about charging strategy for electric buses in public transportation. They applied a decision support model using an Interval-Valued Neutrosophic Set to select the best alternatives among opportunity, depot, inductive, and no-shift. The model was applied to a real case in Turkey. According to the study results, depot charging is the most suitable charging strategy in the selected case. Tang et al. [6] released a study about optimal vehicle types for electric buses regarding shifting departure times. A genetic algorithm has been developed to solve an NP-hard problem. To test the proposed model, a scenario was used in the city of Dandong, China. Regarding results, combining multiple vehicle types for a bus route decreased costs from approximately 11.2% to 14.8% which were for small and large type buses, respectively. The electric bus selection problem was proposed as a case study for Nanjing Bus Company by Gong et al. [7]. The authors presented a test method on the road to select the best bus in Nanjing city of China. Over an 8-week period, 32 test vehicles were tested by 65 drivers. This study will provide decision-makers with a good framework for the selection of pure electric buses. James et al. [8] handled a bus chassis problem for large fleet operators in India using AHP and TOPSIS methods. The criteria were developed by interviews with fleet operators and literature studies. The hybrid methodology was applied to the problem to find the best bus chassis manufacturers for fleet operators. The study will help bus fleet owners to satisfy their requirements.

Though there are many studies about vehicle type selection problems, there is a deficiency in taking into account a holistic perspective that is based on expectation and perception views using the SERVQUAL approach. This study provides to find out improvement areas for the company side. Thus, the study will be a guide to increasing customer satisfaction by discovering the gap between expectations and perceptions.

3 Methodology

3.1 Criteria Definition

Parasuman et al. proposed the SERVQUAL approach with five main criteria which are tangibles, responsiveness, assurance, reliability, and empathy [9]. We presented a novel approach matching SERVQUAL criteria with the bus type selection problem that is shown in Fig. 1. When economic factors are called tangibles, operational issues such as network organization, daily operational constraints, risk, and safety are defined under responsiveness, assurance, and reliability, respectively. Lastly, the empathy criterion matches business and strategy taking into account business continuity, energy and transition strategy, and environmental balance. The criteria definition is presented below.

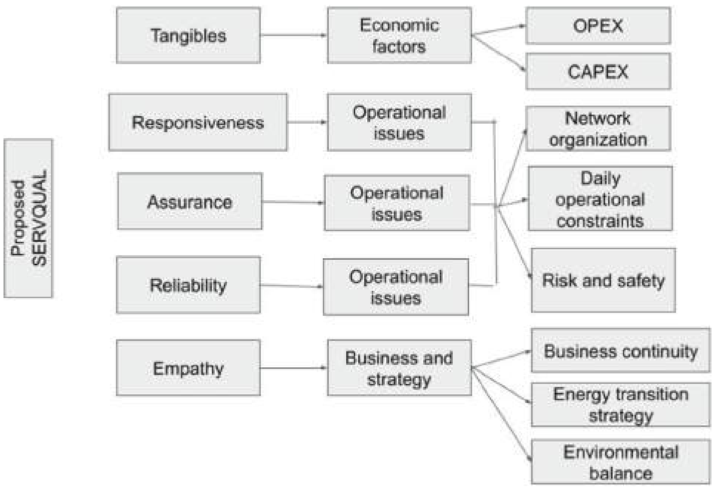


Fig. 1. Developed SERVQUAL approach.

Economic Factors: It includes Operating Expense (Opex) and Capital Expenditure (Capex) which are short-term and long-term perspectives, respectively. Opex: Regarding literature studies, it composes fuel, maintenance, and labor costs for drivers and maintenance staff [10]. Capex: Studies showed that Capex includes bus purchases and infrastructure costs. Infrastructure costs contain pipelines for gases and stations [10]. Business and Strategy: It includes prioritizing public transportation and presenting guidelines for sustainable urban renewal and construction and climate planning [11]. Business Continuity: According to the literature studies, risks are defined such as financial loss, reputational damage, legal action, configuration resilience, obligation, and embeddedness [12] to provide business continuity. Energy Transition Strategy: To provide renewable energy sources, the incumbent energy regime and structural innovation of the energy system should be taken into consideration for energy transformation [13]. Environmental Balance: When fuel consumption is taken into account from an operational perspective, presenting reduction of greenhouse emissions rules and procedures by governments is seen as a policy perspective [14]. Operational Issues: Network systems, operational constraints, risk, and safety issues are seen as operational topics for public transportation [15]. Network organization: Network structure [16], demand [14], depot location [11], risk, and safety [12] are seen as critical issues for network organization. Daily operational constraints: Operational constraints such as capacity regarding passenger demand, depot location, fluctuations in the availability of labor, and emission could be affected offered service [17]. Risks and safety: According to the literature studies death and accident rate are seen as performance indicators [18] for safe transportation.

3.2 Model Design

The bus type selection problem is modeled with two layers with three main and eight subcriteria which is presented in Fig. 2. The alternatives of the proposed model are defined as diesel bus, e-bus overnight charging, e-bus opportunity charging, and fuel cell bus. The explanation of alternatives is as follows.

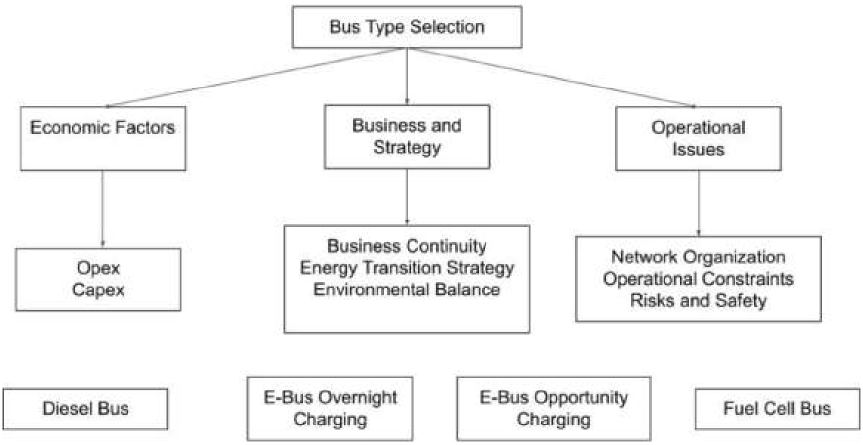


Fig. 2. Proposed model

Diesel bus: Euro I, II, III, IV, V, and VI motors are accepted diesel buses. According to the studies, there is still about 40% of diesel buses in public transportation [19]. Electric bus: It is called an e-bus with different types such as overnight charging and opportunity charging. The key difference between them is battery size with charging time [20]. E-bus overnight charging: It is equipped with plug-in technology. It usually charges in the depot at night. It has a bigger battery size capacity than opportunity charging [20]. E-bus opportunity charging: It is equipped with pantograph-up technology. Not only it provides fast charging but also presents more passenger capacity than overnight charging [20]. Opportunity charging is performed with two different types which are induction and pantograph. When induction technology is equipped under the bus using a wireless system, pantograph technology is located upside of the buses. Fuel cell bus: It is called also the hydrogen bus. Studies show that fuel cell buses reduce nearly 90% of CO2 emissions in comparison to diesel buses [21].

3.3 Proposed Methodology

The bus selection problem is solved with the fuzzy AHP using Spherical Fuzzy Set which was developed by [22]. AHP is mostly preferred method due to ease of use and wider application in the literature [23]. To handle the complexity and vagueness of the problem, Spherical Fuzzy Set is conducted to problem taking into account the SERVQUAL approach.

The methodological framework is presented below in Fig. 3. Firstly, constitute a hierarchical structure with criteria and alternatives. Then, comparison matrices are sent to customers with an explanation of the developed model. Customers assess matrices from both expectation and perception perspectives with a holistic view. When comparison matrices are collected, the consistency ratio is checked. If the consistency ratio is less than 0.1, go further to calculate the local weights of the criteria. Otherwise, researchers go back to step 3 and constitute the pairwise matrices again. To defuzzy local weights, a related score function is used and then the global weights of the criteria are obtained regarding the hierarchical structure of the model. Then, criteria are ordered regarding importance levels and select the best alternative taking into account higher global weights for both views. Lastly, a gap analysis is conducted to discover similarities and differences between customer expectations and perceptions.

4 Application and Findings

To evaluate the proposed model, a case study is conducted with a real case. IETT is a public bus transportation authority in Istanbul, it manages more than sixty thousand buses with carrying more than daily four million passengers [24]. Before green transportation investment, decision-makers want to know customer expectations. At this point, a methodological framework is developed regarding the SERVQUAL approach.

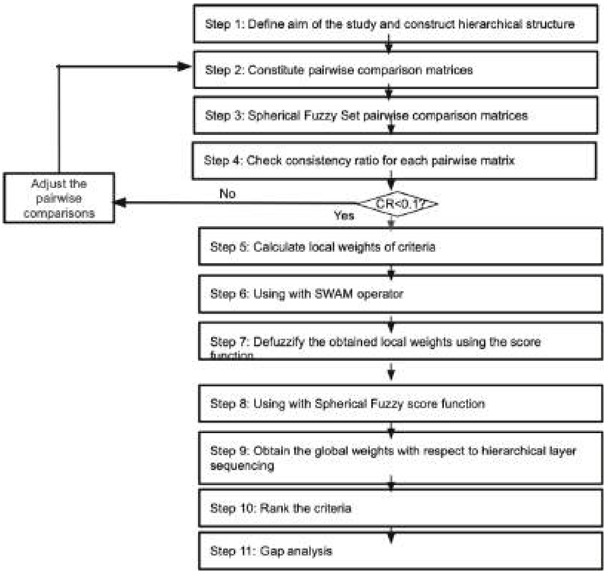


Fig. 3. Methodological framework.

Table 1. Findings for criteria with gap score.

Criteria	Perceived weights	Expected weights	Gap score
Economic factors	0.33	0.41	−0.08
Business and strategy	0.34	0.24	0.10
Operational issues	0.34	0.35	−0.01
Opex	0.14	0.21	−0.07
Capex	0.18	0.20	−0.02
Business continuity	0.14	0.11	0.03
Energy transition strategy	0.10	0.06	0.04
Environmental balance	0.10	0.07	0.03
Network organization	0.08	0.12	−0.04
Daily operational constraints	0.14	0.12	0.02
Risks and safety	0.12	0.11	0.01

Regarding the results are presented for criteria and alternatives in Tables 1 and 2, respectively. According to Table 1, operational, business, and strategy issues are selected as the most important criteria from the customer-perceived view. When we look at the customer-expected view, economic factors come first and operational issues come second in the first level of the model. When capital expenses come first with 18% from the perceived view, operational expenses come first with 21% weight from the expected view. To measure the gap between

Table 2. Findings for alternatives with gap score.

Alternatives	Perceived weights	Expected weights	Gap score
A1	20.85	15.71	5.14
A2	17.44	16.53	0.91
A3	17.95	19.15	−1.21
A4	16.77	16.73	0.04

perception and expectation, a gap analysis is performed [25]. From the gap analysis, there is a −0.08 gap score between customer perception and expectation for economic factors. In addition, there is a 0.01 gap score for operational issues in the first level of the model.

Regarding Table 2, the diesel bus comes first in the perceived view with 21% weight. From the expectation view, e-bus opportunity charging comes first with 19% and there is a 1.21 gap score to meet customer satisfaction.

5 Conclusion and Discussion

A bus selection problem is handled with a hybrid fuzzy approach using Spherical Fuzzy Set. The model is constructed with two layers regarding literature studies. Also, the SERVQUAL approach is conducted to discover the gap between customer expectation and expectation. To validate the proposed model, a real case study is applied in Istanbul. Results show that there are different priorities between customer expectations and perceptions. While economic factors come first for expectations, operational, business, and strategy come first from perception views. Also, there are differences for alternatives, too. From the expectation view, the diesel bus is preferred, from the perception view, the e-bus option comes first with the opportunity charging option. When we look at the gap scores of the criteria, the biggest score is −0.08 between economic factors perceived and expected. And the lowest score is related to operational issues, risks, and safety. Findings show that the biggest gap belongs to the diesel bus in alternatives. To decrease the gap between customer expectations and perceptions customer needs will be analyzed in detail.

The SERVQUAL approach is used effectively to discover improvement areas for increasing customer satisfaction [26,27]. At this point, IETT could use results when new bus purchasing strategies are developed regarding customer expectations. This paper not only focuses on economic and operational issues but also takes into account business and strategy. For this reason, the paper will be a guide for bus purchase decisions for politicians and decision-makers. And also, the paper contributes to the literature by presenting a hybrid fuzzy method based on the SERVQUAL approach for vehicle selection problems.

Different multi-criteria decision making methods could be applied to the proposed problem to get comparative results in the future. In this study, customer expectations and perceptions are analyzed, operator side will be taken into account for the next study to compare similarities and differences.

References

1. United Nations. https://www.un.org/en/about-us/un-and-sustainability#:~:text=The%202020%20Report%20found%20the,came%20from%20staff%20air%20travel. Accessed 2 Apr 2023
2. U.S. Department of Transportation. https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/PublicTransportationsRoleInRespondingToClimateChange2010.pdf. Accessed 2 Apr 2023
3. RATP. https://www.ratpdevusa.com/commitments/sustainable-mobility. Accessed 2 Apr 2023
4. The New York State Department of Transportation (NYSDOT). https://www.dot.ny.gov/programs/greenlites. Accessed 2 Apr 2023
5. Deveci, M., Torkayesh, A.E.: Charging type selection for electric buses using interval-valued neutrosophic decision support model. IEEE Trans. Eng. Manage. 1–14 (2016)
6. Tang, C., Ge, Y.E., Xue, H., Ceder, A., Wang, X.: Optimal selection of vehicle types for an electric bus route with shifting departure times. Int. J. Sustain. Transp. 1–19 (2022)
7. Gong, J., et al.: Road test-based electric bus selection: a case study of the Nanjing bus company. Energies **13**(5), 1253 (2020)
8. James, A.T., Vaidya, D., Sodawala, M., Verma, S.: Selection of bus chassis for large fleet operators in India: an AHP-TOPSIS approach. Expert Syst. Appl. **186**, 115760 (2021)
9. Parasuraman, A., Zeithaml, V.A., Berry, L.L.: A conceptual model of service quality and its implications for future research. J. Mark. **49**(4), 51–50 (1985)
10. Igbojionu, A., Anyadiegwu, C., Anyanwu, E., Obah, B., Muonagor, C.: Technical and economic evaluation of the use of CNG as potential public transport fuel in Nigeria. Sci. Afr. **6**, e00212 (2019)
11. Yim, K.H., Ha, M.C., Jo, C.J., Han, K.M., Baek, J.I., Ban, Y.U.: Strategic planning for the smart-green city through urban governance. Int. J. Built Environ. Sustain. **2**(3), 192–201 (2015)
12. Herbane, B., Elliott, D., Swartz, E.M.: Business continuity management: time for a strategic role? Long Range Plan. **37**(5), 435–457 (2004)
13. Kern, F., Smith, A.: Restructuring energy systems for sustainability? Energy transition policy in the Netherlands. Energy Policy **36**(11), 4093–4103 (2008)
14. Zis, T., North, R.J., Angeloudis, P., Ochieng, W.Y., Bell, M.G.: Environmental balance of shipping emissions reduction strategies. Transp. Res. Rec. **2479**(1), 25–33 (2015)
15. Gray, G.E., Hoel, L.A.: Public Transportation: Planning, Operations, and Management, 2nd edn. OSTI.GOV, Oak Ridge (1979)
16. Soh, H., et al.: Weighted complex network analysis of travel routes on the Singapore public transportation system. Physica A Stat. Mech. Appl. **389**(24), 5852–5863 (2010)

17. Tirachini, A., Hensher, D.A., Rose, J.M.: Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. Transp. Res. Part A Policy Pract. **53**, 36–52 (2013)
18. Porcu, F., Olivo, A., Maternini, G., Barabino, B.: Evaluating bus accident risks in public transport. Transp. Res. Procedia **45**, 443–450 (2020)
19. Cortes, M.C., Acevedo, H.R., Rojas, N.Y.: Diesel emissions from Euro II/III buses of Bogota's urban transport system: the challenge is not reducing mass, but counting particles. WIT Trans. Ecol. Environ. **207**, 157–168 (2016)
20. Al-Saadi, M., et al.: Slow and fast charging solutions for Li-Ion batteries of electric heavy-duty vehicles with fleet management strategies. Sustainability **13**(19), 10639 (2021)
21. Ajanovic, A., Glatt, A., Haas, R.: Prospects and impediments for hydrogen fuel cell buses. Energy **235**, 121340 (2021)
22. Kutlu Gündoğdu, F., Kahraman, C.: Spherical fuzzy sets and spherical fuzzy TOPSIS method. J. Intell. Fuzzy Syst. **36**(1), 337–352 (2019)
23. Buran, B., Erçek, M.: Public transportation business model evaluation with Spherical and Intuitionistic Fuzzy AHP and sensitivity analysis. Expert Syst. Appl. **204**, 117519 (2022)
24. IETT. https://iett.istanbul/icerik/istanbulda-toplu-ulasim. Accessed 2 Apr 2023
25. Shekarchizadeh, A., Rasli, A., Hon-Tat, H.: SERVQUAL in Malaysian universities: perspectives of international students. Bus. Process. Manag. J. **17**(1), 67–81 (2011)
26. Tumsekcali, E., Ayyildiz, E., Taskin, A.: Interval valued intuitionistic fuzzy AHP-WASPAS based public transportation service quality evaluation by a new extension of SERVQUAL Model: P-SERVQUAL 4.0. Expert Syst. Appl. **186**, 115757 (2021)
27. Abari, A.A.F., Yarmohammadian, M.H., Esteki, M.: Assessment of quality of education a non-governmental university via SERVQUAL model. Procedia Soc. Behav. Sci. **15**, 2299–2304 (2011)

Check for updates

Fuzzy Logic Control for Color-Tunable Lighting Systems

Rahib Imamguluyev(✉)

IT and Engineering Department, Odlar Yurdu University, Baku AZ1072, Azerbaijan
rahib.aydinoglu@gmail.com

Abstract. The article provides an overview of the application of fuzzy logic control for color-tunable lighting systems. It begins by introducing color-tunable lighting systems and the need for effective control mechanisms, and then explores how fuzzy logic can be used to address these challenges. The basics of fuzzy logic, including fuzzy sets, membership functions, fuzzy rules, and inference, are explained. The article then delves into the design of color-tunable lighting systems, including color mixing techniques, color temperature, CRI, LED driver, and control circuitry, and feedback mechanisms. The implementation of fuzzy logic control is discussed, including system modeling, rule base development, membership function selection, fuzzy inference engine, and defuzzification techniques. The article also covers performance evaluation, including simulation setup and parameters, and evaluation metrics. Finally, the article explores real-world implementations and benefits, as well as potential future directions for research and development.

Keywords: Fuzzy logic · color-tunable · lighting · CRI · LED

1 Introduction

Color-tunable lighting systems are lighting solutions that offer users the ability to adjust the color temperature and spectral content of the light. These systems typically consist of light sources (such as LEDs) that can emit light at different wavelengths, and control circuitry that enables users to adjust the light output. By changing the color temperature and spectral content of the light, color-tunable lighting systems can create different moods and atmospheres, enhance visual comfort, and provide health benefits. Applications for these systems include residential and commercial lighting, horticulture, and medical and therapeutic lighting [1].

Effective control mechanisms are essential for color-tunable lighting systems to achieve their full potential. Without proper control, these systems may not provide the desired lighting quality, may not be energy-efficient, and may even cause discomfort or health issues for users [2]. Additionally, as the complexity of these systems increases with the number of light sources and control options, traditional control approaches such as PID control may become inadequate. Therefore, advanced control techniques that can handle complex and uncertain information are needed. Fuzzy logic control is one such technique that has been successfully applied to the control of color-tunable lighting systems.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 744–750, 2023.
https://doi.org/10.1007/978-3-031-39777-6_87

1.1 Fuzzy Logic as a Potential Solution

Fuzzy logic as a potential solution.

Fuzzy logic is a control technique that can handle uncertain and imprecise information. Unlike traditional control techniques, which rely on precise numerical values, fuzzy logic can operate on linguistic variables such as "bright" or "warm". This makes it well-suited for applications such as color-tunable lighting systems, where the desired lighting quality may not be easily quantifiable or may vary based on user preferences.

Fuzzy logic control systems typically consist of three main components: a fuzzifier, an inference engine, and a defuzzifier. The fuzzifier converts crisp input values (such as sensor readings or user input) into fuzzy sets, which represent the degree of membership of the input to each linguistic variable. The inference engine applies a set of fuzzy rules to the fuzzy sets to determine the appropriate output values. Finally, the defuzzifier converts the fuzzy output sets into crisp output values (such as control signals for the lighting system) [3].

Fuzzy logic control has been successfully applied to a wide range of control problems, including those in the field of lighting. In the context of color-tunable lighting systems, fuzzy logic control can provide more flexible and intuitive control over the color temperature and spectral content of the light, while also optimizing energy consumption and minimizing user discomfort [4].

2 Fuzzy Logic Control Basics

Fuzzy logic is a control technique that can handle uncertain and imprecise information. Unlike traditional control techniques, which rely on precise numerical values, fuzzy logic can operate on linguistic variables such as "bright" or "warm". This makes it well-suited for applications such as color-tunable lighting systems, where the desired lighting quality may not be easily quantifiable or may vary based on user preferences.

The following are the basic components of a fuzzy logic control system [5]:

- Fuzzy sets and membership functions: Fuzzy sets are sets of values that represent linguistic variables such as "bright" or "warm". These sets are defined by membership functions, which specify the degree of membership of a given input value to each fuzzy set.
- Fuzzy rules: Fuzzy rules are conditional statements that relate input fuzzy sets to output fuzzy sets. For example, a fuzzy rule might state that "if the brightness is high and the color temperature is cool, then the light should be dimmed".
- Fuzzy inference engine: The fuzzy inference engine applies the fuzzy rules to the input fuzzy sets to determine the appropriate output fuzzy sets. This involves combining the input fuzzy sets using fuzzy operators such as AND, OR, and NOT, and then applying the fuzzy rules to the resulting fuzzy sets to determine the output fuzzy sets.
- Defuzzification: The defuzzification process converts the output fuzzy sets into crisp output values. This involves selecting a representative value (such as the centroid or maximum value) from the output fuzzy sets [6].

Fuzzy logic control systems can be designed using various tools and techniques, including MATLAB/Simulink, Fuzzy Control Language (FCL), and rule-based systems.

The effectiveness of a fuzzy logic control system depends on the quality of the fuzzy sets and membership functions, the design of the fuzzy rules, and the selection of appropriate defuzzification methods.

3 Color-Tunable Lighting System Design

The design of a color-tunable lighting system typically involves several stages, including the selection of light sources, the design of the control circuitry, and the implementation of the control algorithms. Fuzzy logic control can be used to optimize the performance of the lighting system and provide a smooth and natural color transition.

- Selection of Light Sources: The first step in designing a color-tunable lighting system is to select the light sources. Typically, color-tunable lighting systems use LEDs, which can produce a wide range of colors and are highly efficient. The LEDs used in the system should have a high color rendering index (CRI) and a wide color gamut to ensure that the colors produced are accurate and vivid.
- Design of Control Circuitry: The control circuitry for a color-tunable lighting system typically consists of a microcontroller or a digital signal processor (DSP) that controls the LEDs. The control circuitry should be designed to provide precise and stable control of the LEDs, and it should be able to communicate with external sensors or user interfaces.
- Implementation of Control Algorithms: The control algorithms for a color-tunable lighting system typically involve the use of feedback control to maintain the desired color output. Fuzzy logic control can be used to optimize the performance of the system and provide smooth and natural color transitions. The fuzzy logic control algorithm can be implemented using a software or hardware solution [7].
- Calibration and Testing: Once the system is designed and implemented, it should be calibrated and tested to ensure that it meets the desired performance criteria. The calibration process involves measuring the color output of the system and adjusting the control parameters to achieve the desired color output.

Overall, the design of a color-tunable lighting system involves careful selection of light sources, precise control circuitry design, and the implementation of an effective control algorithm. Fuzzy logic control can be used to optimize the performance of the system and provide a smooth and natural color transition.

3.1 Color Temperature and CRI

Color temperature and CRI (color rendering index) are important factors to consider in the design of color-tunable lighting systems.

- Color Temperature: Color temperature is a measure of the color appearance of a light source, measured in degrees Kelvin (K). It describes whether a light source appears warm (reddish/yellowish) or cool (bluish). Color temperature is an important consideration in lighting design, as it can affect the mood and ambiance of a space. For example, warm white light (2700–3000 K) is often used in residential spaces to create a cozy and relaxing atmosphere, while cool white light (5000–6500 K) is often used in commercial spaces to create a bright and energizing environment.

- Color Rendering Index: CRI is a measure of a light source's ability to accurately reproduce colors compared to a reference light source. The CRI is a scale from 0 to 100, with 100 being the best possible score. A high CRI indicates that colors appear more natural and vibrant under the light source, while a low CRI can cause colors to appear dull or washed out. In color-tunable lighting systems, a high CRI is desirable to ensure that colors appear as intended, especially in environments where color accuracy is important, such as in art galleries or retail spaces [8].

In summary, color temperature and CRI are important factors to consider in the design of color-tunable lighting systems. The selection of color temperature and CRI should be based on the specific requirements of the lighting application, taking into account factors such as the mood, ambiance, and color accuracy desired.

4 Fuzzy Logic Control Implementation

Implementing a fuzzy logic control system for a color-tunable lighting system involves several steps:

- Define the inputs and outputs: The first step in implementing a fuzzy logic control system is to define the inputs and outputs. In a color-tunable lighting system, the inputs could be the desired color temperature, color rendering index (CRI), and brightness level, while the output would be the control signal for adjusting the intensity of the red, green, and blue LEDs.
- Develop the membership functions: The next step is to develop the membership functions for the inputs and outputs. Membership functions define the degree to which a particular input or output belongs to a fuzzy set. For example, the membership function for the color temperature input might define the degree to which the input belongs to the "cool," "neutral," or "warm" fuzzy set [9].
- Create the fuzzy rules: Fuzzy rules define the relationship between the inputs and outputs. They specify how the input values are combined to determine the output value. For example, a fuzzy rule might state that if the color temperature is "cool" and the CRI is "high," then the intensity of the blue LED should be increased.
- Implement the inference engine: The inference engine is the component of the fuzzy logic control system that evaluates the fuzzy rules to determine the appropriate output value. The inference engine uses the input values and the fuzzy rules to determine the output value based on the degree to which each input value belongs to its fuzzy set [9].
- Test and tune the system: Once the fuzzy logic control system is implemented, it must be tested and tuned to ensure that it is providing the desired output values. This may involve adjusting the membership functions, fuzzy rules, or other parameters to improve the accuracy and responsiveness of the control system [10].

Overall, implementing a fuzzy logic control system for a color-tunable lighting system can provide a flexible and effective way to control the color and brightness of the LEDs. The specific implementation will depend on the requirements of the lighting application and the complexity of the control system.

4.1 Membership Function Selection

Membership function selection is an important step in developing a fuzzy logic control system for color-tunable lighting systems. Membership functions define the degree to which a particular input or output belongs to a fuzzy set, and the selection of appropriate membership functions is critical to the accuracy and responsiveness of the control system [9].

There are several methods for selecting membership functions:

- Expert knowledge: Experts in the field can use their knowledge and experience to define the membership functions for each variable. This approach can be time-consuming and may require extensive experimentation, but it can result in highly accurate and customized control systems [11].
- Data-driven: Membership functions can be selected based on statistical analysis of data. This approach involves collecting data on the system and using clustering or other data analysis techniques to identify the underlying fuzzy sets. This method can be faster than expert knowledge-based approaches, but may not always be applicable when data is not available.
- Trial and error: Membership functions can also be selected through trial and error. This involves selecting a set of initial membership functions and evaluating the performance of the control system. The membership functions can be adjusted based on the performance until the desired output values are achieved [12].
- Regardless of the method used, membership functions should be chosen to reflect the variability and range of the input variables. The shapes of the membership functions should also be selected to capture the characteristics of the variables being modeled, such as symmetry or skewness. Additionally, the number of fuzzy sets should be selected based on the complexity of the system and the accuracy required.

Overall, the selection of appropriate membership functions is a critical step in developing a fuzzy logic control system for color-tunable lighting systems. By carefully selecting membership functions that reflect the characteristics of the input and output variables, it is possible to design a control system that accurately and reliably adjusts the color and brightness of the LEDs.

5 Performance Evaluation

Performance evaluation is an important step in any control system design, including fuzzy logic control for color-tunable lighting systems. There are several metrics that can be used to evaluate the performance of a fuzzy logic control system, including [13, 14]:

- Rise time: The time it takes for the system output to reach a specified percentage of its final value after a step change in the input.
- Settling time: The time it takes for the system output to settle within a specified percentage of its final value after a step change in the input.
- Overshoot: The percentage by which the system output exceeds its final value before settling.

- Steady-state error: The difference between the desired output value and the actual output value when the system has reached a steady-state.
- Tracking performance: The ability of the system to track a desired reference signal.

These metrics can be measured experimentally or using simulation software. In addition to these metrics, other factors such as robustness, stability, and complexity should also be considered when evaluating the performance of a fuzzy logic control system.

In general, the performance of a fuzzy logic control system for color-tunable lighting systems can be evaluated by comparing the actual output of the system to the desired output. By analyzing the performance of the system and making appropriate adjustments, it is possible to optimize the control system and ensure that it meets the desired performance requirements.

6 Conclusion

In conclusion, fuzzy logic control has emerged as a powerful method for controlling color-tunable lighting systems. By leveraging the flexibility and adaptability of fuzzy logic, it is possible to adjust the color and brightness of lighting systems to match changing environmental conditions, user preferences, and task requirements. This can lead to a range of benefits, including energy efficiency, improved visual comfort, increased productivity, enhanced safety and security, and customization and personalization.

Looking ahead, there are several potential avenues for future research and development in the field of fuzzy logic control for color-tunable lighting systems. These include:

- Development of new membership functions: Further research could be done to explore new membership functions and optimize their performance for specific applications.
- Integration with machine learning: Combining fuzzy logic with machine learning techniques could provide even more sophisticated and accurate control mechanisms for color-tunable lighting systems.
- Exploration of new application areas: There is a growing interest in using color-tunable lighting systems in new application areas, such as healthcare, education, and entertainment. Future research could explore the potential benefits and challenges of using fuzzy logic control in these contexts.
- Integration with IoT and other smart technologies: The integration of fuzzy logic control with other smart technologies, such as the Internet of Things (IoT), could enable even more seamless and effective control mechanisms for color-tunable lighting systems.

Overall, the potential benefits of fuzzy logic control for color-tunable lighting systems are significant, and further research and development in this area could help unlock even more possibilities for improving energy efficiency, visual comfort, productivity, safety, and personalization in a range of applications.

References

1. Ruggiero, S., De Masi, R.F., Assimakopoulos, M.N., Vanoli, G.P.: Energy saving through building automation systems: experimental and numerical study of a smart glass with liquid crystal and its control logics in summertime. Energy Build. **273**, 112403 (2022). ISSN 0378-7788. https://doi.org/10.1016/j.enbuild.2022.112403
2. Wagiman, K.R., et al.: Lighting system control techniques in commercial buildings: current trends and future directions. J. Build. Eng. **31**, 101342 (2020). ISSN 2352-7102. https://doi.org/10.1016/j.jobe.2020.101342
3. Pekur, D., Sorokin, V.M., Nikolaenko, Y.: Determination of optical parameters in quasi-monochromatic LEDs for implementation of lighting systems with tunable correlated color temperature. Semicond. Phys. Quant. Electron. Optoelectron. **25**(3), 303–314 (2022). https://doi.org/10.15407/spqeo25.03.303
4. Imam, M.H.T., Afshari, S., Mishra, S.: An experimental survey of feedback control methodologies for advanced lighting systems. Energy Build. **130**, 600–612 (2016). ISSN 0378-7788. https://doi.org/10.1016/j.enbuild.2016.08.088
5. Kahraman, C., Cebi, S., Onar, S.Ç., Öztayşi, B.: Recent developments on fuzzy AHP&ANP under vague and imprecise data: evidence from INFUS conferences. Int. J. Anal. Hierar. Process **14**(2) (2022). https://doi.org/10.13033/ijahp.v14i2.1033
6. Imamguluyev, R., Aliyeva, A.: Analysis of intelligent interfaces based on fuzzy logic in human-computer interaction. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M.B., Sadikoglu, F. (eds.) 15th International Conference on Applications of Fuzzy Systems, Soft Computing and Artificial Intelligence Tools – ICAFS-2022, pp. 720–726. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-25252-5_94
7. Akram, M., Zahid, K., Kahraman, C.: Integrated outranking techniques based on spherical fuzzy information for the digitalization of transportation system. Appl. Soft Comput. **134**, 109992 (2023). ISSN 1568-4946. https://doi.org/10.1016/j.asoc.2023.109992
8. Imamguluyev, R.: Application of fuzzy logic model for correct lighting in computer aided interior design areas. In: Kahraman, C., Cevik Onar, S., Oztaysi, B., Sari, I.U., Cebi, S., Tolga, A.C. (eds.) INFUS 2020. AISC, vol. 1197, pp. 1644–1651. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-51156-2_192
9. Zadeh, L.A., Aliev, R.: Fuzzy logic theory and applications: part I and part II, p. 61 (2018). https://doi.org/10.1142/10936
10. Alekperov, R.B.: Representation of fuzzy queries in the form of a multi-criteria decision matrix using Z-numbers (2021). https://doi.org/10.13140/RG.2.2.28243.91683
11. Abdullayev, T., Imamguluyev, R., Umarova, N.: Application of fuzzy logic model for optimal solution of light reflection value in lighting calculations. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F.M. (eds.) ICSCCW 2021. LNNS, vol. 362, pp. 384–391. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-92127-9_53
12. Valiyev, A., Imamguluyev, R., Ilkin, G.: Application of fuzzy logic model for daylight evaluation in computer aided interior design areas. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F.M. (eds.) ICAFS 2020. AISC, vol. 1306, pp. 715–722. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-64058-3_89
13. Imamguluyev, R., Mikayilova, R., Salahli, V.: Application of a fuzzy logic model for optimal assessment of the maintenance factor affecting lighting in interior design. In: Shakya, S., Ntalianis, K., Kamel, K.A. (eds.) Mobile Computing and Sustainable Informatics: Proceedings of ICMCSI 2022, pp. 463–474. Springer Nature Singapore, Singapore (2022). https://doi.org/10.1007/978-981-19-2069-1_32
14. Akram, M., Zahid, K., Kahraman, C.: New optimization technique for group decision analysis with complex Pythagorean fuzzy sets. J. Intell. Fuzzy Syst. **44**(5), 1–25 (2022). https://doi.org/10.3233/JIFS-220764

Performance Evaluation of Higher Education Institutions Employing Analytical Techniques

Nazli Goker(✉) and Mehtap Dursun

Industrial Engineering Department, Galatasaray University, Ortakoy, Istanbul, Turkey
nagoker@gsu.edu.tr

Abstract. Theoretical enquiries and empirical studies show the significance of performance evaluation in the higher education. Traditionally, performance measurement in universities has an advancing role, helping them to advance their upcoming performance. However, in recent times, higher education has been forced into a highly competitive business environment due to factors such as globalization, the fourth industrial revolution, the high demand for higher education, and increasing competition. The performance of a university has a significant impact on its social influence. It is affected by many factors such as teaching, research and citations, thus it is difficult to optimize all factors because of resource restriction. In this research, data envelopment analysis (DEA) method is used to measure the performance of the best performing 20 universities as of 2023 throughout the world. The data are collected from Times Higher Education website. The universities are evaluated according to five outputs namely "teaching", "research", "citations", "industry income", and "international outlook" and the results are analyzed.

Keywords: Data envelopment analysis · Education · Performance management · University rankings

1 Introduction

Performance management involves many aspects as effectiveness and efficiency measurement, managing stakeholder, organizational culture and motivation management, utilization of financial and non-financial information and making decisions in organizational activities [1].

Recently, with the growing emphasis placed on higher education, it is important to ameliorate university performance management. University performance has crucial effect on their social influence. It is affected by many factors such as teaching, research and citations, thus it is difficult to optimize all factors because of resource restriction.

In many countries, government funding for universities is tied to their performance metrics, including research outputs and negotiated performance targets. This means that universities must prioritize performance measurement to ensure that they meet these targets and secure funding [2].

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 751–757, 2023.
https://doi.org/10.1007/978-3-031-39777-6_88

Moreover, in today's globalized higher education market, universities must compete for both students and staff. Prospective students and faculty members are likely to pay close attention to how different institutions perform, including their rankings, research output, and overall reputation. Universities that perform well in these areas are more likely to attract top talent and secure partnerships with other institutions.

This study aims to measure the performance of universities by employing data envelopment analysis (DEA). The data are collected from Times Higher Education website, which creates university rankings to assess university performance on the global stage. These rankings are based on various performance metrics, such as teaching, research, citations, industry income, and international outlook.

The rest of the study is organized as follows. DEA is explained in Sect. 2. Section 3 illustrates the case study. Conclusions are provided in the last Section.

2 Data Envelopment Analysis

The original DEA model, named as the CCR (Charnes, Cooper, & Rhodes) model, is proposed by Charnes et al. [3]. The model is used to compute the relative efficiency of a decision-making unit (DMU) via maximizing the ratio of its total weighted outputs to its total weighted inputs, under the constraint that the output to input ratio of every DMU be less than or equal to unity.

The DEA model is widely used in performance measurement, particularly in the context of evaluating the efficiency of organizations, which produce multiple outputs from multiple inputs. It is a non-parametric method that does not need any assumptions about the functional form of the production or cost function that makes it a flexible tool for performance measurement.

The conventional DEA formulation can be represented as follows:

$$\begin{aligned}
&\max\ E_{j_0} = \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}} \\
&\text{subject to} \\
&\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \qquad \forall j, \\
&u_r, v_i \ge \varepsilon, \qquad \forall r, i,
\end{aligned} \tag{1}$$

where E_{j_0} is the efficiency score of the evaluated DMU, u_r is the weight assigned to output r, v_i is the weight assigned to input i, y_{rj} is the quantity of output r generated and x_{ij} is the amount of input i consumed by DMU j, respectively, and ε is a small positive scalar.

Formulation (1) possesses non-linear and non-convex properties, however, it can be converted into a linear programming model via a transformation [4]. The linear programming model for computing the relative efficiency of a DMU is as

$$
\begin{aligned}
& \max E_{j0} = \sum_{r=1}^{s} u_r\, y_{rj_0} \\
& \text{subject to} \\
& \sum_{i=1}^{m} v_i x_{ij_0} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \qquad \forall j, \\
& u_r, v_i \ge \varepsilon, \qquad \forall r, i.
\end{aligned}
\tag{2}
$$

Formulation (2) classifies DMUs as efficient and inefficient. The DMUs with efficiency score of one are considered as efficient while DMUs with lower efficiency scores are named as inefficient. In order to ignore the unrealistic weight dispersion and improve the poor discriminating power of DEA, few studies are proposed [5]. Wong and Beasley [6] employed weight restrictions, which enforce some frontiers on weights. Cross-efficiency analysis is another widely used mathematical technique to enhance the discriminating power of DEA [7].

Toloo [8] presented a mixed integer programming approach that minimizes the maximum deviation from efficiency and applied it to 40 professional tennis players problem.

Initially, Toloo [8] developed the following programming model for determining the maximum value of the non-Archimedean infinitesimal.

$$
\begin{aligned}
& \max \varepsilon \\
& \text{subject to} \\
& d_{\max} - d_j \ge 0,\ \forall j, \\
& \sum_{r=1}^{s} u_r y_{rj} + d_j = 1\ \ \forall j\,, \\
& \sum_{j=1}^{n} \theta_j = n - 1, \\
& d_j \le \theta_j\ \ \forall j, \\
& \theta_j \le M d_j\ \ \forall j, \\
& d_j \ge 0\ \ \forall j, \\
& \theta_j \in \{0, 1\}\ \ \forall j, \\
& u_r \ge \varepsilon,
\end{aligned}
\tag{3}
$$

Using the optimal solution obtained from the Formulation (3), Toloo [8] proposed Formulation (4) for ranking the DMUs and indicate the most efficient one.

$$
\begin{aligned}
&\min d_{max} \\
&\text{subject to} \\
&d_{\max} - d_j \geq 0, \quad \forall j, \\
&\sum_{r=1}^{s} u_r y_{rj} + d_j = 1 \;\; \forall j, \\
&\sum_{j=1}^{n} \theta_j = n - 1, \\
&d_j \leq \theta_j \quad \forall j, \\
&\theta_j \leq M d_j \quad \forall j, \\
&d_j \geq 0 \quad \forall j, \\
&\theta_j \in \{0, 1\} \quad \forall j, \\
&u_r \geq \varepsilon,
\end{aligned}
\tag{4}
$$

where M is a sufficiently large positive number, and θ_j is a binary variable. The model guarantees to identify the single best efficient DMU.

3 Case Study

In this section, performance assessment of the best performing 20 universities as of 2023 throughout the world is provided by employing common-weight DEA model addressed in Toloo [9]. The DMUs are evaluated according to five outputs indicated in Times Higher Education, namely "teaching", "research", "citations", "industry income", and "international outlook". The data of 20 universities are given in Table 1 [9].

Table 1. The Data of 20 Universities

DMUs	Name of the University	Teaching	Research	Citations	Industry income	İnternational outlook
1	University of Oxford	92.3	99.7	99.0	74.9	96.2
2	California Institute of Technology	90.9	97.0	97.3	89.8	83.6
3	Harvard University	94.8	99.0	99.3	49.5	80.5
4	Stanford University	94.2	96.7	99.8	65.0	79.8
5	University of Cambridge	90.9	99.5	97.0	54.2	95.8

(*continued*)

Table 1. (*continued*)

DMUs	Name of the University	Teaching	Research	Citations	Industry income	İnternational outlook
6	Massachusetts Institute of Technology	90.7	93.6	99.8	90.9	89.3
7	Princeton University	87.6	95.9	99.1	66.0	80.3
8	University of California, Berkeley	86.4	95.8	99.0	76.8	78.4
9	Yale University	92.6	92.7	97.0	55.0	70.9
10	The University of Chicago	86.5	88.8	97.7	56.2	74.2
11	Columbia University	89.4	87.7	97.1	44.8	79.9
12	Imperial College London	82.8	90.8	98.3	59.8	97.5
13	Johns Hopkins University	79.4	91.5	97.0	89.5	75.3
14	University of Pennsylvania	86.0	88.8	97.0	75.8	71.5
15	ETH Zurich	82.6	95.4	90.7	59.1	97.7
16	Peking University	92.5	96.7	80.4	91.8	65.0
17	Tsinghua University	90.1	97.4	88.0	100.0	40.3
18	University of Toronto	77.3	93.3	92.8	65.5	89.7
19	UCL	74.5	85.4	97.9	44.5	96.7
20	University of California, Los Angeles	80.4	88.9	95.4	58.8	65.0

First, Formulation (3) is solved to obtain the maximum value of non-Archimedean infinitesimal. By utilizing this value, Formulation (4) is employed and the DMUs are arranged according to their deviation scores in descending order. The deviation score of DMU_1, which is University of Oxford, is equal to zero, hence it is the best performing

DMU, followed by Massachusetts Institute of Technology and California Institute of Technology. Full ranking of 20 universities is given in Table 2.

Table 2. Ranking of 20 Universities

DMUs	Name of the University	$d(j)$	Rank
1	University of Oxford	0	1
2	California Institute of Technology	0,015	3
3	Harvard University	0,089	11
4	Stanford University	0,064	5
5	University of Cambridge	0,05	4
6	Massachusetts Institute of Technology	1,00E-06	2
7	Princeton University	0,078	10
8	University of California, Berkeley	0,064	6
9	Yale University	0,126	15
10	The University of Chicago	0,133	18
11	Columbia University	0,139	19
12	Imperial College London	0,066	7
13	Johns Hopkins University	0,073	8
14	University of Pennsylvania	0,103	14
15	ETH Zurich	0,073	9
16	Peking University	0,092	12
17	Tsinghua University	0,13	17
18	University of Toronto	0,092	13
19	UCL	0,128	16
20	University of California, Los Angeles	0,169	20

4 Conclusions

University performance has crucial effect on their social influence. It is affected by many factors such as teaching, research and citations, thus it is difficult to optimize all factors because of resource restriction. In this research, data envelopment analysis is employed to measure the performance of universities. The data are collected from Times Higher Education website. Future research directions may focus on employing other performance measurement techniques to find the rankings of higher education institutions.

Acknowledgement. This work is supported by Galatasaray University Research Fund Project FOA-2022-1092.

References

1. Almohtaseb, A.A., Almahameed, M.A.Y., Shaheen, H.A.K., Al Khattab, M.H.J.: A roadmap for developing, implementing and evaluating performance management systems in Jordan public universities. J. Appl. Res. High. Educ. **11**(2), 325–339 (2019)
2. McCormack, J., Propper, C., Smith, S.: Herding cats? Management and university performance. Econ. J. **124**(578), 534–564 (2014)
3. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision-making units. Eur. J. Oper. Res. **2**, 429–444 (1978)
4. Charnes, A., Cooper, W.W.: Programming with linear fractional functional. Nav. Res. Logist. **9**, 181–186 (1962)
5. Karsak, E.E., Ahiska, S.S.: Pratical common weight multi-criteria decision-making approach with an improved discriminating power for technology selection. Int. J. Prod. Res. **43**, 1537–1554 (2005)
6. Wong, Y.H.B., Beasley, J.R.: Restricting weight flexibility in data envelopment analysis. J. Oper. Res. Soc. **41**, 829–835 (1990)
7. Sexton, T.R., Silkman, R.H., Hogan, A.J.: Data envelopment analysis: critique and extensions. New Dir. Program Eval. **32**, 73–105 (1986)
8. Toloo, M.: The most efficient unit without explicit inputs: an extended MILP-DEA model. Measurement **46**, 3628–3634 (2013)
9. https://www.timeshighereducation.com/world-university-rankings/2022/world-ranking/detailed

Generalized Net Model of Quality of Service Estimation of Overall Telecommunication System Including a Security Stage

Velin Andonov(✉), Stoyan Poryazov, and Emiliya Saranova

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
Acad. G. Bonchev Street, Block 8, 1113 Sofia, Bulgaria
{velin_andonov,stoyan,e.saranova}@math.bas.bg

Abstract. Generalized Nets (GNs) have been successfully used in the modeling of service systems. In recent years, GN models of various compositions of services with intuitionistic fuzzy estimations of uncertainty have been developed. In the present paper, a GN model of Quality of Service (QoS) estimation of an overall telecommunication system including users and a security stage is described. A special naming system for the places and the transitions of the GN is proposed which makes the derivation of analytical expressions easier. The values of some QoS indicators on service stage level and on overall network and system levels are evaluated through the characteristic functions of some special places of the GN. The proposed GN model can be used to define new QoS indicators in overall telecommunication systems. It can also be used for dynamic QoS management of next generation service networks and for development of a framework to assess the QoS and security features of next generation service networks.

Keywords: Generalized nets · Conceptual modeling · Quality of service · Communication security

1 Introduction

The present paper is a continuation of the work of the authors on the modeling of the Quality of Service (QoS) in telecommunication networks. In a series of papers, a methodology for conceptual modeling of service systems based on the Generalized Nets (GNs, see [6]) has been developed. Initially, GN representations of the elements of service systems theory are proposed in [1]. Using these GN

The work of V. Andonov and E. Saranova is supported by the research project Perspective Methods for Quality Prediction in the Next Generation Smart Informational Service Networks (KP-06-N52/2) financed by the Bulgarian National Science Fund. The work of S. Poryazov is partially supported by the Scientific Infrastructure Project (D01-222/22.10.2021), by the Bulgarian Ministry of Education and Science.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 758–771, 2023.
https://doi.org/10.1007/978-3-031-39777-6_89

representations, GN models of QoS estimation of various compositions of services are constructed in [2,3]. Following this, a GN model of traffic quality evaluation of a service stage and a GN model of overall network efficiency evaluation are described in [4,5].

The main goal of the present paper is to construct a GN model of QoS estimation of overall telecommunication system including users and a security stage. For the construction of the GN model, the conceptual model of Information Security service stage and the conceptual model of the contributions to QoS in an overall telecommunication system described in [11] will be used. The GN approach is universal in the sense that it can be applied to all kinds of telecommunication networks, including next generation smart networks where QoS management faces numerous challenges [7,14,15].

2 Preliminaries

We will make use of some basic notions related to the conceptual and analytical modeling of telecommunication networks. They are described in detail in [10,11]. Here, we shall mention briefly the most important ones.

2.1 Base Virtual Service Devices

In the conceptual models of service systems in general, the notion of a base virtual service device plays a central role. It is also the simplest modeling concept. Let us denote by x an arbitrary such device. It has the following parameters: traffic intensity (Yx, measured in Erlang), intensity of the flow of requests (Fx), probability of directing the flow of requests towards the device (Px), mean service time in the device (Tx) and device capacity (Nx). We will use various types of base virtual devices. They are shown in Fig. 1.

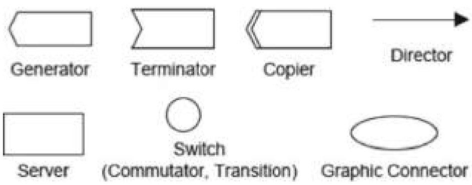


Fig. 1. Base virtual device types.

For detailed description of the function of each type of base virtual device the reader can refer to [10].

2.2 Parameter's Names

For characterization of the traffic the following qualifiers are used: served (srv.), not served (nsr.), offered (ofr.), parasitic (prs.), carried (crr.), demand (dem.), intent (int.), suppressed (sup.), repeated (rep.), additional (add.), security (sec.). The meaning of these qualifiers is explained in detail in [10,11]. For example, $srv.Fx$ is the intensity of the served flow of requests of the device x.

The service stage concept is defined in [12]. Since this notion is important for the present work, we shall give the definition below.

Definition 1. *A service presentation is called a Service stage if it contains:*

– *A service phase realizing one service function.*
– *All supplementary service phases which are not part of the function but support its realization.*

A special naming system is used which enhances the analytical modeling. To illustrate it, let g be a service stage. For example, by $ofr.Fg$ we denote the intensity of the offered flow of requests to the stage.

2.3 Conceptual Model of an Overall Telecommunication System Including a Security Stage

In [11], a conceptual model of an overall telecommunication system including users and a Security stage is described. Its graphical representation is shown in Fig. 2.

The telecommunication network consists of five stages: A-terminal, Dialing (d), Switching (s), B-seizure (z), B-terminal. There are two comprise devices corresponding to the calling (denoted by A) and the called (denoted by B) users. They are denoted by A-User and B-user (see Fig. 2), respectively. The other service stages in the model are: Demanding stage, Offering stage, Finding B-user stage, Information security stage, Target communication stage. For detailed explanation of each of these stages, the reader can refer to [11].

2.4 QoS Indicators on Overall Network and System Level

In [11], on the basis of the conceptual model shown in Fig. 2 many QoS indicators on service stage level, overall network level and overall system level are proposed. Here, we shall present those of them which are used in the GN model described in the next section. The first one is the QoS indicator of the overall network ($Qnet$) which is given by

$$Qnet = \frac{crr.Fis}{ofr.Fa}. \tag{1}$$

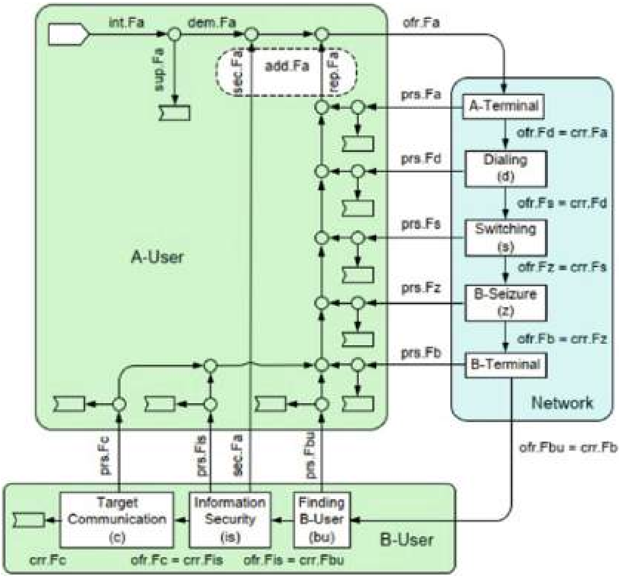


Fig. 2. QoS contributions in an overall telecommunication system including users and a Security stage (see [11]).

It can be also expressed through the QoS of the service stages in the following way:

$$Qnet = Qa\, Qd\, Qs\, Qz\, Qb\, Qbu\, Qis, \tag{2}$$

where Qa is the QoS indicator of the A-Terminal stage; Qd is the QoS indicator of the Dialing stage; Qs is the QoS indicator of the Switching stage; Qz is the QoS indicator of the B-Seizure stage; Qb is the QoS indicator of the B-Terminal stage; Qbu is the QoS indicator of the Finding B-user stage; Qis is the QoS indicator of the Information security stage (see [11]).

The second indicator is the QoS indicator of the overall telecommunication system (denoted by $Qsys$). It is given by

$$Qsys = \frac{crr.Fc}{int.Fa}. \tag{3}$$

Another expression for $Qsys$ derived in [11] is:

$$Qsys = \frac{Adir}{Adem} Qa\, Qd\, Qs\, Qz\, Qb\, Qbu\, Qis\, Qc, \tag{4}$$

where $Adir$ is the performance indicator of the A-User demanding stage (Demand-Intent ratio); $Adem$ is the performance indicator of the A-User offering stage (Demand-Offered ratio); Qc is the QoS indicator of the Communication stage (see [11]).

Finally, we shall demonstrate how the GN model can be used to estimate the QoS at service stage level by evaluating the QoS indicator of the Target communication stage (Qc) given by:

$$Qc = \frac{crr.Fc}{ofr.Fc}. \tag{5}$$

3 Generalized Net Model

In this section, on the basis of the conceptual model shown in Fig. 2 we will construct a GN model of QoS estimation of overall telecommunication system including a security stage. The graphical representation of the GN is shown in Fig. 3.

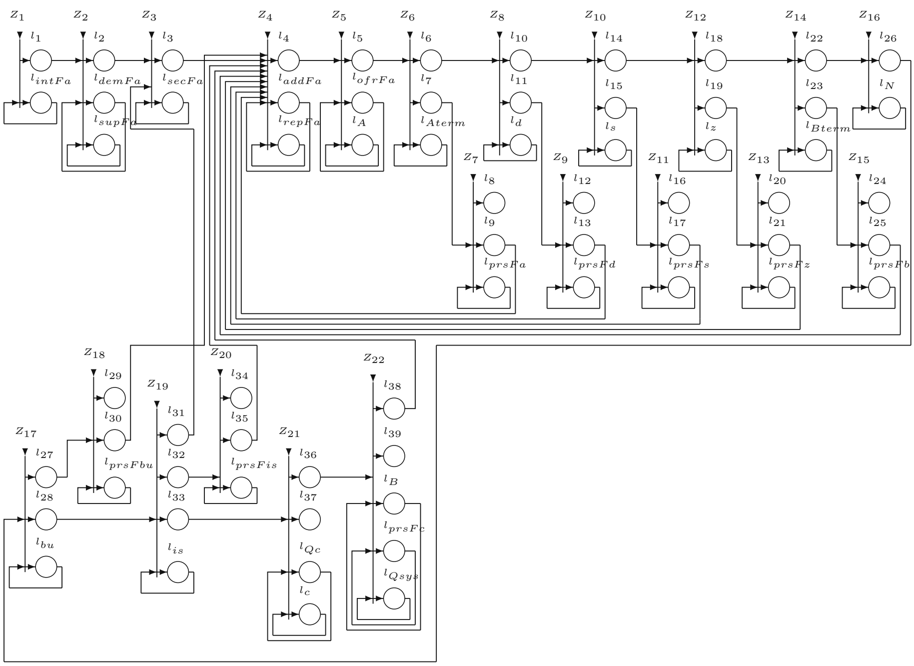


Fig. 3. Graphical representation of the GN model

The GN consists of 22 transitions and 67 places. In addition, 28 types of tokens are used. In some places of the GN, the tokens obtain as characteristic the values of parameters of some device. These places are denoted by l_a where a is either the name of the corresponding virtual device (when the place represents a device) or a parameter's name (when the place is used to represent a parameter) and the symbol '.' is omitted. For example, l_N is the place where the tokens obtain as characteristic the values of the parameters of the comprise Network device. The functions of each transition of the GN are:

- Z_1 represents the function of the generator of the intent call attempts with intensity $int.Fa$.
- Z_2 represents the function of the generator of the suppressed intent call attempts with intensity $sup.Fa$. It is also used to represent the flow intensity of the demand call attempts ($dem.Fa$).
- Z_3 represents the flow intensity of the security checking ($sec.Fa$).
- Z_4 represents the flow intensity of the repeated attempts ($rep.Fa$) and the flow intensity of the additional requests ($add.Fa$).
- Z_5 represents the function of the A-User device and the flow intensity of the offered requests (demand, security and repeated) trying to occupy the A-terminals ($ofr.Fa$).
- Z_6 represents the A-terminal device.
- Z_7 represents the flow intensity of the parasitic served requests by the A-terminal device.
- Z_8 represents the Dialing stage.
- Z_9 represents the intensity of the parasitic call attempts in the Dialing stage.
- Z_{10} represents the function of the Switching stage (the s device).
- Z_{11} represents the intensity of the parasitic call attempts in the s device.
- Z_{12} represents the function of the B-seizure stage.
- Z_{13} represents the intensity of the parasitic call attempts in the z device.
- Z_{14} represents the function of the B-Terminal device.
- Z_{15} represents the intensity of the parasitic call attempts of the B-terminal stage.
- Z_{16} represents the function of the comprise Network device.
- Z_{17} represents the function of the Finding B-User stage.
- Z_{18} represents the intensity of the parasitic call attempts to the B-User device.
- Z_{19} represents the function of the Information security stage.
- Z_{20} represents the intensity of the parasitic call attempts in the Information security stage.
- Z_{21} represents the function of the Target communication stage.

- Z_{22} represents the intensity of the parasitic call attempts in the Target communication stage and also the function of the B-User device. It is also used to estimate the values of the QoS indicators at overall system level.

The 28 types of tokens are denoted by $\alpha, \alpha, \cdots, \omega, \alpha_1, \alpha_2, \alpha_3, \alpha_4$. Initially, the following tokens occupy places of the GN:

- α in place l_{intFa} with initial characteristic "initial value of $int.Fa$".
- α in place l_{demFa} with initial characteristic "initial value of $dem.Fa$".
- γ in place l_{supFa} with initial characteristic "initial value of $sup.Fa$".
- δ in place $l_{sec.Fa}$ with initial characteristic "initial value of $sec.Fa$".
- ϵ in place l_{addFa} with initial characteristic "initial value of $add.Fa$".
- ζ in place $l_{rep.Fa}$ with initial characteristic "initial value of $rep.Fa$".
- η in place $l_{ofr.Fa}$ with initial characteristic "initial value of $ofr.Fa$".
- θ in place l_A with initial characteristic "initial values of the parameters of the A-User device".
- ι in place l_{Aterm} with initial characteristic "initial values of Ya, Fa, Ta, Pa".
- κ in place l_{prsFa} with initial characteristic "initial value of $prs.Fa$".
- λ in place l_d with initial characteristic "initial values of Fd, Yd, Td, Pd".
- μ in place l_{prsFd} with initial characteristic "initial value of $prs.Fd$".
- ν in place l_s with initial characteristic "initial value of Fs, Ys, Ts, Ps".
- ξ in place l_{prsFs} with initial characteristic "initial value of $prs.Fs$".
- o in place l_z with initial characteristic "initial values of Fz, Yz, Tz, Pz".
- π in place l_{prsFz} with initial characteristic "initial value of $prs.Fz$".
- ρ in place l_{Bterm} with initial characteristic "initial values of Fb, Yb, Tb, Pb".
- σ in place l_{prsFb} with initial characteristic "initial value of $prs.Fb$".
- τ in place l_N with initial characteristic "initial values of the parameters of the Network device".
- υ in place l_{bu} with initial characteristic "initial values of Fbu, Ybu, Tbu, Pbu".
- ϕ in place l_{prsFbu} with initial characteristic "initial value of $prs.Fbu$".
- χ in place l_{is} with initial characteristic "initial values of Fis, Yis, Tis, Pis".
- ψ in place l_{prsFis} with initial characteristic "initial value of $prs.Fis$".
- ω in place l_{Qc} with initial characteristic "initial value of the QoS indicator Qc".
- α_1 in place l_c with initial characteristic "initial values of Fc, Yc, Tc, Pc".
- α_2 in place l_B with initial characteristic "initial values of the parameters of the B-User device".
- α_3 in place l_{prsFc} with initial characteristic "initial value of $prs.Fc$".
- α_4 in place l_{Qsys} with initial characteristic "initial value of the QoS indicator $Qsys$".

Below we will describe each of the transitions of the GN.

$$Z_1 = \langle \{l_{intFa}\}, \{l_1, l_{intFa}\}, r_1\rangle,$$

where

$$r_1 = \begin{array}{c|cc} & l_1 & l_{intFa} \\ \hline l_{intFa} & true & true \end{array}.$$

Token α in l_{intFa} splits into two identical tokens. The first one enters place l_1 and receives the characteristic "current value of $int.Fa$". The second one remains in place l_{intFa} with characteristic "current value of $int.Fa$ ".

$$Z_2 = \langle \{l_1, l_{dem.Fa}, l_{supFa}\}, \{l_2, l_{demFa}, l_{supFa}\}, r_2\rangle,$$

where

$$r_2 = \begin{array}{c|ccc} & l_2 & l_{demFa} & l_{supFa} \\ \hline l_1 & true & true & true \\ l_{demFa} & false & true & false \\ l_{supFa} & false & false & true \end{array}.$$

Token α in l_1 splits into three identical tokens. The first one enters place l_2 with characteristic "current value of $dem.Fa$". The second one merges with the α token in place l_{demFa} with characteristic "current value of $dem.Fa$". The third one merges with token γ in place l_{supFa}. Token γ in place $sup.Fa$ obtains the characteristic "current value of $sup.Fa$".

$$Z_3 = \langle \{l_2, l_{31}, l_{sec.Fa}\}, \{l_3, l_{secFa}\}, r_3\rangle,$$

where

$$r_3 = \begin{array}{c|cc} & l_3 & l_{secFa} \\ \hline l_2 & true & false \\ l_{31} & true & true \\ l_{secFa} & false & true \end{array}.$$

Token α from l_2 enters place l_3. The token from place l_{31} splits into two identical tokens. The first one merges with token α in l_3. The second one merges with the δ token in place l_{secFa}. Token δ in place l_{secFa} obtains the characteristic "current value of $sec.Fa$".

$$Z_4 = \langle \{l_3, l_9, l_{13}, l_{17}, l_{21}, l_{25}, l_{35}, l_{38}, l_{addFa}, l_{repFa}\}, \{l_4, l_{addFa}, l_{repFa}\}, r_4\rangle,$$

where

$$r_4 = \begin{array}{c|ccc} & l_4 & l_{addFa} & l_{repFa} \\ \hline l_3 & true & false & false \\ l_9 & false & false & true \\ l_{13} & false & false & true \\ l_{17} & false & false & true \\ l_{21} & false & false & true \\ l_{25} & false & false & true \\ l_{35} & false & false & true \\ l_{38} & false & false & true \\ l_{addFa} & false & true & false \\ l_{repFa} & false & false & true \end{array}.$$

Token α from place l_3 enters place l_4 without obtaining a new characteristic. The tokens from places l_9, l_{13}, l_{17}, l_{21}, l_{25}, l_{35}, l_{38} enter place l_{repFa} where they merge with token ζ obtaining the characteristic "current value of $rep.Fa$". Token ϵ in place l_{addFa} obtains the characteristic "current value of $add.Fa$".

$$Z_5 = \langle \{l_4, l_{ofrFa}, l_A\}, \{l_5, l_{ofrFa}, l_A\}, r_5 \rangle,$$

where

$$r_5 = \begin{array}{c|ccc} & l_5 & l_{ofrFa} & l_A \\ \hline l_4 & true & true & true \\ l_{ofrFa} & false & true & false \\ l_A & false & false & true \end{array}.$$

Token α in l_4 splits into three identical tokens. The first one enters place l_5 with characteristic "current value of $ofr.Fa$". The second merges with token η in place l_{ofrFa}. The third one merges with token θ in place l_A. Token η in place l_{ofrFa} obtains the characteristic "current value of $ofr.Fa$". Token θ in place l_A obtains the characteristic "current values of the parameters of the A-user device".

$$Z_6 = \langle \{l_5, l_{Aterm}\}, \{l_6, l_7, l_{Aterm}\}, r_6 \rangle,$$

where

$$r_6 = \begin{array}{c|ccc} & l_6 & l_7 & l_{Aterm} \\ \hline l_5 & true & true & false \\ l_{Aterm} & false & false & true \end{array}.$$

Token α in l_5 splits into two identical tokens. They are transferred to places l_6 and l_7 respectively, without obtaining a new characteristic. Token ι in place l_{Aterm} obtains the characteristic "current values of Fa, Ya, Ta, Pa".

$$Z_7 = \langle \{l_7, l_{prsFa}\}, \{l_8, l_9, l_{prsFa}\}, r_7 \rangle,$$

where

$$r_7 = \begin{array}{c|ccc} & l_8 & l_9 & l_{prsFa} \\ \hline l_7 & true & true & false \\ l_{prsFa} & false & false & true \end{array}.$$

Token α in l_7 splits into two identical tokens. They are transferred to places l_8 and l_9 respectively, without obtaining a new characteristic. Token κ in place l_{prsFa} obtains the characteristic "current value of $prs.Fa$".

$$Z_8 = \langle \{l_6, l_d\}, \{l_{10}, l_{11}, l_d\}, r_8 \rangle,$$

where

$$r_8 = \begin{array}{c|ccc} & l_{10} & l_{11} & l_d \\ \hline l_6 & true & true & false \\ l_d & false & false & true \end{array}.$$

Token α in l_6 splits into two identical tokens. They are transferred to places l_{10} and l_{11} respectively, without obtaining a new characteristic. Token λ in place l_d obtains the characteristic "current values of Fd, Yd, Td, Pd".

The formal description of transitions Z_9, Z_{11}, Z_{13} and Z_{15} is analogous to that of transition Z_7. Therefore, we shall omit their description.

$$Z_{10} = \langle \{l_{10}, l_s\}, \{l_{14}, l_{15}, l_s\}, r_{10} \rangle,$$

where

$$r_{10} = \begin{array}{c|ccc} & l_{14} & l_{15} & l_s \\ \hline l_{10} & true & true & false \\ l_s & false & false & true \end{array}.$$

Token α in l_{10} splits into two identical tokens. They are transferred to places l_{14} and l_{15} respectively, without obtaining a new characteristic. Token ν in place l_d obtains the characteristic "current values of Fs, Ys, Ts, Ps".

$$Z_{12} = \langle \{l_{14}, l_z\}, \{l_{18}, l_{19}, l_z\}, r_{12} \rangle,$$

where

$$r_{12} = \begin{array}{c|ccc} & l_{18} & l_{19} & l_z \\ \hline l_{14} & true & true & false \\ l_z & false & false & true \end{array}.$$

Token α in l_{14} splits into two identical tokens. They are transferred to places l_{18} and l_{19} respectively, without obtaining a new characteristic. Token o in place l_z obtains the characteristic "current values of Fz, Yz, Tz, Pz".

$$Z_{14} = \langle \{l_{18}, l_{Bterm}\}, \{l_{22}, l_{23}, l_{Bterm}\}, r_{14} \rangle,$$

where

$$r_{14} = \begin{array}{c|ccc} & l_{22} & l_{23} & l_{Bterm} \\ \hline l_{18} & true & true & false \\ l_{Bterm} & false & false & true \end{array}.$$

Token α in l_{18} splits into two identical tokens. They are transferred to places l_{22} and l_{23} respectively, without obtaining a new characteristic. Token ρ in place l_{Bterm} obtains the characteristic "current values of Fb, Yb, Tb, Pb".

$$Z_{16} = \langle \{l_{22}, l_N\}, \{l_{26}, l_N\}, r_{16} \rangle,$$

where

$$r_{16} = \begin{array}{c|cc} & l_{26} & l_N \\ \hline l_{22} & true & true \\ l_N & false & true \end{array}.$$

Token α in place l_{22} splits into two identical tokens. The first one enters place l_{26} without obtaining a new characteristic. The second one enters place l_N where it merges with token τ. Token τ in place l_N obtains the characteristic "current values of the parameters of the Network device; current value of the QoS indicator on network level ($Qnet$)".

$$Z_{17} = \langle \{l_{26}, l_{bu}\}, \{l_{27}, l_{28}, l_{bu}\}, r_{17} \rangle,$$

where

$$r_{17} = \begin{array}{c|ccc} & l_{27} & l_{28} & l_{bu} \\ \hline l_{26} & true & true & false \\ l_{bu} & false & false & true \end{array}.$$

Token α in l_{26} splits into two identical tokens. They are transferred to places l_{27} and l_{28} respectively, without obtaining a new characteristic. Token ν in place l_{bu} obtains the characteristic "current values of Fbu, Ybu, Tbu, Pbu".

$$Z_{18} = \langle \{l_{27}, l_{prsFbu}\}, \{l_{29}, l_{30}, l_{prsFbu}\}, r_{18} \rangle,$$

where

$$r_{18} = \begin{array}{c|ccc} & l_{29} & l_{30} & l_{prsFbu} \\ \hline l_{27} & true & true & false \\ l_{prsFbu} & false & false & true \end{array}.$$

Token α in l_{27} splits into two identical tokens. They are transferred to places l_{29} and l_{30} respectively, without obtaining a new characteristic. Token ϕ in place l_{prsFbu} obtains the characteristic "current value of $prs.Fbu$".

$$Z_{19} = \langle \{l_{28}, l_{is}\}, \{l_{31}, l_{32}, l_{33}, l_{is}\}, r_{19} \rangle,$$

where

$$r_{19} = \begin{array}{c|cccc} & l_{31} & l_{32} & l_{33} & l_{is} \\ \hline l_{28} & true & true & true & false \\ l_{is} & false & false & false & true \end{array}.$$

Token α in l_{28} splits into three identical tokens. They are transferred to places l_{31}, l_{32} and l_{33} respectively. They do not receive a new characteristic. Token χ in place l_{is} obtains the characteristic "current values of Fis, Yis, Tis, Pis".

$$Z_{20} = \langle \{l_{32}, l_{prsFis}\}, \{l_{34}, l_{35}, l_{prsFis}\}, r_{20} \rangle,$$

where

$$r_{20} = \begin{array}{c|ccc} & l_{34} & l_{35} & l_{prsFis} \\ \hline l_{32} & true & true & false \\ l_{prsFis} & false & false & true \end{array}.$$

Token α in l_{32} splits into two identical tokens. They are transferred to places l_{34} and l_{35} respectively, without obtaining a new characteristic. Token ψ in place l_{prsFis} obtains the characteristic "current value of $prs.Fis$".

$$Z_{21} = \langle \{l_{33}, l_{Qc}, l_c\}, \{l_{36}, l_{37}, l_{Qc}, l_c\}, r_{21} \rangle,$$

where

$$r_{21} = \begin{array}{c|cccc} & l_{36} & l_{37} & l_{Qc} & l_c \\ \hline l_{33} & true & true & false & false \\ l_{Qc} & false & false & true & false \\ l_c & false & false & false & true \end{array}.$$

Token α in place l_{33} splits into two identical tokens. They are transferred to places l_{36} and l_{37} respectively, without obtaining a new characteristic. Token ω in place l_{Qc} obtains the characteristic "current value of the QoS indicator of the Target communication stage (Qc)". Token α_1 in place l_c obtains the characteristic "current values of Fc, Yc, Tc, Pc".

$$Z_{22} = \langle \{l_{36}, l_B, l_{prsFc}, l_{Qsys}\}, \{l_{38}, l_{39}, l_B, l_{prsFc}, l_{Qsys}\}, r_{22} \rangle,$$

where

$$r_{22} = \begin{array}{c|ccccc} & l_{38} & l_{39} & l_B & l_{prsFc} & l_{Qsys} \\ \hline l_{36} & true & true & false & false & false \\ l_B & false & false & true & false & false \\ l_{prsFc} & false & false & false & true & false \\ l_{Qsys} & false & false & false & false & true \end{array}.$$

Token α in l_{36} splits into two identical tokens. They are transferred to places l_{38} and l_{39} respectively, without obtaining a new characteristic. Token α_2 in place l_B obtains the characteristic "current values of the parameters of the B-user device". Token α_3 in place l_{prsFc} obtains the characteristic "current value of $prs.Fc$". Token ω in place l_{Qsys} obtains the characteristic "current value of the QoS indicator on overall system level ($Qsys$)".

4 Conclusion

For the first time, a GN model of QoS estimation of overall telecommunication system including users and a security stage is proposed. The estimations of the QoS of the Target communication stage, the QoS on overall network level and the QoS on overall system level are obtained through the characteristic function of some special places of the GN. The GN approach allows to evaluate other QoS indicators just by changing the characteristic function related to the place where the tokens obtain as characteristic the QoS indicator values. This can be done without changing the other components of the GN and, in particular, the graphical representation of the GN remains the same.

The proposed model can be used for monitoring prediction and management of QoS in overall telecommunication system. It allows for implementation of dynamic pricing policies. The approach can be also applied to next generation informational service networks for construction of a large scale QoS guarantee architecture [8,9].

In future, the proposed GN model of QoS estimation in overall telecommunication system including a security stage will be extended through the inclusion of some of the intuitionistic fuzzy estimations of uncertainty proposed in [13].

References

1. Andonov, V., Poryazov, S., Saranova, E.: Generalized net representations of elements of service systems theory. Adv. Stud. Contemp. Math. **29**(2), 179–189 (2019)
2. Andonov, V., Poryazov, S., Saranova, E.: Generalized net model of a serial composition of services with intuitionistic fuzzy estimations of uncertainty. In: Kahraman, C., Tolga, A.C., Cevik Onar, S., Cebi, S., Oztaysi, B., Sari, I.U. (eds.) Intelligent and Fuzzy Systems, INFUS 2022. LNNS, vol. 504, pp. 616–623. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09173-5_71
3. Andonov, V., Poryazov, S., Saranova, E.: Generalized net model of a parallel composition of services with intuitionistic fuzzy estimations of uncertainty. In: 2022 IEEE 11th International Conference on Intelligent Systems (IS), Warsaw, Poland, pp. 1–6 (2022)
4. Andonov, V., Poryazov, S., Saranova, E.: Generalized net models of traffic quality evaluation of a service stage. In: Proceedings of the Twentieth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, Warsaw, Poland, 14 October 2022. (in press)
5. Andonov, V., Poryazov, S., Saranova, E.: Generalized net model of overall network efficiency evaluation. In: Proceedings of the Twentieth International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, Warsaw, Poland, 14 October 2022. (in press)
6. Atanassov, K.: On Generalized Nets Theory. Prof. M. Drinov Academic Publishing House, Sofia (2007)
7. Gao, F.: Modelling, planning and adaptive implementation of semantic event service. Doctoral dissertation, University of Galway, Ireland (2016)
8. Ghazel, C., Saidane, L.: An improved model for next generation networks with guaranteed end-to-end quality of service. In: 2007 15th IEEE International Conference on Networks, Adelaide, SA, Australia, pp. 112–117 (2007)
9. Natkaniec, M., Gozdecki, J., Sargento, S.: A QoS architecture integrating AdHoc and infrastructure in next generation networks. In: Bestak, R., Simak, B., Kozlowska, E. (eds.) PWC 2007. ITIFIP, vol. 245, pp. 25–37. Springer, Cham (2007). https://doi.org/10.1007/978-0-387-74159-8_3
10. Poryazov, S., Saranova, E.: Models of Telecommunication Networks with Virtual Channel Switching and Applications. Prof. Marin Drinov Academic Publishing House, Sofia (2012). (in Bulgarian)
11. Poryazov, S., Progonov, D., Saranova, E.: Quality of telecommunications as a composition of qualities of subservices, including security and trusted third parties. In: 2017 25th Telecommunication Forum (TELFOR), Belgrade, Serbia, pp. 1–4 (2017)
12. Poryazov, S., Saranova, E., Ganchev, I.: Scalable traffic quality and system efficiency indicators towards overall telecommunication system's QoE management. In: Ganchev, I., van der Mei, R., van den Berg, H. (eds.) Autonomous Control for a Reliable Internet of Services. LNCS, vol. 10768, pp. 81–103. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-90415-3_4

13. Poryazov, S., Andonov, V., Saranova, E.: Three intuitionistic fuzzy estimations of uncertainty in service compositions. In: Atanassov, K.T. et al. (eds.) Uncertainty and Imprecision in Decision Making and Decision Support: New Advances, Challenges, and Perspectives. IWIFSGN 2020, BOS/SOR 2020. LNNS, vol. 338, pp. 72–84. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-95929-6_6
14. Saibharath, S., Mishra, S., Hota, C.: Joint QoS and energy-efficient resource allocation and scheduling in 5G Network Slicing. Comput. Commun. **202**, 110–123 (2023)
15. Yerima, S.Y., Parr, G.P., McClean, S.I., Morrow, P.J.: Adaptive measurement-based policy-driven QoS management with fuzzy-rule-based resource allocation. Future Internet **2012**(4), 646–671 (2012)

A New Method of Smart City Modeling Using Big Data Techniques

Muhammet Sıddık Emek[1,2] and Ismail Burak Parlak[1,2](✉)

[1] Department of Computer Engineering, Galatasaray University, Ciragan Cad. No: 36, Ortakoy 34349, Istanbul, Turkey
[2] NLPLAB, Ciragan Cad. No: 36, Ortakoy 34349, Istanbul, Turkey
muhammetsiddik.emek@ogr.gsu.edu.tr, bparlak@gsu.edu.tr

Abstract. The classification of city activities on social media is a complex problem based on semantic analysis for smart city modeling. Urban area characteristics are analyzed through hierarchical tasks where location properties are gathered along with municipality records and city demographics. Actually, urban locations are graded in social media since multimodal spatiotemporal features incorporating real-time topics in user activities are indexed. Their popularity, their quality, their customer based interaction are evaluated and shared as city insights in different platforms. As a data source, social communication leads to a gap between unstructured information and online city trends. People who share their activities in social media bring their sentiments as multimodal opinions in several formats which result in a complex big data. City semantics are coupled with city demographics and citizen feelings in urban locations. Therefore, new methods are required to support municipality records in smart city classification. In this study, we have focused on smart city classification tasks in two metropolitan areas; Istanbul and Ankara. We have fetched five different types of locations from Foursquare and Twitter platforms. A pre-trained Turkish language model has been used to perform the microblogging scoring. After the preprocessing steps, we have applied the tuned language model onto our Turkish dataset. Five different language models have been compared using statistical evaluation metrics. Our results show that location microblogging would be a premising benchmark in city semantics. The accuracy rates were found above 90% in three different classes. We conclude that a spatial distribution of user reviews leads to offer new metrics in smart city measurement.

Keywords: Social Media · Text Classification · Natural Language Processing · Machine Learning · BERT

1 Introduction

The social interactions in mobile technologies have brought new perspectives in Internet to monitor social activities occurring within a city. Activity analysis is mirrored in virtual environments where social feedbacks reveal the conjunctions in city based interactions. Thus, social activities leads to generate dynamic

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 772–779, 2023.
https://doi.org/10.1007/978-3-031-39777-6_90

relations in smart city analysis. The social data collected in social platforms offer significant insights of individual activities related to the behavioral trends in cities. In general, these activities are considered as an integral part of user daily life. Moreover, the data refers to social big data comprising vast amount of information. The temporal activity is modeled in social networks as digital twins of physical locations. Therefore, social data enables individuals to infer city metrics about the formation of activity patterns within the city through a bottom-up approach.

The enhanced accessibility of location information allows individuals to integrate location aspects into their profiles on social networks Since users can upload location-tagged photographs, provide comments on microblogs, they provide a heterogeneous data regarding the event location, Thus, Foursquare and Twitter platforms create a social digital twin for physical locations to share the current whereabouts by providing social city semantics. Furthermore, the narratives on microblogs provide a location based context related to the previous experiences of users in social network. Consequently, social platforms function as a city based tool that promotes the exchange of personal ideas while browsing a physical location.

In this study, we have collected our dataset through Foursquare and Twitter social platforms There are several alternatives to access data from Foursquare and Twitter. The Foursquare Places API is a popular option that enables developers to create location-based applications and services by accessing the information on users, venues, check-ins, photos, and tips. We have evaluated new big data classification algorithms for Turkish language in two metropolitan areas such as Ankara and Istanbul. The main challenge was the multiclass modeling of location based user comments. Thus, we divided the study into two parts; pretrained data analysis and classification. We trained our initial dataset with BERT Turkish language model and we measured the performance of five different classification techniques which are proposed as a new tool in smart city modeling. We aimed to combine two social platforms and to collect and measure user comments in smart city analysis.

We have organized this study as it follows. The second section deals with the smart city classification problem through machine learning techniques in recent approaches. The third section details the fetching in social media and the main steps of corpus preprocessing, training, classification and evaluation process. The fourth reviews our classification in three classes where the performance has been presented through the accuracy scores in validation datasets. Consequently, the final section highlights the key points and our future steps in city classification problem.

2 Related Works

A location-based social network (LBSN) adds a location dimension to an existing social network and creates a new social structure based on the interdependence of individuals People in LBSN are gathered from their physical coordinates in the world, as well as multimodal media content with location tags. The physical location is determined by the individual's instantaneous location at a particular time and location history over a certain period. Interdependence refers to the coexistence of two people in the same coordinates or sharing similar location histories and information such as common activities, likes, comments and behaviors derived from the individual's location and location-tagged data. LBSN data with geolocation information is a significant component of smart city analysis. However, traditional social network analysis methods often fail to incorporate location information when generating recommendation services.

In the last decade, the development of big LBSN has provided a huge amount of data in city based activities and recommendations. Social, cultural, natural and historical aspects of a city are combined in the exploration of tourism. The city power is measured intrinsically in the development strategy of tourism. The multilingual reviews, recommendations, photos and videos in LBSN create an unstructured data pool for multimodal insights. Therefore, it becomes crucial to infer and predict city metrics in smart city growth. In traditional studies, researches were focused on quantitative measurements through demographics. In big data era, these multimodal data sources present new social insights to explore large scale activities in metropolitan area. Activity visualization, interaction, scheduling, recommendation are the most important keypoints in city infographics. Especially, the spatial and temporal distribution in activities becomes appropriate tools in tourist recommendations. On the other hand, the flows between touristic activities are generally ignored or not visualized in depth. LBSN provides spatial and temporal network insights in city analysis for touristic recommendation. Thus, the functional relation in touristic landmarks might be predicted through LBSN data. These relations are provided in previous experiences and lead to quantify city dynamics in social activities.

Foursquare becomes a popular LBSN by enabling users to search for and find the suitable places based on previous experiences, check-ins, ratings and comments. Moreover, it provides several features such as search places, and explore other people's recommendations. Silva et al. [1] used transition graphs to analyze the social activities within a Personal Social Network (PSN) across eight different cities with Foursquare. They achieved to summarize the dynamics of each city by visualizing the network coverage and seasonality. Assem et al. [2] involved the analysis and the visualization of the temporal functional regions that were generated for a 4-hour time interval in New York City. They collected the dataset on Foursquare to perform spatiotemporal analysis and location clustering. Gallegos et al. [3] revealed that venues with check-ins in Los Angeles are inherently distinct from other areas using Twitter and Foursquare. They observed the locations featuring a diverse range of activities and establishments, such as restaurants, gyms, and beaches since they tend to attract a larger number of visitors. They measured the human mobility to underline the geomapping and sentiment strength in social platforms. In a similar study, De et al. [4] study presented various aspects related to spatio-temporal and composite lifestyles in Foursquare. They derived lifestyle patterns to extract hidden spatio-temporal behavior patterns.

In Turkish language, there are few studies related to spatiotemporal human activity classification in smart city. Aksu and Karaman [5] labeled manually the reviews retrieved from Tripadvisor, Google Maps and Foursquare. They compared the performance of machine learning approaches in sentiment analysis for touristic information. Yuksel and Tan [6] developed a decision support system for social networks in Turkish and English languages. They labeled manually Foursquare comments in text mining and sentiment analysis.

3 Methodology

In this study, we have created both manually and automatically annotated corpuses. Firstly, Foursquare tips were pulled out on Twitter to obtain venue homepage URL information on Foursquare. Twint API was preferred instead of Twitter API to accomplish this task. Twint is a Python programming language library for tweet scraping through different parameters from user profiles. There are different alternatives to accessing Foursquare data, as such in Twitter. The Foursquare Places API, provided location-based experiences with a variety of information about venues, users, photos, check-ins, and tips, is the first can be used. All these information can be obtained through different services with special usage fees. In this study, we have combined Beautiful Soup and Selenium tools to retrieve our dataset with physical locations as it is represented in Fig. 1. Selenium achieved all the automated web browser interactions, then the physical location and user comments were extracted by Beautiful Soup.

14221 venue information was extracted from Foursquare, along with significant attributes including venueId, venue name, venue actual url, location, category list, tip count, rating, geolocation and address. On the other hand, an unique identifier has been assigned to each venue to associate user tips. Then the deduplication of the venue information according to the venue homepage url which is unique on Foursquare has been applied. Thus, it has been prevented that the comments of the same place are fetch more than once. In addition, some venues were removed because they were not located in Türkiye. Finally, 2043 venue information in Istanbul and 671 venue information in Ankare were located as unique physical locations.

The final dataset prepared by this study consists of 2 csv files which are the Venue Information and the Tips of the users about the relevant venue. All these information were gathered from Foursquare web page by using the method clarified before. In venue information file, there are 2741 unique records with many significant attributes like name, rating, address, location and categories. On the other hand, more than 1000 different categories give information about the type of place, and these categories are divided into 15 groups in total. The distribution of the venues according to the city and category group are shown in Fig. 2.

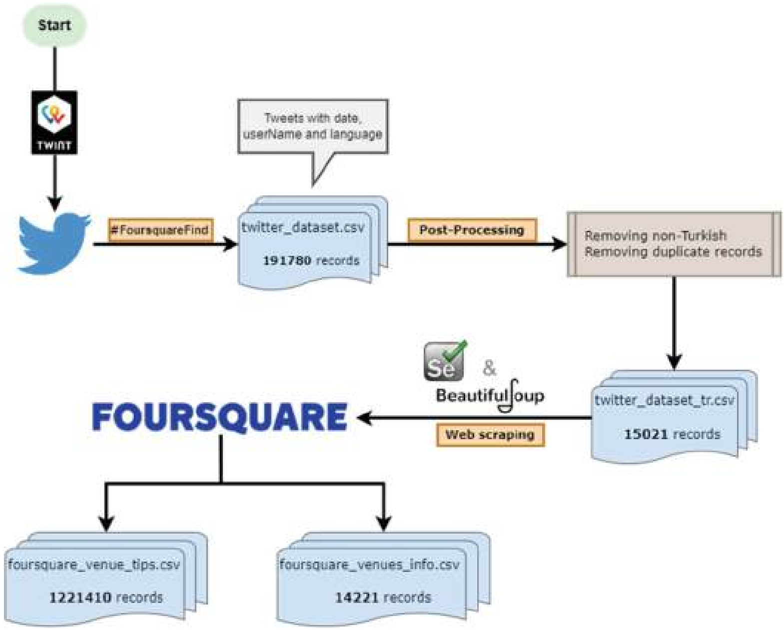


Fig. 1. The flowchart of the methodology.

A new generation big data language model has been developed by Google in 2018. Bidirectional Encoder Representations from Transformers (BERT) is an open sourced data model which is designed to pretrain bidirectional representations from unlabeled microblogs. As a result, the pre-trained BERT model can be finetuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task specific architecture modifications. It has 24 transformer blocks, 1024 hidden layers, and 340M parameters for BERT Large model and 12 transformer blocks, 768 hidden layers, and 110M parameters for BERT Base model. Pre-trained with large text data, BERT allows it to be transferred for other small language processing tasks as needed. In its research stages, the framework achieved groundbreaking results in 11 natural language understanding tasks, including sentiment analysis, semantic role labeling, sentence classification and the disambiguation of polysemous words, or words with multiple meanings. A brief information regarding BERT models [7] are given below.

- BERT Base Turkish Cased is trained on a filtered and sentence segmented version of the Turkish OSCAR corpus, a recent Wikipedia dump, various OPUS corpora and a special corpus provided by Kemal Oflazer and is case sensitive
- BERT Base Turkish Uncased is trained on same dataset as used BERT Base Turkish Cased but is not case sensitive.

- BERT Base Turkish 128k Cased is trained on larger Turkish data compare to model and is case sensitive.
- BERT Base Turkish 128k Uncased is trained on same dataset as used BERT Base Turkish Cased but is not case sensitive.
- DistilBERT Turkish Cased is trained on 7 GB of the original training data BERTurk [7] using the cased version.

We have analyzed the user reviews in three categories; Positive, Neutral and Negative. All classes have been evaluated through statistical metrics such as Precision (PR), Recall (REC), F1 and Accuracy (ACC).

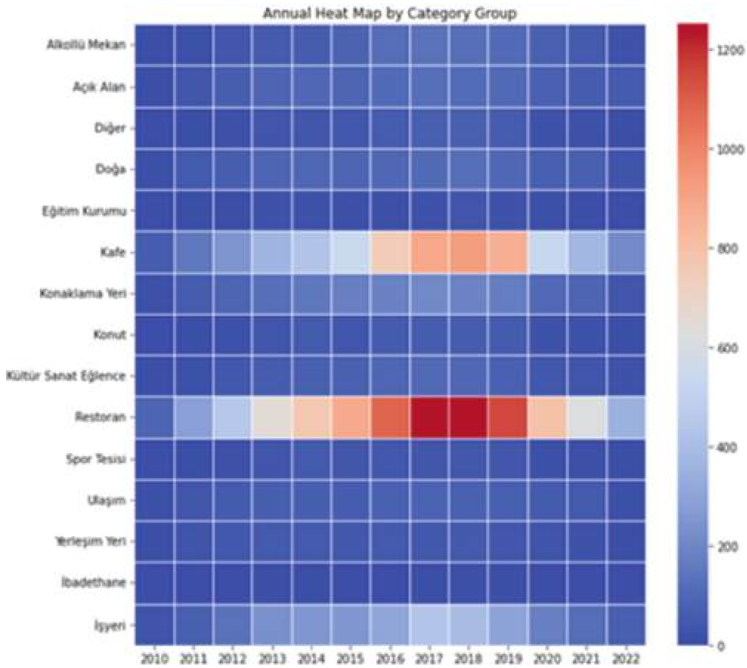


Fig. 2. The annual distribution of city labels in our corpus.

4 Results

Accuracy is a measure for evaluating the quality of the classification model. This measure can be defined as the ratio of the number of correctly classified examples according to the total number of classified examples. Although, this metric is useful in data analysis, it does not reflect the overall quality of the model. Therefore, we need other evaluation tools such as precision, recall, F-score. We have compared five different big data methods in city modeling such as BERT Turkish Cased, BERT Turkish Uncased, BERT Turkish 128k Cased, BERT Turkish 128k Uncased, DistilBERT Turkish Cased. Classification of the Foursqare tips are categorized in three different classes; negative, neutral or positive. They have been evaluated on four different statistical metrics such as Precision (PR), Recall (REC), F1 and Accuracy (ACC). The results are given in Table 1. We have noted that the accuracy scores are closer. Thus, we have interpreted the performance using PR, REC and F1. We have concluded that BERT Turkish 128k Cased represented best classification scores in smart city modeling on social media.

Table 1. Evaluation metrics of classifiers.

	Negative			Neutral			Positive			
	PR	REC	F1	PR	REC	F1	PR	REC	F1	ACC
BERT Turkish Cased	93	93	93	89	89	89	94	95	94	92
BERT Turkish Uncased	90	92	91	87	85	86	92	93	93	90
BERT Turkish 128k Cased	93	94	94	90	89	90	95	95	95	93
BERT Turkish 128k Uncased	91	93	92	88	85	86	93	93	93	90
DistilBERT Turkish Cased	93	92	88	87	88	88	94	93	93	91

5 Discussion and Conclusion

The social interactions in mobile technologies have brought new perspectives in Internet to monitor social activities occurring within a city. Activity analysis is mirrored in virtual environments where social feedbacks reveal the conjunctions in city based interactions. Thus, social activities leads to generate dynamic relations in smart city analysis. The social data collected in social platforms offer significant insights of individual activities related to the behavioral trends in cities. In general, these activities are considered as an integral part of user daily life. On the other hand, the places such as restaurants, public transport and accommodations comprise the functional bases of tourism in a city. Online social media sites, which have become a platform that facilitates the sharing of people's experiences, have a very important place in this mission. As a data source, social communication leads to a gap between unstructured information and online city trends. People who share their activities in social media bring their sentiments as multimodal opinions in several formats which result in a complex big data. City semantics are coupled with city demographics and citizen feelings in urban locations. Therefore, new methods are required to support municipality records in smart city classification. In this study, we have focused on smart city classification tasks in two metropolitan areas; Istanbul and Ankara. We have fetched five different types of locations from Foursquare and Twitter platforms. Sentiment Analysis is the most common text classification tool that analyses this big and unstructured data and tells whether the underlying sentiment is positive, negative our neutral. It helps us better understand how other people feel about it.

In this study, valuable data that can be used for sentiment analysis on user comments were gathered, thus creating a medium-sized Turkish data set. 14221 venue information was extracted from Foursquare, along with significant attributes including venueId, venue name, venue actual url, location, category list, tip count, rating, geolocation and address. 2043 venue information in Istanbul and 671 venue information in Ankare were located as unique physical locations. After preprocessing steps explained in detail, pre-trained BERT models have been fine-tuned with manually labelled data. We have compared five different big data methods in city modeling such as BERT Turkish Cased, BERT Turkish Uncased, BERT Turkish 128k Cased, BERT Turkish 128k Uncased, DistilBERT Turkish Cased BERT Turkish 128k Cased represented best classification scores in smart city modeling on social media For the future works, it is planned to work with various Machine Learning algorithms and models on the Auto-Annotated Dataset obtained in order to visualize social activities through several aspects.

References

1. Silva, T.H., de Melo, P.O.V., Almeida, J.M., Salles, J., Loureiro, A.A.: Visualizing the invisible image of cities. In: 2012 IEEE International Conference on Green Computing and Communications, pp. 382–389. IEEE (2012)
2. Assem, H., Xu, L., Buda, T. S., O'Sullivan, D.: Spatio-temporal clustering approach for detecting functional regions in cities. In: 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 370–377. IEEE (2016)
3. Gallegos, L., Lerman, K., Huang, A., Garcia, D.: Geography of emotion: where in a city are people happier?. In: Proceedings of the 25th International Conference Companion on World Wide Web, pp. 569–574 (2016)
4. De, S., Jassat, U., Grace, A., Wang, W., Moessner, K.: Mining composite spatio-temporal lifestyle patterns from geotagged social data. In: 2022 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), pp. 444–451. IEEE (2022)
5. Aksu, M.Ç., Karaman, E.: Analysis of Turkish sentiment expressions about touristic sites using machine learning. J. Intell. Syst. Theory Appl. **4**(2), 103–112 (2021)
6. Sinan Yüksel, A., Gülşah Tan, F.: A real-time social network-based knowledge discovery system for decision making. Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije **59**(3–4), 261–273 (2018)
7. Schweter, S.: BERTurk - BERT models for Turkish, Zenodo (2020). https://doi.org/10.5281/zenodo.3770924

High-Tech Means and Implements of Crime: Challenging the Security of Sustainable Urban Development

Murad M. Madzhumayev(✉)

RUDN University, Moscow 117198, Russia
murad.mad@outlook.com

Abstract. Cutting-edge technological advancements, e.g. 3D printers and unmanned aerial vehicles (UAVs), may become high-tech means and implements of offences posing enormous security threats to sustainable urban development. The relevance of the subject matter arises out of the nexus between unrest and urban security as it manifests on two levels, on the one hand, urban planning can contribute to disturbances, and on the other hand, there is an actual threat caused by riots to urban facilities. 3D printers and UAVs can potentially be exploited for tactical purposes, including supplying crowds with 3D-printed weapons, transporting them, and obstructing law enforcement officers. The elimination or minimization of outlined misuses of these tools of the Industry 4.0, augments urban security in a fundamental way. The main result of the study is that the abovementioned tools cannot be asserted as determinants of crime, they only play a peripheral, auxiliary role in the commission of criminal acts. Thus, it is concluded that these instruments do need to be regulated through an intelligent control approach. Research outcomes and conclusions were based on general research methods including logical, systemic, sociological, as well as special research methods such as formal-legal, interpretative, statistical, etc.

Keywords: Urban Security · Riots · Mob · 3D Printers · Unmanned Aerial Vehicles

1 Introduction

Inclusiveness, safety, resilience and eco-sustainability of cities and communities are all integral pillars of the 17 Sustainable Development Goals promoted at the UN Summit on Sustainable Development in New York held in September 2015, "Transforming our World: The 2030 Agenda for Sustainable Development" [1]. These are associated with urbanization processes, the attractiveness of economic upsurge as a means of poverty abolition, employment opportunities, societal cohesion, etc. Sustainable development notion emerged intrinsically in urban design, as referred in several instances in the New Urban Agenda (Declaration on Sustainable Cities and Human Settlements for All and New Urban Agenda Implementation Plan) adopted by the UN Conference on Housing and Sustainable Urban Development (Habitat III) in 2016, which was endorsed by UN General Assembly resolution 71/256 of December 23, 2016.

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 780–787, 2023.
https://doi.org/10.1007/978-3-031-39777-6_91

The quintessential essence of sustainable urban development and planning is the state of its secureness vis-a-vis hierogeneous menaces and risks. Civil disorders (riots) appear to be representative in this regard. Thus, the urban planning interaction with riots has two dimensions:

1. As determinants: Improper planning in conjunction with various social, economic, institutional, or political issues might lead to violent unrest. For instance, ghettos in the US in the 1960s triggered the most devastating turmoil of that decade.
2. As a consequence: Besides the human casualties, there is significant damage caused precisely as a result of the civil disturbances to the city itself. During mayhem, historical and cultural monuments, the business establishment can be destroyed, buildings and structures may be set on fire, and there can be excess vandalism of city facilities.

Civil disorders are an unruly, dynamic, diverse, and heterogeneous mob formation that in modern history has led to and remains capable of initiating radical (not always positive) transformations worldwide. Along with conventional means like distribution of leaflets, media announcements, utilization of stones, truncheons and explosive devices, organizers and active rioters may also potentially weaponize 3D-printed items, and unmanned aerial vehicles (UAVs) for their vindictive intentions.

Technological innovations and digitalization shaped the advent of the fourth industrial revolution (Industry 4.0) [2]. One of the mega-trends of Industry 4.0 is additive manufacturing [2]. Sophistication of aviation industry along with technological advances in this direction in general provides rioters additional opportunities to use non-traditional techniques and means of committing the specified offense. In this regard, here the author is predisposed to refer such phenomena as civil disorders 2.0.

Legal scholars, criminologists, social scientists, and remaining researchers still have not comprehensively studied whether information and communication technologies, 3D printers, and UAVs constitute per se the root causes of such phenomena as civil disorders.

Researchers like A. T. Little, M. Ananyev, G. Zudenkova, and M. Petrova state that contemporary information and communication technologies are contributing to protest movements which challenge authoritarian regime stability [3, 4]. Through structural (technological, political, socioeconomic contexts) and instrumental (claims, traditional media, suppression of demonstrations, international resonance) factors, C.D.M. Soares and L.A. Joia addressed social networks' influences on social movements [5]. A study undertaken by V. N. Grigoryev outlines distinctive peculiarities of criminal prosecution of civil disorders with application of information technologies and quadcopters, particularly in the following directions: investigation management of riots; recording and tracking of primary and criminal case files on disorders; documenting crime scene circumstances; evidencing participation in civil disturbances [6].

Proponents of techno-determinism D. Christensen and F. Garfias argue in favor of the preeminent role of technology in collective actions since civil unrest is a violent form of the latter [7].

In contrast, C. Fuchs qualifies the assumption of technology as the catalyst of mass disorder as a fetishism of things, considering only social interactions to be primordial [8].

The Arab spring, England riots of 2011, the "Yellow Vest" movement in France, RickyLeaks in Puerto Rico, Iranian demonstrations of 2019, Hong Kong disturbances of

2019–2020, and the activity of the BLM movement in 2020–2021 all raise relevance of research on the probable impact of ICTs on civil disorders. No research at all is available regarding such an impact of 3D printers and unmanned aerial vehicles. The prevalence and accessibility of the latter tools constitutes a strong rationale in favor of such a study.

It seems essential to expand the research on the use of 3D printers and UAVs in mass disorders within a sustainable urban development framework due to the lack of sufficient research. The objective pursued in this research is to examine the determinacy of the above instruments in mass disorders. For achieving stated purpose was set a task as follows: elucidating the potentially possible ways of their exploitation in riots. The author argues for the indispensable role of intelligent control, referring to activities allowing the assessment and interaction with the business community and accordingly tailoring production, thereby minimizing the risks that they will become high-tech means and implements of crime.

2 Research and Results

Taking into consideration the statistical data given in Figs. 1–2, made certain inferences, on the reasonability of the research regarding the possible influences of 3D printers and UAVs on the course of collective action development, especially in the civil disorders.

By 2020, an estimated global market of 3D printing products and services was worth approximately \$12.6 billion (Fig. 1). A compound annual growth rate of the industry estimated to grow about 25 percent between 2020 and 2026, where calculations as:

$$\textit{Compound annual growth rate} = \left(\frac{FV}{PV}\right)^{\left(\frac{1}{Years}\right)} - 1$$

Start or Present Value, PV: 12,6
Ending or Future Value, FV: 44,5
Years: 6
Growth Rate, (FV-PV)/PV: (44,5–12,6)/12,6 = 253,174

$$\textit{Compound annual growth rate} = \left(\frac{12,6}{44,5}\right)^{\left(\frac{1}{6}\right)} - 1$$

CAGR Using Formula: 25,317%

Such industrial advancement as well as the household use of 3D printers will lead to their exploitation by organizers or abettors of criminal offenses. Ergo, these trends in combination with the development of Computer-aided design (CAD) have a potential impact on the printing of weaponry, and/or objects suitable for use as weapons in riots.

Analogously, UAVs also present an opportunity of being used both for tactical purposes and in confronting law enforcement through hacking operations targeting police drones during mass disorders. The FAA reported a total of 1 436 500 registered drones (up to 55 lb in weight) in civilian use by 2020 (Fig. 2). Yet it is also essential to consider the latent state of usage of such technologies.

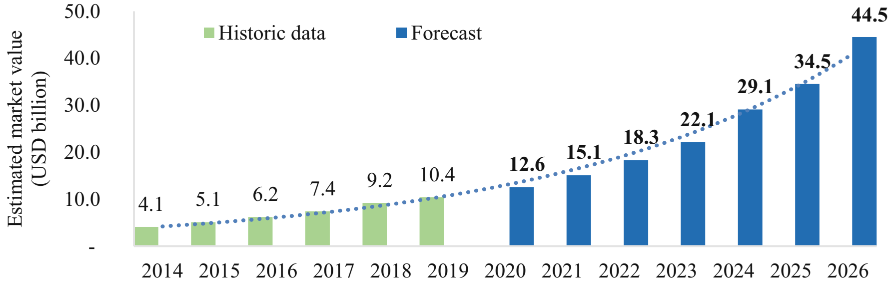


Fig. 1. Additive manufacturing trend report 2021: 3D printing market growth in the year of the COVID-19, Hubs A Protolabs Company, Source: https://www.hubs.com/get/trends/

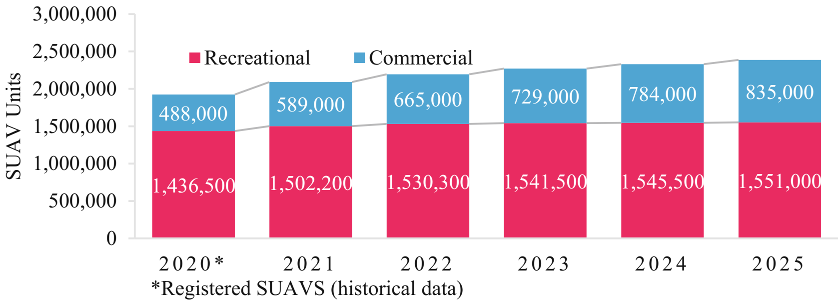


Fig. 2. FAA Aerospace Forecast Fiscal Years 2021–2041, U.S. Federal Aviation Administration, Source: https://www.faa.gov/data_research/aviation/aerospace_forecasts/

3 Discussion

Mob formation is a complex, rapid process, composed of both group and individual behavior. Meanwhile, the nature of such process can be spontaneous or, on the contrary, organized with its commencement, denouement, and culmination.

Violence is the dominant feature of the considered above phenomenon. Therefore, it is vital to emphasize a distinction between such offenses and peaceful assemblies, rallies, and marches, through which people collectively express their dissatisfaction with existing social ills. Joint participation of a considerable number of persons in commission of such grave crimes is inherent to this offence. Those eligible, alongside (co)perpetrators, may be organizers, instigators, or aiders and abettors.

3.1 3D Printers

High-tech advancements and their integration into diverse areas of public life can jeopardize public safety and public order. Indicative of this was the first “successful” firing of a plastic gun in 2013, developed via a computer-aided design (CAD) system that was printed on a 3D printer [9]. In many cases, malefactors are in possession of such technology. This obviously simplifies perpetration of certain criminal offenses. A remarkable example would be the use of 3D-printed weaponry in the commission of civil disorders.

Three-dimensional printing allows transforming a drawing created with a CAD software into a physical object [10]. Basically, 3D printer forms an object layer by layer from separate fragments of material [10]. Due to this feature, they can form objects with internal moving parts [10]. It excludes the necessity to manually merge the printed parts of the object, since the printed item is assembled. Meanwhile, the possibility of printing individual parts of the desired object is not precluded. Moreover, some 3D printers are capable of printing a significant number of their own parts, which practically allows them to self-replicate [10].

In May 2013, appeared a video displaying the shooting from a completely 3D-printed plastic firearm. Soon the digital design (drawing) of the weapon was posted online and downloaded more than 100,000 times until the website was blocked by U.S. authorities [10]. The culmination was the comprehension of the potential threat of manufacturing lethal firearms circumventing regulated gun production.

3D-printed shotguns are functionally operative weapons which often fail to be detected with standard metal detectors [11], because they are fabricated not from metallic materials (e.g., plastics). It is also untraceable due to the missing serial numbers. Anyone willing with an affordability to buy a 3D-printer and access to drawing files will be able to print such a weapon, own it or transfer it to third parties - even those who have no legal right to possess firearms like convicted felons, mentally unstable people, etc. As source materials for 3D-printing can be used metal, ceramics, wood, plastic, molecules and even living cells [12].

Pursuant to Title 18 U.S.C. § 232 (Definitions), a firearm is understood to be any weapon which is designed or intended to be fired with the energy of an explosive. Explosive or incendiary device" means (A) dynamite and other forms of explosives; (B) a bomb, grenade, rocket, or similar device; (C) any incendiary ammunition, bottle, or similar device whose composition (i) includes a destructive container filled with a flammable liquid or combustible compound, a wick which is capable of being ignited and (ii) is capable of being carried or ejected by a person acting alone.

The availability of 3D printers to a large part of the population, including criminals, is not a far-off prospect and is no longer solely a prediction of futurologists. It is an actual, predictable fact based on available data. Despite the imperfection and costliness of 3D-printed weapons, it can be reasonably predicted that in time the above drawbacks will be eliminated. All this makes it urgent to develop a theoretical, normative and ideological basis for the use of such tools in routine life.

3.2 Unmanned Aerial Vehicle (UAV)

The exploitation of unmanned aerial vehicles (UAVs) during civil disorders manifests itself in two ways: directly during the riots and in obstructing the activities of law enforcement agencies involved in dispersing violent crowds.

The drones are equipped with a remote-control system and range from small models for private use up to commercial and military aircraft [13].

As a matter of fact, organizers of mass disorders may use drones in tactical purposes. This device can be used to locate law enforcement officers (riot squads), re-routing the mob to a different site for acts of rioting. They are crucial in terms of coordination and distribution of resources in such social activism.

Eventually, drones might be able to carry weapons during unrest. Indeed, video and photography are not the only capacities of UAVs [14]. They can transport "cargo" and deploy it in the right place at the right time. Besides, they are also notoriously challenging to detect. Section 231(a)(2) of the U.S.C. states the criminalization of the manufacture for the purpose of transporting or carrying firearms, explosives, or incendiary devices for their unlawful use in furtherance of a civil disturbance.

Noteworthy applications of drones by law enforcement agencies for surveillance of territories and infrastructure; border and highway patrols; detection, identification, and surveillance of persons; search for missing persons; investigation of crimes; search and rescue operations, etc. [15, 16]. This practice was observed during riots in the US and England. Most drones can be hacked, considering the specifics of C + or C + + programming [17]. By hacking drones, it is possible to block the ability to properly respond to the operator's requirements [17]. As well, the control of the UAV can be established through hacking.

In order to minimize the illegal use of UAVs in criminal activities during civil unrest, the author sets forth his following assumptions and reflections:

1. Need for classification of UAVs, by designation (military, commercial, private, etc.); by weight (micro *(<5 kg)*, light *(5–50 kg)*, medium *(50–200 kg)*, heavy *(200–2000 kg)*, superheavy *(>2000 kg)*) by intensity and range (low *(<5 h, < 100 km)*, medium *(5–24 h, 100–400 km)*, high *(>24 h, > 1500 km)*) by maximum flight altitude (low *(<1000 m)*, medium *(1000–10000 m)*, high *(>10000 m)*), etc. [18].
2. Establishment of registers of unmanned aerial vehicles, based on the compiled classification. The foundation of such registers may be on the premises of the authorized state body in the field of civil aviation and air transport. In doing so, the registry may not be in the public domain.
3. Implementation of mandatory registration procedures for drones. It will enable tracking of drones in civil circulation and their owners. The registration procedure consists systematic record of certain information about the device, the operator, and the owner, followed with the assignment of an identification number for marking the drone.
4. Licensing of pilots/users of drones serves an additional milestone in the safe use of the state's airspace. Attaching the license of drone operators to the registry will make it possible to identify the person who violated the established rules and (or) used the specified devices for illegal purposes, in a mass disturbance.
5. Compulsory technical requirement to install GPS at the manufacturing stage of drones can be set. Despite the impossibility of determining the altitude and actual speed of flight with this system, nevertheless, the movement / relocation of them can be useful at the stage of investigation of specific crimes. The speed and altitude of the UAV can be determined using a complex radar system.
6. Flight data recorder systems for drones can also be handy in identifying the operator of the vehicle. It is a system of on-board recorders, an electronic device placed on board and used as the primary method of collecting flight data from the vehicle's sensors [19].
7. Conceptualizing drone ownership would be effective in establishing liability.

4 Conclusions

Summing up, the attributes of the Industry 4.0 (3D-printers), and aviation industry (UAVs) though auxiliary and peripheral, are not a direct determinant of civil disorders. Policymakers are unlikely intent on a complete ban of these technologies, considering their general purpose with many other socially beneficial applications [20]. Stepping aside from the techno-determinist model, the author leans toward the views of Professor C. Fuchs, who considers exclusively social relations as triggers of a conflict atmosphere in society.

Still, each of them remains available to be used by the organizers or active participants of mass disturbances. One way or another, their utilization areas range from informational interaction, engagement of individuals to mob violence, organization, allocation of roles, coordination, illegal printing weapons and explosive devices in a 3D printer, and their transportation using drones to target locations. All of these pose an imminent threat to public order, public safety, and to sustainable urban development. Indeed, the potential harm, damage as a result of riots can vary from the destruction of urban housing, commercial facilities, transportation infrastructure, historical and cultural monuments, to increasing crime rates, becoming a global and local unattractive city for investors, tourists and residents.

Comprehensive study in the perspective of foreign experience is presumed to be more than justified, relevant in terms of establishing the penal and criminological characteristics of disorders to ensure public safety generally, and urban security in particular. The minimum necessary measures to minimize the exploitation of drones in riots have been proposed supra. Follow-up in-depth examinations to prevent the use of 3D printers and ICTs as a security threat (especially the intermediary liability of Internet service providers as access providers, hosting providers, caching service providers, backbone providers, and last mile providers) would be meritorious.

References

1. Transforming our world: the 2030 Agenda for Sustainable Development, United Nations General Assembly Resolution A/RES/70/1 from 25 September 2015 https://sdgs.un.org/2030agenda, (Accessed 5 Feb 2023)
2. Schwab, K.: The Fourth Industrial Revolution. World Economic Forum, Cologny/Geneva, Switzerland, pp. 18–19 (2016)
3. Little, A.T.: Communication technology and protest. J. Politics **78**(1), 152–166 (2016)
4. Ananyev, M., Xefteris, D., Zudenkova, G., Petrova, M.: Information and communication technologies, protests, and censorship. Protests, and Censorship (2019). /https://doi.org/10.2139/ssrn.2978549, (Accessed 5 Feb 2023)
5. Soares, C.D.M., Joia, L.A.: The influence of social media on social movements: an exploratory conceptual model. In: Tambouris, E., et al. (eds.) ePart. LNCS, vol. 9249, pp. 27–38. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22500-5_3
6. Grigoryev, V.N.: Information technology in the investigation of Riots. Vestnik Saint Petersburg Univ. Law **12**(2), 334–355 (2021)
7. Christensen, D., Garfias, F.: Can you hear me now? How communication technology affects protest and repression. Q. J. Political Sci. **13**(1), 89–117 (2018)
8. Fuchs, C.: Social media, riots, and revolutions. Cap. Class **36**(3), 383–391 (2012)

9. Daly, A., Mann, M., Squires, P., Walters, R.: 3D printing, policing and crime. Polic. Soc. **31**(1), 40 (2021)
10. Van den Berg, B., Van der Hof, S., Kosta, E. (eds.).: 3D printing: Legal, philosophical and economic dimensions. Information Technology and Law Series, vol. 26. Springer. IX, T.M.C. Asser Press, The Hague (2015)
11. Chase, R.J., LaPorte, G.: The next generation of crime tools and challenges: 3D printing. Nat. Instit. Justice J. **279**, 6 (2018)
12. Polubinskaya, S.V., Klebanov, L.R.: Criminal threats of 3D-printing (dark side of hi-tech). Vestnik Tomskogo Gosudarstvennogo Universiteta-Pravo (34), 113 (2019)
13. Sehrawat, V.: Legal status of drones under LOAC and international law. Penn State J. Law Internat. Affairs **5**(1), 171–172 (2017)
14. Hassanalian, M., Abdelkefi, A.: Classifications, applications, and design challenges of drones: A review. Prog. Aerosp. Sci. **91**, 111 (2017)
15. Hodgkinson, D., Johnston, R.: Aviation law and drones: Unmanned aircraft and the future of aviation. Routledge, London (2018)
16. Tumbarska, A.: Remotely controlled non-lethal weapon systems in the context of law enforcement. Sec. Future **2**(3), 103 (2018)
17. Mohan, M.: Cybersecurity in drones (Order No. 10109635). Ph.D. dissertation, Fac. Utica College, Utica, NY, USA. Available from ProQuest Dissertations & Theses Global. (1796055579), p. 26 (2016)
18. Arjomandi, M., Agostino, S., Mammone, M., Nelson, M., Zhou, T.: Classification of unmanned aerial vehicles. Report for Mechanical Engineering class, University of Adelaide, Adelaide, Australia, p. 8 (2006)
19. Huang, S., Wang, Z., Zhang, C., Fang, Y.: Exploiting aerial heterogeneous network for implementing wireless flight recorder. In: Chellappan, S., Cheng, W., Li, W. (eds.) WASA 2018. LNCS, vol. 10874, pp. 189–200. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-94268-1_16
20. Thierer, A., Marcus, A.: Guns, limbs, and toys: What future for 3D printing. Minnesota J. Law, Sci. Technol. **17**, 853 (2016)

Author Index

C. Kahraman et al. (Eds.): INFUS 2023, LNNS 759, pp. 789–791, 2023.
https://doi.org/10.1007/978-3-031-39777-6