231654

OFFICE OF THE PRESIDENT
RECEIVED
Date: JUL 18 2023
Time: 6:20 pm
By:

SLSU OFFICE OF RESEARCH SERVICES

Office of
Research Services

18 July 2023

Dr. DORACIE B. ZOLETA-NANTES
University President

Thru: **Dr. MARISSA C. ESPERAL**
Vice President for Research, Extension, Production,
Development and Innovation

Dear Mesdames,

The CHED-DBM Joint Circular 3, s. 2023 (commonly known as the new instrument for Faculty Re-Classification) requires that Peer Reviewer engagement of faculty members in academic journals receive proper authorization from the President or the concerned Vice President. However, these guidelines were issued towards the end of the coverage period of the 1st Cycle of the Joint Circular (July 1, 2019–July 31, 2023).

As you may be aware, peer review requests from academic journals are normally directly communicated by editors to the peer reviewer and not through the institution where s/he may be affiliated. In consultation with the Institutional Evaluation Committee, I was informed that the CHED provides leeway for additional evidence for Peer Reviewer engagement – that a list of institutionally-recognized peer reviewer engagement would be enough as additional evidence for this cycle.

In this regard, I wish to respectfully seek your **approval in principle** of the participation of faculty members listed in the attached file. Rest assured that the ORS thoroughly screened these reported Peer Reviewer engagement of our faculty members to include only those done with reputable journal publications and book publishers.

We look forward to your usual support on this matter as this will contribute greatly to the career development of our dedicated faculty researchers.

Thank you very much!

Very truly yours,

NICANOR L. GUINTO, PhD
Director, Office of Research Services

Recommending Approval:

MARISSA C. ESPERAL, PhD
Vice President for Research, Extension,
Production, Development and Innovation

APPROVED / DISAPPROVED

Doracie B. Zoleta-Nantes, PhD
University President

JUL 19 2023

SOUTHERN LUZON STATE UNIVERSITY
Office of Research Services

C E R T I F I C A T I O N

This is to certify that the **peer reviewer engagement** of the personnel named below are approved in principle as they have been invited to review journal articles and/or book proposals while being affiliated with the University. For having been directly contacted by Editors of reputable journals and book publishers, their recognized expertise and leadership in their respective areas of research specialization contributed significantly to building the good name of Southern Luzon State University in local and international academic circles.

Name	Academic Rank	College/ Campus	Area of Research Specialization	Journal Name/ Book Publisher that made the request	Coverage/ Readership	Indexed/ Published by:	Tentative Title of the Article/ Book Proposal reviewed	Date when the invitation is received:	Date when the review was sent back to the editor:
AGUDILLA, MARY ANN R.	ASSOCIATE PROFESSOR 4	CAG	BIODIVERSITY, INSECTS, ECOSYSTEM VALUATION	PHILIPPINE JOURNAL OF SCIENCE	International	Scopus	SETTING THE INITIAL CARBON TAX RATE FOR THE CARBON TAX POLICY IN THE PHILIPPINES THROUGH THE SOCIAL COSTS OF CARBON AND WILLINGNESS TO PAY METHODS, AND THE CORRESPONDING BENEFIT-COST ANALYSIS	12/11/2022	1/2/2023
AGUDILLA, MARY ANN R.	ASSOCIATE PROFESSOR 4	CAG	BIODIVERSITY, INSECTS, ECOSYSTEM VALUATION	ACADEMIA-BIOLOGY	International	Academia Publishing	TREE HEIGHT, CANOPY COVER AND LEAF LITTER PRODUCTION OF RHIZOPHORA APICULATA IN BAGANGA, DAVAO, ORIENTAL, PHILIPPINES	1/11/2023	1/27/2023
Alinea, Jess Mark L.	Assistant Professor I	Lucena Campus	TVET, Technical Teacher Education, Curriculum and Instruction	Journal of Technical Education and Training	International	Scopus	The Role of Al-Balqa Applied University in Developing Vocational Education in Jordan	10/26/2021	11/2/2021
Alinea, Jess Mark	Assistant Professor I	Lucena Campus	TVET, Technical Teacher Education,	Journal of Technical	International	Scopus	Training-based Assessment of Employees Performance: A Case Study of Bahir Dar	12/27/2021	1/5/2022

			Communication	Applied Linguistics					
Guinto, Nicanor L.	Associate Professor III	College of Arts and Sciences	Sociolinguistics, Discourse Analysis, Communication	rEFLections	International	Scopus/ King Mongkut's University, Thailand	Filipino Non-Native English-Speaking Teachers and the Bias in Their Own Backyard	07/10/2023	07/19/2023
Maaliw, Renato III R.	Associate Professor II	CEN	Computer Vision, Machine Learning, Data Analytics	Cogent Engineering	International	Scopus, Web of Science, ASEAN Citation Index	Integrating Video Feedback Into Architectural Design Education to Engage Diverse Learning Styles	3/27/2023	4/20/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Machine Learning, Computer Vision, Data Analytics	Healthcare Analytics (Elsevier)	International	Scopus, Web of Science, ASEAN Citation Index	Prediction of Systolic and Diastolic Blood Pressures Using Machine Learning	5/4/2023	5/16/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics	Engineering (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Using ARIMA to Predict the Growth in the Subscriber Data Usage	11/4/2022	11/14/2022
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Analytics	Sensors (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Missing Traffic Data Imputation with a Linear Model Based on Probabilistic Principal Component Analysis	12/2/2022	12/10/2022
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Sensors (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Using Machine Learning on V2X Communications Data for VRU's Collisions Predictions	12/23/2022	12/26/2022
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics	Applied Science (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Performance Predictions of Sci-Fi Films via Machine Learning	1/31/2023	2/5/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Sustainability (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Thermal Images Classifications of Solid Wastes with Deep Convolutional Neural Networks	2/15/2023	2/25/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Sustainability (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	Static Evaluation of a Midimew Connected Torus Network for Next Generation Supercomputers	3/2/2023	3/13/2023
Maaliw,	Associate	College of	Computer Vision,	Journal of	International	Scopus, Web of	Machine-Learning-Based Composition	3/23/2023	4/1/2023

Renato III R.	Professor II	Engineering	Machine Learning, Data Analytics, Computer Engineering	Nuclear Engineering (MDPI)		Science, ASEAN Citation Index	Analysis of the Stability of V–Cr–Ti Alloys		
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Mathematics (MDPI)	International	Scopus, Web of Science, ASEAN Citation Index	A Federated Personal Mobility Service in Autonomous Transportation	5/19/2023	5/29/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	IJERPH (MDPI)	International	Scopus, Web of Science,	Machine Learning in Predicting Severe Acute Respiratory Infection	6/6/2023	6/11/2023
Maaliw, Renato III R.	Associate Professor II	College of Engineering	Computer Vision, Machine Learning, Data Analytics, Computer Engineering	Journal of Theoretical and Applied Electronic Commerce Research	International	Scopus, Web of Science, ASEAN Citation Index	Unveiling the Power of ARIMA, Support Vector Machine and Random Forest Regressors for the Future of Dutch Employment Market	6/14/2023	6/23/2023
Mabunga, Zoren P.	Instructor 1	College of Engineering	Artificial Intelligence, Electronics and Communication Engineering, Internet of Things	2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA 2022)	International	Scopus	Semi Autonomous Detection of Bite Points for a Surgical Needle	2/24/2022	3/7/2022
Mabunga, Zoren P.	Instructor 1	Engineering	Artificial Intelligence, Electronics and Communication Engineering, Internet of Things	IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC-2021)	International	Scopus	1. A Survey of Vulnerability Management Using Machine Learning Techniques, 2. An Adaptive Algorithm based on Interference Aware Cooperative Energy Efficiency Maximization for 5G UltraDense Networks, 3. GRAMIN GENIE-A SMART KIOSK, 4. An Automated Deep Learning Model for Detecting Sarcastic Comments,	7/2/2021	8/12/2021

YAO, CLAIRE ANN M.	ASSISTANT PROFESSOR IV	CABHA MAIN	BUSINESS ENTREPRENEURSHIP, PRODUCT DEVELOPMENT, TOURISM, LEISURE, AND HOSPITALITY	PATHWAY TO REFEREED JOURNAL PUBLICATION IN THE FIELD OF BUSINESS	Local	INSTITUTIONAL	PROBLEMS ENCOUNTERED BY MSME'S IN TAGUIG CITY AND THE ACTION TO COUNTER THE POSSIBLE EFFECTS OF ASEAN INTEGRATION: A SITUATION ANALYSIS	3/24/2020	4/4/2020
YAO, CLAIRE ANN M.	ASSISTANT PROFESSOR IV	CABHA MAIN	BUSINESS ENTREPRENEURSHIP, PRODUCT DEVELOPMENT, TOURISM, LEISURE, AND HOSPITALITY	PATHWAYS TO REFEREED JOURNAL IN THE FIELD OF BUSINESS	Local	INSTITUTIONAL	MANYAMAN MANGAN QUENI (DELICIOUS TO EAT HERE):SUCCESS FACTORS OF SELECTED RESTAURANT ENTREPRENEURS IN PAMPANGA	4/16/2020	4/21/2020

Issued this 19th day of July 2023 at Southern Luzon State University, Lucban, Quezon.

NICANOR L. GUINTO, Ph.D.
Director, Office of Research Services

MARISSA C. ESPERAL, Ph.D.
Vice President, REPDI

DORACIE B. ZOLETA-NANTES, Ph.D.
University President

RM

Renato Racelis Maaliw

(Maaliw III, Renato R. R.)

Southern Luzon State University

Web of Science ResearcherID: EXW-3524-2022

Published names	Maaliw, Renato R., III Maaliw, Renato R. Maaliw, Renato Racelis, III Maaliw Iii, R. R. Maaliw Iii, Renato R. R.
Published Organizations	Southern Luzon State University, Southern Luzon Univ
Subject Categories	Computer Science; Telecommunications; Engineering; Environmental Sciences & Ecology; Materials Science

Documents | Peer Review

Verified peer reviews

4	Sensors
3	Sustainability
2	Applied Sciences
2	Eng
2	Journal of Theoretical and Applied Electronic Commerce Research
1	Cogent Engineering
1	International Journal of Environmental Research and Public Health
1	International Journal of Retail and Distribution Management
1	Journal of Nuclear Engineering
1	Mathematics

Metrics

Open dashboard

Profile summary

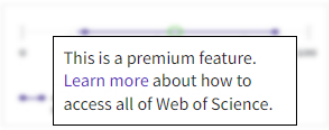
20	Total documents
15	Web of Science Core Collection publications
0	Preprints
18	Verified peer reviews
0	Verified editor records

Web of Science Core Collection metrics

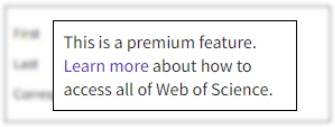
3 H-Index	**15** Publications in Web of Science
17 Sum of Times Cited	**10** Citing Articles
0 Sum of Times Cited by Patents	**0** Citing Patents

View citation report

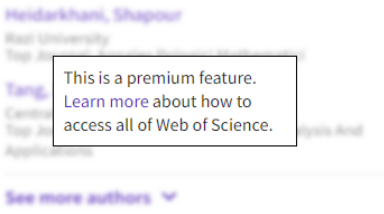
Author Impact Beamplot Summary



Author Position



You might be interested in...



See more authors

Co-authors

[Applied Sciences] Manuscript ID: applsci-2166653 - Review Request

External | Research Reviews (Journals) ×

Applied Sciences Editorial Office <applsci@mdpi.com>
to me, Applied, Robin

Tue, Jan 31, 3:38 PM (3 days ago)

Dear Dr. Maaliw,

We have received the following paper, submitted to Applied Sciences (https://www.mdpi.com/journal/applsci/).

Type of manuscript: Article
Title: Performance Predictions of Sci-Fi Films via Machine Learning
Special Issue: Advanced Computing and Neural Networks Applied in Learning Systems
https://www.mdpi.com/journal/applsci/special_issues/L658L07IMU

We kindly invite you to review this paper and evaluate its suitability for publication in Applied Sciences. The article abstract is available at the end of this message.

If you choose to accept this invitation, we would appreciate receiving your comments within 1 week. Please let us know if you are likely to need more time to complete your review.

Please click on the link below to let us know if you will be able to provide a review and access the full manuscript and review report form.

https://susy.mdpi.com/user/review/review/35943130/mVh4Z13I

In recognition of the contribution of reviewers, for thorough and timely review reports we provide discount vouchers for Article Processing Charges (APCs) applicable for manuscripts accepted for publication after peer review in any MDPI journal. Advice for completing your review can be found at: https://www.mdpi.com/reviewers.

Please disclose any potential conflicts of interest you might have concerning the manuscript's contents or the authors.

If you are not able to review this manuscript, we kindly ask you to decline by clicking on the above link such that we can continue processing this submission. We would also appreciate any feedback you can provide at this time (i.e., your general impression regarding the quality of this manuscript) and any suggestions for alternative expert reviewers.

Applied Sciences is an open access journal of MDPI. Thank you very much for your consideration and we look forward to hearing from you.

Kind regards,
Mr. Robin Hou
E-Mail: robin.hou@mdpi.com

--
MDPI Office

MDPI Applied Sciences Editorial Office
St. Alban-Anlage 66, 4052 Basel, Switzerland
E-Mail: applsci@mdpi.com
http://www.mdpi.com/journal/applsci

Manuscript details:
Journal: Applied Sciences
Manuscript ID: applsci-2166653
Type of manuscript: Article
Title: Performance Predictions of Sci-Fi Films via Machine Learning
Authors: Amjed Subhi Hassan Al Fahoum *, Tahani Ghobon
Submitted to section: Computing and Artificial Intelligence,
https://www.mdpi.com/journal/applsci/sections/computing_artificial_intelligence

Special Issue: Advanced Computing and Neural Networks Applied in Learning Systems
https://www.mdpi.com/journal/applsci/special_issues/L658L07IMU

Abstract: The actions of adolescents are greatly influenced by the films they

watch. After viewing a film starring an actor with a particular social habit or personality feature, viewers, especially young ones, may attempt to emulate the actor's conduct. This study presents an algorithm-based method for predicting the market potential of forthcoming science fiction films. Each year sees the release of several new science fiction films, and working in the film industry is both profitable and rewarding. Prior to the film's release, it is essential to undertake research and make informed forecasts regarding its success. In this analysis, multiple MATLAB-implemented machine-learning algorithms are investigated to classify and predict the financial success of movies. Utilizing fourteen machine learning algorithms, it was possible to forecast how individuals would vote on science fiction films. The fine, medium, and weighted KNN algorithms were given additional weight due to their superior performance. The results demonstrated that the KNN-adopted algorithms provide more precision (0.89–0.93), recall (0.88–0.92), accuracy (90.1–93.0%), a fast execution rate, and more robust estimations, as well as a shorter execution time compared to previous research. The tabulated results of this investigation demonstrate that the weighted KNN algorithm is efficient and reliable. The film industry and its global expansion can benefit from accurate and consistent prediction outcomes if many KNN algorithms targeting specific viewer behavior are logically coupled. This study demonstrates how future-focused data analytics may benefit the film business. Studying a film's box office performance and critical reception makes it feasible to construct a model that forecasts its success, effect, and social behavior.

Keywords: Movie Success; Artificial Intelligence; Machine learning; KNN Algorithms

1 of 2,877

[Applied Sciences] Manuscript ID: applsci-2166653- Review Request Confirmation

External Inbox ×

robin.hou@mdpi.com <robin.hou@mdpi.com> 2:28 PM (2 hours ago)
to me, applsci

Dear Dr.Maaliw,

Thank you very much for agreeing to review the manuscript:applsci-2166653

We have just received enough reports from other referees and would like to send them back to the authors to make any necessary revisions. We would like to know if you have started the review report. If yes, could you send us the current comments to us ? If no, would you mind if we cancel our invitation this time? My apologies for any inconvenience this may cause you.

Best regards,

Robin Hou
Assistant Editor

MDPI
Applied Sciences Editorial Office
St. Alban-Anlage 66, 4052 Basel, Switzerland
E-mail: applsci@mdpi.com
http://www.mdpi.com/journal/applsci/

* Applied Sciences 2021 CiteScore - 3.7 (Q2 in Engineering)
* Applied Sciences 2021 Impact Factor - 2.838 (Q2 in "Physics Applied" and "Engineering, Multidisciplinary")

Call for Contributions:
http://susy.mdpi.com/user/manuscripts/upload/applsci

Special Issue Proposal Application:
http://www.mdpi.com/journalproposal/sendproposalspecialissue/applsci

Special Issue Mentor Program - an opportunity for early career scientists to work closely with journal and gain more editorial experience:
https://www.mdpi.com/about/announcements/3896

Applied Sciences 2022 Early Career Investigator Award (1000 CHF) - Nomination deadline 28 February 2023:
https://www.mdpi.com/journal/applsci/awards

LinkedIn: https://www.linkedin.com/company/68370265/admin/
Twitter: https://twitter.com/Applsci

Reply Reply all Forward

[Applied Sciences] Manuscript ID: applsci-2166653 - Acknowledgement - Review Received

External Inbox ×

applsci@mdpi.com
to me, Applied, Robin

4:34 PM (4 minutes ago)

Dear Dr. Maaliw,

Thank you for submitting your review of the following manuscript:

Manuscript ID: applsci-2166653
Title: Performance Predictions of Sci-Fi Films via Machine Learning
Authors: Amjed Subhi Hassan Al Fahoum *, Tahani Ghobon

Our Editorial Office and Academic Editors will contact you if they have any questions about your review report. We ask that you remain available, as far as possible, during the peer-review process in case of follow-up questions. To help us improve our services, we kindly ask you to fill in our online survey on the peer-review process at
https://www.surveymonkey.com/r/reviewerfeedbackmdpi

We encourage you to register an account on our submission system and bind your ORCID account (https://susy.mdpi.com/user/edit). You are able to deposit the review activity to your ORCID account manually via the below link:
https://susy.mdpi.com/user/reviewer/status/finished

We also invite you to contribute to Encyclopedia (https://encyclopedia.pub), a scholarly platform providing accurate information about the latest research results. You can adapt parts of your paper to provide valuable reference information for others in the field.

Kind regards,
Applied Sciences Editorial Office
Postfach, CH-4020 Basel, Switzerland
Office: St. Alban-Anlage 66, CH-4052 Basel
Tel. +41 61 683 77 34 (office)
E-mail: applsci@mdpi.com
https://www.mdpi.com/journal/applsci/

*** This is an automatically generated email ***

Review Report Form

Journal: Applied Sciences (ISSN 2076-3417)
Manuscript ID: applsci-2166653
Type: Article

Title: Performance Predictions of Sci-Fi Films via Machine Learning

Authors: Amjed Subhi Hassan Al Fahoum * , Tahani Ghobon

Section: Computing and Artificial Intelligence
Special Issue: Advanced Computing and Neural Networks Applied in Learning Systems

Abstract: The actions of adolescents are greatly influenced by the films they watch. After viewing a film starring an actor with a particular social habit or personality feature, viewers, especially young ones, may attempt to emulate the actor's conduct. This study presents an algorithm-based method for predicting the market potential of forthcoming science fiction films. Each year sees the release of several new science fiction films, and working in the film industry is both profitable and rewarding. Prior to the film's release, it is essential to undertake research and make informed forecasts regarding its success. In this analysis, multiple MATLAB-implemented machine-learning algorithms are investigated to classify and predict the financial success of movies. Utilizing fourteen machine learning algorithms, it was possible to forecast how individuals would vote on science fiction films. The fine, medium, and weighted KNN algorithms were given additional weight due to their superior performance. The results demonstrated that the KNN-adopted algorithms provide more precision (0.89–0.93), recall (0.88–0.92), accuracy (90.1–93.0%), a fast execution rate, and more robust estimations, as well as a shorter execution time compared to previous research. The tabulated results of this investigation demonstrate that the weighted KNN algorithm is efficient and reliable. The film industry and its global expansion can benefit from accurate and consistent prediction outcomes if many KNN algorithms targeting specific viewer behavior are logically coupled. This study demonstrates how future-focused data analytics may benefit the film business. Studying a film's box office performance and critical reception makes it feasible to construct a model that forecasts its success, effect, and social behavior.

Thank you for contributing to the review process, your comments have been successfully submitted.

- see your review history
- download a letter confirming your review activity

Other reviewers' comments

Reviewer 1	Review Report (round1) (Reconsider after major revision (control missing in some experiments))
Reviewer 2	Review Report (round1) (Reject (article has serious flaws, additional experiments needed, research not conducted correctly))
Reviewer 3	Review Report (round1) (Reconsider after major revision (control missing in some experiments))
Reviewer 4	Review Report (round1) (Reconsider after major revision (control missing in some experiments))

Review Report Form

Reviewer's Information (will not be revealed to authors)

Name: Dr. Renato Maaliw
Email: rmaaliw@slsu.edu.ph
Website: https://renatomaaliw3.github.io/
Affiliation: Southern Luzon State University, Lucban, Quezon, Philippines

Research Keywords: data; machine learning; computer vision; deep learning

Report 1 Hide Report and Author Response [-]

High Average Low No Overall Recommendation

	High	Average	Low	No Answer
Originality / Novelty	( )	( )	(x)	( )
Significance of Content	( )	(x)	( )	( )
Quality of Presentation	( )	(x)	( )	( )
Scientific Soundness	( )	(x)	( )	( )
Interest to the readers	( )	(x)	( )	( )
Overall Merit	( )	(x)	( )	( )

Overall Recommendation

( ) Accept in present form
( ) Accept after minor revision (corrections to minor methodological errors and text editing)
(x) Reconsider after major revision (control missing in some experiments)
( ) Reject (article has serious flaws, additional experiments needed, research not conducted correctly)

English language and style

( ) English very difficult to understand/incomprehensible
( ) Extensive editing of English language and style required
(x) Moderate English changes required
( ) English language and style are fine/minor spell check required
( ) I don't feel qualified to judge about the English language and style

	Yes	Can be improved	Must be improved	Not applicable
Does the introduction provide sufficient background and include all relevant references?	( )	( )	(x)	( )
Are all the cited references relevant to the research?	( )	(x)	( )	( )
Is the research design appropriate?	( )	( )	(x)	( )
Are the methods adequately described?	( )	(x)	( )	( )
Are the results clearly presented?	( )	( )	(x)	( )
Are the conclusions supported by the results?	( )	(x)	( )	( )

Comments and Suggestions for Authors

1. What is the optimization process or methods used in the research?

2. What is the novelty of this work? Explain in detail based from the gaps found in the literature

3. The feature selection technique is not specified for data reduction

4. When comparing data, it will be much better to tell the story into graphs than tables (Table 7). It is hard to grasp the information in a tabular form

5. What are the lessons for the readers or researchers of this study? What new information can be extracted from the knowledge gained? How is this study useful? State clearly in the conclusions.

6. Fix grammar and structure of sentences.

7. Why do you think the Fine, Weighted, and Medium KNN types outperform other machine learning? What properties does it possess that proved its advantages?

8. Is there a data scaling methods used to mitigate the possibility of under or overfitting?

Less...

	Yes	No
Do you have any potential conflict of interest with regards to this paper?	( )	(x)
Did you detect plagiarism?	( )	(x)
Did you detect inappropriate self-citations by authors?	( )	(x)
Do you have any other ethical concerns about this study?	( )	(x)

REVIEW CONFIRMATION CERTIFICATE

We are pleased to confirm that

Renato Maaliw

has reviewed 9 papers for the following MDPI journals in 2023:

Mathematics, Journal of Theoretical and Applied Electronic Commerce Research, International Journal of Environmental Research and Public Health, Journal of Nuclear Engineering, Sustainability, Applied Sciences, Sensors

Shu-Kun Lin

Dr. Shu-Kun Lin, Publisher and President
Basel, 30 June 2023

MDPI

MDPI is a publisher of open access, international, academic journals. We rely on active researchers, highly qualified in their field to provide review reports and support the editorial process. The criteria for selection of reviewers include: holding a doctoral degree or having an equivalent amount of research experience; a national or international reputation in the relevant field; and having made a significant contribution to the field, evidenced by peer-reviewed publications.

applied sciences

MDPI

Article

Performance Predictions of Sci-Fi Films via Machine Learning

Amjed Al Fahoum1* and Tahani A. Ghobon 2

1Biomedical Systems and Informatics Engineering Dept., Hijjawi Faculty for Engineering Technology, Yarmouk University, Irbid, 21163, Jordan

2School of Engineering Technology, Al Hussein Technical University, Jordan

* Correspondence: Author; Amjed Al Fahoum: afahoum@yu.edu.jo

Abstract: The actions of adolescents are greatly influenced by the films they watch. After viewing a film starring an actor with a particular social habit or personality feature; viewers; especially young ones; may attempt to emulate the actor's conduct. This study presents an algorithm-based method for predicting the market potential of forthcoming science fiction films. Each year sees the release of several new science fiction films; and working in the film industry is both profitable and rewarding. Prior to the film's release; it is essential to undertake research and make informed forecasts regarding its success. In this analysis; multiple MATLAB-implemented machine-learning algorithms are investigated to classify and predict the financial success of movies. Utilizing fourteen machine learning algorithms; it was possible to forecast how individuals would vote on science fiction films. The fine; medium; and weighted KNN algorithms were given additional weight due to their superior performance. The results demonstrated that the KNN-adopted algorithms provide more precision (0.89–0.93); recall (0.88–0.92); accuracy (90.1–93.0%), a fast execution rate; and more robust estimations; as well as a shorter execution time compared to previous research. The tabulated results of this investigation demonstrate that the weighted KNN algorithm is efficient and reliable. The film industry and its global expansion can benefit from accurate and consistent prediction outcomes if many KNN algorithms targeting specific viewer behavior are logically coupled. This study demonstrates how future-focused data analytics may benefit the film business. Studying a film's box office performance and critical reception makes it feasible to construct a model that forecasts its success; effect; and social behavior.

Keywords: Movie Success; Artificial Intelligence; Machine learning; KNN Algorithms

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Appl. Sci.* **2021**, *11*, x. https://doi.org/10.3390/xxxxx

Received: date
Accepted: date
Published: date

1. Introduction

As a significant industry in its own right, the film industry also plays a significant role in international trade and marketing. Movies are powerful weapons that can change people's minds and alter our lives. The movies that teens watch significantly impact their conduct, both positively and negatively, according to a study published in [1]. A study published in [2] found that people are more likely to smoke after watching a movie with a smoking character.

Some filming locations are fascinating to viewers and will stay that way for a long time, no matter the time of year or season. This factor is there because they have been in popular movies. As a result, movies are why vacation bookings are up [3]. Movies in the science fiction (sci-fi) genre get high marks from critics. There are several visually and technically impressive science fiction films. Award ceremonies often include competition from films in different categories, including science fiction. According to Table 1, SF movies can take several forms, including drama, romance, horror, thriller, adventure, and action. A significant impact of science fiction on culture and technology has been observed. The research found that showing middle school kids sci-fi movies helped them think of new ways to solve problems and better understand how

technology works [4]. More research supports the idea that science fiction films inspire people to think about what they could create.

Moreover, they will pique people's curiosity about robotics and AI [5]. One study shows that after watching the film "The Day After Tomorrow," whose plot revolves around the sudden transition of Earth's climate into a new ice age, viewers' perspectives are changed, at least temporarily. Everyone who saw the film understood their role in maintaining a healthy ecosystem and a consistent climate [6].

When deciding whether or not to make a film, a filmmaker must consider several factors, the most prominent of which is the film's potential box office take. The film's success can be measured by how well it does at the box office. Many things affect a movie's performance at the box office, such as its genre, writers, directors, actors, length, year of release, producers, and production company. The firms that sell, promote, and market the film are crucial to its overall success [7]. There are too many variables at play for anyone to accurately predict the health of the market or the fortunes of a particular film. Many scientific papers have used machine learning algorithms to predict how a movie will do at the box office, how popular it will be, and how it will be rated. In [8], the researchers suggested a methodology for predicting movie success that incorporates machine learning, social network analysis, and text mining approaches. Their model extracts many sorts of characteristics, such as "who," "what," "when," and "hybrid." They examined movie success from three perspectives: audience, release, and movie. Their data comes from IMDB and Box Office Mojo. In [9]. the researchers compared the performance and outcomes of three supervised learning approach models to forecast revenue: the linear regression model, the logistic regression model, and the support vector machine regression model. Their data comes from IMDB, Rotten Tomatoes, and Wikipedia. The researchers suggested machine learning strategies to predict movie popularity [10]. They used and compared the following classifiers: Logistic Regression, Simple Logistic, Multilayer Perceptron, J48, Nave Bayes, and PART. Their data source is IMDB, and they categorized the movies as Terrible, Poor, Average, and Excellent based on their IMDB rating, which is the same technique used by [11]. In [11], they offered a machine learning strategy based on intrinsic features employing C4.5, PART, and correlation coefficient to predict movie popularity categorization, whereas [12] proposed a data mining approach to assess and forecast movie ratings. Using data mining techniques like neural networks, regression, and decision trees, researchers devised a way to predict how much money a movie will make at the box office and estimate its profit [13]. Opus Data gathered the dataset used in this article. The dataset was then supplemented with information from IMDB, IDMB, Metacritic, and Mojo 2016a. They created a mathematical model [14] to predict movie ratings and success for Hollywood and Bollywood films. The films were categorized as Flop, Neutral, or Hit. For classification, the KNN method was utilized. The information was gathered from IMDB and social media networks such as Facebook, Instagram, and Twitter. In [15], the writers projected a film's profitability to aid early-stage production finance decisions. Using social network analysis and text mining techniques, the team presents a system that can automatically extract information on a movie's cast, narrative, and release date from several data sources. In comparison to other methods, their data confirmed the system's effectiveness. The feature selection strategy they employed considerably improved the forecast's accuracy. Through an analysis of the industry's most influential driving forces, they want to create a decision-making tool that may be utilized. Using a mathematical model, in [16], they could accurately anticipate the performance of upcoming films at the box office. Budget, cast, director, producer, set location, narrative writer, movie release date, competing movies released simultaneously, music, release location, and target audience are just a few factors considered when evaluating a film's success. They developed a model by evaluating how distinct traits interact with one another. Each important component was assigned a weight, and the prediction was developed based on them. In addition, they illustrated the approach's prescriptive potential by demonstrating how it may be used to select a set of income-maximizing ac-

tors. Instead of depending on the opinions of critics and others, researchers [17] provide a method for forecasting a film's performance at the box office before its release. This paper presents the approach for estimating an IMDb rating using the IMDb Movie Dataset. Several algorithms were included in the study's analysis, but Random forest generated the most accurate predictions. It was discovered that the number of individuals who voted for the picture, the number of critics who reviewed it, the number of Facebook likes it earned, the film's running length, and its overall box office profits all had a substantial impact on the film's IMDb rating. Dramas and biopics are often the highest-quality examples of their respective film genres. In [18], researchers employed regression techniques to develop a model that considers multiple factors, assigns a weight to each element, and predicts the success or failure of forthcoming films based on the factor's value. In [19], models were created to forecast the performance and ratings of a new film before its premiere. A revenue threshold was established based on heuristics to classify the film as successful or unsuccessful. The comments on teasers and trailers were collected from YouTube since they are quite helpful when rating a film. Natural Language Processing (NLP) was used to extract keywords from user evaluations, and those reviews were evaluated as positive or negative based on emotional analysis. A detailed comparison of the various machine learning models used to estimate the success rate of a movie was made in [20]. These models' accuracy and statistical significance were examined to determine which is the best predictor. There are also some observations about aspects that impact the success of movies. The analyzed models include regression models, machine learning models, a time series model, and a neural network, with the neural network exhibiting the highest accuracy at about 86%. In addition, as part of the testing, statistics regarding 2020's film releases are evaluated. The authors of [21] made sure that the success of casting a movie depends on many things, such as the type of movie, when it comes out, who is in it, and how much money it makes in total. Understanding the risks associated with a film's release, which might affect its success or failure, can be an important step in expanding the film industry. As a result, they offered an ensemble learning technique to assess such comprehension, in which predictions from previously guided attribute computations may be utilized to improve future success/failure accuracy. Diverse methods are employed in the literature to examine and compare the generated data. The dataset is used with the machine learning algorithms SVM, KNN, Naive Bayes, Boosting Ensemble Technique, Stacking Ensemble Technique, Voting Ensemble Technique, and MLP Neural Network to guess how well a movie will do at the box office. In [21], they used many algorithms and their trends to predict the outcome of a movie and showed that the suggested method is better than the current research when these algorithms and trends are used.

Table 1. Science Fiction Movies Genres.

Movie	Genres
About Time	Comedy, Drama, Fantasy, Romance, Sci-Fi
Alien	Horror, Sci-Fi
A Clockwork Orange	Crime, Drama, Sci-Fi
Boss Level	Action, Mystery, Thriller, Sci-Fi
Coherence	Drama, Horror, Mystery, Thriller, Sci-Fi
Dune	Action, Adventure, Sci-Fi

Inception	Action, Adventure, Thriller, Sci-Fi
Iron Sky	Action, Adventure, Comedy, Sci-Fi
Mac and Me	Adventure, Family, Fantasy, Sci-Fi
Maggie	Drama, Horror, Sci-Fi
Megaforce	Action, Sci-Fi
Okja	Action, Adventure, Drama, Sci-Fi
Only	Drama, Romance, Sci-Fi
The Day	Action, Drama, Horror, Sci-Fi, Thriller
The Martian	Adventure, Drama, Sci-Fi

2. Methodology

As depicted in Fig. 1, the suggested model in this work will undergo the following processes to increase its ability to forecast a film's success.

Extraction and reduction of data
Data processing and analysis
Algorithms for machine learning
Evaluate the performance of various algorithms

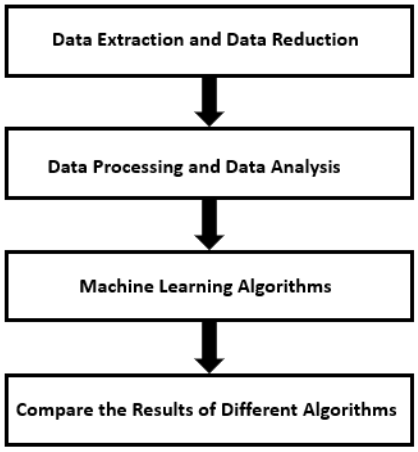


Figure 1: Methodology

2.1. Data Extraction and Data Reduction

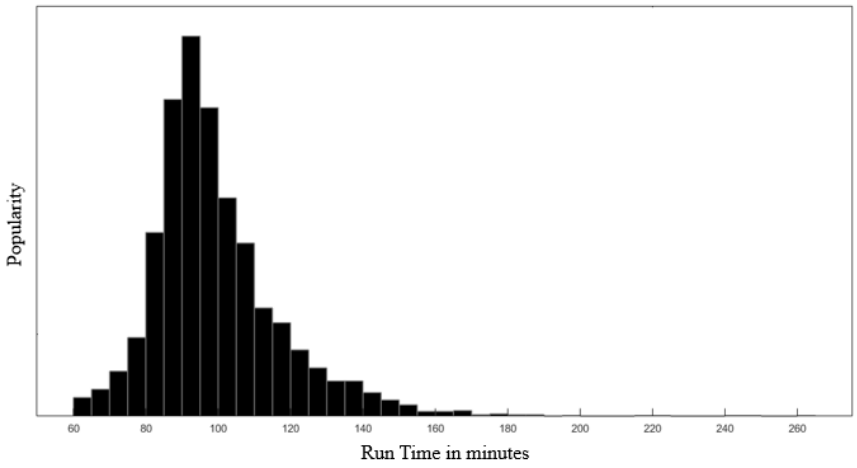
This work's data was sourced entirely from IMDB. IMDB is an online database that provides all data about movies, television, programs, home videos, video games, and streaming content, such as names and ratings of actors, directors, writers, plot summaries, production crew, movie ratings, box office revenues, trailers, and more. Every movie on IMDB has a rating from 1 to 10. After extracting the data, all the unwanted files, movies of other genres other than Sci-Fi movies, TV series, movies with a running time of less than 1 hour (60 minutes), and movies rated less than 1000 votes were excluded.

3.2. Data Processing and Data Analysis

For analysis and classification purposes, the ratings are classified into four classes: flop, below average, average, and hit, as represented in Table 2. The same approach used in [22] is used in this paper. Data analysis and the study of the relationship between the factors that affect the movie's success are critical. From the following four figures, we can extract meaningful data. Various plots were made from the data we extracted about sci-fi movies, and further data was extracted from these plots. The relationships between run time in minutes and popularity, IMDB rating and the number of votes, and IMDB rating and run time in minutes are shown in figures 2, 3, and 4, respectively. Moreover, figure 5 shows the box office by month.

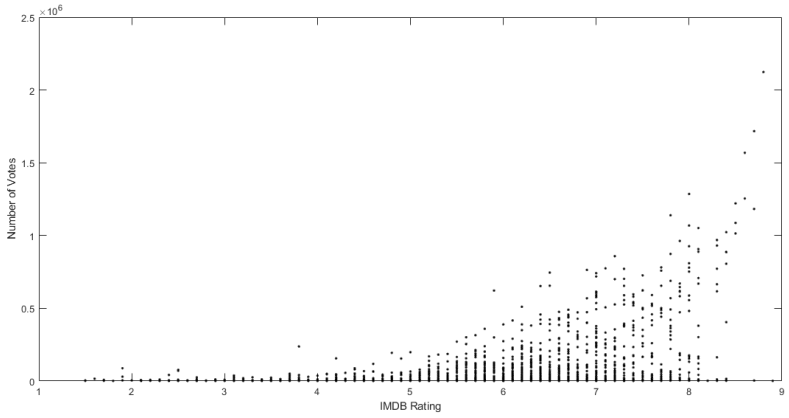
Table 2. Rating Classification.

Range of Rating	Class
0.0 – 3.5	Flop
3.6 – 5.8	Below Average
5.9 – 7.4	Average
7.5 – 10.0	Hit



.

Figure 2: The relationship between Run Time in Minutes and Popularity.



.

Figure 3: The relationship between the IMDB Rating and the Number of Votes.

The highest-grossing science fiction movies of all time are shown in table 3.

3.3. Machine Learning Algorithms

Numerous sectors have utilized information technology (IT) to create records, files, articles, photographs, scientific data, and other data types. To make better decisions based on the data collected by multiple apps, one must have a strategy for mining large datasets for insights. Researchers can extract significant insights from massive datasets by utilizing "knowledge discovery in databases" (KDD). Data mining (DM) uses various tools and algorithms to uncover and extract meaningful patterns from recorded data. Several disciplines have incorporated data mining approaches, including statistics, machine learning, pattern reorganization, artificial intelligence, and computational capabilities. Several techniques, including decision trees, neural networks, Naive Bayes, and K-Nearest Neighbor, are utilized in educational data mining (EDM). Examples of information discovery methods include classifications, association rules, and cluster analysis. One may use the collected data to forecast films' impact, performance, and other results. Since its inception, nearest neighbor (NN) categorization has achieved widespread practical and academic use [23].

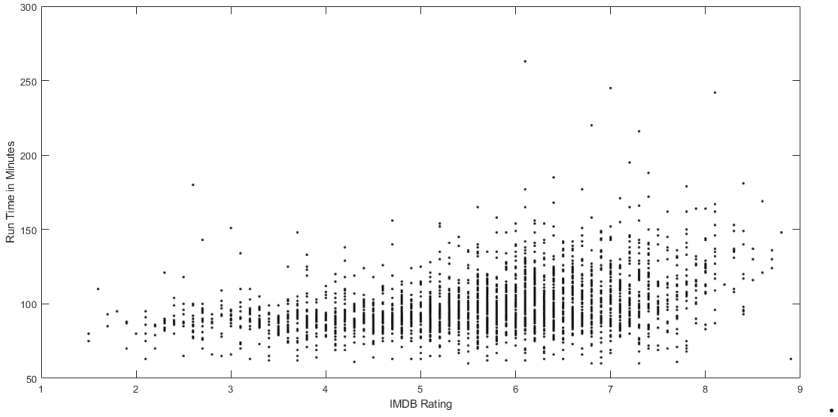


Figure 4: The Relationship Between IMDB Rating and Run Time in Minutes.

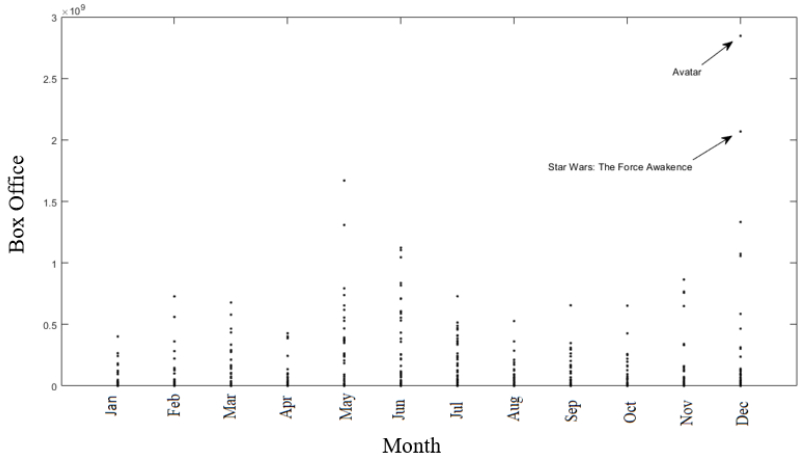


Figure 5: The Box Office by Month.

Table 3: The Highest-Grossing Sci-Fi Movies.

Movie	Month	Box Office
Avatar	12	2,847,246,203 $
Star Wars: The Force Awakence	12	2,068,223,624 $
Jurassic World	5	1,670,516,444 $
Star Wars: The Last Jedi	12	1,332,539,889 $
Jurassic World: Fallen Kingdom	5	1,308,467,944 $
Transformers: Dark of the Moon	6	1,123,794,079 $
Transformers: Age of Extinction	6	1,104,054,072 $
Star Wars: The Rise of Skywalker	12	1,074,144,248 $
Rogue One: A Star Wars Story	12	1,056,057,273 $
Jurrasic Park	6	1,033,928,303 $
Star Wars: Episode I – The Phantom Menace	5	1,027,082,707 $

Its widespread popularity results from its ease of use and the quality of the programs it generates. Ultimately, the search demands considerable memory and computations, yet these factors rose and decreased after employing the NN method. NN classifiers identify an input based on its proximity to a training point neighbor (denoted X'). The K Nearest Neighbors methods are widely used in data mining and machine learning [24] due to their high precision. In addition to being one of the top ten DM techniques, this algorithm is effective and favorable in DM [23], pattern recognition [25], and machine learning [26]. The KNN classifier identifies the K training cases (Xr, r = 1...k) that are most similar to x and then classifies x based on the consensus of the k nearest neighbors. When working with massive amounts of training data, it may be essential to perform an excessive calculation due to the necessity of computing and ranking the distance between each training point and each new data point. Only the number of neighbors defines the complexity of KNN. The greater k, the smoother the extension of the classification border.

Consequently, ask increases, the complexity of KNN decreases. Supervised learning can increase KNN models' categorization performance. The Nearest Prototype Classifier is a classification model that assigns observational class labels based on the training samples whose mean is closest to the observation. In addition, many NN classifiers can be created from the KNN classifier. The distance metric applied and the number of neighbors, represented by the K value, vary among these classifiers. Extracting relevant information or knowledge from raw data takes much work because data science was developed; consequently, it is desirable to use customized algorithms to process data.

3. KNN Algorithms

A vector represents each training sample in a multidimensional feature space and its corresponding class label. The training step of the method consists of just storing the feature vectors and class labels of the training instances. A user-defined constant k is utilized throughout the classification process to determine which label is more prevalent among the k training samples nearest to a query point. The Euclidean distance is commonly employed as a distance metric for continuous variables [36]. For discrete variables, such as text classification, an alternative metric, such as the overlap metric, may be used

(or Hamming distance). In combination with Pearson and Spearman correlation coefficients, k-NN has been used to assess microarray data on gene expression. When the distance metric is learned with particular methods such as large margin nearest neighbor or neighboring component analysis, the accuracy of k-classification NN is frequently significantly enhanced [27]. When there is a disparity in the distribution of classes, the majority voting method fails. In other words, examples of a more common class tend to dominate the prediction of a new instance due to their dominance among the k nearest neighbors. One such option is incorporating a distance metric between the test site and its k nearest neighbors into the classification weighting procedure. Each of the k nearest points has its class (or value, in regression problems) enhanced by a weight proportional to the inverse distance between it and the test point. Abstraction in data representation is an additional method for addressing skew. In a self-organizing map (SOM), each node serves as a representation (center) of a cluster of similar points, regardless of their density in the original training data. The K-NN may then employ the SOM [28]. The optimal value of k will vary from dataset to dataset, but in general, more significant values of k lessen the influence of the noise on the classification at the expense of fewer clearly defined class borders. Multiple heuristics exist for choosing an appropriate k. The nearest neighbor approach is used in the rare scenario when the predicted class is the same as the class of the nearest training sample (i.e., when k = 1).

Noise, irrelevant features, or incongruent feature scales can significantly impact the k-NN algorithm's performance. Selecting or scaling characteristics to enhance categorization has been the subject of extensive study. Using evolutionary algorithms to fine-tune feature scaling is a common strategy. Another standard method is scaling features using the mutual information between the training data and the training classes. When solving a binary (two-class) classification issue, it is preferable to have k be an odd integer, so there are no ties. The bootstrap approach is often used to select the one that is best in this circumstance from an empirical perspective. [29,30]. The kind of nearest neighbor classifier that assigns a point x to the same category as its closest neighbor in the feature space is the most intuitively understood. The one nearest neighbor classifier guarantees an error rate no worse than twice the Bayes error rate even as the size of the training data set approaches infinity (the minimum achievable error rate given the data distribution). Based on these considerations, we wish to categorize the science fiction films' dataset and how their success may be scored, ascertain the classification structure, and display the relationships between the data item sets. This can aid in the early prediction of success rates and assist producers in discovering films with poor audience success and finding methods to improve them. The KNN has supervised learning; it is one of the simplest machine learning algorithms based on feature similarity. It can be used for both classification and regression problems. KNN is a classical non-parametric algorithm [31]; it preserves and uses the training data during the test point classification process. The distance between the training points and the test point is calculated, and then the KNN algorithm classifies the test point based on the closest neighbors of that point.

There are many distance formulas; the formula used in this paper is Euclidean distance, which is the most commonly used [32]. It is a particular case of the Minkowski distance where p = 2, cosine, and cubic Minkowski, which is also a case of the Minkowski where p = 3, depending on the algorithms.

Minkowski distance:

$$d(x_i, y_i) = \sqrt[p]{|x_i - y_i|^p} \quad (1) \qquad (1)$$

Euclidean distance:

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (2) \qquad (2)$$

Cosine distance:

$$d(x_i, y_i) = 1 - \frac{x_i y_i}{\sqrt{(x_i x_i)(y_i y_i)}} \quad (3) \qquad (3)$$

Cubic Minkowski:

$$d(x_i, x_j) = \sqrt[3]{\sum_{i=1}^{n}(x_i - y_i)^3} \quad (4) \qquad (4)$$

Where x_i and x_j are the test point and training point respectively.

This study employed multiple KNN variants, including fine KNN, medium KNN, and weighted KNN. Table 4 displays the number of neighbors, the distance metric formula, and the distance weight for each type. The fine KNN has one neighbor, while the medium KNN has two, and the weighted KNN has 10. Fine KNN, medium KNN, and weighted KNN all employ Euclidean distance for their distance metrics. The equal distance (no distance weight) is used for fine KNN, medium KNN, and weighted KNN, while the squared inverse is used for weighted KNN. Also, other machine learning algorithms were implemented to compare the results and deduce certain features of the corresponding algorithm that may be considered for future adoption.

Table 4: KNN Algorithms

Algorithm	**Number of Neighbors**	**Distance Metric**	**Distance Weight**	**Standardize Data**
Fine KNN	1	Euclidean	Equal	True
Medium KNN	10	Euclidean	Equal	True
Weighted KNN	10	Euclidean	Squared Inverse	True

4. Results

The Receiver Operator Characteristic (ROC) is a probability curve representing the true positive rate versus the false positive rate. The area under the curve (AUC) is the measure of the ability of the classifier to distinguish between classes (separability). The confusion matrix and the AUC-ROC curve were analyzed to compare the results, in addition to accuracy, prediction speed, and training time. The confusion matrix is an N x N matrix used to evaluate the classification model's performance. Where N is the number of classes, in our example, N is 4. where the actual classes are on the y-axis, and the predicted classes are on the x-axis. The confusion matrix extracted each class's true positive and false positive rates. The confusion matrix can generate three performance accuracy, precision, and recall measures [33].

A summary of the obtained results for different utilized KNN classifiers is provided in table 5:

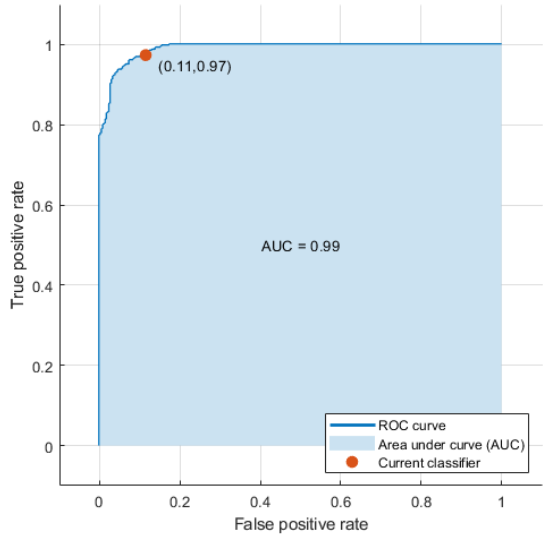
Table 5: Summary of the results for different KNN algorithms.

KNN Classifier	**Accuracy (%)**	**Prediction Speed (obs/sec)**	**Training Time**	**Class True Positive Rate**			
				Average	**Below**	**Flop**	**Hit**

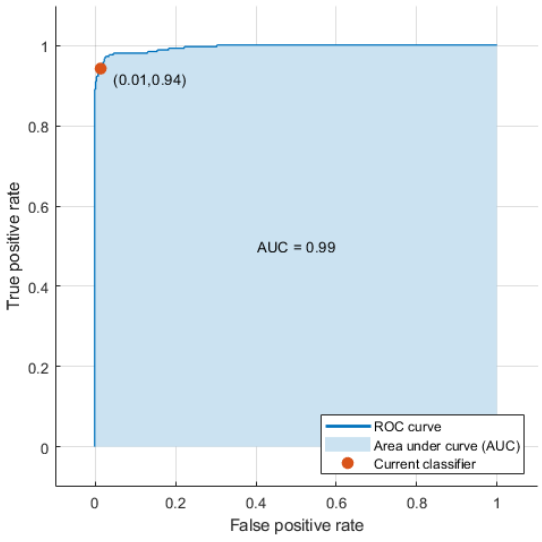
					Average		
Fine KNN	93.0	~31000	0.6328	96%	93%	100%	77%
Medium KNN	90.3	~24000	0.21746	98%	89%	73%	48%
Weighted KNN	92.9	~24000	0.21903	97%	94%	93%	60%

Fig. 6 shows the AUC-ROC curves for one of the weighted KNN methods. Each subgraph represents the relationship between the true positive rate and the positive rate for the viewer's opinion. Three main parameters are extracted from each graph: the ROC curve behavior, the AUC, and the accuracy of each classifier. The area under the receiver operating characteristic curves (AUC-ROC) for one of the weighted KNN techniques is displayed in Fig. 6. As you move down the graph, it can be seen how the percentage of "yes" answers reflects the viewer's confidence. Each graph is analyzed for its ROC curve behavior, AUC, and classifier accuracy, three key metrics.

An AUC of 0.99, a ROC point of 0.11, and a ROC that plateaus at a 1% true positive rate for a 0.18 percent false positive rate are all easily discernible in the graphic. However, the AUC for the below-average opinion is 0.99, the ROC point is (0.01, 0.94), and the ROC reaches one true positive rate at a false positive rate of 0.32, suggesting a more significant error is accomplished to guarantee the one true positive rate. Similarly, the flop opinion has an AUC of 1, the ROC point is (0.00, 0.93), and the ROC reaches a true positive rate of 1 at a false positive rate of 0 quickly. The ROC point for the hit opinion is (0.01, 0.6), and it achieves a true positive rate of 1 at a false positive rate of 0.2%, indicating that the rate of false positives is moderate.



Average



Below Average

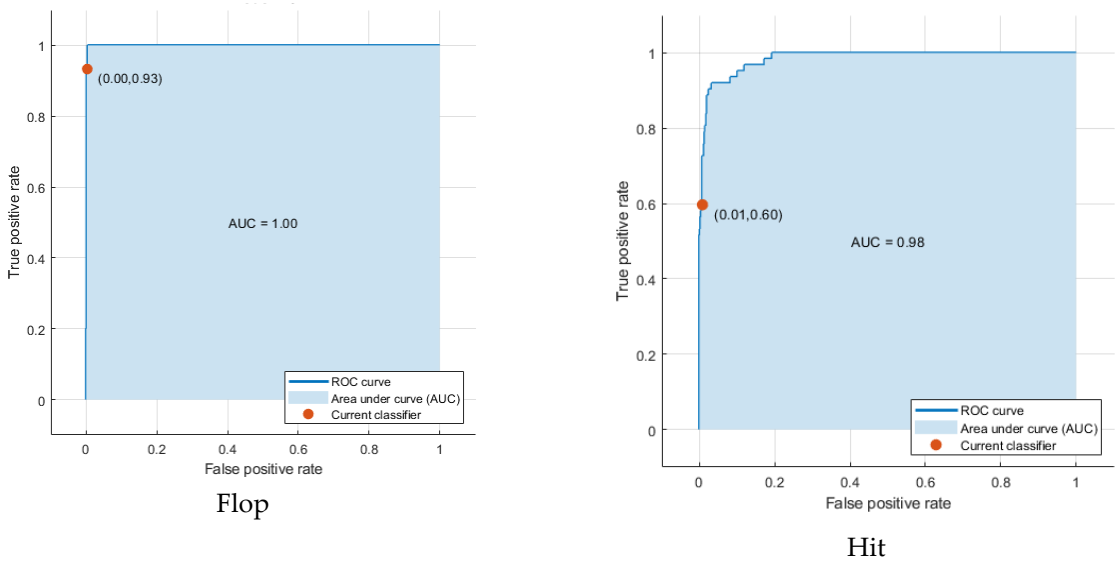


Figure 6: Weighted KNN ROC

Table 5 summarizes the results obtained when different KNN algorithms are implemented. The Fine KNN algorithm provides the highest accuracy of 93%, but it has the lowest speed among the three classes. On the other hand, the Medium KNN classifier has the highest accuracy for the Average class with a value of 98% but the lowest for the Hit class with a value of 48%. The most consistent classifier is the weighted KNN classifier, which provides consistent estimates for the average, below average, and flop at rates of 97%, 94%, and 93%, respectively. However, it did not perform as well with the Hit class as it did with others. The three classifiers provide somehow excellent accuracy, ranging from 90.3–93%.

Table 6 summarizes the results of various classification techniques. The results demonstrate their applicability and credibility by contrasting the provided algorithms' findings with those in [38,40] and their citations.

Table 7. The Accuracy, precision, and recall for various machine learning and deep learning algorithms.

Algorithm	Accuracy	Precision	Recall
Multinomial Regression	**0.29**	**0.23**	**0.29**
Fine KNN	**0.93**	**0.91**	**0.91**
Medium KNN	**0.90**	**0.89**	**0.88**
Weighted KNN	**0.93**	**0.93**	**0.92**
Bi-Directional LSTM	**0.79**	**0.75**	**0.77**
LSTM	**0.78**	**0.80**	**0.78**
GRU	**0.77**	**0.79**	**0.77**
Logistic regression	**0.89**	**0.88**	**0.88**
Naive Bayes	**0.81**	**0.80**	**0.81**
Decision tree	**0.87**	**0.86**	**0.87**
Gradient Boosting	**0.87**	**0.86**	**0.86**
Random Forest	**0.85**	**0.84**	**0.84**
SVM linear kernel	**0.73**	**0.72**	**0.71**

SVM RBF kernel	**0.75**	**0.74**	**0.73**

The weighted KNN algorithm shows the highest and most robust results among the considered classifiers.

5. Discussion

This paper provides an alternate approach to open space risk reduction that employs k-nearest neighbor approaches to discover discriminative characteristics. In order to study the role of k, we conduct contrastive learning with a variety of k values (within a given range) and compare the model's performance in detecting out-of-domain (OOD) intents across the board (while other hyper-parameters remain constant). The results demonstrate that as the value of k is increased, the model's performance (cosine-based) decreases across datasets. Indeed, this was to happen as k increased, and the possibility that an OOD sample would be incorrectly identified as in-domain (IND) decreased because of the decreased proportion of open space at the outset of the process. Later, as the IND semantic space shrank, a more significant percentage of intra-OOD samples were labeled as IND, and this trend tends to stabilize. Adopting this strategy also shows that our method is better than other methods for lowering the open-space risk, which is the risk that all IND intents from the same class will collide.

As listed in [15]–[21], the best-achieved accuracy was very low compared to the tabulated results in this work. Sivakumar, Pirunthavi, et al trained an IMDb-extracted random forest algorithm to predict the success of a film. In addition, the Naive Bayer model was trained using YouTube user evaluations to estimate the rating of a movie. The accuracy of the models was evaluated based on their performance on real-world datasets. Two results have been reached: the rating of a new film cannot be anticipated in advance using the comments on its trailers and teasers on YouTube, while the success of a new film may be predicted in advance using online data or characteristics. The overall accuracy of 70% is achieved; however, they claimed that the success prediction model might be used as an early assessment tool for movies, which would benefit the movie business and its audience. In [21], gradient boosting was the most effective method in their analysis, with an 84.1287% success rate. According to the summary of the linked study, machine learning algorithms could help predict how well a movie will do by considering everything that could affect its success rate. Because of this, it might be possible to build on past results and make an algorithm that improves on the tested methods and gets better results. To further compare the results of this work, several works were summarized in [34]–[35]. In [34], three machine learning algorithms were analyzed and tabulated: Naive Bayes with 80.68% accuracy, Decision Trees C4.5 with 86.47% accuracy, and Logistic Regression with 89.98% accuracy. To predict movie success, the authors used a data set from IMDB, and they classified the movies according to their IMDB rating into five classes: flop, average, good, hit, and super hit. In [35], they proposed a model to predict whether the movie would be a hit or flop before it was released using machine learning techniques and algorithms; they used the IMDB dataset in their paper. They imported six algorithms: Nave Bayes with 72.18% accuracy, Decision Tree with 81.04% accuracy, K-Nearest Neighbor with 73.76% accuracy, Support Vector Machine with 72.68% accuracy, Logistic Regression with 73.26% accuracy, and Random Forest with 85.2% accuracy.

The tabulated findings of this study indicate that the Weighted KNN generates accurate and robust outcomes.

6. CONCLUSION

The article explains the optimization target for non-domain-specific features. We look at the problems with current methods, and then we offer a simple yet effective method for acquiring discriminative semantic features. Our method groups domain intents with their k-nearest neighbors and separates them from samples of other classes to

lessen empirical and open space risks. Extensive testing on challenging datasets confirms the consistency and reliability of our method without imposing artificial constraints on the distribution of features. Forecasts of future interest in SF films were made using several different KNN algorithm implementations. Films were classified as flip, below average, average, or Hit, all based on their rating on IMDB. The Fine, Weighted, and Medium KNN types outperform other machine learning types regarding accuracy. Each type of KNN takes less than 0.7 seconds to make a prediction.

In the future, other machine learning approaches will be applied and compared, and new criteria from sources other than IMDB will be used to improve the prediction model's accuracy. Examples of such sources include Rotten Tomatoes and Wikipedia. Intuitively combining more than KNN algorithm targeting particular viewer behavior may produce consistent and authentic predictive results that support the film industry and its worldwide spread.

Acknowledgment: The dataset's developers are appreciated by the authors for making it accessible online. They also like to express their gratitude to the anonymous reviewers who helped them to make this paper better.

Funding Statement: This research isn't funded.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] Udofia, Nsikak-Abasi, and Joy Stephen Anyim. "Assessing the Impact of Modern Movies on Students—A Prospective Study." Journal of Culture, Society and Development 31 (2017): 1-11.

[2] Lochbuehler, Kirsten, et al. "Effects of smoking cues in movies on immediate smoking behavior." Nicotine & Tobacco Research 12.9 (2010): 913-918.

[3] Vagionis, Nikolaos, and Maria Loumioti. "Movies as a tool of modern tourist marketing." Tourismos 6.2 (2011): 353-362.

[4] Lin, Kuen-Yi, et al. "Effects of a science fiction film on the technological creativity of middle school students." Eurasia Journal of Mathematics, Science and Technology Education 9.2 (2013): 191-200.

[5] Lorenčík, D., M. Tarhaničová, and P. Sinčák. "Influence of sci-fi films on artificial intelligence and vice-versa." 2013 IEEE 11th international symposium on applied machine intelligence and informatics (SAMI). IEEE, 2013.

[6] Lowe, Thomas, et al. "Does tomorrow ever come? Disaster narrative and public perceptions of climate change." Public understanding of science 15.4 (2006): 435-457.

[7] Stimpert, J. L., et al. "Factors influencing motion picture success: Empirical review and update." Journal of Business & Economics Research (JBER) 6.11 (2008).

[8] Lash, Michael T., and Kang Zhao. "Early predictions of movie success: The who, what, and when of profitability." Journal of Management Information Systems 33.3 (2016): 874-903.

[9] Nithin, V. R., et al. "Predicting movie success based on IMDb data." International Journal of Data Mining Techniques and Applications 3 (2014): 365-368.

[10] Latif, Muhammad Hassan, and Hammad Afzal. "Prediction of movies popularity using machine learning techniques." International Journal of Computer Science and Network Security (IJCSNS) 16.8 (2016): 127.

[11] Asad, Khalid Ibnal, Tanvir Ahmed, and Md Saiedur Rahman. "Movie popularity classification based on inherent movie attributes using C4. 5, PART and correlation coefficient." 2012 International Conference on Informatics, Electronics & Vision (ICIEV). IEEE, 2012.

[12] Saraee, Mohamad, Sean White, and J. Eccleston. "A data mining approach to analysis and prediction of movie ratings." WIT Transactions on information and communication technologies 33 (2004).

[13] Galvão, Marta, and Roberto Henriques. "Forecasting movie box office profitability." Journal of Information Systems Engineering & Management 3.3 (2018): 1-9.

[14] Gaikar, Dipak, et al. "Movie success prediction using popularity factor from social media." International Research Journal of Engineering and Technology 6.4 (2019): 5184-5190.

[15] Lash, Michael T., and Kang Zhao. "Early predictions of movie success: The who, what, and when of profitability." Journal of Management Information Systems 33.3 (2016): 874-903.

[16] Ahmad, Javaria, et al. "Movie success prediction using data mining." 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2017.

[17] Dhir, Rijul, and Anand Raj. "Movie success prediction using machine learning algorithms and their comparison." 2018 first international conference on secure cyber computing and communication (ICSCCC). IEEE, 2018.

[18] Chakraborty, Partha, Md Zahidur Rahman, and Saifur Rahman. "Movie success prediction using historical and current data mining." International Journal of Computer Applications 178 (2019): 1-5.

[19] Sivakumar, Pirunthavi, et al. "Movie Success and Rating Prediction Using Data Mining Algorithms." (2020).

[20] Agarwal, Manav, et al. "A Comprehensive Study on Various Statistical Techniques for Prediction of Movie Success." arXiv preprint arXiv:2112.00395 (2021).

[21] Gupta, Vedika, et al. "Predicting attributes based movie success through ensemble machine learning." Multimedia Tools and Applications (2022): 1-30.

[22] Bristi, Warda Ruheen, Zakia Zaman, and Nishat Sultana. "Predicting imdb rating of movies by machine learning techniques." 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2019.

[23] Papadopoulos, Apostolos N., and Yannis Manolopoulos. Nearest Neighbor Search:: A Database Perspective. Springer Science & Business Media, 2006.

[24] Kononenko, Igor, and Matjaz Kukar. Machine learning and data mining. Horwood Publishing, 2007.

[25] Shakhnarovich, Gregory, Trevor Darrell, and Piotr Indyk. "Nearest-neighbor methods in learning and vision." IEEE Trans. Neural Networks 19.2 (2008): 377.

[26] Wu, Xindong, et al. "Top 10 algorithms in data mining." Knowledge and information systems 14.1 (2008): 1-37.

[27] Jaskowiak, Pablo A., and R. J. G. B. Campello. "Comparing correlation coefficients as dissimilarity measures for cancer classification in gene expression data." Proceedings of the Brazilian symposium on bioinformatics. Brazil: Brasília, 2011.

[28] Burgot, G., F. Auffret, and J-L. Burgot. "Determination of acetaminophen by thermometric titrimetry." Analytica chimica acta 343.1-2 (1997): 125-128.

[29] Nigsch, Florian, et al. "Melting point prediction employing k-nearest neighbor algorithms and genetic parameter optimization." Journal of chemical information and modeling 46.6 (2006): 2412-2422.

[30] Hall, Peter, Byeong U. Park, and Richard J. Samworth. "Choice of neighbor order in nearest-neighbor classification." the Annals of Statistics 36.5 (2008): 2135-2152.

[31] Zhang, Chao, et al. "Hybrid Metric K-Nearest Neighbor Algorithm and Applications." Mathematical Problems in Engineering 2022 (2022).

[32] Jawthari, Moohanad, and Veronika Stoffová. "Predicting students' academic performance using a modified kNN algorithm." Pollack Periodica 16.3 (2021): 20-26.

[33] Uddin, Shahadat, et al. "Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction." Scientific Reports 12.1 (2022): 1-11.

[34] Jaiswal, M., Prasad, A., Srivastava, A., & Siddqui, T. J. Prediction and Analysis of Movie Success with Machine Learning Approach "(2020).

[35] Sadashiv, Sagar, Sween Seli, and Sankruth Sreram. "MOVIE SUCCESS PREDICTION USING MACHINE LEARNING."(2021).

IMPACT FACTOR 2.838 CITESCORE 3.7

applied sciences

Submit to *Applied Sciences*

Review for *Applied Sciences*

Journal Menu

- *Applied Sciences* Home
- Aims & Scope
- Editorial Board
- Reviewer Board
- Topical Advisory Panel
- Instructions for Authors
- Special Issues
- Topics
- Sections & Collections
- Article Processing Charge
- Indexing & Archiving
- Editor's Choice Articles
- Most Cited & Viewed
- Journal Statistics
- Journal History
- Journal Awards
- Society Collaborations
- Conferences
- Editorial Office

Journal Browser

volume

issue

Go

Forthcoming issue

Current issue

Vol. 13 (2023) Vol. 12 (2022) Vol. 11 (2021) Vol. 10 (2020) Vol. 9 (2019) Vol. 8 (2018) Vol. 7 (2017)

Vol. 6 (2016) Vol. 5 (2015) Vol. 4 (2014) Vol. 3 (2013) Vol. 2 (2012) Vol. 1 (2011)

Editorial Board

- Applied Biosciences and Bioengineering Section
- Applied Dentistry and Oral Sciences Section
- Biomedical Engineering Section
- Applied Microbiology Section
- Applied Neuroscience and Neural Engineering Section
- Materials Science and Engineering Section
- Chemical and Molecular Sciences Section
- Nanotechnology and Applied Nanosciences Section
- Surface Sciences and Technology Section
- Environmental Sciences Section
- Green Sustainable Science and Technology Section
- Earth Sciences and Geography Section
- Energy Science and Technology Section
- Optics and Lasers Section
- Acoustics and Vibrations Section
- Applied Physics Section
- Quantum Science and Technology Section
- Applied Thermal Engineering Section
- Fluid Science and Technology Section
- Computing and Artificial Intelligence Section
- Mechanical Engineering Section
- Civil Engineering Section
- Electrical, Electronics and Communications Engineering Section
- Applied Industrial Technologies Section
- Food Science and Technology Section
- Robotics and Automation Section
- Aerospace Science and Engineering Section
- Marine Science and Engineering Section
- Transportation and Future Mobility Section
- Agricultural Science and Technology Section
- Additive Manufacturing Technologies Section

Members (2196)

Search by first name, last name, affiliation, interest...

Prof. Dr. Takayoshi Kobayashi Website SciProfiles

Editor-in-Chief

Advanced Ultrafast Laser Research Center, The University of Electro-Communications, 1-5-1, Chofugaoka, Chofu, Tokyo 182-8585, Japan

Interests: ultrafast and nonlinear optical processes; quantum optics; quantum photobiology

Special Issues, Collections and Topics in MDPI journals

Prof. Dr. Francesco Cappello * Website SciProfiles

Associate Editor

Department of Biomedicine, Neuroscience and Advanced Diagnostics, University of Palermo, 90127 Palermo, Italy

Interests: human anatomy, histology and embryology; pathological anatomy; neuroscience; advanced diagnostics; carcinogenesis, immunomorphology; extracellular vesicles; exosomes; heat shock proteins; molecular chaperons; chaperonins; chaperonopathies; chaperonotherapy

* Section: "Applied Biosciences and Bioengineering"

Special Issues, Collections and Topics in MDPI journals

Dr. Francesca Intonti * Website SciProfiles

Associate Editor

European Laboratory for Non-Linear Spectroscopy, Department of Physics and Astronomy, University of Firenze, Via N. Carrara 1, 50019 Sesto Fiorentino, Italy

Interests: nanophotonics; photonic bandgap materials; mie resonators; semiconductor nanostructures; high spatial resolution imaging; scanning near-field optical microscopy

* Section: "Optics and Lasers"

Prof. Dr. Dimitris Mourtzis * Website SciProfiles

Associate Editor

Laboratory for Manufacturing Systems and Automation (LMS), Department of Mechanical Engineering and Aeronautics, University of Patras, 26504 Rio Patras, Greece

Interests: simulation; design; planning and control of manufacturing systems and networks; robotic systems; automation; augmented, mixed, and virtual reality in manufacturing; manufacturing process modeling; cloud technologies; Internet of Things (IoT); digital twin; 5G; artificial intelligence; product-service systems (PSS); industry 4.0; industry 5.0; society 5.0; metaverse; blockchain; adaptive scheduling

* Section: "Robotics and Automation"

Special Issues, Collections and Topics in MDPI journals

Prof. Dr. Alexander N. Pisarchik Website SciProfiles

Associate Editor

Center for Biomedical Technology, Technical University of Madrid, Campus Montegancedo, Pozuelo de Alarcón, 28223 Madrid, Spain

Interests: complex systems; bioinformatics; mathematical and computational biology; optics and photonics; biological physics; cognitive neuroscience

Special Issues, Collections and Topics in MDPI journals

Prof. Dr. Ilaria Cacciotti Website SciProfiles

Section Associate Editor

Dipartimento di Ingegneria, Università di Roma 'Niccolò Cusano', INSTM RU, 00166 Rome, Italy

Interests: biomaterials; bioceramics; biopolymers; biocomposites; ecosustainable materials; scaffold; spheres; fibers; coatings; sol–gel processes; valorization of agro-food waste extracts and by-products; electrospinning; additive manufacturing; physicochemical characterization; microstructure; thermal and mechanical properties; tissue engineering/regenerative medicine; drug delivery

Special Issues, Collections and Topics in MDPI journals

Dr. Nunzio Cennamo * Website SciProfiles

Section Associate Editor

Department of Engineering, University of Campania Luigi Vanvitelli, Via Roma 29, 81031 Aversa, Italy

Interests: optical sensors; biosensors and chemical sensors; optical fiber sensors and optoelectronic devices

* Section: "Electrical, Electronics and Communications Engineering"

Special Issues, Collections and Topics in MDPI journals

Prof. Dr. Tomislav Dragicevic * Website

Section Associate Editor

Department of Electrical Engineering, Center for Electric Power and Energy, Technical University of Denmark - DTU, 2800 Kgs, Lyngby, Denmark

Interests: renewable energy; power electronics; energy storage; e-mobility; artificial intelligence; IoT-driven