

Navigating techniques in job recommender systems on internship profile matching: a systematic review

Flordeliza P. Poncio

*Faculty of Information and Computer Technologies,
Paragon International University, Phnom Penh, Cambodia and
ICS, University of the Cordilleras, Baguio City, Philippines*

Abstract

Purpose – This review article is focused on the following research questions: RQ1: What are the methods used by authors to collect data in order to evaluate one's profile? RQ2: What are the classification algorithms and ranking metrics used to give suggestions to users? RQ3: How effective are these algorithms and metrics identified in RQ2?

Design/methodology/approach – There are four major systematic review phases being carried out in this survey, namely the formulation of research questions, conducting the review, which includes the selection of articles and appraising evidence quality, data extraction and narrative data synthesis.

Findings – Collecting from primary sources is more personalized and relevant. Embedded skill sets that have a considerable impact on one's career aspirations could be mined from secondary sources. A hybrid recommender system helped mitigate the limitations of both. The effectiveness of the models depends not only rely on the filtering techniques used but also on the metrics used to measure similarity and the frequency of words or phrases used in a document.

Research limitations/implications – The study benefits internship program coordinators of a university aiming to develop a recommender or matching system platform for their students. The content of the study may shed a light on how university decision-makers can explore options on what are the techniques or algorithms to be integrated. One of the advantages of internship or industrial training programs is that they would help students align them with their career goals. Research studies have discussed other RS filtering techniques apart from the three major filtering techniques.

Practical implications – The outcome of the study, which is a recommendation system to match a student's profile with the knowledge and skills being sought by organizations, may help ease the challenges encountered by both parties. The study benefits internship coordinators of a university who are planning to create a recommendation system, an innovative project to be used in teaching and learning.

Social implications – Internship programs can help a student grow personally and professionally. A university student looking for internship opportunities can find it a daunting task to undertake, as there is a vast pool of opportunities offered in the market. The confidence levels needed to match their knowledge, skills and career goals with the job descriptions (JDs) could be challenging. The same holds with companies, as

© Flordeliza P. Poncio. Published in *Journal of Research in Innovative Teaching & Learning*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

The author wished to acknowledge the people who helped one way or another to make this research paper possible. To Dr Thelma D. Palaoag, Dr Meirambek Zhaparov and to the JRITL team, I am indebted to you.

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work, the author used AI-assisted technologies. The assistance sought was synonyms of significant words and to clarify differences and similarities of words or phrases or basis for paraphrasing. Suggestions from the AI tool were also sought. After using this tool/service, the author reviewed and edited the content as needed and took full responsibility for the content of the publication.



finding the right people for the right job is a tough endeavor. The main objective of conducting this study is to identify models implemented in recommendation systems to give and/or rank suggestions given to users.

Originality/value – While surveys regarding recommender systems (RS) exist, there are gaps in the presentation of various data collection methods and the comparison of recommendation filtering techniques used for both primary and secondary sources of data. Most recommendation systems for internship programs are intended for European universities and not much for Southeast Asia. There are also a limited number of comparative studies or systematic review articles related to recommendation systems for internship programs offered in an Southeast Asian landscape. Systematic reviews on the usability of the proposed recommendation systems are also limited. The study presents reviews of articles, from data collection and techniques used to the usability of the proposed recommendation systems, which were presented in the articles being studied.

Keywords Recommender systems, Internship, Job placement, Filtering techniques, Systematic review

Paper type Literature review

1. Introduction

Recommendation platforms are efficient tools for giving relevant suggestions to users across different domains, be it in the field of e-commerce, e-learning platforms, the entertainment industry and research databases to name a few. Most e-commerce companies have benefited from recommender systems (RS), such as book, and movie and gadget recommendations (Wang *et al.*, 2018).

The major techniques implemented by RS are collaborative filtering (CF), content-based filtering (CBF), and hybrid RS. CF is a technique where suggestions to users have a similarity of preferences from past activities or other users. Therefore, selecting the appropriate similarity is fundamental to a system's performance in implementing this technique (Al-Bashiri *et al.*, 2019). Several methods have been introduced including Pearson's correlation coefficient (PCC) and cosine similarity (Kirch, 2008; Ken, 2023). CBF suggests elements for users that are related to the items that were previously chosen. First, the relationship between the object and its properties is established in the matrix term, and by using different mathematical functions, the most identical items to the target item will be selected (Nallamala *et al.*, 2020). A hybrid recommender system is a combined collaborative and CBF technique. The main objective of an RS is to predict users' preferences and match them with a large number of choices according to their liking (Afoudi *et al.*, 2021).

As the industrial revolution is moving towards digitalization, learners are called to be adaptive and ever-ready to meet market demands (Adeosun *et al.*, 2021). To enhance their future career goals, and to foster university students to acquire employability skills, internship programs and/or work-integrated learning courses are offered by universities (Saeed *et al.*, 2023).

There are four principles of learning by doing (Echarcharqy, 2020), one of which is experiential learning like internship programs. It allows students to apply, develop and practice their skills in a professional setting (Thompson *et al.*, 2021).

Looking for open internship opportunities can be a strenuous task to undertake due to the diversity of job seekers' profiles (Mhamdi *et al.*, 2020). Matching a user's profile with the existing job descriptions (JDs) in the market could be a laborious task. Finding the right people for the right job can be challenging for organizations (Velciu, 2018) in terms of technical skills and soft skills. To enhance career goals and foster students to acquire employability skills, internship programs and/or work-integrated learning courses are offered by universities (Saeed *et al.*, 2023).

A typical approach to creating job recommendations involves calculating the similarity between the candidate's profile and the job description. LinkedIn aims for results that are tailored to the individual, where algorithms need to be scalable (de Ruijt and Bhulai, 2021).

A review article by [Al-walidi et al. \(2020\)](#) identified problems that were addressed by RS. The corresponding solutions to these problems were also enumerated like accuracy issues. The digital era is heading toward knowledge-intensive technologies, which shaped the way people do things, be it in the workforce, business or even daily activities ([Lobanova, 2021](#); [Kraus et al., 2021](#));

The objectives of the literature review are (1) to investigate systematic data collection methods aligning student career goals with internship opportunities; (2) to explore innovative techniques used in RS giving suggestions to users or matching one's profile and (3) to identify and evaluate the classification algorithms or ranking metrics used by authors. There are existing systematic reviews, but data collection, innovative techniques and metrics used were not given emphasis. Systematic reviews about job or internship RS used in Southeast Asia are also limited.

The study addresses two of the seventeen sustainable development goals (SDGs) of the United Nations, namely, goal number four, which is to ensure inclusive and equitable quality education and promote lifelong opportunities for all, and goal number nine, which is to build resilient infrastructure and foster innovation ([UN, 2023](#)).

2. Methods

There are four major systematic review phases ([Kitchenham and Charters, 2007](#); [Shepperd, 2013](#)) being carried out in this survey, namely the formulation of research questions, conducting the review which includes selection of articles and appraising evidence quality, data extraction and narrative data synthesis.

2.1 Research questions

The review article is focused on the following research questions:

- RQ1. What are the methods used by authors to collect data to evaluate one's profile?
- RQ2. What are the classification algorithms and ranking metrics used to give suggestions to users, and
- RQ3. How effective are these algorithms and metrics identified in [RQ2](#)?

2.2 Conducting the review

2.2.1 Selection of articles. To have a well-structured systematic review, the inclusion criteria were identified. (1) The articles to be examined must have implemented CF, CBF and hybrid recommendation filtering techniques. (2) The articles are pertinent to internship or job placement programs in the information technology domain. (3) The locale of the study in the article is Southeast Asia and Nigeria, which are facing both challenges and favorable circumstances regarding job placements or internship opportunities. (4) Articles that extract data from web career portals through web scraping. (5) The articles were published between 2019 and 2022 and are Scopus-indexed. (6) The sources of data are primary or secondary research. The exclusion criteria were articles not written in English, articles that do not have open access privileges and articles not having a reasonable number of citations ([Aksnes et al., 2019](#)). The summary of the articles being explored was presented in [Table 1](#) in [Section 4](#).

2.2.2 Appraising evidence quality. Five studies for each filtering technique were initially considered, but some articles were published in non-Scopus or non-open access journals. The number of articles was down to five, two for CF, two for CBF and only one for hybrid, as the discussions addressed all the research questions.

Article ID	Title	Filtering technique	Publication
1	A recommender system for selecting potential industrial training organizations	Collaborative Filtering	Ogunde, A., and Idialu, J. (2019). "A recommender system for selecting potential industrial training organizations". <i>Engineering Reports</i> , Vol. 1 No. 3. https://doi.org/10.1002/eng2.12046
2	Job matcher: A web application job placement using collaborative filtering recommender system		Mendez, J., and Bulanadi, J. (2020). "Job matcher: A web application job placement using collaborative filtering recommender system". <i>International Journal of Research Studies in Education</i> , Vol. 9 No. 2, pp 103–120. https://doi.org/10.5861/ijrse.2020.5810
3	Student Career Recommendation System Using Content-Based Filtering Method	Content-Based Filtering	Rashid, A., Mohamad, M., Masrom, S., and Selamat, A. (2022). "Student Career Recommendation System Using Content-Based Filtering Method". <i>International Conference on Artificial Intelligence and Data Sciences</i> . Pp 60–65. DOI: 10.1109/AiDAS56890.2022.9918766
4	Job Recommendation Based on Extracted Skill Embeddings		Kara, A., Daniş, F.S., Orman, G.K., Turhan, S.N., and Özlü, Ö.A. (2023). "Job Recommendation Based on Extracted Skill Embeddings. In: Arai, K. (Ed.s), <i>Intelligent Systems and Applications. IntelliSys 2022. Lecture Notes in Networks and Systems</i> . DOI: https://doi.org/10.1007/978-3-031-16075-2_35
5	Technical Job Recommendation System Using APIs and Web Crawling	Hybrid of both	Kumar, N., Gupta, M., Sharma, D., and Ofori, I. (2022). "Technical Job Recommendation System Using APIs and Web Crawling". <i>Computational Intelligence and Neuroscience</i> . Retrieved from https://www.hindawi.com/journals/cin/2022/7797548/

Source(s): Tables by author

Table 1.
List of explored articles

2.3 Data extraction for comparative examination

The relevant information that was evaluated from the research articles was (1) the type of data being collected, (2) procedures undertaken to clean or pre-process the data were identified, (3) the methods used to implement the studies were evaluated and (4) the application of the RS filtering techniques was discussed, including the metrics in measuring relationships of variables. The test conducted on the accuracy of the model or algorithm was also examined, including the measurement of the reliability of the accuracy model. Another set of data being extracted from the articles is the significant details when the user acceptance test (UAT) (Techtarget) was conducted. The discussions about the techniques were presented in Section 3. The data collection methods and models used by the authors were presented in Tables 2 and 3, respectively.

2.4 Narrative data synthesis summary

The discussions on the filtering techniques and the algorithms were summarized in a tabular form. Significant formulas from the articles were also noted in this paper.

In the conduct of the study, generative AI tools were used.

Article ID	Data	Methods	Source or locale
1	<ul style="list-style-type: none">- Demographic Data- About the company- About Industrial Training Experience- Preferences	Google Forms	Nigeria
2	<ul style="list-style-type: none">- Job search methods- Practices when finding a job- Need for a web app RS	Questionnaire using jotform.com forms	Philippines
3	Available careers from careers website	Unstructured interview	Malaysia-based
4	<ul style="list-style-type: none">- Turkish and English JDs related to IT- CVs of those who are employed in the IT industry	Web scraping from career recommendation websites	Kariyer.net
5	Available job databases	Web crawling from a job recruiting platform	
		Web Crawling APIs using submodules <ul style="list-style-type: none">- Company fetching module- Company detail gathering- Job listing fetching module- Job listing platform crawling- Data fields unifying module	Career portals of numerous company websites

Table 2.
Data acquisition

Source(s): Tables by author

Article ID	Tests conducted	Implementation	Accuracy of the model	Reliability measure
1	Training dataset: 70% of data Testing dataset: 30% of data	C4.5 algorithm through WEKA's J48	78.84% were correctly classified and 21.16% were incorrectly classified	Kappa Statistics: 0.7839 Mean Absolute Error: 0.0058 (means the magnitude of errors is almost insignificant)
4	TF-IDF with cosine similarity as the benchmark	Precision at $k = \{1,3,5,10\}$ Skill embeddings and WMB, with $k = 1$	95% Generated results vary based on methods and k-values	
5	Comparison of the system's efficiency with and without pre-processing the JDs with respect to time taken to generate	Proposed hybrid RS algorithm giving significance to the processing of JDs The relevance of unified database having the same schema	Processing JDs increased the chances of the user of getting matched by 103.44% Increased in match score for raw and processed JDs	

Table 3.
Accuracy of the model

Source(s): Tables by author

3. Comparative examination based on matching typologies

3.1 Collaborative filtering

The first article being investigated is a recommender system to match the requirements of companies offering industrial training to students (Ogunde and Idialu, 2019), where CF was deployed using the C4.5 algorithm, a method to generate a decision tree based on a set of data (Liu and Yang, 2022). The data were collected through a questionnaire using Google Forms, containing demographic data, about the company and their industrial training experience. The data were collected from students of Lagos State, Nigeria.

The data generated by the decision trees will form a knowledge base, an essential application for finding answers to questions (Zeng *et al.*, 2023) for the system. The end user will then interact with future predictions through the inference engine. J48 is the direct implementation of the C4.5 algorithm in WEKA, which contains a collection of machine learning algorithms (Panigrahi and Borah, 2018; Hall *et al.*, 2009), and was used to build a model using the data obtained from previous industrial training students.

Another article that implemented C4.5 is a recommendation system using artificial intelligence for an internship placement based on competence (Permana and Pradnyana, 2019). The variables being considered are the student's grades in courses related to programming or networking, their grade point average (GPA), and the results of an inventory personal survey (IPS). The IPS is based on Holland's personality traits (Savickas and Savickas, 2017) and is given to students who will take internship courses. It reflects and evaluates one's character, abilities, interests and behavior.

The second article that implemented a CF technique is a web application named Job Matcher by Mendez and Bulanadi (2020). The article discussed the efficiency of the proposed system and its usability to its intended audiences. Data were collected from on-the-job training (OJT) students and human resource representatives of the university's partner companies via jotform.com and an unstructured interview. The system implemented a user-based technique – a memory-based method used in collaborative RS (Atoum, 2020). A preliminary evaluation of the possibility of an RS was conducted. Online job search, followed by searching on company websites, is the most common practice when looking for a job or internship opportunities. One of the suggestions stated was to include a module that would match the applicant's profile with the currently existing job vacancies, considering that a single company needs distinctive skill sets.

3.2 Content-based filtering

CBF in recommendation systems analyzes the description of items to build a model of new items for other users. CBF extracts conclusions from the metadata of products, which are usually provided by product companies (Papadakis *et al.*, 2023).

The first article is a student career RS from Rashid *et al.* (2022). Web scraping, a technique used to mine online data (Lotfi *et al.*, 2022), was implemented to extract data from career recommendation websites like Jobstreet. To extract meaningful information, Python's BeautifulSoup is used to parse collected Hypertext Markup Language (HTML) and Extensible Markup Language (XML) files. Pandas library and MS Excel are used to analyze and clean data, respectively.

The types of CBF techniques carried out in the article's proposed system are term frequency-inverse document frequency (TF-IDF) and vector space method (VSM). TF-IDF is used for extracting core words from documents, deciding search ranking and the relevance of the words (Rashid *et al.*, 2022; Kim and Gil, 2019). The three equations shown in Figure 1 are the TF-IDF equations used to determine the relevance of the words to a given document.

Figure 1.
TF-IDF equation

$$TF(t) = \frac{\text{(Number of times term } t \text{ appears in a document)}}{\text{(Total number of terms in the document)}} \quad (\text{EQ 1})$$

$$IDF(t) = \log_e \left(\frac{\text{Total number of documents}}{\text{Number of documents with term } t \text{ in it}} \right) \quad (\text{EQ 2})$$

$$W_{x,y} = tf_{x,y} \times \log \left(\frac{N}{df_x} \right) \quad (\text{EQ 3})$$

$tf_{x,y}$ = frequency of x in y
 df_x = number of documents containing x
 N = total number of documents

TF – IDF
Term x within document y

Note(s): A word phrase is considered essential if the TF-IDF score is higher
Source(s): Figure courtesy of Rashid *et al.*, 2022, p. 63

The data model used in designing the recommendation engine is the VSM, which represents documents and queries as vectors in a multi-dimensional space, where a term is a dimension (Gao *et al.*, 2023). To match an “item” and a “user,” the cosine similarity was calculated, where the cosine of the angle between two vectors is measured to determine their similarities. The “item” may refer to the available careers in the dataset, and the “user” may refer to the user profile.

The second content-based article is a job recommendation system based on embedded skills extracted from a job recruiting platform (Kara *et al.*, 2023). The article proposed a framework that may be compatible with unstructured curriculum vitae (CVs) and job JDs. The study collected JDs written in Turkish and English languages, where 450,000 were used to train a word2vec model, a model used to represent words in vector form (Jatnika *et al.*, 2019), and others were used for testing. The 7,700 CVs, which are also considered unstructured documents, were collected from those who are employed in the IT industry.

The skill extraction module in the proposed RS extracts all the skills, be they single-phrased or multi-word phrases, from the CVs and then matches them with the JDs using Skill Dictionary. The study trained only JDs in the word2vec model as CVs have more noise (Gupta and Gupta, 2019), and affect how the semantic relationship of the words were captured. The study implemented the continuous bag-of-words (CBOW) model, a technique to represent text in a computer-readable format and then convert the skills into vectors. The word2vec model was preferred over other language models since, in the article, the essence of a technical skill keyword is not dependent on the context (Sevastjanova *et al.*, 2021), and not all skill keywords are regarded as skills. Skill embeddings whose cosine similarity is less than a threshold (Rekabsaz *et al.*, 2017) are deleted from the extracted skill set. The article mentioned two approaches to assign similarity scores for each JD, namely the Word Mover’s Distance (WMD) approach, a technique for measuring the similarity of two documents (Sato *et al.*, 2022) and the cosine similarity approach.

3.3 Hybrid approach

The article to be explored is a recommendation system using Puppeteer and Representational State Transfer web crawling application programming interface (API) from Kumar *et al.* (2022). The data were collected from third-party job placement platforms and were divided into submodules. These submodules are company fetching modules, job listing fetching

modules and job listing platform crawling, which was initiated from a set of Uniform Resource Locators (URLs), and the breadth-first search (BFS) technique, which represents ranking references (Pedronette *et al.*, 2021) was used to extend the URL retrievals.

Other crawling techniques like ontology-based, API-based and HTML crawlers were used to fetch data from companies who do not post their job listings through third-party portals. Apart from the aforementioned modules, the article included a “data fields unifying module” to address the variations of data collected regarding the terminology used by various companies about JDs. Some companies used the term “job description”, while others would use “qualifications”, and a common schema was obtained.

The CBF part was implemented in two ways, where jobs are recommended in descending order based on the user’s profile. This was made possible through profile-description matching (Kaushal, 2020) and keyword-based searching (Rezaei and Fränti, 2014). For new accounts, profiles were built by explicitly asking them about their preferences. The documents are pre-processed, and the document vectors are computed using TF-IDF. The TF-IDF weight determination is based on the user’s interaction with jobs. The content score is calculated using cosine similarity.

For the CF part, similar users are classified into clusters. The system recommends to each user based on the preferences of its respective cluster. Jobs will be suggested to the active user based on the correlation and similarity between the two users. There are two approaches used in CF, user-based and item-based. In the user-based approach, users with collective preferences were grouped as they had common choices (Atoum, 2020; Liu *et al.*, 2022). The PCC measures the mutuality of the relationship between the user and the other users, where values are between zero and one. A zero signifies that it is not rated, and one is recommended. In the CF algorithm, a matrix indicating user-job interactions is created to measure similarity between the user and other users. The algorithm used is the Jaccard Similarity Coefficient or Tanimoto Coefficient, used to calculate similarity between two sets, and it effectively generates job recommendations. The equation is represented in Figure 2.

The average of the content score and collaborative score is computed as the total hybrid score. The combination of CBF and CF alone does not suffice when giving the best suggestions to the user. The article devised a hybrid recommendation system where a final recommendation score or the total hybrid score was generated for each job and then presented job rankings to users.

4. Narrative synthesis summary

The highlights of the study are summarized in the following tables.

The list of articles explored in the study is outlined in Table 1.

To address RQ1, the data collection methods implemented in each article were identified. Table 2 shows that Articles 1 and 2 collected data from primary sources, where questionnaires were distributed to their intended users. Collecting data from primary sources facilitates software developers to create a more personalized recommendation system. Articles 3, 4 and 5 have collected data from job recruiting platforms or career websites through web crawling. A web crawler is a script designed to retrieve web pages in a systematic, automated manner (Kausar *et al.*, 2013).

$$J(A, B) = \frac{A \cap B}{A \cup B}$$

Note(s): Measures the degree of similarities between two datasets, like collective preferences

Source(s): Figure courtesy of Kumar *et al.*, 2022, p. 7

Figure 2.
Mathematical
representation of
Jaccard similarity

Article 5 created submodules in its data collection process. These are (1) the company fetching module, where a list of top business organizations is being fetched; (2) the company detail gathering module, where parameters to filter out based on the funds, investors, size of the organization, unicorn status and their latest technology stack; (3) the job listing fetching module, where the sources of the job openings of shortlisted companies were harvested and aggregated in the database, and (4) job listing platform crawling to extend the URL retrievals. One of the benefits of web crawling is it will fetch or extract information from various websites in a customized way (Boppana and Sandhiya, 2021).

To address RQ2, Table 3 presents a synopsis of how suggestions are given to users. The implementation could be a classification algorithm, such as C4.5 (Alghobiri, 2018), or a ranking metric, such as Precision@K (P@K), which measures the degrees of relevance of documents being examined (Jairvelin and Kekailainen, 2017). Article 5 implemented hybrid filtering and utilized several metrics. For the CF algorithm, the PCC calculates the similarity of mutual users, and the Jaccard Similarity Index reveals the similarity between two sample sets of data. For the CBF algorithm, TF-IDF is measured to count the occurrences of terms in the document and cosine similarity to calculate the degree of similarity among the vectors.

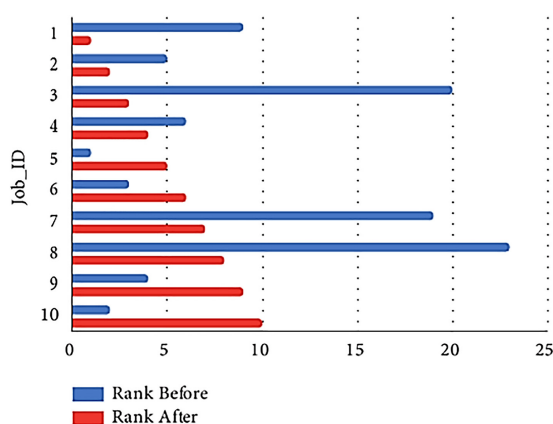
To address RQ3 in this study, the accuracy of the models being carried out was also tested, and the reliability of how accurate the models were also measured. For recommendation systems, a reliability test is used to weigh the confidence of users and information being presented and to establish a recommendation measure to improve accuracy (Bobadilla *et al.*, 2018). The dataset in Article 1 was categorized, where 70% were training datasets and 30% were model testing. The accuracy of the model was measured using the J48 decision tree classifier, and it yielded 78.84%. To test its reliability, a simulation error was computed from model testing using Kappa statistics (McHugh, 2012), and it resulted in 0.7839, which signifies reliability. If the value of Kappa is 1.0, then there is no discrepancy between the two measurements, suggesting a high level of accuracy.

The CV-JD pair in Article 4 was labeled as one or zero to approximate recommendations based on the job seeker's skills and skills stated in the JD. The suggestions given to each CV were the top ten JDs generated based on the P@K (Sujatha and Dhavachelvan, 2011). The P@K was generated using two methods at different k-values. The combined methods were TF-IDF with cosine similarity, skill embeddings with cosine similarity and skill embeddings with WMB. Based on the results, the combination of skill embeddings and WMD (a technique for measuring the similarity of two documents) surpasses the others as it generates a P@K of 0.95, having 1 as the k value. The semantic meaning of the skill phrases was captured using the word2vec model, and the ranking metrics are that the WMD model outperforms both cosine similarity-based RS and TD-IDF-based RS.

Article 5 presented a comparison between the existing models for job RS in the market and two other research papers. It also summarized top jobs being liked by particular users using CBF and hybrid scores. One of the significant results presented in the article is the evaluation of ranks after pre-processing the JDs, as it enhanced the probability of the user being matched to the desired job by 103.44%.

Hybrid filtering techniques surmount the weaknesses of the other and increase the ranking efficiency. It also eliminates cold start problems or sparsity of user-item ratings (Duricic *et al.*, 2018), specifically in CF techniques. Overall, the article realized the consolidation of quality recommendation systems through web crawling purposefully. As shown in Figure 3, the significant increase in the match score is evident in the comparison of job rankings when using content-based only and using the hybrid recommendations presented in the article.

While there could be several metrics to measure a system's reliability, assessment from the end user is also paramount. Article 2 implemented the International Organization for Standardization's (ISO's) product quality model, where different evaluators were given



Source(s): The figure represents the job rankings before and after implementing the proposed hybrid recommendations

Source(s): Figure courtesy of Kumar *et al.*, 2022, p. 7

Figure 3.
Comparison of job
rankings

different test case scenarios and were rated based on a Likert scale. Article 3 implemented a usability tool (Brooke, 2023).

The UAT implemented in Article 2 was based on the systems and software quality requirements and evaluation (SQuARE), which is an ISO 25010 framework on product quality model. The properties of a product quality model are based on functional suitability, performance efficiency, compatibility, usability, reliability, security, portability and maintainability, which are the fundamental principles of a product quality evaluation system (ISO, 2023). These properties, except maintainability, were measured by collecting data from significant representatives. Based on the interpretation in the descriptive statistics, where responses were coded after collecting the dataset, an average weighted mean was generated (Poncio, 2023). Based on the result, the functions in the system performed to a very great extent, given specific conditions at a certain period.

Article 3 used the system usability scale (SUS) (Brooke, 1995). The test garnered an average score of 81.25%, which implies that the system has a high degree of usability (Sauro, 2011).

5. Implications

The study benefits internship program coordinators of a university aiming to develop a recommender or matching system platform for their students. The content of the study may shed light on university decision-makers, explore options on what are the techniques or algorithms to be integrated into such platforms, conduct significant tests and provide keys to measure the level of accuracy and reliability of the models as presented in Tables 3 and 4.

One of the advantages of internship or industrial training programs is that they would help students align them with their career goals. Collecting from primary sources is more personalized and customized as to what motivates them to take internships, identifying areas for growth, courses they are good at, or results of psychological tests to evaluate their personalities. For secondary sources, data-fetching models could determine relevant embedded skill sets that have a considerable impact on one's career aspirations.

Table 4.
Usability quality

Article ID	Standards	Evaluators	Result	Remarks
2	ISO25010s SQuaRE framework	<ul style="list-style-type: none">• OJT Students• HR Supervisors• Practicum coordinator• IT experts	Average weighted mean is greater than 4 on a 1–5 Likert scale	Very good extent
3	System Usability Scale (SUS)	Final Year Computer Science students	Average SUS score of 81.25%	High degree of usability

Source(s): Tables by author

A hybrid recommender system may not eliminate the shortcomings of collaboration and CBF, but it could help mitigate the limitations of both. To effectively leverage the knowledge of students or job seekers, it is imperative to devise separate modules for the data collection and pre-processing. Analyzing transcripts of records or grades for major courses will also play a vital role in internship profile matching, as it leverages one’s skills and intellect. The effectiveness of the models does not only rely on the filtering techniques used but also on the metrics used to measure similarity, frequency of words or phrases used in a document or relationship of variables to name a few. The implementation of several usability quality tests will help improve further the functionality of the proposed system if the feedback of the evaluators is well addressed.

Internship programs can help a student grow personally and professionally. A university student looking for internship opportunities can find this a daunting task to undertake as there is a vast pool of opportunities offered in the market. The confidence levels to match their knowledge, skills and career goals with the JDs could be challenging. The same holds with companies, as finding the right people for the right job is a tough endeavor. The main objective of conducting this study is to identify models implemented in recommendation systems to give and/or rank suggestions given to users.

The study benefits internship coordinators of a university who are planning to create a recommendation system for internship programs, an innovative project to be used in teaching and learning.

6. Limitations

The study is focused on the three major filtering techniques used in recommendation systems. There are a few techniques like knowledge-based systems and other classification algorithms and ranking metrics.

Based on this paper’s observation, there are two missing key points. The first is statistics on the efficiency of the rankings provided by the recommender system by identifying the number of online users who clicked or accepted the suggestions given by the system, and second, the cost of conducting data extraction via web crawling or archiving these data.

7. Future work

Future work could be to determine what other underlying factors affect the level of accuracy of a model. Aside from assigning different values in a given set of parameters, host organizations may also receive suggestions, where “users” are JDs and the items to be recommended are CVs. Research studies have discussed other RS filtering techniques apart from the three major filtering techniques.

8. Conclusions

The existence of RS is becoming more relevant in disparate business workflows of any domain. In the academe alone, RS can be used to give suggestions of class schedules, majors to take for incoming freshmen or internship opportunities (de Silva *et al.*, 2023). The articles in this study demonstrated thought-provoking discussions about the recommendation techniques.

The CF technique provides predictions about the users' preferences through user and item entities. The predicted preferences, or items, may be given as suggestions to the user. The basis could be items preferred by other users. In this case, the cold start problem arises as there will be no basis or even point of reference for suggestions to be given. The issue about this particular filtering technique is data sparsity, affecting the accuracy of predictions (Shen and Jiamthapthaksin, 2016) as the amount of data increases. In CBF, the items are analyzed and classified based on their features or characteristics and then built a model to represent a user's data profile. One of the problems could be the relevance of the features being extracted. During the pre-processing, significant words might be removed from the corpus. This affects the efficiency of the model's learning process, which in turn may not provide personalized recommendations. Kumar *et al.* (2022) mentioned that hybrid recommendation systems surpass the limitations of both, but both filtering techniques have similar problems like the cold start. The article addressed this by explicitly asking the new user's preference.

In light of the techniques presented, the hybrid recommendation systems outdo the limitations and challenges of content-based and collaborative techniques. The effectiveness of a recommender system does not only rely on the techniques alone but also on other factors like data collection methods and identifying the relevant data to be collected. Pre-processing of data helps in the efficiency of the algorithms to be implemented as well. In conjunction with these three major filtering techniques are the implementation accuracy models like C4.5, a classification algorithm and P@K, a ranking metric. It is also imperative to measure the reliability of the accuracy models. The integration of web crawling techniques can also increase the match scores between JDs and user profiles.

References

- Adeosun, O., Shittu, A. and Owolabi, T. (2021), "University Internship Systems and preparation of young people for world of work in the 4th industrial revolution", *Ragagiri Management*, Vol. 22 No. 2, pp. 164-179, doi: [10.1108/ramj-01-2021-0005](https://www.emerald.com/insight/content/doi/10.1108/ramj-01-2021-0005/full/html), available at: <https://www.emerald.com/insight/content/doi/10.1108/ramj-01-2021-0005/full/html>
- Afoudi, Y., Lazaar, M. and Al Achhab, M. (2021), "Hybrid recommendation system combined content-based filtering and collaborative prediction using artificial neural network", *Simulation Modelling Practice and Theory*, Vol. 113, 102375, doi: [10.1016/j.simpat.2021.102375](https://doi.org/10.1016/j.simpat.2021.102375).
- Aksnes, D., Langfeldt, L. and Wouters, P. (2019), "Citations, citation indicators, and research quality: an overview of basic concepts and theories", *SAGE Open*, Vol. 9 No. 1, doi: [10.1177/2158244019829575](https://doi.org/10.1177/2158244019829575).
- Al-Bashiri, H., Kahtan, H., Abdulgabber, Romli, A. and Adam, M. (2019), "Memory-based collaborative filtering: impacting of common items on the quality of recommendation", *IJACSA*, Vol. 10 No. 12, doi: [10.14569/IJACSA.2019.0101218](https://doi.org/10.14569/IJACSA.2019.0101218).
- Al-walidi, N., Khamis, A. and Ramadan, N. (2020), "A systematic literature review of recommender systems for requirements engineering", *International Journal of Computer Applications*, Vol. 175 No. 14, pp. 31-41, doi: [10.5120/ijca2020920630](https://doi.org/10.5120/ijca2020920630).
- Alghobiri, M. (2018), "A comparative analysis of classification algorithms on diverse datasets. Engineering", *Technology and Applied Science Research*, Vol. 8 No. 2, pp. 2790-2795, doi: [10.48084/etasr.1952](https://doi.org/10.48084/etasr.1952).

- Atoum, I. (2020), "A novel framework for measuring software quality-in-use based on semantic similarity and sentiment analysis of software reviews", *Journal of King Saud University - Computer and Information Sciences*, Vol. 32 No. 1, pp. 113-125, doi: [10.1016/j.jksuci.2018.04.012](https://doi.org/10.1016/j.jksuci.2018.04.012).
- Bobadilla, J., Gutierrez, A., Ortega, F. and Zhu, B. (2018), "Reliability quality measures for recommender systems", *Information Sciences*, Vols 442-443, pp. 145-157, doi: [10.1016/j.ins.2018.02.030](https://doi.org/10.1016/j.ins.2018.02.030).
- Boppana, V. and Sandhya, P. (2021), "Web crawling based context-aware recommender system using optimized deep recurrent neural network", *Journal of Big Data*, Vol. 8, 144, doi: [10.1186/s40537-021-00534-7](https://doi.org/10.1186/s40537-021-00534-7).
- Brooke, J. (1995), "SUS: a quick and dirty usability scale", *Usability Evaluation in Industry*, Vol. 189, available at: https://www.researchgate.net/publication/228593520_SUS_A_quick_and_dirty_usability_scale
- Brooke, J. (2023), "SUS: a retrospective", *Journal of User Experience*, Vol. 8 No. 2, pp. 29-40, available at: <https://uxpajournal.org/sus-a-retrospective/>
- de Ruijt, C. and Bhulai, S. (2021), "Job recommender systems: a review", arXiv Labs, Vol. 2111, 13576, doi: [10.48550/arXiv.2111.13576](https://doi.org/10.48550/arXiv.2111.13576).
- de Silva, F., Slodkowski, B., da Silva, K. and Cazella, S. (2023), "A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations, and opportunities", *Education and Information Technologies*, Vol. 28 No. 3, pp. 3289-3328, doi: [10.1007/s10639-022-11341-9](https://doi.org/10.1007/s10639-022-11341-9).
- Duricic, T., Lacic, E., Kowald, D. and Lex, E. (2018), "Trust-based collaborative filtering: tackling the cold start problem using regular equivalence", *Association for Computing Machinery*, pp. 446-450, doi: [10.1145/3240323.3240404](https://doi.org/10.1145/3240323.3240404).
- Echarcharqy, S. (2020), "Learning by doing: an innovative method of teaching management disciplines in a higher business school", *International Journal of Scientific and Engineering Research*, Vol. 11 No. 9, pp. 34-40, available at: <https://www.ijser.org/onlineResearchPaperViewer.aspx?Learning-by-doing-an-innovative-method-of-teaching-management-disciplines-in-a-higher-business-school.pdf>
- Gao, L., Liu, Y., Chen, Q., Yang, H., He, Y. and Wang, Y. (2023), "A user-knowledge vector space reconstruction model for the expert knowledge recommendation system", *Information Sciences*, Vol. 632, pp. 358-377, doi: [10.1016/j.ins.2023.03.025](https://doi.org/10.1016/j.ins.2023.03.025).
- Gupta, S. and Gupta, A. (2019), "Dealing with noise problem in machine learning data-sets: a systematic review", *Procedia Computer Science*, Vol. 161, pp. 466-474, doi: [10.1016/j.procs.2019.11.146](https://doi.org/10.1016/j.procs.2019.11.146).
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. (2009), "The WEKA data mining software: an update", *SIGKDD Explorations Newsletter*, Vol. 11 No. 1, pp. 10-18, doi: [10.1145/1656274.1656278](https://doi.org/10.1145/1656274.1656278).
- ISO/IEC 25010:2023(en) (2023), "Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Product quality model", available at: <https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-2:v1:en>
- Jairvelin, J. and Kekailainen, J. (2017), "IR evaluation methods for retrieving highly relevant documents", *ACM SIGIR Forum*, Vol. 51 No. 2, pp. 41-48, available at: <https://sigir.org/wp-content/uploads/2017/06/p243.pdf>
- Jatnika, D., Bijaksana, M.A. and Suryani, A.A. (2019), "Word2Vec model analysis for semantic similarities in English words", *Procedia Computer Science*, Vol. 157, pp. 160-167, doi: [10.1016/j.procs.2019.08.153](https://doi.org/10.1016/j.procs.2019.08.153).
- Kara, A., Daniş, F.S., Orman, G.K., Turhan, S.N. and Özlü, Ö.A. (2023), "Job recommendation based on extracted skill embeddings", in Arai, K. (Ed.), *Intelligent Systems and Applications. IntelliSys 2022. Lecture Notes in Networks and Systems*. doi: [10.1007/978-3-031-16075-2_35](https://doi.org/10.1007/978-3-031-16075-2_35).
- Kausar, M., Dhaka, V. and Singh, V. (2013), "Web crawler: a review", *International Journal of Computer Applications*, Vol. 63 No. 2, pp. 31-36, doi: [10.5120/10440-5125](https://doi.org/10.5120/10440-5125).
- Kaushal, R. (2020), "User identity linkage: data collection, DataSet biases, method, control and application", Doctoral dissertation, Delhi, India, available at: <https://repository.iiitd.edu.in/jspui/bitstream/handle/123456789/831/Rishabh%20Kaushal%20thesis.pdf?sequence=2&isAllowed=y>

- Ken, G. (2023), "Testing and validating the cosine similarity measure for textual analysis", SSRN. doi: [10.2139/ssrn.4258463](https://ssrn.com/abstract=4258463), available at: <https://ssrn.com/abstract=4258463> (accessed 25 October 2022).
- Kim, S. and Gil, J. (2019), "Research paper classification systems based on TF-IDF and LDA schemes", *Human-Centric Computing and Information Sciences*, Vol. 9 No. 30, doi: [10.1186/s13673-019-0192-7](https://doi.org/10.1186/s13673-019-0192-7).
- Kirch, W., (eds) (2008), "Pearson's correlation coefficient", in *Encyclopedia of Public Health*, Springer, Dordrecht, available at: https://link.springer.com/referenceworkentry/10.1007/978-1-4020-5614-7_2569
- Kitchenham, B. and Charters, S. (2007), "Guidelines for performing systematic literature reviews in software engineering", EBSE Technical Report, available at: https://legacyfileshare.elsevier.com/promis_misc/525444systematicreviewsguide.pdf
- Kraus, S., Jones, P., Kailer, N., Weinmann, A., Chaparro-Banegas, N. and Roig-Tierno, N. (2021), "Digital transformation: an overview of the current state of the art of research", *SAGE Open*, Vol. 11 No. 3, doi: [10.1177/21582440211047576](https://doi.org/10.1177/21582440211047576).
- Kumar, N., Gupta, M., Sharma, D. and Ofori, I. (2022), "Technical job recommendation system using APIs and web crawling", *Computational Intelligence and Neuroscience*, Vol. 2022, pp. 1687-5265, doi: [10.1155/2022/7797548](https://doi.org/10.1155/2022/7797548), available at: <https://www.hindawi.com/journals/cin/2022/7797548/>
- Liu, Y. and Yang, S. (2022), "Application of decision tree-based classification algorithm on content marketing", *Journal of Mathematics*, Vol. 2022, pp. 1-10, doi: [10.1155/2022/6469054](https://doi.org/10.1155/2022/6469054).
- Liu, C., Kong, X., Li, X. and Zhang, T. (2022), "Collaborative filtering recommendation algorithm based on user attributes and item score", *Scientific Programming*, Vol. 2022, pp. 1-7, doi: [10.1155/2022/4544152](https://doi.org/10.1155/2022/4544152), available at: <https://www.hindawi.com/journals/sp/2022/4544152/>
- Lobanova, Y. (2021), "Relevance, design experience and analysis of the results of point implementation of interdisciplinary courses in the educational process at the university", *International Journal of Education and Information Technologies*, Vol. 15, pp. 237-244, doi: [10.46300/9109.2021.15.24](https://doi.org/10.46300/9109.2021.15.24).
- Lotfi, C., Srinivasan, S., Ertz, M. and Latrous, I. (2022), "Web scraping techniques and applications: a literature review", Raju Pal and Praveen Kumar Shukla (eds), *SCRS Conference Proceedings on Intelligent Systems*, India, SCRS, pp. 381-394, doi: [10.52458/978-93-91842-08-6-38](https://doi.org/10.52458/978-93-91842-08-6-38).
- McHugh, M. (2012), "Interrater reliability: the kappa statistic", *Biochem Med (Zagreb)*, Vol. 22 No. 3, pp. 276-282, doi: [10.11613/bm.2012.031](https://doi.org/10.11613/bm.2012.031), available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/>
- Mendez, J.S. and Bulanadi, J.D. (2020), "Job matcher: a web application job placement using collaborative filtering recommender system", *International Journal of Research Studies in Education*, Vol. 9 No. 2, pp. 103-120, doi: [10.5861/ijrse.2020.5810](https://doi.org/10.5861/ijrse.2020.5810).
- Mhamdi, D., Moulouki, R., El Ghoumari, M.Y., Azzouazi, M. and Moussaid, L. (2020), "Job recommendation based on job profile clustering and job seeker behavior", *Procedia Computer Science*, Vol. 175, pp. 695-699, doi: [10.1016/j.procs.2020.07.102](https://doi.org/10.1016/j.procs.2020.07.102).
- Nallamala, S.H., Bajjuri, U.R., Anandarao, S., Prasad, D. and Mishra, P. (2020), "A brief analysis of collaborative and content-based filtering algorithms used in recommender systems", *IOP Conference Series: Materials Science and Engineering*, Vol. 981 No. 2, 022008, doi: [10.1088/1757-899X/981/2/022008](https://doi.org/10.1088/1757-899X/981/2/022008).
- Ogunde, A.O. and Idialu, J.O. (2019), "A recommender system for selecting potential industrial training organizations", *Engineering Reports*, Vol. 1 No. 3, doi: [10.1002/eng2.12046](https://doi.org/10.1002/eng2.12046).
- Panigrahi, R. and Borah, S. (2018), "Rank allocation to J48 group of decision tree classifiers using binary and multiclass intrusion detection datasets", *Procedia Computer Science*, Vol. 12, pp. 323-332, doi: [10.1016/j.procs.2018.05.186](https://doi.org/10.1016/j.procs.2018.05.186).
- Papadakis, H., Papagrigoriou, A., Kosmas, E., Panagiotakis, C., Markaki, S. and Fragopoulou, P. (2023), "Content-based recommender systems taxonomy", *Foundations of Computing and Decision Sciences*, Vol. 48 No. 2, pp. 211-241, doi: [10.2478/fcds-2023-0009](https://doi.org/10.2478/fcds-2023-0009).

- Pedronette, D., Valem, L. and Torres, R. (2021), "A BFS-Tree of ranking references for unsupervised manifold learning", *Pattern Recognition*, Vol. 111, 107666, doi: [10.1016/j.patcog.2020.107666](https://doi.org/10.1016/j.patcog.2020.107666).
- Permana, A.A.J. and Pradnyana, G.A. (2019), "Recommendation systems for internship place using artificial intelligence based on competence", *Journal of Physics: Conference Series*, Vol. 1165, 012007, doi: [10.1088/1742-6596/1165/1/012007](https://doi.org/10.1088/1742-6596/1165/1/012007).
- Poncio, F. (2023), "An investigation of the gender gap in the information technology and engineering programs through text mining", *Decision Analytics Journal*, Vol. 6, 100158, doi: [10.1016/j.dajour.2022.100158](https://doi.org/10.1016/j.dajour.2022.100158).
- Rashid, A., Mohamad, M., Masrom, S. and Selamat, A. (2022), "Student career recommendation system using content-based filtering method", *International Conference on Artificial Intelligence and Data Sciences*, pp. 60-65, doi: [10.1109/AiDAS56890.2022.9918766](https://doi.org/10.1109/AiDAS56890.2022.9918766).
- Rekabsaz, N., Lupu, M. and Hanbury, A. (2017), "Exploration of a threshold for similarity based on uncertainty in word embedding", *Advances in Information Retrieval*, pp. 396-409, doi: [10.1007/978-3-319-56608-5_31](https://doi.org/10.1007/978-3-319-56608-5_31).
- Rezaei, M. and Fränti, P. (2014), "Matching similarity for keyword-based clustering", in *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 193-202, doi: [10.1007/978-3-662-44415-3_20](https://doi.org/10.1007/978-3-662-44415-3_20).
- Saeed, K., Keat, O.B. and Than, J. (2023), "Does internship moderate the relationship between critical thinking skills and graduate employability?", *Journal of Data Acquisition and Processing*, Vol. 38 No. 3, pp. 488-500, doi: [10.5281/zenodo.7922924](https://doi.org/10.5281/zenodo.7922924).
- Sato, R., Tamada, M. and Kashima, H. (2022), "Re-Evaluating word mover's distance", *Proceedings of the 39th International Conference on Machine Learning*, doi: [10.48550/arXiv.2105.14403](https://doi.org/10.48550/arXiv.2105.14403).
- Sauro, J. (2011), "Measuring usability with the system usability scale (SUS)", available at: <https://measuringu.com/sus/>
- Savickas, M. and Savickas, S. (2017), "Vocational psychology, overview", in *Reference Module in Neuroscience and Biobehavioural Psychology*, doi: [10.1016/B978-0-12-809324-5.05746-1](https://doi.org/10.1016/B978-0-12-809324-5.05746-1).
- Sevastjanova, R., Kalouli, A., Beck, C., Schäfer, H. and El-Assady, M. (2021), "Explaining contextualization in language models using visual analytics", *Association for Computational Linguistics*, Vol. 1, pp. 464-476, doi: [10.18653/v1/2021.acl-long.39](https://doi.org/10.18653/v1/2021.acl-long.39).
- Shen, F. and Jiamthapthaksin, R. (2016), "Dimension independent cosine similarity for collaborative filtering using MapReduce", *8th International Conference on Knowledge and Smart Technology*, Chiang Mai, Thailand, pp. 72-76, doi: [10.1109/KST.2016.7440484](https://doi.org/10.1109/KST.2016.7440484).
- Shepperd, M. (2013), "Combining evidence and meta-analysis in software engineering", in *In Book: Software Engineering*, Brunel University, London, pp. 46-70, doi: [10.1007/978-3-642-36054-1_2](https://doi.org/10.1007/978-3-642-36054-1_2).
- Sujatha, P. and Dhavachelvan, P. (2011), "Precision at K in multilingual information retrieval", *International Journal of Computer Applications*, Vol. 24 No. 9, p. 40, doi: [10.5120/2990-3929](https://doi.org/10.5120/2990-3929), available at: <https://www.ijcaonline.org/volume24/number9/pxc3873929.pdf>
- Thompson, M., Perez-Chavez, J. and Fetter, A. (2021), "Internship experiences among college students attending an HBC: a longitudinal grounded theory exploration", *Journal of Assessment*, Vol. 29 No. 4, pp. 589-607, doi: [10.1177/1069072721992758](https://doi.org/10.1177/1069072721992758).
- United Nations (2023), "The 17 goals", available at: <https://sdgs.un.org/goals>
- Velciu, M. (2018), "Matching skills and jobs: experience of employees in Romania", *New Trends and Issues Proceedings on Humanities and Social Sciences*, Vol. 4 No. 8, pp. 200-204, doi: [10.18844/prosoc.v4i8.3032](https://doi.org/10.18844/prosoc.v4i8.3032).
- Wang, D., Liang, Y., Xu, D., Feng, X. and Guan, R. (2018), "A content-based recommender system for computer science publications", *Knowledge-Based Systems*, Vol. 157, pp. 1-9, doi: [10.1016/j.knosys.2018.05.001](https://doi.org/10.1016/j.knosys.2018.05.001).
- Zeng, Z., Li, Y., Yong, J., Tao, X. and Liu, V. (2023), "Multi-aspect attentive text representations for simple question answering over knowledge base", *Natural Language Processing Journal*, Vol. 5, 100035, doi: [10.1016/j.nlp.2023.100035](https://doi.org/10.1016/j.nlp.2023.100035).

Further reading

- Gillis, A. (2022), "User acceptance testing", available at: <https://www.techtarget.com/searchsoftwarequality/definition/user-acceptance-testing-UAT>
- ISO/IEC 25010 (2022), available at: <https://iso25000.com/index.php/en/iso-25000-standards/iso-25010>
- Jaccard Score (2024), available at: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.jaccard_score.html
- Kappa Statistics (2021), available at: <https://www.sciencedirect.com/topics/medicine-and-dentistry/kappa-statistics>
- Min-max normalization (2024), available at: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781788393485/fd5b8a44-e9d3-4c19-bebb-c2fa5a5ebfee.xhtml>
- Cosine similarity (2012), available at: <https://www.sciencedirect.com/topics/computer-science/cosine-similarity>
- Puppeteer (2024), available at: <https://pptr.dev/>

Corresponding author

Flordeliza P. Poncio can be contacted at: exflore.it@gmail.com