

Predicting Graduate Students' Industry Readiness Courses Using Big Data and a Recommendation Engine*

Aravinda C V^{1,*†}, Mohamed Samiulla Khan^{2,‡}, Anjum Zameer Bhat^{3,§},
Dhanalakshmi Venugopal^{4,¶} and Dhivya Bino^{5,||}

¹Middle East College Oman, Affiliated to Coventry University, U.K , Faculty of Computer and Electronics Engineering Department

²Middle East College Oman, Affiliated to Coventry University, U.K , Faculty of Computer and Electronics Engineering Department

³Middle East College Oman, Affiliated to Coventry University, U.K , Faculty of Computer and Electronics Engineering Department

⁴Middle East College Oman, Affiliated to Coventry University, U.K , Faculty of Computer and Electronics Engineering Department

⁵Middle East College Oman, Affiliated to Coventry University, U.K , Faculty of Computer and Electronics Engineering Department

Abstract

Assistance machines have risen as a response to added information, which has become a challenge in terms of value and volume in terms of the number of hours spent on their search and the quantity of knowledge acquired. As mentioned, the significance of recommended educational resources plays a significant role in improving a student's learning process, highlighting the need to guarantee relevant, valuable information precisely and consistently as part of education recommendations. Big data analytics methodologies are currently being used to evaluate these educational data and provide various recommendations and suggestions for students, instructors, and universities. This research recommends education courses acquired from previous grades utilizing collaborative filtering-based recommendation algorithms. The outcomes from this study could be used by schools, colleges, and universities to advise alternative optional courses to students. It is the primary purpose of this methodical review to examine previous work on recommendations infrastructures aimed at supporting pedagogical approaches to discover what topics and types of education are addressed, what developmental approach is utilized, and what materials are recommended. It also aims to identify any gaps that are present for future courses that will serve professionals.

Keywords

Methodical review, Suggestion systems Education, Machine learning, Elective courses, Higher education,

1. Introduction

This study focusing on students pursuing graduate study abroad after completing their undergraduate studies is significant. A fully funded opportunity is an extremely competitive process, and many students do not get the opportunity to pursue graduate studies. Academic history and standardized testing scores are used to select students for admission to universities throughout the world. Students are accepted to the institution based on their academic performance. Since the introduction of smart devices, education has been reformed into new marketplaces centered on mobile commerce. The volume of data that can be gathered has significantly increased, so users have far more options for gaining several types of information. The wide information resulted from the rapid expansion of the social network. People face difficulty with what exactly they are seeking the data information. Because they participated in social surveys on educational websites, students have recently been able to actively share their reviews and obtain discounts. Given the recommender system, proposes items are most likely to be of interest to individuals [1]. They're a useful method to help users search through a wide

The 7th International Symposium on Advanced Technologies and Applications in the Internet of Things (ATAIT 2025), September 10-11, 2025, Kusatsu, Japan

*Aravinda C V.

†These authors contributed equally.

✉ aravinda@mec.edu.om (A. C. V); mkhan@mec.edu.om (M. S. Khan); azameer@mec.edu.om (A. Z. Bhat);
dhanalakshmi@mec.edu.om (D. Venugopal); dhivya@mec.edu.om (D. Bino)

🌐 <https://mec.edu.om/en//> (A. C. V); <https://mec.edu.om/en//> (M. S. Khan); <https://mec.edu.om/en//> (A. Z. Bhat);
<https://mec.edu.om/en//> (D. Venugopal); <https://mec.edu.om/en//> (D. Bino)

🆔 0000-0003-4588-4105 (A. C. V)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

range of options to find the ones which are most likely to be selected. The field of recommender systems is rapidly expanding, and these devices are extensively being used in a wide range of disciplines [2]. They utilize algorithms that consider a variety of factors, including the user's browsing habits, searches, transactions, and preferences. [3]. Collaborative filtering is a technique that suggests items based on similarity. Collaborative filtering comes in two ways: (I) User-based and (II) item-based. The user-based filtering algorithm allows the consumers to see interesting information they would like by presuming they would approve. To find the user's neighbors based on user similarity, the algorithm will use either supervised or k-nearest or unsupervised learning such as the k-means technique. [4][5]. The main goal of the study is to create a system of recommendations for graduate admissions candidates that will look at past data from graduate students who are already enrolled and use that information to assist candidates in selecting the best university for them based on their academic record, academic standing, and test scores. Recommender systems can help mentors in identifying solutions to their recorded requirements in a professional development system where teachers are responsible for their personal development, as described further below. Instructors sought forces and people who other teachers found interesting and useful, as well as recommendations for things that fit their interests and aims [6]. There 1 shows the process of self-training methods as shown in Figure 1.

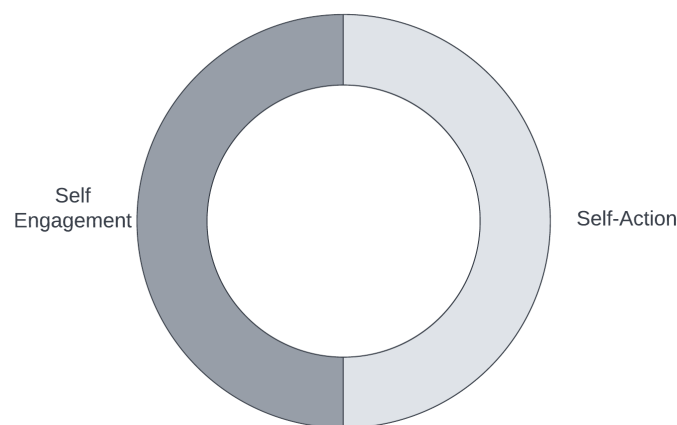


Figure 1: Self-Training Process

2. Literature Suvery

The method of entry has been the subject of many studies, but few have employed the Machine Learning domain to aid in the decision-making process for admissions to universities. There is a demand for a recommendation system that can aid students in choosing the best graduate school for their academic and career goals. Learning path recommendations are one example: although agency portrays the individual as competent and in control of his or her learning and asked if learning path suggestions would hinder or encourage agency . The most popular ones are hybrid recommendation systems, collaborative filtering, and content-based systems. Content-based recommendation algorithms promote products that the consumer had also previously enjoyed. The very first believes that if a user likes the item, he or she will prefer related objects, while the second recommends items based on features that match the user profile. To create a suggestion for a specific user, collaborative filtering-based recommender systems use the preferences of other users. Collaborative systems get their name from the fact that they consider two items (books, movies, etc.) to be linked because many other users appreciate them, rather than evaluating all the items' attributes. Artificial intelligence approaches such as Bayesian algorithms, artificial neural networks, and machine learning algorithms, as well as genetic algorithms and fuzzy set algorithms, have been frequently employed. According to, employing these tactics to

usher in the Big Data era is a potential solution [7]. As a result, educational recommendation systems face a challenging problem in deciding users' interests because they are strongly dependent on the area in which they work [8]. The use of RS in education, including teaching and academic advisory services, is the focus of this systematic review [9][10]. Burke classified the several sorts of collaboration as follows in an overview of six types of suggestion techniques: knowledge-based filtering, contextual filtering, demographic filtering, and hybrid filtering [11]. Customary search engines are trying to keep up with the growing number of educational materials by serving the needs of students surfing the web for products and services while learning [12]. The study proves how deeply ingrained collective filtering is, which involves evaluating items while considering the opinions of others. [13][14][15]. By selecting the most significant information based on a user's preferences and interests, the recommender system sorts through a huge amount of information. It decides whether a user and an item are compatible, as well as the similarities between users and objects, to make recommendations [16][17]. Recent work in that area encouraged to conduct of research on graduate school suggestion systems, which will deeply aid learners in their search for graduate programs. Irrespective of whether collaborative, content-based, or hybrid approach recommendation filtering is used, different vectors or mathematical rules and formulas are used for the data [18][19]. The suggestions are the conclusion of this process. Suggestion engines are a strong promoting tool with the talent to increase proceeds, press-through rates, conversions, and even customer pleasure. Learn how to use recommendations to effectively upsell, cross-sell, and grow your business.

3. Phases of the recommendation process

1. **Phase 1: Information Gathering:** The most valuable type of feedback involves explicit feedback, which is the learner's specific input on the value of the item, or strategy feedback, which is implicitly calculated and evaluated by observing the behavior of the user. An E-learning platform's user profile is a collection of personal information about the user that does not consist only of explicit feedback but also of implicit feedback.
2. **Phase 2: Explicit Criticism:** A user's layout will be developed by the computer using the framework interface, which the user is asked to input. How many reviewers are assigned to a consumer will determine how consistent the input is. It is believed that unambiguous input requires more user effort, but because it does not require expecting behaviors, it provides more trustworthy data.
3. **Implicit Feedback:** Based on the user's previous actions, including history, browsing history, time spent on particular web pages, user-followed links, channeled links, and email information, the system automatically calculates and predicts the user's preferences. An implicit user relieves the user's burden by inferring what the user expects from the software based on their actions. As a result, it is less effective since there is no human involvement.
4. **Phase of Instruction:** The features that were gathered from user feedback during the data collection phase will be filtered and developed during this phase by a learning algorithm.
5. **Phase of Recommendation :** The client's final product selection is hinted at or implied. The data gathered during the data collection process either influence how the user perceives the device's activity or how the user reacts to the data gathered during the data collection process. Figure.2 depicts the phase of suggestion and Figure 3 indicates the functional flow of the model.

4. Recommendation System

1. **Filtering based on content:** A content-based method is a domain-specific method that focuses on examining object attributes to produce predictions. In most cases, records such as web pages, journals, and news can be screened most effectively. Users' profiles and past ratings are used for content-based filtering, which makes decisions based on their characteristics.

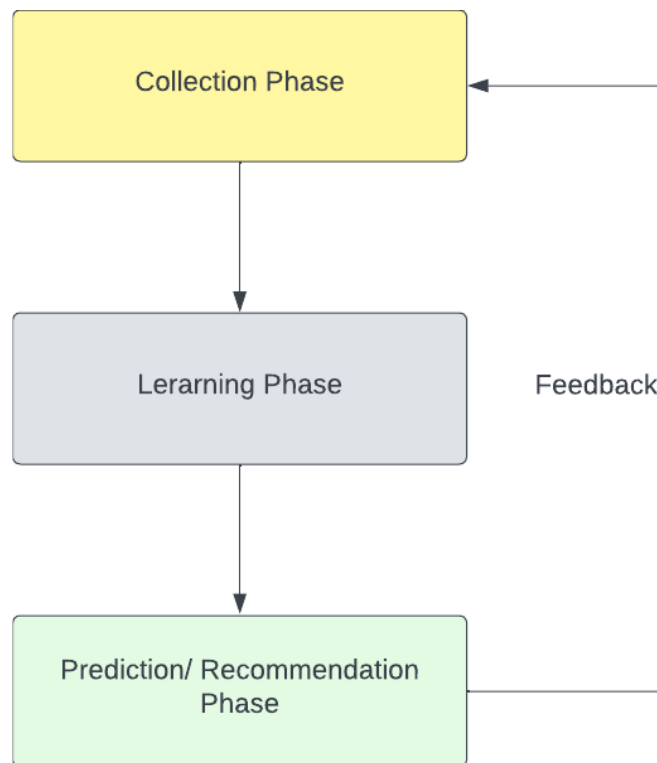


Figure 2: Personalized collection process

2. **Content-based filtering benefits and drawbacks:** CBF solutions can help resolve your CF issues. There is still a possibility of getting new things even though buyers don't reveal their scores. The quality of the suggestions does not change even if the database is empty of user preferences.
3. **Collaborative Filtering:** - Recommendation systems bring a significant benefit to businesses and users. They offer reliable recommendations tailored to suit users' needs. E-commerce sites have utilized this filtering to customize every browser's experience by collecting their recommendation system. As depicted in Table:1, this filtering is commonly used as an approach in recommendation systems.
4. **Advantages and Disadvantages of Collaborative Filtering Techniques:** In comparison to CBF, collaborative filtering works well in contexts where there are fewer information connections with objects, and objects are difficult to understand, such as beliefs and ideals. The CF approach can give precise and correct recommendations without knowing the user's profile information, which means it can suggest things that are deemed relevant to the user.
5. **Collaborative filtering based on user input:** The collaborative filtering strategy based on user input aims to predict the path of a target user who has previously expressed interest, as well as additional users who are similar to the target user.

5. Methodology

- **An Overview of Collaborative Filtering:** In general, collaborative filtering consists of three steps: gathering the user rating data matrix, picking similar neighbors based on rating similarity, and finally generating predictions.
- **Data Input User Rating Score:** User data is typically represented by rows and columns of a recommendation system based on CF technology. Users are rows, and they count lists representing sent the columns' reach as mentioned in Table: 2

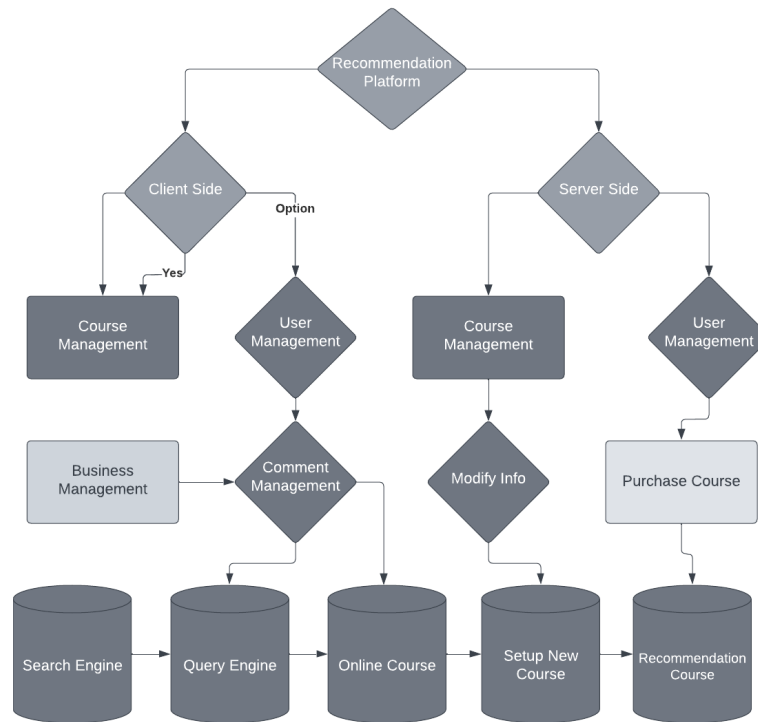


Figure 3: Flow of System Functional Model

Table 1

Essential knowledge and skills

Programs	Data Analyst	M.L	Data-Eng	Data-Scientist
Tools	Top-ranking	Top-ranking	Top-ranking	Top-ranking
Visualization	Top-ranking	Moderate ranking	Moderate ranking	Top-ranking
Intuition	Low-ranking	Top-ranking	Moderate ranking	Moderate ranking
Statistics	Low-ranking	Top-ranking	Moderate ranking	Top-ranking
ML	Low-ranking	Top-ranking	Low-ranking	Top-ranking
Calculus	Low-ranking	Top-ranking	Low-ranking	Moderate ranking
Software.	Engineering	Low-ranking	Moderate ranking	Low-ranking

Table 2

Essential knowledge and skills

Course/Users	C1	C2	C3	...	CS
U1	R1,1	R1,2	R1,3		R1,n
U2	R2,1	R2,2	R2,3		R2,n
U3	R3,1	R3,2	R3,3		R3,n
Ur	Rur,1	Rur,2	Rur,3		Rur,n

- **Subject-based collaborative filtering (SBCF):** In this filtering technique, the approach is objective to predict the subjects based on the similarities they are to the other courses which are familiar to the other students as shown in Figure 4

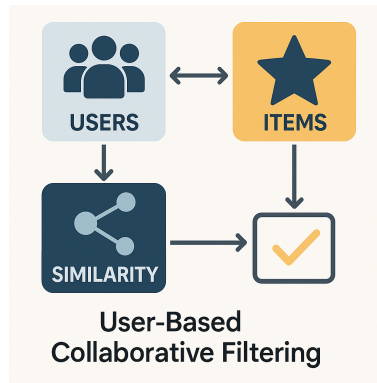


Figure 4: User-based collaborative filtering

5.1. Process of Collaborative Filtering:

This processing phase is mainly divided into three steps: (I) Collecting the data (ii) Finding the similarities (iii) implementing the collaborative filtering. Creating prediction and the recommendation flow are shown in Figure 5 and the process of content-based filtering is shown in Figure 6

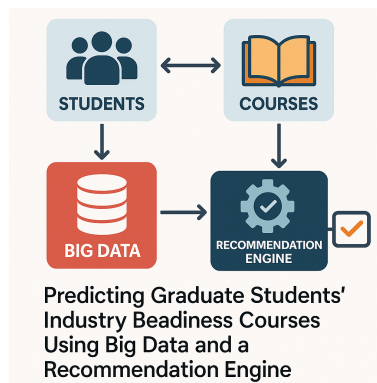


Figure 5: Process of Collaborative Filtering

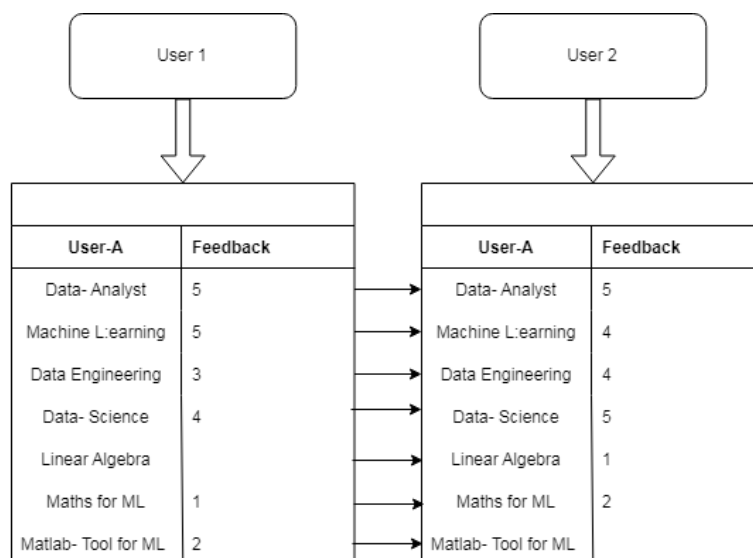


Figure 6: Process of Feedback Collaborative Filtering

5.2. Cognitive strategy:

This strategy is also considered a cognitive technique because all this information is taken into consideration when suggesting or recommending products or services based on an earlier search, and explicit feedback.

6. Proposed System:

This paper suggests that the Hadoop Framework could be used to create an automated recommendation system based on the amount of data manually collected from existing circuit branches from the fifth, sixth, seventh, and eighth semesters and from the passed-out classes of about 1000 students in terms of ratings, reviews, opinions, complaints, remarks, feedback, and comments about the classes. (Regular-course, placement-helpful, individual, and other) in the form of ratings, reviews, opinions, complaints, remarks, feedback, and comments about any courses (regular-course, placement-helpful, individual, and other). A combination sorting strategy was suggested to filter various kinds of evaluations, opinions, notes, complaints, and so on. We investigated a recommendation based on numerical factors, such as rating system and ranks, for a wide range of courses, even though different reviewers generate recommendations based on different factors, including such ratings, ranks, content, reviewer activity, and review timing and the flow is as shown in Figure 7.

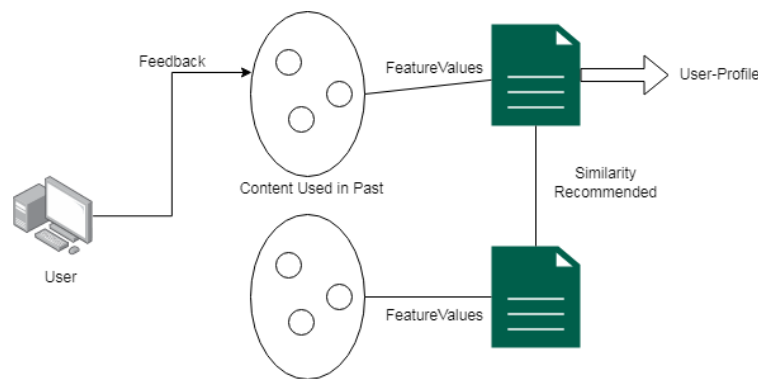


Figure 7: Process of Content-Based Filtering

6.1. Recommendation System Processes. ETL: Extract, Transform and Load

- " Key-point Creation "
- " Algorithms techniques for Recommendation /Model Generation "
- " Flows and Scheduling"
- " Model testing assistance "
- "Evaluates Effectiveness'

6.2. Hadoop Framework:

Hadoop is a framework for storing, managing, and retrieving substantial amounts of data, which is referred to as Big Data. This framework proves how the Hadoop framework may be used to integrate a variety of high-level languages, predictive analysis algorithms, and other tools for various tasks. Our recommendation database is configured with Mahout and Hive within the Hadoop framework as specified processes. HDFS is a distributed file system suitable for a variety of hardware platforms. It is fault-tolerant and can use on low-cost hardware. MapReduce is a programming model that divides the work into a series of independent tasks to manage massive amounts of data in parallel. Hive is a Hadoop-based data warehousing application. HiveQL, a SQL-like query language, is available. Apache

Table 3

Essential knowledge and skills

Prediction-based	Information-Retrieval
Absolute-Difference-Evaluating	Generic-Recommender
The method used: evaluate ()	The method used: evaluate ()
Data Model-Training set	Data Model-Relevance
Training Set size 70%	Threshold (mean + standard deviation)

Table 4

Essential knowledge and skills

Filename	File-Size	CPU-Time-Consumed
The datasets	Generic-Recommender	
The method used: evaluate ()	255bytes	168403 ms
C1.csv	25MB	631717 ms
C2.csv	105MB	1408669 ms
ratings.csv	125MB	1684505 ms

Mahout is an open machine learning toolkit that is designed to help analyze copious amounts of data in a distributed manner using machine learning and data mining tools.

6.3. Similarity and Neighborhood Measures

- ‘Categorization ‘,
- ‘Partitioning ‘,
- ‘Recommender / Collaborative ‘,
- ‘Filtering ‘,
- ‘Sympatric speciation of Algorithms ‘,
- ‘Pattern Mining ‘,
- ‘Regression ‘,
- ‘Dimension reduction ‘,
- ‘Similarity Vectors’,
- ‘Similarity Measures ‘,
- ‘Pearson Correlation ‘,
- ‘Spearman Correlation ‘,
- ‘Euclidean Distance ‘,
- ‘Log-Likelihood Similarity ‘,
- ‘Neighborhood Measures ‘,
- ‘Nearest N Users Algorithm’

The best part of Mahout is that it supplies a standard interface for the evaluation of a recommendation system. Evaluation of different implementations is very time-consuming the result is shown in Table 3

7. Experimental Analysis:

In the experimental study, we developed a recommendation system based on a dataset of course ratings provided by different users, which was analyzed by a Mahout and analyzed using smaller size file sets and the result is as shown in Figure 8 and 9 and in Table 4 the CPU execution time with different dataset sizes is mentioned.

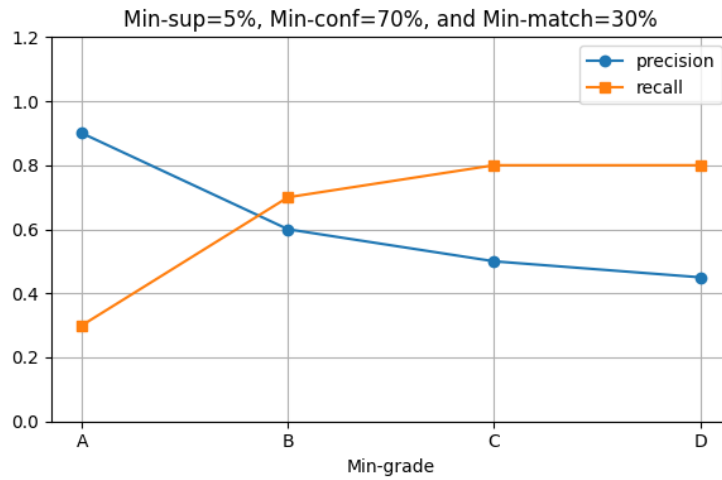


Figure 8: Separating match values focused on the smallest match outcomes

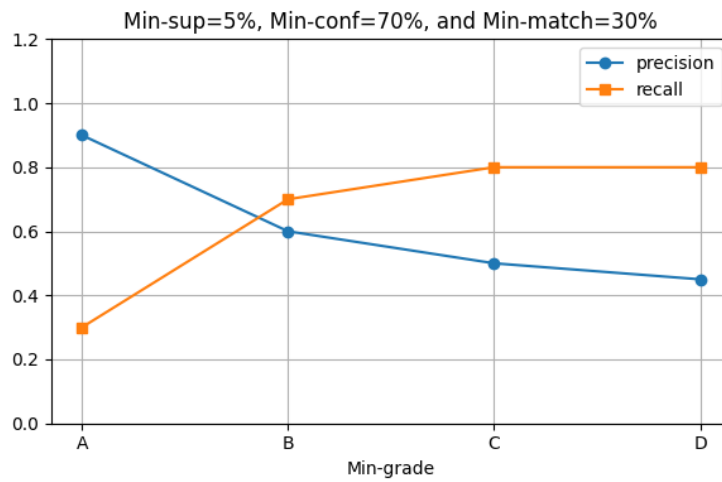


Figure 9: Separating the smallest score values used to assess performance

Table 5

A Percentage number of courses according to the type of platform

Type of platform	Course list	Total-Course in %
Online	50	50.45%
Chatbot	1	1.0%
Mobile	1	1.5%
Moodle	50	50.55%

7.1. Recommendation systems use a variety of technologies

The data was used from an online platform for implementation purposes for the studies reviewed, and we used the internet, chatbot, and mobile as data sources, as shown in Table 5.

7.2. Metrics of quality:

The matching evaluation was performed for each of the courses using the defined indicators to decide the assessment connected to the quality metrics, as given in Table 6.

Table 6

A Percentage number of courses according to the type of platform

Semester	Metrics
Final year	Accuracy/Recall
Pre-final-year	Accuracy/Recall
5th-sem	Mean absolute error
3rd-sem	The harmonic means between precision and recall

8. Conclusion and Future work

Reviews, opinions, feedback, remarks, and complaints are treated as Big Data in the recommendation system and cannot be analyzed directly. These data must be filtered/translated first. We implemented a recommendation system for the course dataset using the Hadoop framework and the industry-ready course dataset, employing filtering techniques and analyzing data of varying sizes. We propose to apply a weighting on the rating of items using review summaries and opinions as a future enhancement of this work. There are several approaches to improve the Next Framework within which effort from involving participation. Users could filter their courses based on their career levels, such as graduate or undergraduate. In contrast, if we kept a record of the topics a student already has taken, we could eliminate those from the recommendations. From a planning perspective, it would be beneficial to develop course patterns based on recommendations. We could also improve the framework of our recommendations so that the user can identify which recommendations are based on skills extracted from context and which are based on relevant professional emerging skills. This module converts a basic course recommendation system into a tool to find both required skills and programs. It educates students with the knowledge they have to make an informed decision about their academic objectives and to understand exactly what is needed of them to place the career of their ambitions.

9. Conflict of Interest

The authors do not have any conflict of interest.

Acknowledgment

This work was supported by the Middle East College, Oman, affiliated with Coventry University, UK, Faculty of Computer and Electronics Engineering, Department.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] H. Drachsler, K. Verbert, O. C. Santos, N. Manouselis, Panorama of recommender systems to support learning, in: *Recommender Systems Handbook*, Springer US, Boston, MA, 2015, pp. 421–451.
- [2] M. D. Ekstrand, Collaborative filtering recommender systems, *Found. Trends® Hum.–Comput. Interact.* 4 (2011) 81–173.
- [3] H. Drachsler, H. G. K. Hummel, R. Koper, Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model, *Int. J. Learn. Technol.* 3 (2008) 404.

- [4] G. Sarwar, J. Kaypi, J. Konstan, Item-based collaborative filtering recommendation algorithms, in: At the 10th International World Wide Web Conference, 2001, pp. 285–295.
- [5] S. Gong, A collaborative filtering recommendation algorithm based on user clustering and item clustering, *J. Softw.* 5 (2010).
- [6] J. Tang, S. Wu, J. Sun, H. Su, Cross-domain collaboration recommendation, in: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '12, ACM Press, New York, New York, USA, 2012.
- [7] U. Nadine, H. Cao, J. Deng, Competitive recommendation algorithm for e-commerce, in: 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), IEEE, 2016.
- [8] A. H. Alharbi, A. C. V, M. Lin, B. Ashwini, M. Y. Jabarulla, M. A. Shah, Detection of peripheral malarial parasites in blood smears using deep learning models, *Comput. Intell. Neurosci.* 2022 (2022) 3922763.
- [9] L.-Z. Cui, F.-L. Guo, Y.-J. Liang, Research overview of educational recommender systems, in: Proceedings of the 2nd International Conference on Computer Science and Application Engineering - CSAE '18, ACM Press, New York, New York, USA, 2018.
- [10] M. Hassan, M. Hamada, Smart media-based context-aware recommender systems for learning: A conceptual framework, in: 2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET), IEEE, 2017.
- [11] G. Padmajadevi, K. S. Aprameya, Comparison of offline signature verification systems with support vector machines and neural network classifiers using geometrical (2017) 16181–16190.
- [12] W. Yuan-hong, T. Xiao-qiu, A real-time recommender system based on hybrid collaborative filtering, in: 2010 5th International Conference on Computer Science & Education, IEEE, 2010.
- [13] C. V. Aravinda, An approach for signature recognition using the contours-based technique, in: “2019 International Conference on Advanced Mechatronic Systems (ICAMechS), 2019, pp. 2019–2027.
- [14] C. V. Aravinda, L. Meng, K. R. Uday Kumar Reddy, A. Prabhu, Signature recognition and verification using multiple classifiers combination of hu's and HOG features, in: 2019 International Conference on Advanced Mechatronic Systems (ICAMechS), IEEE, 2019.
- [15] C. V. Aravinda, Offline signature recognition and verification using ORBKey point matching techniques”, *Adv. Sci. Technol. Eng. Syst. J* 5 (2020) 1–7.
- [16] D. Singh, C. V. Aravinda, M. Kaur, M. Lin, J. Shetty, V. R. Reddicherla, H.-N. Lee, Dknet: Deep kuzushiji characters recognition network, *IEEE Access* 10 (2022) 75872–75883. doi:10.1109/ACCESS.2022.3191429.
- [17] N. Nojiri, Z. Meng, K. Saho, Y. Duan, K. Uemura, C. V. Aravinda, G. A. Prabhu, H. Shimakawa, L. Meng, Apathy classification based on doppler radar image for the elderly person, *Front. Bioeng. Biotechnol.* 8 (2020) 553847.
- [18] A. H. Alharbi, C. V. Aravinda, J. Shetty, M. Y. Jabarulla, K. B. Sudeepa, S. K. Singh, Computational models-based detection of peripheral malarial parasites in blood smears, *Contrast Media Mol. Imaging* 2022 (2022) 9171343.
- [19] C. V. Aravinda, R. Bhatnagar, V. R. Reddicherla, Automated quality control framework for solanum lycopersicum using image processing, in: 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, 2022.