Received 14 September 2023, accepted 1 November 2023, date of publication 24 November 2023,
date of current version 8 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3336676

RESEARCH ARTICLE

Developing a Personalized E-Learning and MOOC Recommender System in IoT-Enabled Smart Education

SAMINA AMIN[1], M. IRFAN UDDIN[1], (Member, IEEE), WALI KHAN MASHWANI[2], ALA ABDULSALAM ALAROOD[3], ABDULRAHMAN ALZAHRANI[4], AND AHMED OMAR ALZAHRANI[5]

[1]Institute of Computing, Kohat University of Science and Technology (KUST), Kohat 26000, Pakistan
[2]Institute of Numerical Sciences, Kohat University of Science and Technology (KUST), Kohat 26000, Pakistan
[3]College of Computer Science and Engineering, University of Jeddah, Jeddah 21493, Saudi Arabia
[4]Department of Information Systems and Technology, College of Computer Science and Engineering, University of Jeddah, Jeddah 21493, Saudi Arabia
[5]Faculty of Computer Science and Engineering, University of Jeddah, Jeddah 21493, Saudi Arabia

Corresponding author: M. Irfan Uddin (irfanuddin@kust.edu.pk)

This work was supported by the University of Jeddah, Jeddah, Saudi Arabia, under Grant UJ-22-DR-63.

ABSTRACT Smart strategies and intelligent technologies are enabling the designing of a smart learning environment that successfully supports the development of personalized learning and adaptive learning. This trend towards integration is in line with the growing prevalence of Internet of Things (IoT)-enabled smart education systems, which can leverage Machine Learning (ML) techniques to provide Personalized Course Recommendations (PCR) to students. Current recommendation techniques rely on either explicit or implicit feedback, often failing to capture changes in learners' preferences effectively as they integrate both types of feedback. This paper proposes a new model for personalized learning and PCR that is enabled by a smart E-Learning (EL) platform. The model aims to gather data on students' academic performance, interests, and learning preferences, using this information to recommend the most beneficial courses for each student. Our approach suggests courses based on the learner's interactions with the system and the cosine similarity of related content, combining explicit (user ratings) and implicit (views and behavior) methodologies. The method employs various ML algorithms and an EL Recommender System (RecSys) based on Collaborative Filtering (CF), including Random Forest Regressor (RFR), Decision Tree Regressor (DTR), K-Nearest Neighbors (KNN), Singular Value Decomposition (SVD), eXtreme Gradient Boosting Regressor (XGBR), and Linear Regression (LR). To evaluate our proposed solution, we benchmark it against existing approaches in terms of predictive accuracy and running time. Experiments are conducted using two benchmark datasets from Coursera and Udemy. The proposed model outperforms existing top-K recommendation techniques in terms of accuracy metrics such as precision@k, Mean Average Precision (MAP)@k, recall@k, Normalized Discounted Cumulative Gain (NDCG)@k, Mean Squared Error (MSE)@k, Root Mean Squared Error (RMSE)@k, and Mean Absolute Error (MAE)@k for PCR. The results show that SVD performs particularly well, demonstrating higher precision, recall, MAP, and NDCG along with lower MAE, RMSE, and MSE values compared to other proposed algorithms. This success can be attributed to SVD's ability to capture complex interactions between students and courses. Our proposed solutions exhibit promise across two datasets and can be applied to various RecSys domains.

INDEX TERMS Machine learning, collaborative filtering, recommender system, e-learning, Internet of Things, smart education, MOOC.

I. INTRODUCTION

The purpose of education has been to gradually enhance the atmosphere of smart learning and techniques throughout

The associate editor coordinating the review of this manuscript and approving it for publication was Anandakumar Haldorai.

the last few years. Another element that accelerates learning is the use of different techniques. Students' progression is boosted by ongoing performance evaluation and monitoring employing a variety of sensing devices and Internet-connected devices. Everyone has several possibilities for achieving high standards and novel developments due to smart education. Furthermore, IoT advancements are expected to transform the traditional educational system into a smart learning environment. IoT has the power to compel the interconnection of items and the use of sensed data in smart education. The use of IoT can improve the learning involvement of students, enabling them to study quickly, successfully, and pleasantly [1], [2], [3].

Technology has become an essential component of the education sector, contributing to the enhancement of educational quality and the promotion of knowledge sharing among students worldwide via various platforms [4], [5], [6]. EL, a form of education delivered through the Internet, has emerged as an effective means of information exchange and learning. EL platform enables learners to enhance their knowledge and skills without needing to attend physical classes or be present at an educational institution due to factors such as the inaccessibility of institutions or their remote locations [7]. However, IoT-enabled smart education systems can be incredibly useful in providing PCR to students using ML techniques. The system can gather data about students' academic performance, interests, and learning preferences, and use this data to suggest the most appropriate courses for them. In real-time data collection, IoT can facilitate remote learning by enabling students to access educational content and resources from anywhere. This includes online courses, video conferencing tools, and virtual laboratories. In the proposed work, we use online course data extracted from MOOC platforms, including Udemy and Coursera.

The Massive Open Online Courses (MOOC) platform's learning materials are growing daily as a result of the digitization and network exchange of educational materials [8]. On the one hand, learners can always find the curriculum resources they actually need when it comes to various educational materials; due to learners' varying levels of expertise, it might be challenging for them to accurately recognize the content of the materials [9]. Because they cannot quickly eliminate the learning resources they are interested in, students get lost or make ill-informed decisions wasting extra time. The MOOC system places a lot of emphasis on the course or video suggestion because it enables students to find materials for their learning needs. The MOOC platforms can successfully assist learners in coping with issues like resource overload and knowledge fragmentation by recommending instructional materials that match their preferences and requirements.

Conventional EL platforms rely on static content and assume that all learners are homogeneous, making it impossible for them to meet the requirements of every single student. These platforms are less adaptable, and if a platform supporting a specific method has been created and put into place, it is less likely to evolve in response to how students engage and what they prefer [10].

The quick transition to smart education, especially in the age of pandemics, makes the use of EL, remote learning platforms, and hybrid models necessary. It becomes necessary to develop adaptive and personalized education to lessen the drawbacks of distance learning and keep high standards of accomplishment. ML, the Internet of behavior, and explainable AI are just a few of the technologies influencing the direction of smart education in the era of EL and MOOC through personalization and customization [11].

Due to the emergence of COVID-19 and the subsequent closure of educational institutions, authorities proposed alternative methods for learning in emergencies to ensure students continue their education and prevent the spread of the virus [12]. Traditional educational methods were replaced by EL, also known as online learning, as social gatherings in educational institutions were deemed a high-risk activity for the spread of the virus. EL has gained significant attention in recent years and has become a crucial means of acquiring knowledge, not only for students in schools and universities but also for lifelong learners seeking to improve their social and workplace skills [13]. Coursera is an example of a MOOC platform that offers courses, degrees, programs, and credentials from renowned institutions and businesses including Stanford, Harvard, Google, and IBM. MOOC offers online courses and materials spanning a variety of disciplines like computer science, mathematics, and business. Learners must choose the courses and materials that will help them advance their careers and remain competitive in the global economy from the many options accessible [14]. Udemy[1] is a MOOC platform that offers a wide range of courses on various topics, including business, technology, personal development, and the arts, among others. It enables instructors to build and distribute courses on the platform, and it also allows learners to access courses on demand. Udemy's courses are designed for learners of all levels, from beginner to advanced, and they offer various formats, such as video lectures, quizzes, and assignments. Additionally, Udemy offers a certificate of completion for many courses upon finishing them. Therefore, it is important to create RecSys to encourage learners to choose academic courses or learning contents in EL.

RecSys are the main information filtering solutions that are efficient for increased internet availability, personalization tendencies, and evolving computer user activities. Although current RecSys are effective at delivering reasonable suggestions, they continue to struggle with issues including accuracy, scalability, and cold-start [15]. Finding precise information from data sources is becoming more challenging due to the student population's ever-growing access to EL resources. Personalization systems like adaptive EL and RecSys simplify the process of finding information from data

[1] https://www.udemy.com/

pools, but they encounter challenges like data scarcity, cold-start, scalability, time consumption, and accuracy [16].

Personalized RecSys [17], [18] are capable of offering engaging content that matches users' interests and contributes to alleviating the problem of information overload. Most frequently, recommendation algorithms present customers with possible items based on a range of data. RecSys uses the history of user-item interactions to suggest items in real-world situations, and it then solicits user feedback to improve those suggestions. In other ways, RecSys also makes recommendations for products based on interactions with users to learn about those users' preferences. The initial recommendation study mainly concentrates on designing content-based filtering and CF techniques to achieve this [19]. Personalized learning and PCR are important features for EL platforms to enhance the learning experience of their users. CF is a popular technique used to build such models [20], [21].

Numerous colleges and universities are recognizing this idea and approach. They create a learning environment online and open their courses for free on the MOOC system to make it easier for students to learn online and share information. Due to the abundance of network resources at this time, students are having difficulties swiftly and precisely locating the right course materials. Therefore, the current research priorities in curriculum suggestion are how to distribute knowledge, convey relevant knowledge to people who need it, reduce redundant knowledge generation, and encourage the widespread utilization of knowledge more effectively.

Integrating MOOC resources into the learning process is crucial, and the internet and big data analysis technology offer significant benefits for learners. However, finding suitable MOOCs has become complex due to their variety and easy access to platforms such as Udacity, Coursera, Udemy, and edX, and learners can be overwhelmed by the vast amount of information offered. As distance learning demand grows, RecSys solutions have emerged to enhance course selection quality and help learners navigate the overwhelming amount of course information available. MOOC-based RecSys provides suggested quality courses to learners. One way that IoT can be used to improve education is through the development of personalized EL and MOOC RecSys. These systems can collect data from students, such as their learning styles, interests, and progress, to recommend the most relevant and engaging content. This can help students learn more effectively and efficiently. Personalized EL and MOOC RecSys can be valuable tools for improving smart education. By providing students with the most relevant and engaging content, these systems can help students learn more effectively and efficiently.

According to our knowledge, many students continue to be curious about how to select the most suitable courses to take during the learning experience. Students frequently select appropriate courses according to their preferences, interests, assistance from accompanying experts, etc. The proposed study aims to develop a PCR using ML and CF techniques to predict students' academic performance and recommend appropriate courses. The study will compare the effectiveness of CF and ML techniques. The results could provide insights into developing an effective PCR, but ethical considerations must be considered to ensure unbiased recommendations. The material/content of the MOOCs regarding the students' profiles will be classified into beginners, intermediate, expert, difficulty, and all levels to make recommendations. By applying explicit (user ratings) and implicit (views and behavior e.g., the video lecture the students have watched, the articles they have read, or the exams/assignments they have done, to identify patterns) techniques to make recommendations based on the learner's interactions with the system and the cosine similarity in the similar contents. Additionally, the proposed model will combine temporal course correlation explicitly and learners' preferences and interest in course content implicitly. To this end, the Udemy data is utilized for learner's implicit behavior, and Coursera data is utilized to make recommendations based on explicit behavior. Then the explicit and implicit feedback are combined by designing a matrix for each learner group. CF and ML algorithms are applied to the matrix to generate recommendations for each learner.

To ensure that learning materials are personalized to each learner, smart educational systems should emerge. This work aims to create a novel PCR that provides learners with the best learning materials based on their choices, preferences, interests, past performance, learning style, contextual knowledge, and information-storage abilities. For EL in smart education, a novel RecSys based on CF using ML approaches is introduced. To deliver personalized learning materials, this strategy is integrated into a smart EL environment. Through the design, implementation, analysis, and assessment of a personal learning framework, we demonstrate the effectiveness of the proposed RecSys for ranking the top@k courses. Based on the insights obtained from data analysis, It should also consider the student's IoT-enabled devices and their usage patterns to personalize the recommendations further. Personalization and adaptability are currently crucial components of technology integration due to the quick development of smart EL systems. For instance, personalized learning environments are created to give students control over their learning environment and experience.

The main contributions of the proposed PCR RecSys are summarized as follows.

1. To provide a paradigm for smart education and assess whether current educational systems effectively meet the needs of learners, this study examines information on student preferences.
2. Based on the data collected and analyzed, the system will generate PCR for each student. The recommendation engine employs ML and CF algorithms to suggest courses that match the student's academic performance, interests, learning styles, and learning preferences.

3. To explore CF and ML algorithms, including RFR, KNN, DTR, SVD, XGBR, and LR to suggest top@k courses that match the student's academic and learning preferences.
4. To evaluate the performance of the proposed system for top-k recommendations in terms of accuracy metrics such as recall@k, precision@k, NDCG@k, MAP@k, MAE@k, MSE@k, and RMSE@k are employed.
5. The proposed system can be integrated with the learning management system to provide students with convenient access to the recommended courses. The system will deliver feedback on the courses taken by the students, which can be used to enhance the PCR.
6. The system provides real-time feedback to the students regarding their performance and recommends the most relevant courses. This recommendation can help students stay motivated and engaged in their learning.

The rest of the article is categorized into subsequent sections: related work is conducted in section II while section III illustrates the proposed course recommendation model. Section IV reports experimental analysis and results to validate the proposed objectives and section V discusses the results and provides suggestions for future research directions. Finally, section VI concludes the proposed work.

II. LITERATURE REVIEW

Recently, ML for education has drawn a lot of interest [1], [2], [22]. Prior studies have concentrated on topics like grade prediction [23], drop-out prediction [24], individualized teaching methods and resources [25], evaluating students' conceptual understanding of a subject matter [26], multimedia and cooperative learning [27], interpretation of students' performance [28], MOOC recommendations [29], etc. The vital but under-researched issue of (tailored) course sequence recommendation is examined in this work. Shortening the time required for students to graduate is one of the potentially big effects of addressing this problem. The solutions for resolving this issue can then be integrated with other approaches to offer a complete collection of resources for individualized learning.

Numerous approaches have been established in recent years and evaluated against publicly accessible datasets. The RecSys problem has been approached in a variety of ways in literature. These techniques are divided into content-based filtering [30], [31], CF [32], [33], and hybrid [34] algorithms. Popular buckets incorporate heuristic approaches, matrix factorization-based CF approaches [35], [36], neighborhood-based CF approaches, and ML approaches.

There is a lot of research has been conducted on PCR. For instance, Chen et al. [37] presented a hybrid RecSys. To acquire content-related items, it leverages item-based CF. These items are then filtered using a sequential pattern mining technique following typical learning sequences. To predict course trends and carry out rating predictions following them, Symeonidis and Malakoudis [38] designed a multi-dimensional matrix factorization framework coupled with the CF technique that makes use of data from external resources (such as users' skills, course characteristics, etc.). Additionally, Wan and Niu [39] developed a learning object self-organization-based RecSys in which the learning object is cumulatively modeled as an intelligent entity applying the self-organization concept and communicates with each other independently. As time passes, immediately after the learning process begins, the students typically put in a lot of effort on their assignments and video watching, and they are highly motivated to finish the course [40]. But as the course goes on, the challenge grows, and students get more dissatisfied [41].

According to Halawa et al. [42], students typically leave school because they are not satisfied with their education. Better satisfaction promotes MOOC persistence. A high score helps with contentment. Regular MOOC recommendations are typically based on interest. Thus, better features should be considered. Before attempting learning modules in an EL setting, RecSys can propose familiarization exercises. RecSys creates "user rating resource matrices" and provides individualized suggestions for learning resources according to users' learning preferences and learning paths in EL systems [16], [43]. More and more articles are being submitted to computer science journals and conferences as computer science and information technology have a wide and profound impact on our daily lives. Wan et al. [31] presented the content-based filtering Journals and Conferences RecSys on computer science, together with its online service, to aid authors in picking where to publish their studies. Depending on a manuscript's abstract, this method prioritizes conferences or journals that would be appropriate. To do this, a web crawler is leveraged to keep updating the training set and the learning method to keep up with the rapidly evolving field of computer science and technology. They provided an effective hybrid approach based on chi-square feature selection and SoftMax regression to produce an interactive online response.

Amane et al. [44] presented a dynamic ontology-based EL RecSys. In order to get the top recommendations applying clustering approaches, their suggested strategy characterizes course and learner linguistically. This information will be merged into collaborative and content-based filtering processes. The experiments' measurements were conducted by employing the Coursera dataset with the USMBA dataset from the university. Yang et al. [45] created a max cost flow structure to handle the limitations after developing a context-aware matrix factorization framework for predicting students' opinions over questions. A question recommendation with limitations for MOOC was the suggestion made in their work. Lazarevic et al. [46] established a mechanism for recommending courses based on commonalities between courses that are driven by ML. The suggested method uses a variety of data mining approaches to highlight course similarities. The experimental portion of this study demonstrated that Cosine metrics are the most accurate for calculating these values. The approach suggested in this work is built on rankings according to study areas. Jiang et al. [47] employed

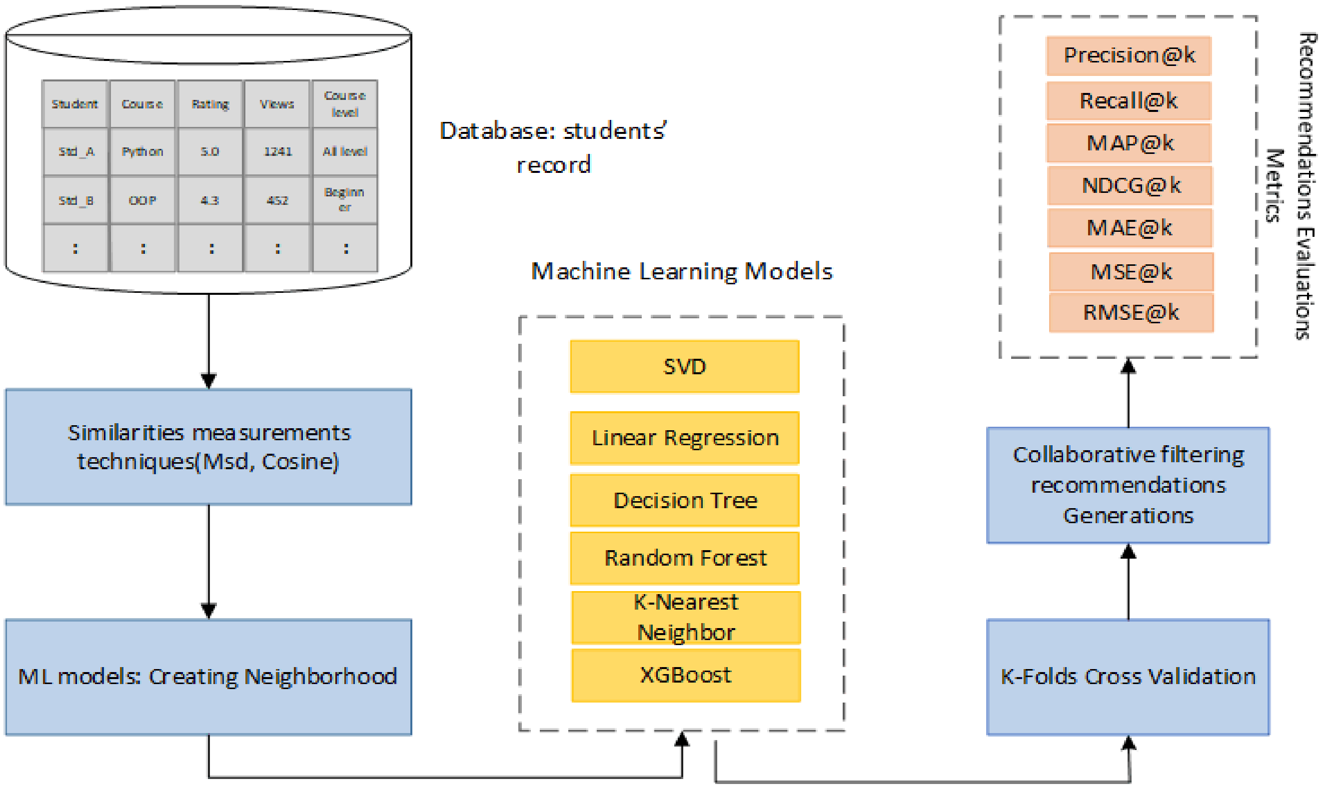


FIGURE 1. **Proposed MOOC recommendations framework.**

a Latent Dirichlet Allocation (LDA) user interest approach to construct an online education course RecSys. LDA is an unsupervised ML technique. The user's preference for subjects, their interest in online education courses, and the completion of the recommendation of online education courses depending on this are assessed using the LDA user interest paradigm. The proposed strategy is then assessed using data from EL websites.

For the movie RecSys, several modifications of the KNN method with various similarity metrics have been presented in [48]. These various KNN method modifications have been applied to real data from the MovieLens dataset. Ahuja et al. [49] built a movie RecSys by leveraging the K-Means clustering and KNN methods. Gupta et al. [50] designed CF and K-NN approaches that are mainly used to show how they can improve the model's accuracy when compared to content-based filtering. By combining a CF strategy with cosine similarity and KNN, this method improves upon the drawbacks of content-based filtering. Kumar and Bhasker [51] created a novel Deep Learning (DL) hybrid RecSys to close the gaps in CF structures and attain the greatest projected accuracy doable using DL. The suggested method uses a decreasing learning rate along with increasing weight decay, with the values cycled over epochs, to further boost accuracy. The suggested solution is compared to existing approaches based on projected accuracy and computational effectiveness. Extensive study has been conducted on several datasets, comprising MovieLens 100K, FilmTrust, Book-Crossing, and MovieLens 1-M. Tahmasebi et al. [52], introduced a hybrid social RecSys that makes use of a deep autoencoder structure. Their proposed framework takes advantage of social influence from users in addition to CF and content-based filtering. Social influence is defined for each user based on their social traits and online behaviors using Open Movie and MovieTweetings datasets.

In the current studies, a course may only be taken once all its required courses have been completed and passed. Traditional RecSys typically focus on recommending items or sets of items but often do not consider prerequisites when making recommendations. Therefore, recommending a course to a student who has not completed the prerequisite courses is not meaningful.

To the best of our knowledge, it appears that no research has been conducted on a MOOC RecSys specifically utilizing CF and ML techniques like RFR, KNN, DTR, SVD, XGBR, and LR, along with two different similarity measurements (cosine and Mean Square Difference (MSD) baseline similarities) for neighborhood and top-k recommendations calculation. This statement suggests that there is a gap in the existing literature regarding the specific combination of KNN variants and similarity measurements for movie recommendation systems. Conducting research in this area could contribute to the understanding and improvement of movie recommendation algorithms. To know how to place

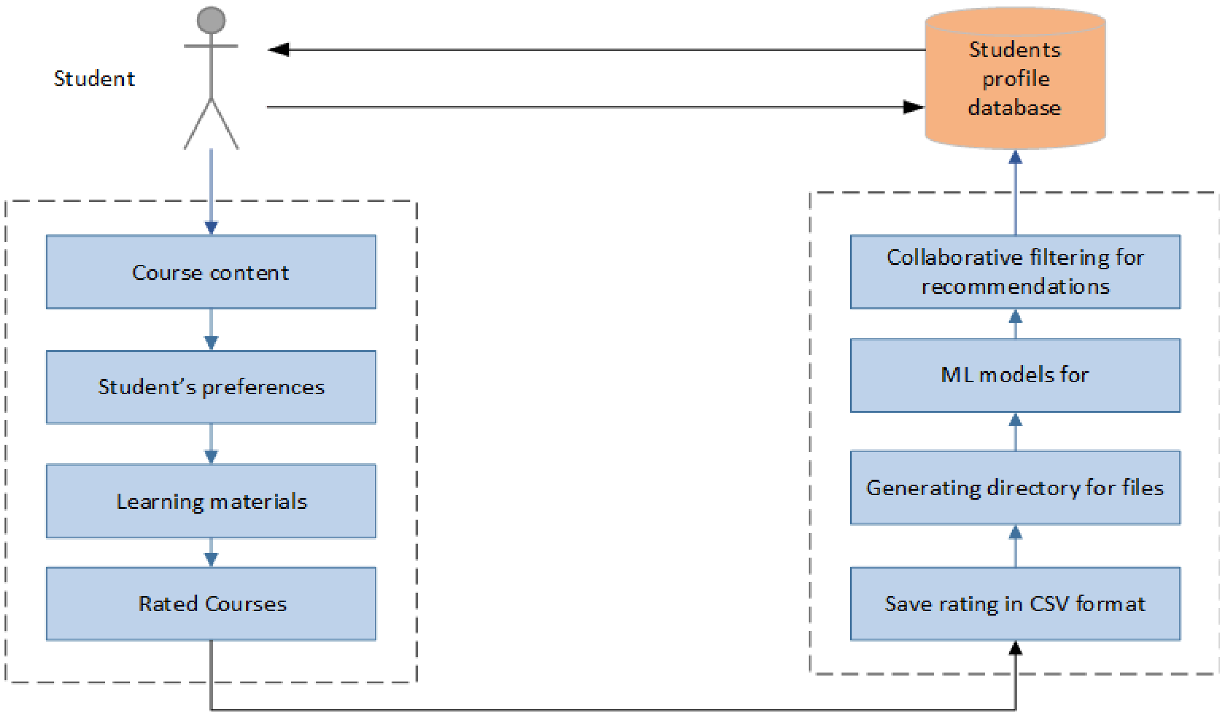


FIGURE 2. A CF-based PCR architecture for MOOC.

the learners' importance on the learning context by including the course's goal and learning preferences. It is necessary to make distinctions between suggestions for full-time students and those who study part-time.

III. PROPOSED SYSTEM ARCHITECTURE

Figure 1 demonstrates the framework designed for the proposed personalized EL and MOOC RecSys. In this case, MOOC datasets were employed to provide recommendations. The user-item rating matrix is used to determine how similar different students are to one another. For a particular user-item rating matrix, two types of similarities are calculated: cosine, and MSD baseline. To provide a course recommendation, a form of KNN-based CF and ML recommendation algorithms are employed with five-fold cross-validation techniques. Metrics like precision@k, MAP@k, recall@k, NDCG@k, MSE@k, RMSE@k, and MAE@k for top-k recommendations are evaluated for the achieved results. According to our knowledge, it appears that no research has been conducted on a MOOC RecSys specifically utilizing CF and ML techniques including RFR, KNN, DTR, SVD, XGBR, and LR, along with two different similarity measurements (cosine and MSD baseline similarities) for neighborhood and top-k recommendations calculations.

A. MOOC DATASET

Due to the COVID-19 pandemic, MOOCs have seen a boom in 2021. It is difficult for students to learn new skills when there are so many paid and free resources available online. Because of this, RecSys may be developed using this dataset to suggest courses to students based on their skills and level of difficulty. Moreover, a Course link is provided, which RecSys may make available for simple access. For the model training, two benchmark datasets are used during the experiment. The well-known MOOC datasets, including Coursera and Udemy, are used in the experiments and evaluations. The following subsections provide a detailed description of the MOOC datasets.

B. UDEMY DATASET

The Udemy course dataset is a collection of data that includes information on courses available on the Udemy platform. The dataset contains over 130,000 courses and includes various attributes or parameters such as course title, description, instructor name, price, rating, number of subscribers, and number of reviews.[2] It can be used for various purposes such as data analysis, ML, RecSys, and NLP. It can provide valuable insights into the most popular courses, the highest-rated courses, and the courses with the most subscribers. This dataset can be used by scholars, investigators, and analysts to better understand the Udemy platform and its users and to build online course creation and promotion strategies. Table 1 reports the features of the benchmark Udemy dataset.

The Udemy dataset is made up of 3678 rows and 12 columns. This dataset includes 3678 entries from Udemy courses in the following 4 categories: business finance, graphic design, musical instruments, and web design.

[2] https://www.kaggle.com/datasets/andrewmvd/udemy-courses

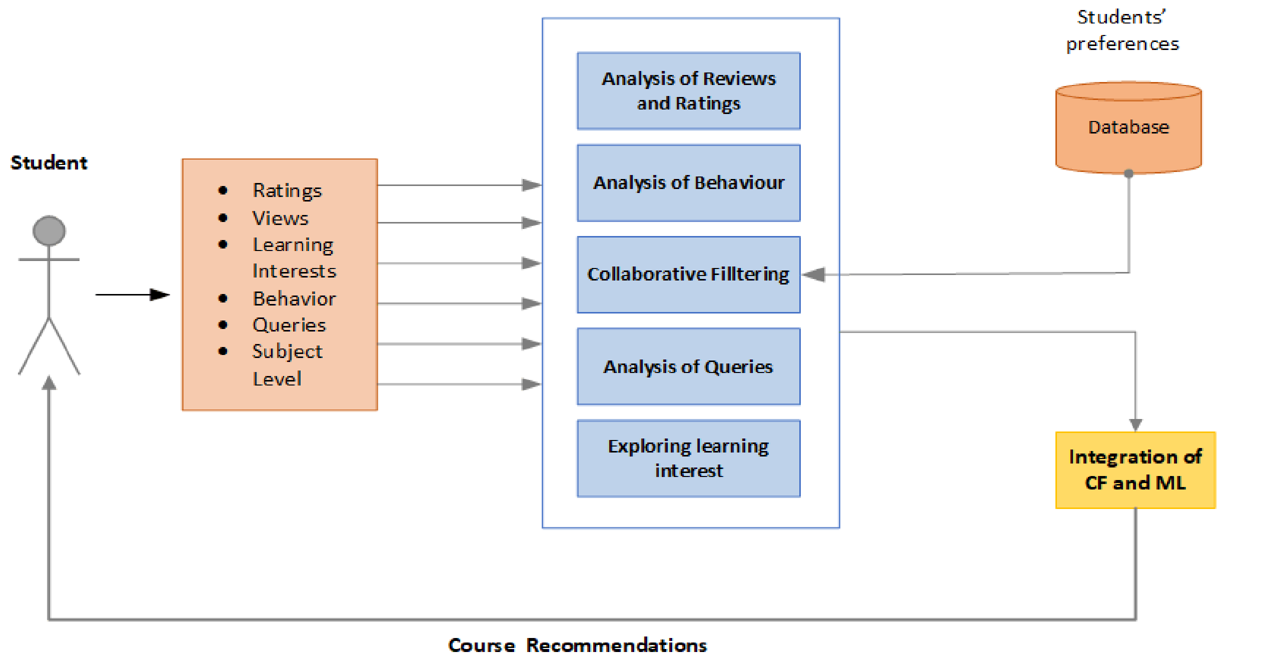


FIGURE 3. Framework for feature representations.

A MOOC platform termed Udemy provides both free and paid courses. For simplicity of usage, this version alters the column names eliminates empty columns, and consolidates everything into a single CSV file [53] after preprocessing.

C. COURSERA DATASET

The Coursera course dataset is a collection of data that includes information on courses available on the Coursera platform. The dataset comprises more than 3,900 courses with different parameters like duration, language, rating, instructor name, course title, description, and enrolled students. The dataset is available for download[3] and can be used for various purposes such as data analysis, ML, RecSys, and NLP. When it comes to the most well-liked, highly rated, and enrolled courses, it can offer insightful information. Researchers, analysts, and educators can leverage this dataset to better understand the Coursera platform and its users, and to develop strategies for creating and promoting online courses. The Coursera course dataset is frequently updated to include new courses and to reflect changes in the platform.

This dataset was scraped off the publicly available information on the Coursera website in September 2021 and manually entered in the case where the data was improperly scraped. It can be used in RecSys to promote Coursera courses based on the Difficulty Level and the Skills needed.

Course dataset scrapped from Coursera website. The dataset comprises six columns and 3522 course data. The detailed illustrations of the data are reported in Table 2:

[3]https://www.kaggle.com/datasets/khusheekapoor/coursera-courses-dataset-2021

TABLE 1. Representation of the parameters/features of the Udemy dataset.

Index	Description
Course ID	id field for courses
Course title	title field for courses
Number of subscribers	demand field for each course
Number of reviews	review number for each course
Number of lectures	lecture per course
Level	course level by trainee experience
Content duration	course duration in hours
Subject	course type field

D. COLLABORATIVE FILTERING FOR MOOC RECOMMENDATION

CF is a popular technique in RecSys that leverages the preferences and behaviors of a group of users to make recommendations for an individual user. It is based on the idea that users with similar tastes and preferences in the past (courses previously selected and/or numerical ratings given to those courses) are likely to have similar preferences in the future. This model is then used to predict courses (or ratings for items) that the user may have an interest in.

In this work, we used CF based model SVD to suggest courses that match the student's academic performance, interests, and learning preferences. Figure 2 depicts a CF-based PCR architecture for MOOC recommendations in EL, as developed in [54]. The architecture consists of two components including user knowledge and Interactions and application server. Here students will be able to rate any course they study on a scale of 1-5. This rating will be saved in

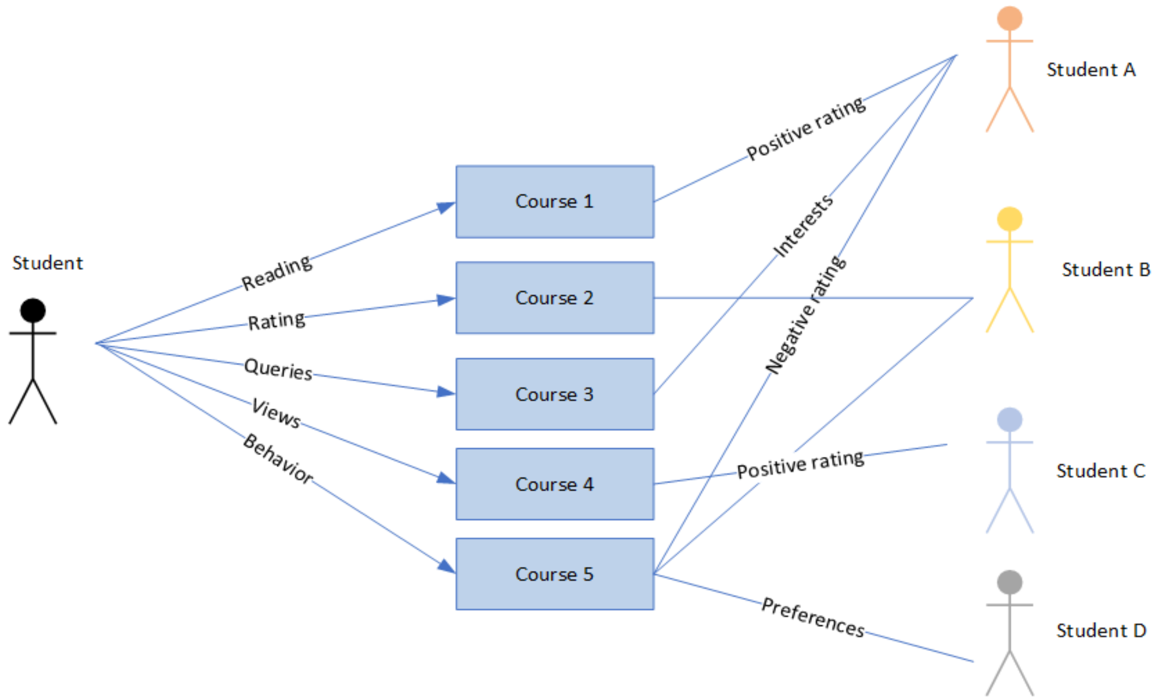


FIGURE 4. Use (student) case diagram for MOOC RecSys based on CF.

TABLE 2. Descriptions of the Coursera dataset.

Index	Description
Course Name	Contains the course title.
Organization	Identifies the institution running the courses.
Course Description	Provide specifics on certifications that are offered in the courses.
Course Rating	Includes the ratings assigned to each course.
Difficulty Level	Describes the degree of difficulty or course level.
Coursestudentsenrolled	Includes the total number of students registered for the course.

a CSV format and later exported to a directory created using a terminal. The model will access this file, after which we will generate recommendations using CF-based algorithms and apply existing ML prediction functions. In addition, Figure 3 depicts a framework for feature representations. In order to determine the model parameters, the first stage involves acquiring data regarding the attributes of the student profile, such as subject level, learning interests, behavior, queries, ratings, views, and more. The suggested model will then examine these parameters by integrating CF and ML by designing a matrix for each learner group. CF and ML algorithms are applied to the matrix to generate recommendations for each student. By using student feedback in the form of ratings or reviews, the model can improve these parameters during the recommendation process. Furthermore, a use case diagram for MOOC RecSys based on CF is shown in Figure 4. The fundamental idea is to consider the opinions and views of students on different courses and suggest the best course for each student based on their past rankings and the opinions of students who are similar to them.

E. RECOMMENDATIONS MODELS

The suggested research offers a cutting-edge context-aware MOOC RecSys for developing EL courses that are built on ML and CF. In this research, ML-based models are applied for MOOC RecSys, including Random Forest Regressor (RFR) [55]; K-Neighbors Regressor (KNN) [20]; Decision Tree Regressor (DTR) [53]; eXtreme Gradient Boosting Regressor (XGBR) [56]; Linear Regression (LR) [56]; and Singular Value Decomposition (SVD) [57] as follows:

RFR [55] utilizes large datasets effectively. It is capable of successfully predicting the missing data, even when there are no pre-processing and significant amounts of missing data. It combines a random feature with bagging. Decision trees paired with individual learners are found in random forests.

In the suggested PCR, first, the recommendation problem has been transformed into a regression task, where the goal is to predict a learner's rating or preference for a course. Next, an RFR model is utilized to predict learner ratings and reviews. Then, a feature matrix has been designed where each row represents a student-course pair, and each column represents a feature related to students, course, and their interactions (see Table 1 and Table 2). The datasets were split into a train set and a test set. The model receives the training set as input from the feature matrix. It is an ensemble of decision trees, and it can capture complex relationships between features and ratings.

DTR [53] is used for classification problems. In classification jobs, where the objective is to group data points into groupings or categories, decision trees are most frequently utilized. However, by properly defining the problem, DTR can be modified for recommendation purposes. In the proposed PCR, first, we need to characterize the issue as a regression task in order to use a DTR. For instance, a continuous rating score for each student-course pair can be predicted. Then, a feature matrix has been designed where each row shows a student-course pair, and each column shows a feature related to students, course, and their interactions as reported in Table 1 and Table 2. The datasets were split into a train set and a test set. The model receives the training set as input from the feature matrix such as using features as input and the learner's ratings as the target variable.

XGBR [56] is a scalable ML model designed for tree boosting. In the regression task, it is referred to as a gradient gradient-boosted regression tree, while in classification, it is also known as a gradient gradient-boosted decision tree. We adapted XGBR for RecSys, particularly in scenarios where we have student-course interactions and additional features that can be used for prediction. In the proposed PCR, to estimate the likelihood of each sample and select a suitable threshold for course rating, we build a sophisticated model based on the XGBR. Then, a feature matrix has been designed where each row represents a student-course pair, and each column represents a feature related to students, course, and their interactions (see Table 1 and Table 2). The datasets were split into a train set and a test set. The model receives the training set as input from the feature matrix.

LR [56] is employed to identify correlations between input data and a target variable and to predict continuous values. For PCR RecSys, the main objective is to recommend course content to students.

In the proposed PCR RecSys, first, the course recommendation problem has been transformed into a regression task with a continuous variable reflecting learners' preference such as course ratings as the target variable. Then, a feature matrix has been designed where each row shows a student-course pair, and each column shows a feature related to students, course, and their interactions (see Table 1 and Table 2). The datasets were split into a train set and a test set. The model receives the training set as input from the feature matrix such as using features as input and the learner's ratings as the target variable.

KNN [20] is used to identify user clusters based on shared course ratings and predict results using the average rating of the top k nearest neighbors. As an illustration, we first display the ratings in a matrix, which has a row for each course and a column for each student.

RecSys is often based on the KNN model and generates suggestions based on shared similar features. However, both classification and regression issues can be solved using KNN. The objective of using KNN-based RecSys is to predict the top K courses based on a given course. We set K = 5, 10, 15, and 20. The nearest Neighbors mechanism and neighbors' model are employed to recommend the topmost closely related courses.

SVD [58] is a matrix factorization method in CF to factorize the rating matrix into separate user matrix and item matrix, respectively. Although Euclidean distance is employed as a similarity measure and SVD is utilized to reduce dimensionality. However, the anticipated ranking result cannot be predicted with sufficient accuracy as the SVD does not hold additional information from the user and item. To improve the accuracy of rating predictions by utilizing more features, we employ a redistribution feature of the user matrix (learner matrix in our case) and the item (course) matrix through SVD. This helps in overcoming the problem of sparse data.

F. EVALUATION METRIC FOR MOOC RECOMMENDATIONS

Evaluation metrics for recommendations are used to measure the effectiveness of RecSys in providing personalized recommendations to users. The main goal of these metrics is to evaluate the quality of the recommendations provided by the system and to identify areas for improvement. Some common evaluation metrics for recommendations include [59], [60] as follows:

Precision@k is an evaluation statistic that determines the percentage of suggested courses in the top k that are pertinent to the student's preferences or needs. It is used to assess the correctness of a RecSys. Relevant courses are those that the student enjoys, engages with, or finds beneficial. The precision@k is 0.4, for instance, if a RecSys suggests 10 courses to a student and 6 of those courses are pertinent to the student, the system is assessed at k=5. This is because 6 pertinent courses were among the top 5 recommendations. Precision@k is a useful metric for evaluating the proposed PCR system's top recommendations since it takes into account both the relevance and the recommendation order of the courses.

Recall*@k* measures the proportion of relevant courses that were recommended in the top k courses. A course is considered relevant if it is something the student would have interacted with or found interesting if it had been suggested. For instance, the student's Recall@5 is 0.3 if they have interacted with 10 courses, a RecSys recommends 5 courses and 3 of the recommended courses are pertinent to them.

MAP@k determines the proportion of pertinent courses that each student should take out of the top k suggested courses, then averages these percentages for all students. For instance, if a RecSys is evaluated at k=5, and the MAP@k is 0.8, it means that on average, 80% of the relevant courses for each student are among the top 5 recommended courses. MAP@k assesses the proposed PCR system where the total number of pertinent courses is unknown among students and takes into account both the precision and recall of the RecSys. However, metrics like NDCG tackle the problem of distinguishing between relevant courses that are listed higher or lower in the recommendation list.

NDCG@k is employed to evaluate the effectiveness of RecSys by measuring the ranking quality of the

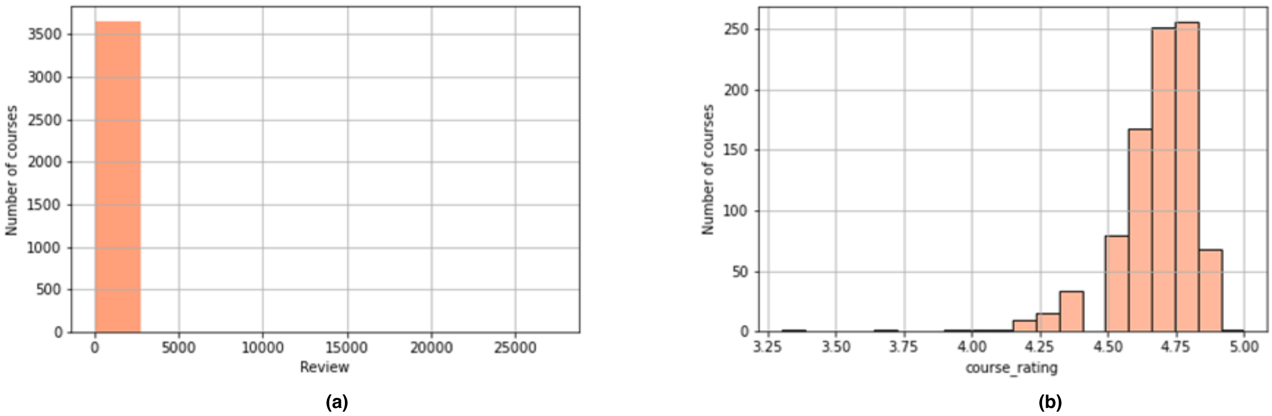


FIGURE 5. Distribution of (a) review of Udemy and (b) ratings of Coursera.

recommended courses. NDCG@k is a normalized version of Discounted Cumulative Gain (DCG@k), which calculates the sum of relevance scores of the top k recommended courses, with a logarithmic discount applied to the positions of the courses. The logarithmic discount is employed to provide higher weight to courses that are ranked higher in the list.

NDCG@k is designed as the ratio of DCG@k and the ideal DCG@k, which is the DCG@k that would be achieved if all the relevant courses were ranked at the top of the list. For example, if a RecSys is evaluated at k=5, and the NDCG@k is 0.9, it means that on average, the recommended courses are ranked at positions that are 90% as good as the ideal ranking.

RMSE@k measures the average deviation of the predicted ratings from the actual ratings across all the courses in the test set. It is calculated as the square root of the mean of the squared differences between the predicted ratings and the actual ratings. For recommendations, RMSE@k is determined by only reflecting the top-k courses in the recommendation list. This signifies that only the predicted ratings and actual ratings of the top-k courses are used to calculate the RMSE. For example, if a RecSys is calculated at k=10, and the RMSE@10 is 1.5, it implies that on average, the predicted ratings of the top 10 recommended courses differ from the actual ratings by 1.5 units.

MAE@k measures the average absolute deviation of the predicted ratings from the actual ratings across all the courses in the test set. In the proposed study, MAE@k is calculated by suggesting the top-k courses in the recommendation catalog. This indicates that only the predicted ratings and actual ratings of the top-k course are employed to determine the MAE. For example, if a RecSys is evaluated at k=10, and the MAE@10 is 1.2, it means that on average, the predicted ratings of the top 10 recommended courses deviate from the actual ratings by 1.2 units.

MSE@k is computed for the top k recommendations. MSE prioritizes the accuracy of the most important recommendations rather than considering all recommendations. MSE@k is helpful in assessing RecSys' effectiveness when the objective is to give the student a condensed collection of high-quality recommendations. A low MSE@k indicates that the system can accurately predict the student's preferences for the top 'k' recommendations, which can improve the overall user experience and satisfaction with the system.

IV. EXPERIMENTAL ANALYSIS AND RESULTS

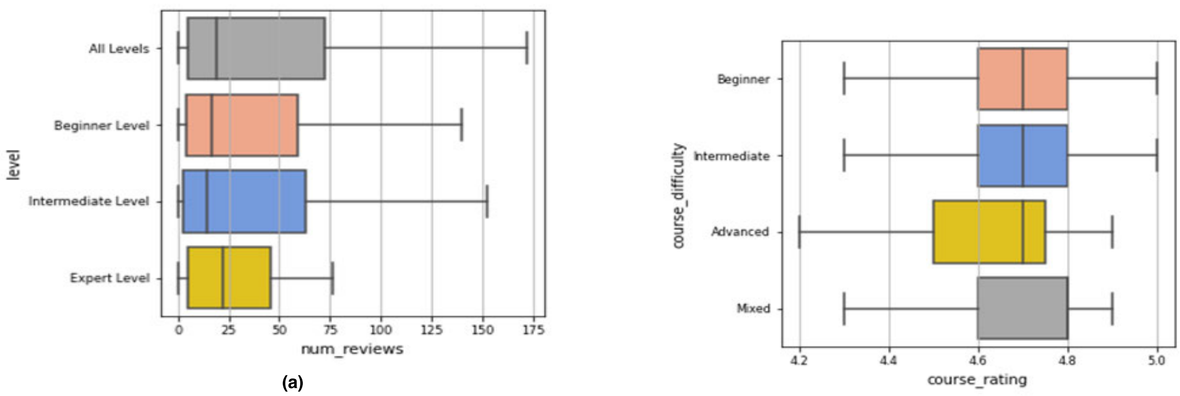
This section demonstrates the experimental data analysis and results achieved during the experiments. It also evaluates and compares the results using the proposed algorithms.

A. EXPLORATORY DATA ANALYSIS

Exploratory data analysis, also known as exploratory data processing, is the critical method of presenting early analyses of data to identify patterns, detect anomalies, test hypotheses, and double-check suppositions using statistical and graphical representation tools. Figure 5 shows a distribution of ratings of (a) review of Udemy and (b) ratings of the Coursera dataset and Figure 6 represents the average distributions of (a) review of Udemy courses and (b) Ratings of Coursera courses. Figure 7 depicts the classification of learning levels (a) Udemy courses and (b) Coursera courses. The visualizations indicate that courses with higher difficulty levels tend to receive lower ratings. This may be due to either the lower quality of difficult courses or the challenges faced by creators when teaching advanced courses online (as shown in Figure 6). However, the pie chart depicts the courses for advanced levels are poorly presented on the platform. Fewer courses can result in fewer courses with good qualities and, consequently, an overall lower ranking (Figure 7). More data is needed to get deeper into the analysis.

B. EVALUATION OF THE PROPOSED RECOMMENDATIONS MODELS

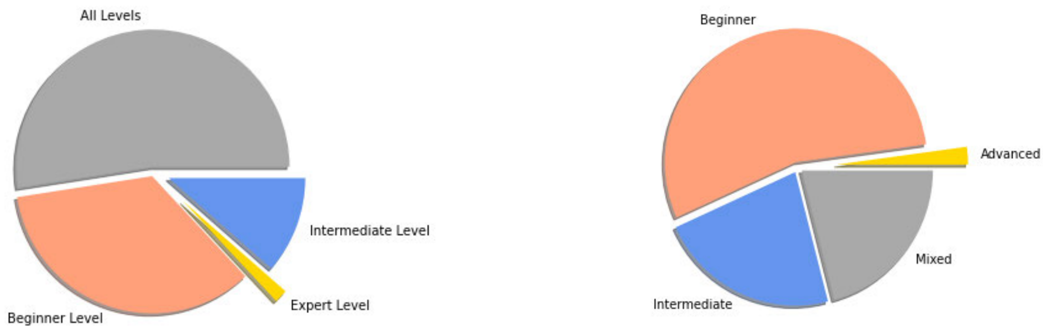
For explicit and implicit-based MOOC RecSys, we have implemented CF and ML RecSys. Six models are applied including SVD, KNN, RFR, LR, DTR, and XGBR. Each of these six methods have been evaluated using the neighborhood calculation's similarity measures, including cosine and MSD similarity. The MSE, RMSE, and MAE metrics are used



(a)

(b)

FIGURE 6. Average distribution of (a) review of Udemy courses and (b) Ratings of Coursera courses.



(a)

(b)

FIGURE 7. Classification of learning level (a) Udemy coursers and (b) Coursera courses.

to assess how well certain similarity measures perform when applied to various ML techniques. For top-k recommendations, precision@k, recall@k, NDCG@k, and MAP@k are also determined.

The performance of each comparison approach for each of the four measures, Precision@k, MAP@k, recall@k, and NDCG@k, on each of the two datasets (Udemy and Coursera), is shown in Table 3 and Table 4, respectively. Finally, Recall at top@k is depicted as a sign of the true label in a top@k recommendations catalog. For this, the recall index through K = 5, 10, 15, and 20 is compiled. The extent of the recommendation list is shown by the fact that k denotes the top course recommendations for the proposed PCR. For top k, we used the top @5, @10, @15, and @20 course recommendations, and the key results are listed in Table 3 and Table 4.

1) PERFORMANCE COMPARISON OF RECOMMENDATION MODELS ON UDEMY DATASET

Table 3 outlines the results for top courses in terms of evaluation metrics using the Udemy dataset. With a Precision of 0.8496, the SVD model has the best Precision@5, meaning that 84.96% of its top 5 recommendations are applicable. Moreover, other models with comparatively good Precision@5 scores include KNN and XGBR (Table 3 and Figure 8 (a)). Again, SVD has the highest precision values, with a Precision@10 of 0.8478 and Precision@15 of 0.8398, suggesting that 84.78% and 83.98% of the top 10 and top 15 recommendations are relevant. SVD continues to have the highest Precision@20, with a Precision of 0.8596, suggesting that 85.96% of the top 20 recommendations are relevant. XGBR and KNN are also competitive with Precision@20 of 0.7422 and 0.7538, respectively Figure 8 (a). Similarly for Recall@k, SVD stands out as the top-performing model with a Recall@5, Recall@10, Recall@15, and Recall@20 of 0.8123, 0.8477, 0.8501, and 0.8642, respectively. These results indicate that SVD successfully captures 81.23%, 84.77%, 85.01%, and 86.42% of the relevant courses within the top 5, 10, 15, and 20 recommendations, respectively. Other models like KNN and XGBR also have relatively high Recall@k scores (see Figure 8 (b)). Furthermore, SVD also leads with an NDCG@5, NDCG@10, NDCG@15, and NDCG@20 of 0.8202, 0.8490, 0.8601, and 0.8721, respectively, suggesting high-quality recommendations. An NDCG@20 score of 87.21% shows that SVD excels at providing relevant and well-ranked recommendations (see Figure 8 (d)). For MAP@, SVD is also the top-performing model with MAP@5, MAP@10, MAP@15, and MAP@20

TABLE 3. Performance of all the recommendation models for comparison on the Udemy dataset: bold indicates the best performance value in the metric in a comparison among proposed models.

Metrics	Top @k recommendations	Models					
		RFR	LR	XGBR	DTR	KNN	SVD
Precision@k (higher is better)	@5	0.6048	0.5953	0.7103	0.5881	0.6843	0.8496
	@10	0.5937	0.5123	0.7198	0.5608	0.7231	0.8478
	@15	0.5821	0.5241	0.7210	0.5498	0.7001	0.8398
	@20	**0.6691**	**0.6178**	**0.7538**	**0.6872**	**0.7422**	**0.8596**
Recall@k (higher is better)	@5	0.5963	0.5901	0.7012	0.6034	**0.7202**	0.8123
	@10	0.6234	0.6098	0.7194	0.6262	0.7031	0.8477
	@15	0.6704	0.6823	0.7831	0.6504	0.7110	0.8501
	@20	**0.6895**	**0.7002**	**0.7996**	**0.6790**	0.6943	**0.8642**
NDCG@k (higher is better)	@5	0.6985	0.6985	0.7304	0.6910	0.7432	0.8202
	@10	0.6956	0.7023	0.7276	0.7001	0.7131	0.8490
	@15	**0.7074**	0.7098	0.7295	0.7200	0.7110	0.8601
	@20	0.7049	**0.7103**	**0.7883**	**0.7234**	**0.7302**	**0.8721**
MAP@k (higher is better)	@5	0.6911	0.6993	0.7299	0.6986	0.7982	0.8306
	@10	0.6988	0.7021	0.7208	0.6890	0.7931	0.8580
	@15	0.7104	0.7098	0.7202	0.6987	0.7810	0.8610
	@20	**0.7197**	**0.7302**	**0.7984**	**0.7369**	**0.8243**	**0.8678**
MAE@k (lower is better)	@5	0.4811	0.4403	0.3221	0.4210	0.4010	0.3021
	@10	0.4612	0.4421	0.3204	0.4601	0.4101	0.2510
	@15	0.4329	0.4230	0.3321	0.4820	0.4220	0.2820
	@20	**0.4225**	**0.4221**	**0.3203**	**0.4200**	**0.4100**	**0.2001**
MSE@k (lower is better)	@5	0.4665	0.4843	0.3021	0.4331	0.4231	0.2671
	@10	0.4612	0.4471	0.3104	0.4500	0.4200	0.2521
	@15	0.4546	0.4340	0.3011	0.4562	0.4362	0.2562
	@20	**0.4304**	**0.4243**	**0.2934**	**0.4301**	**0.4101**	**0.2212**
RMSE@k (lower is better)	@5	0.5210	0.4896	0.3542	0.4731	0.3411	0.3211
	@10	0.5012	0.4747	0.3682	0.4679	0.3642	0.3342
	@15	0.5120	0.4720	0.3702	0.4783	0.3793	0.3693
	@20	**0.5010**	**0.4643**	**0.3522**	**0.4451**	**0.3263**	**0.3203**

of 0.8306, 0.8580, 0.8610, and 0.8678, respectively. These scores suggest that SVD produces high-quality recommendations that are well-ranked across the top 5, 10, 15, and 20 positions (see Figure 8 (c)).

Furthermore, among all the models (RFR, SVD LR, DTR, XGBR, and KNN), evaluated consistently exhibits the lowest MAE@k values, including MAE@5, MAE@10, MAE@15, and MAE@20, with respective scores of 0.3021, 0.2510, 0.2820, and 0.2001 (see Table 3 and Figure 10 (a)). This indicates that, in the context of the top @k recommendations, the predicted ratings by SVD are the closest to the actual ratings. While SVD excels in terms of MAE@k, other models such as XGBR, DTR, and KNN also perform reasonably well in providing accurate predictions. Similarly, SVD continues to lead with the lowest MSE@5, MSE@10, MSE@15, and MSE@20 of 0.2671, 0.2521, 0.2562, and 0.2212, respectively. These results demonstrate that SVD with MSE@20 performs exceptionally well. KNN, XGBR, and LR also perform well in terms of MSE@ (see Table 3 and Figure 10 (b)). Furthermore, SVD exhibits the lowest RMSE@5, RMSE@10, RMSE@15, and RMSE@20 of 0.3211, 0.3342, 0.3693, and 0.3203, respectively. These RMSE@k results suggest that predicted ratings of SVD are highly accurate within the top 20 recommendations. Other models, such as LR and KNN, also perform reasonably well in terms of RMSE@20 (see Table 3 and Figure 10 (c)).

2) PERFORMANCE COMPARISON OF RECOMMENDATION MODELS ON COURSERA DATASET

Table 4 presents the evaluation results for top courses using the Udemy dataset. The Precision metric, which measures the relevance of recommendations as follows:

SVD consistently performs exceptionally well, with the highest Precision@5, Precision@10, Precision@15, and Precision@20. For example, it achieves a Precision@20 of 0.8621, indicating that 86.21% of its top 20 recommendations are relevant. Other models like KNN and XGBR also show relatively high Precision scores (Figure 9(a)). As shown in Table 4 and Figure 9(b), SVD excels in Recall@k as well, leading with the highest scores across all values of k. For instance, it achieves a Recall@20 of 0.8764, indicating that 87.64% of the relevant courses are captured in the top 20 recommendations. KNN and XGBR also perform well in terms of Recall@k.

TABLE 4. Performance of all the recommendation models for comparison on the Coursera dataset: bold indicates the best performance value.

Metrics	Top k recommendations	Models					
		RFR	LR	XGBR	DTR	KNN	SVD
Precision@k (higher is better)	@5	0.6126	0.6153	0.7212	0.6482	**0.7682**	0.8529
	@10	0.6098	0.6382	0.7279	0.6296	0.7629	0.8445
	@15	0.6347	0.6554	0.7402	0.6012	0.7601	0.8478
	@20	**0.6731**	**0.6673**	**0.7701**	**0.6673**	0.7667	**0.8621**
Recall@k (higher is better)	@5	0.6876	0.6153	0.7403	0.6343	0.7634	0.8342
	@10	0.6532	0.6201	0.7453	0.6563	0.7656	0.8545
	@15	0.6867	0.6653	0.7876	0.6654	0.7665	0.8498
	@20	**0.6978**	**0.6909**	**0.7897**	**0.6734**	**0.7673**	**0.8764**
NDCG@k (higher is better)	@5	0.6805	0.6905	0.7333	0.6900	0.7690	0.8202
	@10	0.6997	0.6992	0.7198	0.6990	0.7699	0.8490
	@15	0.7079	0.7005	0.7289	0.7296	0.7729	0.8601
	@20	0.7103	**0.7202**	**0.7903**	**0.7369**	**0.7736**	**0.8731**
MAP@k (higher is better)	@5	0.6876	0.7045	0.7390	0.6821	0.7682	0.8365
	@10	0.6823	0.7234	0.7432	0.6976	0.7697	0.8238
	@15	0.7058	0.7154	0.7452	0.6798	0.7779	0.8365
	@20	**0.7097**	**0.7403**	**0.7898**	**0.7023**	**0.7793**	**0.8534**
MAE@k (lower is better)	@5	0.4875	0.4987	0.3657	0.3974	0.4374	0.3776
	@10	0.4645	0.5500	0.3565	0.5321	0.4321	0.3686
	@15	0.5023	0.5223	0.3562	0.4986	0.3826	0.3542
	@20	**0.4333**	**0.4767**	**0.3503**	**0.4321**	**0.3621**	**0.3325**
MSE@k (lower is better)	@5	0.4756	0.5324	0.4093	0.4703	0.4832	0.2897
	@10	0.4986	0.5378	0.3955	0.4887	0.4887	0.3452
	@15	0.4445	0.5673	0.3769	0.4986	0.4986	0.3543
	@20	**0.4685**	**0.4979**	**0.3753**	**0.4832**	**0.4703**	**0.3501**
RMSE@k (lower is better)	@5	0.5112	0.4676	0.3876	0.4932	0.3932	0.3565
	@10	0.4981	0.4707	0.3854	0.5001	0.4001	0.3431
	@15	0.4805	0.4983	0.3687	0.4987	0.3987	0.3552
	@20	**0.4801**	**0.4564**	**0.3632**	**0.4623**	**0.3623**	**0.3297**

SVD maintains its position as the top-performing model with the highest MAP@k values, indicating the quality of its recommendations in the top positions (as listed in Table 4 and Figure 9 (c)).

According to Table 4 and Figure 9(d), SVD continues to lead with the highest NDCG@k scores, indicating the quality and ranking of its recommendations. For example, it achieves an NDCG@20 of 0.8731, suggesting that its recommendations are well-ranked and highly relevant.

In Table 4 and Figure 11 (a) and (b), SVD consistently outperforms other models, showcasing the lowest MAE@20 and MSE@20 of 0.3325 and 0.3501, respectively. These results imply that the predicted ratings generated by SVD are the closest to the actual ratings within the top-k recommendations. While SVD excels in providing accurate predictions, it is worth noting that other models like XGBR, DTR, and KNN also demonstrate competence in delivering accurate recommendations.

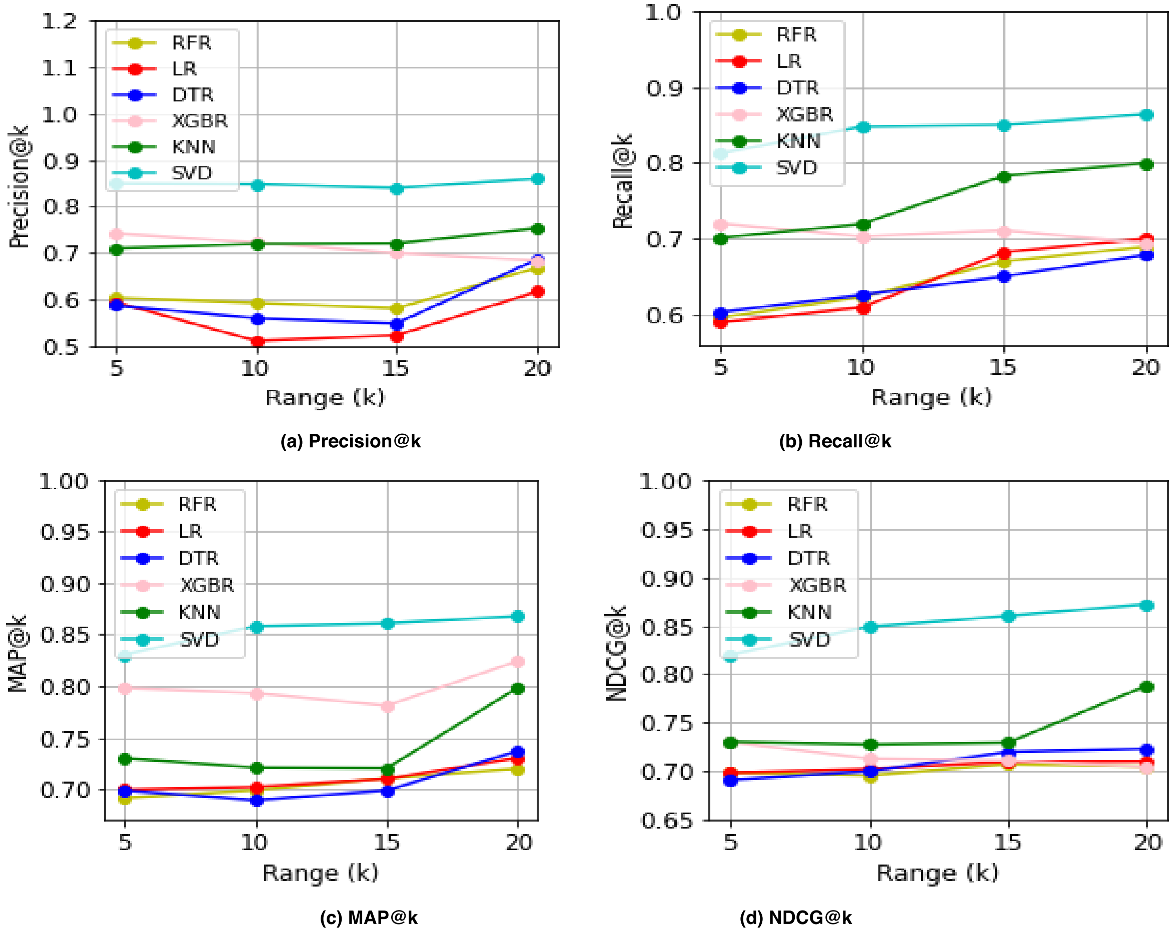
As shown in Table 4 and Figure 11 (c), SVD leads with the lowest RMSE@k values, indicating the accuracy of its predicted ratings within the RMSE@20 of 0.3297. LR, XGBR, and KNN also perform well in terms of RMSE@k using the Udemy dataset.

The suggested approach employs an intensive ML framework to analyze selected courses based on user learning activities and experiences. This proposed method enhances the precision of identifying relevant courses to recommend to students and those likely to be taught or learned. The information used for evaluation encompasses implicit and explicit learner characteristics from online MOOC platforms such as Coursera and Udemy. For both the suggested approach and other ML algorithms, we assessed various evaluation metrics, including Precision@k, Recall@k, MAP@k, NDCG@k, MAE@k, MSE@k, and RMSE@k. When compared to other ML methods, SVD consistently achieved the lowest error rate and significantly improved the prediction accuracy of the PCR RecSys across multiple evaluation metrics. This demonstrates its ability to deliver high-quality, relevant, and well-ranked recommendations, while other models also excelled in different aspects of recommendation quality.

V. DISCUSSION

A. RESULTS DISCUSSION AND ANALYSIS

The study demonstrates seven major evaluation metrics to compare the performance of the proposed models giving



(a) Precision@k (b) Recall@k

(c) MAP@k (d) NDCG@k

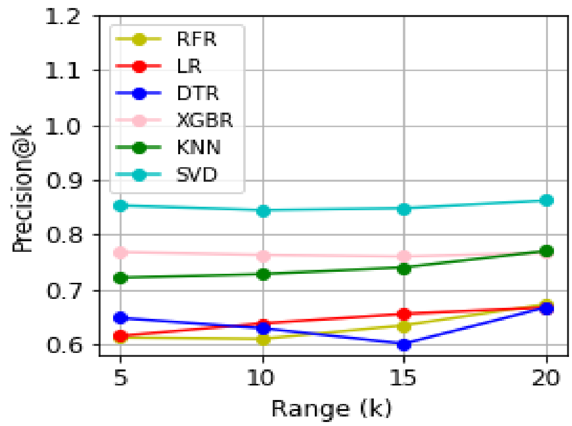
FIGURE 8. Accuracies of all the recommendation models for comparison measured by (a) precision@k, (b) recall@k, (c) MAP@k, and (d) NDCG@k on the Udemy dataset.

multiple perspectives to analyze and to distinguish which one is more appropriate. The results demonstrate that SVD and KNN are better models as compared to others such as XGBR, RFR, DTR, and LR because they are better suited to capture complex student-course interactions.

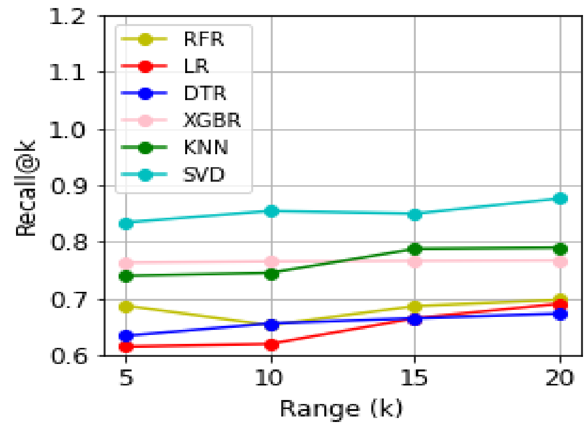
The results of the proposed course recommendation methods are compiled in Table 3 and Table 4. In both the Udemy and Coursera datasets, as revealed in Table 3 and Table 4, respectively, the recommendation performances achieved by SVD outperform those of other competitor approaches in terms of several evaluation metrics. For instance, when compared to the XGBR, LR, DTR, RFR, and KNN models, SVD enhances course recommendation performances on the Udemy dataset by 2.63% to 3.28% in terms of MAP and 2.30% to 4.03% in terms of NDCG, and it even enhances recommendation accuracy on the Coursera dataset by 5.12% to 6.04% in terms of MAP and 3.24% to 3.84% in terms of NDCG. Moreover, on Udemy and Coursera, it can be found that SVD saves 1.50 sec and 1.90 sec every episode, respectively. These findings lead to the following deductions.

For instance, we achieved an NDCG score of 20, which is quite impressive. In other words, out of a list of 20 options, the course that 87.21% of the students actually interacted with was recommended to them, as reported in Table 3 and Table 4. Thus, it can be stated that a higher error leads to less accurate recommendations. Conversely, a higher predicted rating makes it more likely for a course to be recommended to the student.

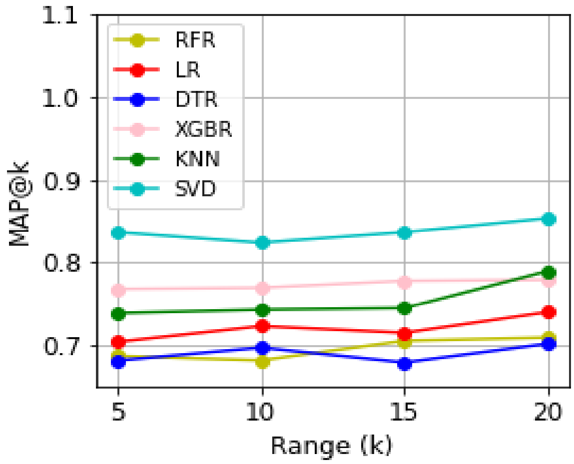
Figure 8 and Figure 9 compared the proposed models using a variety of evaluation metrics on the Udemy and Coursera datasets, respectively. We conducted experiments and evaluated the proposed models on these benchmark datasets. To evaluate the performance, we leveraged a set of metrics including Precision@k, recall@k, MAP@k, and NDCG@k. For performance evaluation, we utilized a set of metrics, including Precision@k, Recall@k, MAP@k, and NDCG@k. In addition, we considered MAE@k, MSE@k, and RMSE@k, where 'k' represents the number of top courses in our experiments. The number of closest neighbors is a parameter that can significantly impact the performance of a RecSys. Therefore, we assessed the outcomes under various nearest-neighbor settings for the top@k courses, as depicted in Figures 10 and 11 and Tables 3 and 4.



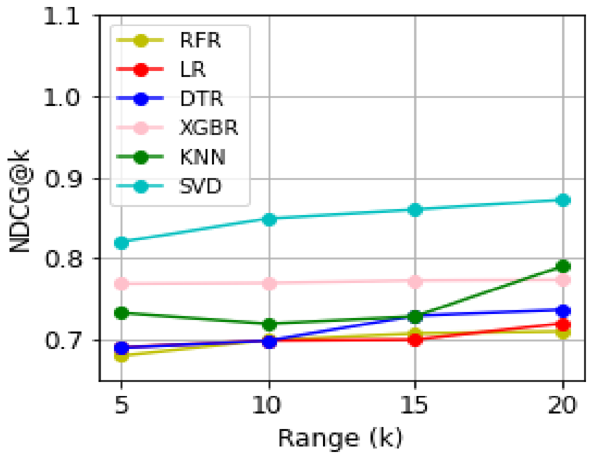
(a) Precision@k



(b) Recall@k

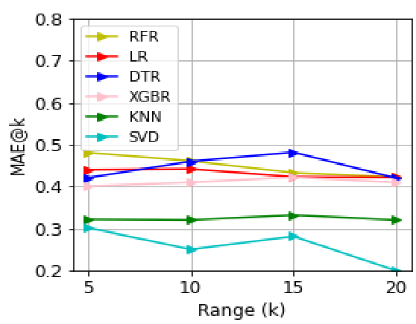


(c) MAP@k

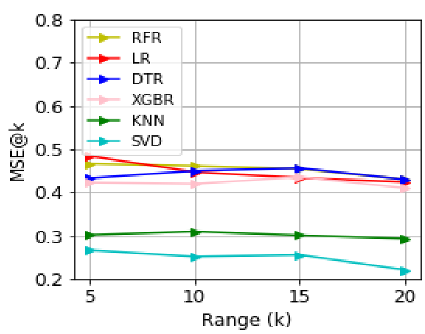


(d) NDCG@k

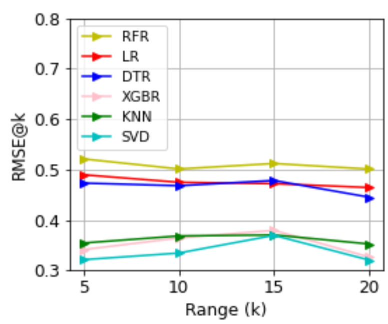
FIGURE 9. Accuracies of all the recommendation models for comparison measured by (a) precision@k, (b) recall@k, (c) MAP@k, and (d) NDCG@k on the Coursera dataset.



(a) MAE@k



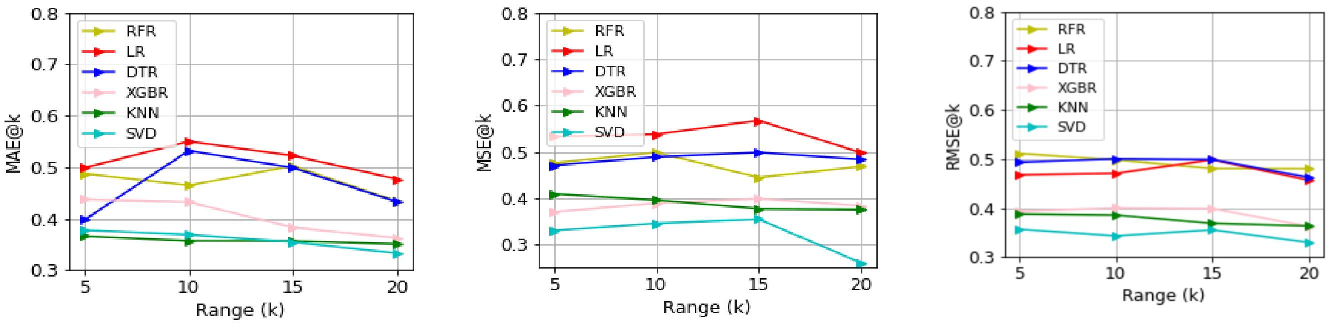
(b) MSE@k



(c) RMSE@k

FIGURE 10. Mean errors of all the recommendation models for comparison measured by (a) MAE@k, (b) MSE@k, and (c) RMSE@k, on the Udemy dataset.

Moreover, common error metrics, such as RMSE, MSE, and MAE, have proven to be stable in our analysis. As a result, we have determined that a neighborhood size (K) of 20 is optimal for our dataset. Achieving a score of @20, our proposed metrics have performed quite well. To put this in perspective, it means that more than 86% of students were recommended specific courses from a list of 20 choices, and they indeed engaged with these courses. We have also explored how the choice of K affects various metrics, including precision, recall, MAP, NDCG, MAE, MSE, and RMSE. For improved performance in the field of PCR, it is advisable to target higher precision, recall, MAP, and NDCG scores, while simultaneously reducing MAE, MSE, and RMSE scores.



(a) MAE@k (b) MSE@k (c) RMSE@k

FIGURE 11. Mean errors of all the recommendation models for comparison measured by (a) MAE@k, (b) MSE@k, and (c) RMSE@k, on the Coursera dataset.

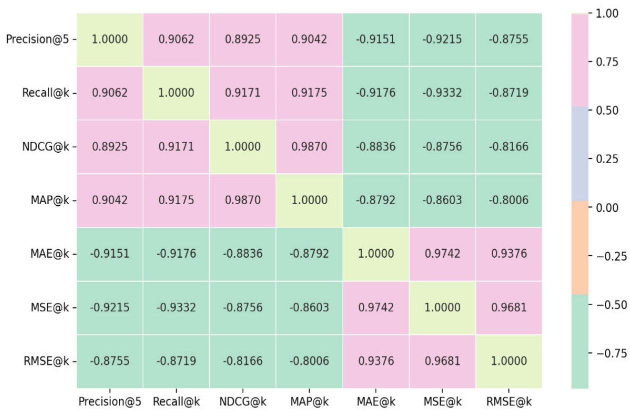


FIGURE 12. Pairwise Kendall's correlation of Precision@k, Recall@k, NDCG@k, MAP@k, MAE@k, MSE@k, and RMSE@k on the Udemy dataset. Pink signifies higher correlation values whereas light green indicates lower correlation.

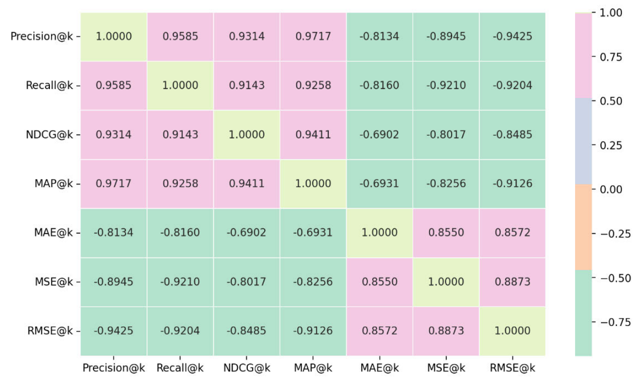


FIGURE 13. Pairwise Kendall's correlation of Precision@k, Recall@k, NDCG@k, MAP@k, MAE@k, MSE@k, and RMSE@k on the Coursera dataset. Pink signifies higher correlation values whereas light green indicates lower.

Furthermore, the accuracy of recommendations is evaluated using MAE, MSE, and RMSE as can be seen in Figure 10 and Figure 11. The results show that in both non-cold start and cold start situations, the proposed method outperforms the alternatives [20], [61]. Based on its higher performance on two different datasets, the proposed solution has a general methodology that may be applied to new RecSys rating prediction tasks.

B. CORRELATION AMONG METRICS

In this section, the relationships (correlations) between the model configurations based on several rating metrics are explored. Figure 12 and Figure 13 display Kendall's correlation among metrics on the Udemy and Coursera datasets, respectively. All correlations on the Udemy dataset are more than 0.9, indicating that the metrics yield nearly equal ranks. Including some correlations below 0.8, the other Coursera dataset shows stronger disparities.

It can be observed that, particularly for the Udemy dataset, NDCG@k and MAP@k stand out significantly from the rest of the metrics. These two measurements are related to one another. We believe that the extremely skewed long-tailed rating distribution across the dataset's courses may be the cause of this. On the two datasets, MAP is substantially connected with NDCG. Nonetheless, the rankings generated by the other metrics demonstrate a rather substantial correlation between them.

C. IMPLICATIONS

The significance and the novelty of the work lie in the fact that the current studies did not consider the Coursera and Udemy data for top K course suggestions by developing MOOC RecSys. This study advances the domain by combining ML and CF techniques to provide top-k course recommendations, which can offer more accurate, diverse, and fair suggestions. The research considers recent advancements and future perspectives in IoT-based smart EL technology, ensuring that recommendations align with the evolving landscape. It contributes to ongoing academic discussions in this smart EL domain, acknowledging both its limitations and opportunities, and promoting a collaborative environment for intellectual advancement. This study is inspired by the need to make students independent learners by limiting information overload in MOOC scenarios, which aligns with the concept of ''smart learning''. This recommendation can help students stay motivated and engaged in their learning. These promising initial findings provide a possible solution to assess this challenge further in future work. In addition, the contributions of this study allow the exploration of innovative intelligent systems for IoT-enabled smart EL education and for more potential applications in RecSys.

Moreover, as with any comprehensive study, this research is not without its limitations. In the following sections, we go through the limitations and restrictions of our approach as well as potential directions for further research based on the results of this study. These analyses open the door for further research in this area and offer a deeper view of the broader implications of our research.

D. LIMITATIONS

The proposed method's drawback is that while it worked well for the Coursera and Udemy datasets, it has not been compiled on other similar datasets. The second drawback of our technique is that, as it performs well for small datasets, it may have run-time and storage implications for larger datasets.

Another drawback is regarding the privacy issue; there are concerns about the use of personal information on EL platforms, and the potential misuse of this information for recommendations purposes. Developing privacy-preserving methods regarding the use of personal information on MOOCs can be addressed for data collection and analysis. Furthermore, investigating online recommendation approaches and implementing data privacy and security measures in real-time can effectively tackle data security, and privacy issues [62], [63].

E. FUTURE RESEARCH DIRECTIONS

In the future, to address the aforementioned drawbacks by utilizing a more accurate distance metric, such as Mahalanobis distance, which uses the variance of the data distribution to assess the distance from all points in addition to the distance between any two spots, the proposed method can be further enhanced. Other methods based on DL like deep collaborative neural networks, BERT, ANN, and RNN, can be used to enhance the performance of the suggested method. It is possible to use other MOOC datasets like edX and Udacity to improve performance and make it more generalized. Real-time data collection from learner activities is now possible due to the increasing popularity of IoT solutions. Using body area sensor networks and data compression techniques is a useful way to collect data from students' learning activities.

Finally, utilizing the rating-based feedback and learning from the sentiment scores, we want to construct a live demo of the proposed PCR and assess its quality in an online learning environment with real students and learners. This can make it easier to evaluate the effectiveness of the discussed PCR method in various situations and application domains.

VI. CONCLUSION

By combining IoT technologies with RecSys, personalized EL and MOOC RecSys can provide learners with tailored educational experiences, improving engagement and enhancing learning outcomes within the context of IoT-enabled smart education. To this end, this research introduced a novel system for personalized learning and MOOC RecSys, enabled by a smart EL platform. To validate the model, we conducted experiments using real-world benchmark datasets from Coursera and Udemy. Furthermore, this research addressed a significant challenge in EL systems and recommendation systems for new learners. We proposed mitigating this challenge by incorporating input from other learners in the recommendation mechanism, so this complexity can be minimized.

The primary goal of this research was to develop an adaptive PCR-based EL RecSys. To generate the top K recommendations using similarity techniques, our suggested strategy characterizes the course and learner in a semantically rich manner. The method is combined with CF and ML techniques. To assess the effectiveness of each technique, we used seven major evaluation metrics, including Precision@k, recall@k, MAP@k, NDCG@k MAE@k, MSE@k, and RMSE@k. Better ranking accuracy is typically indicated by higher metric values. The proposed approach produced the lowest error rate and improved the prediction accuracy of the PCR, as evidenced by the evaluation metrics.

ACKNOWLEDGMENT

The authors thank the University of Jeddah for its technical and financial support.

REFERENCES

[1] S. Tahir, Y. Hafeez, M. A. Abbas, A. Nawaz, and B. Hamid, "Smart learning objects retrieval for E-learning with contextual recommendation based on collaborative filtering," *Educ. Inf. Technol.*, vol. 27, no. 6, pp. 8631–8668, Jul. 2022, doi: 10.1007/s10639-022-10966-0.

[2] S. H. H. Madni, J. Ali, H. A. Husnain, M. H. Masum, S. Mustafa, J. Shuja, M. Maray, and S. Hosseini, "Factors influencing the adoption of IoT for E-learning in higher educational institutes in developing countries," *Frontiers Psychol.*, vol. 13, Jul. 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.915596

[3] D. Pandey, N. Singh, V. Singh, and M. W. Khan, *Paradigms of Smart Education With IoT Approach*. Cham, Switzerland: Springer, 2022.

[4] S. Amin, M. I. Uddin, A. A. Alarood, W. K. Mashwani, A. Alzahrani, and A. O. Alzahrani, "Smart E-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning," *IEEE Access*, vol. 11, pp. 89769–89790, 2023, doi: 10.1109/access.2023.3305584.

[5] K. Dahdouh, A. Dakkak, L. Oughdir, and A. Ibriz, "Large-scale e-learning recommender system based on spark and Hadoop," *J. Big Data*, vol. 6, no. 1, pp. 1–23, Dec. 2019, doi: 10.1186/s40537-019-0169-4.

[6] W. Deng, P. Zhu, H. Chen, T. Yuan, and J. Wu, "Knowledge-aware sequence modelling with deep learning for online course recommendation," *Inf. Process. Manage.*, vol. 60, no. 4, Jul. 2023, Art. no. 103377, doi: 10.1016/j.ipm.2023.103377.

[7] S. Ali, Y. Hafeez, M. Humayun, N. S. M. Jamail, M. Aqib, and A. Nawaz, "Enabling recommendation system architecture in virtualized environment for e-learning," *Egyptian Informat. J.*, vol. 23, no. 1, pp. 33–45, Mar. 2022.

[8] R. Chaker, F. Bouchet, and R. Bachelet, "How do online learning intentions lead to learning outcomes? The mediating effect of the autotelic dimension of flow in a MOOC," *Comput. Hum. Behav.*, vol. 134, Sep. 2022, Art. no. 107306, doi: 10.1016/j.chb.2022.107306.

[9] B. Li, G. Li, J. Xu, X. Li, X. Liu, M. Wang, and J. Lv, "A personalized recommendation framework based on MOOC system integrating deep learning and big data," *Comput. Electr. Eng.*, vol. 106, Mar. 2023, Art. no. 108571, doi: 10.1016/j.compeleceng.2022.108571.

[10] S. Benhamdi, A. Babouri, and R. Chiky, "Personalized recommender system for e-Learning environment," *Educ. Inf. Technol.*, vol. 22, no. 4, pp. 1455–1477, Jul. 2017, doi: 10.1007/s10639-016-9504-y.

[11] O. Embarak, "An adaptive paradigm for smart education systems in smart cities using the Internet of behaviour (IoB) and explainable artificial intelligence (XAI)," in *Proc. 8th Int. Conf. Inf. Technol. Trends (ITT)*, May 2022, pp. 74–79, doi: 10.1109/ITT56123.2022.9863950.
[12] A. M. Maatuk, E. K. Elberkawi, S. Aljawarneh, H. Rashaideh, and H. Alharbi, "The COVID-19 pandemic and E-learning: Challenges and opportunities from the perspective of students and instructors," *J. Comput. Higher Educ.*, vol. 34, no. 1, pp. 21–38, Apr. 2022.
[13] M. A. Almaiah, A. Al-Khasawneh, and A. Althunibat, "Exploring the critical challenges and factors influencing the E-learning system usage during COVID-19 pandemic," *Educ. Inf. Technol.*, vol. 25, no. 6, pp. 5261–5280, Nov. 2020.
[14] Q. Zhang, J. Lu, and G. Zhang, "Recommender systems in E-learning," *J. Smart Environ. Green Comput.*, vol. 1, pp. 76–89, 2022, doi: 10.20517/jsegc.2020.06.
[15] Z. Batmaz, A. Yurekli, A. Bilge, and C. Kaleli, "A review on deep learning for recommender systems: Challenges and remedies," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 1–37, Jun. 2019.
[16] S. S. Khanal, P. W. C. Prasad, A. Alsadoon, and A. Maag, "A systematic review: Machine learning based recommendation systems for e-learning," *Educ. Inf. Technol.*, vol. 25, no. 4, pp. 2635–2664, Jul. 2020, doi: 10.1007/s10639-019-10063-9.
[17] Z. Huang, X. Xu, H. Zhu, and M. Zhou, "An efficient group recommendation model with multiattention-based neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4461–4474, Nov. 2020, doi: 10.1109/TNNLS.2019.2955567.
[18] S. Deng, L. Huang, G. Xu, X. Wu, and Z. Wu, "On deep learning for trust-aware recommendations in social networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 5, pp. 1164–1177, May 2017, doi: 10.1109/TNNLS.2016.2514368.
[19] Z. Y. Khan, Z. Niu, S. Sandiwarno, and R. Prince, "Deep learning techniques for rating prediction: A survey of the state-of-the-art," *Artif. Intell. Rev.*, vol. 54, no. 1, pp. 95–135, Jan. 2021, doi: 10.1007/s10462-020-09892-9.
[20] K. K. Jena, S. K. Bhoi, T. K. Malik, K. S. Sahoo, N. Z. Jhanjhi, S. Bhatia, and F. Amsaad, "E-learning course recommender system using collaborative filtering models," *Electronics*, vol. 12, no. 1, p. 157, Dec. 2022.
[21] E. Aeiad and F. Meziane, "An adaptable and personalised E-learning system applied to computer science programmes design," *Educ. Inf. Technol.*, vol. 24, no. 2, pp. 1485–1509, Mar. 2019, doi: 10.1007/s10639-018-9836-x.
[22] M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, p. 11, Dec. 2022, doi: 10.1186/s40561-022-00192-z.
[23] R. Costa-Mendes, T. Oliveira, M. Castelli, and F. Cruz-Jesus, "A machine learning approximation of the 2015 Portuguese high school Student grades: A hybrid approach," *Educ. Inf. Technol.*, vol. 26, no. 2, pp. 1527–1547, Mar. 2021.
[24] N. S. Sani, A. Fikri, Z. Ali, M. Zakree, and K. Nadiyah, "Drop-out prediction in higher education among B40 students," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 11, pp. 550–559, 2020.
[25] Y. Ping, "Research on computer professional process and individualized teaching reform based on big data thinking," in *Proc. Int. Conf. Reform, Technol., Psychol. Educ. (ICRTPE)*, 2019, pp. 397–400.
[26] F. Cruz-Jesus, M. Castelli, T. Oliveira, R. Mendes, C. Nunes, M. Sa-Velho, and A. Rosa-Louro, "Using artificial intelligence methods to assess academic achievement in public high schools of a European union country," *Heliyon*, vol. 6, no. 6, Jun. 2020, Art. no. e04081.
[27] F. M. A. Khan and M. Masood, "The effectiveness of an interactive multimedia courseware with cooperative mastery approach in enhancing higher order thinking skills in learning cellular respiration," *Proc.-Social Behav. Sci.*, vol. 176, pp. 977–984, Feb. 2015.
[28] M. Adnan, M. I. Uddin, E. Khan, F. S. Alharithi, S. Amin, and A. A. Alzahrani, "Earliest possible global and local interpretation of students' performance in virtual learning environment by leveraging explainable AI," *IEEE Access*, vol. 10, pp. 129843–129864, 2022, doi: 10.1109/ACCESS.2022.3227072.
[29] B. Wu and L. Liu, "Personalized hybrid recommendation algorithm for MOOCs based on learners' dynamic preferences and multidimensional capabilities," *Appl. Sci.*, vol. 13, no. 9, p. 5548, Apr. 2023, doi: 10.3390/app13095548.
[30] P. B. Thorat, R. M. Goudar, and S. Barve, "Survey on collaborative filtering, content-based filtering and hybrid recommendation system," *Int. J. Comput. Appl.*, vol. 110, no. 4, pp. 31–36, Jan. 2015.
[31] D. Wang, Y. Liang, D. Xu, X. Feng, and R. Guan, "A content-based recommender system for computer science publications," *Knowl.-Based Syst.*, vol. 157, pp. 1–9, Oct. 2018.
[32] J. Bobadilla, A. Gutiérrez, S. Alonso, and Á. González-Prieto, "Neural collaborative filtering classification model to obtain prediction reliabilities," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 7, no. 4, pp. 18–26, 2022.
[33] Y. Koren, S. Rendle, and R. Bell, "Advances in collaborative filtering," in *Recommender Systems Handbook*. New York, NY, USA: Springer, 2021, pp. 91–142.
[34] G. Geetha, M. Safa, C. Fancy, and D. Saranya, "A hybrid approach using collaborative filtering and content based filtering for recommender system," *J. Phys., Conf. Ser.*, vol. 1000, Apr. 2018, Art. no. 012101.
[35] D. Bokde, S. Girase, and D. Mukhopadhyay, "Matrix factorization model in collaborative filtering algorithms: A survey," *Proc. Comput. Sci.*, vol. 49, pp. 136–146, Jan. 2015.
[36] S. Rendle, W. Krichene, L. Zhang, and J. Anderson, "Neural collaborative filtering vs. matrix factorization revisited," in *Proc. 14th ACM Conf. Recommender Syst.*, Sep. 2020, pp. 240–248.
[37] W. Chen, Z. Niu, X. Zhao, and Y. Li, "A hybrid recommendation algorithm adapted in e-learning environments," *World Wide Web*, vol. 17, no. 2, pp. 271–284, Mar. 2014, doi: 10.1007/s11280-012-0187-z.
[38] P. Symeonidis and D. Malakoudis, "Multi-modal matrix factorization with side information for recommending massive open online courses," *Expert Syst. Appl.*, vol. 118, no. 15, pp. 261–271, Mar. 2019.
[39] S. Wan and Z. Niu, "An e-learning recommendation approach based on the self-organization of learning resource," *Knowledge-Based Syst.*, vol. 160, pp. 71–87, Nov. 2018, doi: 10.1016/j.knosys.2018.06.014.
[40] S. Zheng, M. B. Rosson, P. C. Shih, and J. M. Carroll, "Understanding Student motivation, behaviors and perceptions in MOOCs," in *Proc. 18th ACM Conf. Comput. Supported Cooperat. Work Social Comput.*, Feb. 2015, pp. 1882–1895.
[41] R. F. Kizilcec, C. Piech, and E. Schneider, "Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses," in *Proc. 3rd Int. Conf. Learn. Anal. Knowl.*, Apr. 2013, pp. 170–179.
[42] S. Halawa, D. Greene, and J. Mitchell, "Dropout prediction in MOOCs using learner activity features," in *Proc. 2nd Eur. MOOC Stakehold. Summit*, 2014, vol. 37, no. 1, pp. 58–65.
[43] A. Klašnja-Milićević, M. Ivanović, B. Vesin, and Z. Budimac, "Enhancing e-learning systems with personalized recommendation based on collaborative tagging techniques," *Int. J. Speech Technol.*, vol. 48, no. 6, pp. 1519–1535, Jun. 2018, doi: 10.1007/s10489-017-1051-8.
[44] M. Amane, K. Aissaoui, and M. Berrada, "ERSDO: E-learning recommender system based on dynamic ontology," *Educ. Inf. Technol.*, vol. 27, no. 6, pp. 7549–7561, Jul. 2022, doi: 10.1007/s10639-022-10914-y.
[45] D. Yang, D. Adamson, and C. P. Rosé, "Question recommendation with constraints for massive open online courses," in *Proc. 8th ACM Conf. Rec. Syst.*, Oct. 2014, pp. 49–56, doi: 10.1145/2645710.2645748.
[46] S. Lazarevic, T. Zuvela, S. Djordjevic, S. Sladojevic, and M. Arsenovic, "Machine learning driven course recommendation system," in *Proc. 21st Int. Symp. Infoteh-Jahorina (INFOTEH)*, Mar. 2022, pp. 1–5, doi: 10.1109/INFOTEH53737.2022.9751282.
[47] X. Jiang, L. Bai, X. Yan, and Y. Wang, "LDA-based online intelligent courses recommendation system," *Evol. Intell.*, vol. 16, no. 5, pp. 1619–1625, Oct. 2023, doi: 10.1007/s12065-022-00810-2.
[48] S. Airen and J. Agrawal, "Movie recommender system using K-nearest neighbors variants," *Nat. Acad. Sci. Lett.*, vol. 45, no. 1, pp. 75–82, Feb. 2022, doi: 10.1007/s40009-021-01051-0.
[49] R. Ahuja, A. Solanki, and A. Nayyar, "Movie recommender system using K-means clustering AND K-nearest neighbor," in *Proc. 9th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2019, pp. 263–268.
[50] M. Gupta, A. Thakkar, Aashish, V. Gupta, and D. P. S. Rathore, "Movie recommender system using collaborative filtering," in *Proc. Int. Conf. Electron. Sustain. Commun. Syst. (ICESC)*, Jul. 2020, pp. 415–420, doi: 10.1109/ICESC48915.2020.9155879.
[51] R. Kiran, P. Kumar, and B. Bhasker, "DNNRec: A novel deep learning based hybrid recommender system," *Expert Syst. Appl.*, vol. 144, Apr. 2020, Art. no. 113054, doi: 10.1016/j.eswa.2019.113054.
[52] H. Tahmasebi, R. Ravanmehr, and R. Mohamadrezaei, "Social movie recommender system based on deep autoencoder network using Twitter data," *Neural Comput. Appl.*, vol. 33, no. 5, pp. 1607–1623, Mar. 2021, doi: 10.1007/s00521-020-05085-1.

[53] S. Amin, M. I. Uddin, H. H. Al-Baity, M. A. Zeb, and M. A. Khan, ''Machine learning approach for COVID-19 detection on Twitter,'' *Comput., Mater. Continua*, vol. 68, no. 2, pp. 2231–2247, 2021, doi: 10.32604/cmc.2021.016896.
[54] D. F. O. Onah and J. E. Sinclair, ''Collaborative filtering recommendation system: A framework in massive open online courses,'' in *Proc. INTED*, 2015, pp. 1249–1257.
[55] J. Herce-Zelaya, C. Porcel, J. Bernabé-Moreno, A. Tejeda-Lorente, and E. Herrera-Viedma, ''New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests,'' *Inf. Sci.*, vol. 536, pp. 156–170, Oct. 2020, doi: 10.1016/j.ins.2020.05.071.
[56] Z. Shahbazi, D. Hazra, S. Park, and Y. C. Byun, ''Toward improving the prediction accuracy of product recommendation system using extreme gradient boosting and encoding approaches,'' *Symmetry*, vol. 12, no. 9, p. 1566, Sep. 2020, doi: 10.3390/sym12091566.
[57] T. Widiyaningtyas, M. I. Ardiansyah, and T. B. Adji, ''Recommendation algorithm using SVD and weight point rank (SVD-WPR),'' *Big Data Cognit. Comput.*, vol. 6, no. 4, p. 121, Oct. 2022.
[58] K. Vahidy Rodpysh, S. J. Mirabedini, and T. Banirostam, ''Model-driven approach running route two-level SVD with context information and feature entities in recommender system,'' *Comput. Standards Interfaces*, vol. 82, Aug. 2022, Art. no. 103627, doi: 10.1016/j.csi.2022.103627.
[59] T. Silveira, M. Zhang, X. Lin, Y. Liu, and S. Ma, ''How good your recommender system is? A survey on evaluations in recommendation,'' *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 5, pp. 813–831, May 2019, doi: 10.1007/s13042-017-0762-9.
[60] W. Krichene and S. Rendle, ''On sampled metrics for item recommendation,'' *Commun. ACM*, vol. 65, no. 7, pp. 75–83, Jul. 2022.
[61] A. Klasnja-Milicevic and D. Milicevic, ''Top-N knowledge concept recommendations in MOOCs using a neural co-attention model,'' *IEEE Access*, vol. 11, pp. 51214–51228, 2023.
[62] M. Slokom, A. Hanjalic, and M. Larson, ''Towards user-oriented privacy for recommender system data: A personalization-based approach to gender obfuscation for user profiles,'' *Inf. Process. Manage.*, vol. 58, no. 6, Nov. 2021, Art. no. 102722, doi: 10.1016/j.ipm.2021.102722.
[63] B. Cheng, P. Chen, X. Zhang, K. Fang, X. Qin, and W. Liu, ''Personalized privacy protection-preserving collaborative filtering algorithm for recommendation systems,'' *Appl. Sci.*, vol. 13, no. 7, p. 4600, Apr. 2023, doi: 10.3390/app13074600.

SAMINA AMIN received the M.S. degree in computer science from the Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan, in 2020, where she is currently pursuing the Ph.D. degree. She has published several articles in reputable journals. Her research interests include machine learning, deep learning, data science, data mining, natural language processing, social media analysis, neural networks, recurrent neural networks, convolutional neural networks, image processing, health informatics, recommender systems, ensemble learning, and reinforcement learning. She has received the Gold Medal from the Kohat University of Science and Technology. She serves as a reviewer for different journals.

M. IRFAN UDDIN (Member, IEEE) received the degree in computer science. He was a Researcher in European Union-funded projects. He is currently a Faculty Member with the Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan. He is involved in teaching and research activities related to different diverse topics of computer science and has more than 18 years of teaching plus research experience. He has been actively involved in organizing national and international seminars, workshops, and conferences. He has published more than 100 research papers in reputed international journals and conferences. His research interests include machine learning, data science, artificial neural networks, deep learning, convolutional neural networks, recurrent neural networks, attention models, reinforcement learning, generative adversarial networks, computer vision, image processing, machine translation, natural language processing, speech recognition, big data analytics, parallel programming, multi-core, many-core, and GPUs. He is a member of ACM and HiPEAC.

WALI KHAN MASHWANI received the M.Sc. degree in mathematics from the University of Peshawar, Khyber Pakhtunkhwa, Pakistan, in 1996, and the Ph.D. degree in mathematics from the University of Essex, U.K., in 2012. He is currently an Associate Professor of mathematics and the Director of the Institute of Numerical Sciences, Kohat University of Science and Technology (KUST), Khyber Pakhtunkhwa. He has published more than 100 academic papers in peer-reviewed international journals and conference proceedings. His research interests include evolutionary computation, hybrid evolutionary multi-objective algorithms, decomposition-based evolutionary methods for multi-objective optimization, mathematical programming, numerical analysis, and artificial neural networks.

ALA ABDULSALAM ALAROOD received the Bachelor of Computer Science and Master of Computer Science degrees and the Ph.D. degree in computer science from the University of Technology Malaysia (UTM), in 2017. He is currently an Associate Professor with the Department of Data Science, College of Computer Science and Engineering, University of Jeddah. His research interests include data science, artificial intelligence, the Internet of Things (IoT), and cybersecurity.

ABDULRAHMAN ALZAHRANI is currently an Assistant Professor with the Department of Information Systems and Technology, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia. He is also the Head of the Computer Network Department. He teaches several courses to bachelor's and master's students. He teaches e-commerce and data visualization classes. His research is focused on the field of information technology and innovations in the areas of data science, machine learning, and health informatics.

AHMED OMAR ALZAHRANI received the Master of Computer Science degree from California Lutheran University and the master's and Ph.D. degrees in information systems and technology from Claremont Graduate University. He is currently an Assistant Professor with the Faculty of Computer Science and Engineering, University of Jeddah. His research interests include geographic information systems (GIS), energy informatics, remote sensing, and machine learning.