Research Article
Construction of Personalized Learning Platform Based on Collaborative Filtering Algorithm

Qian Zhang

Henan Institute of Technology, Xinxiang, HeNan, 453000, China

Correspondence should be addressed to Qian Zhang; xixigongzhu@hait.edu.cn

Received 19 January 2022; Revised 16 February 2022; Accepted 17 February 2022; Published 18 March 2022

Academic Editor: Xin Ning

Copyright © 2022 Qian Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

On the network service platform for vocational education, there are currently over 10,000 online courses. Learners face a challenge in selecting interesting courses from the vast resources available. Learners’ urgent need for personalized learning is becoming more apparent as educational informatization progresses. Personalized recommendation (PR) technology can aid personalized learning and increase learners’ learning efficiency significantly. This paper constructs a smart classroom model based on AI (artificial intelligence) by studying the connotation and characteristics of smart classroom in light of the current research status and trend of smart classroom at home and abroad. The merits of the recommendation system are determined by the recommendation algorithm used by PR system. This paper primarily focuses on developing a personalized learning platform based on the CF (collaborative filtering) algorithm, as well as conducting system requirements analysis, database design, functional module design, implementation, and testing on this foundation. Experiments are carried out to see if the optimized PR algorithm in the network learning platform is effective.

1. Introduction

With the full implementation of vocational education and the rise of online vocational education model, the knowledge system of curriculum construction is becoming more and more abundant [1]. In the information-based teaching, students can acquire knowledge not only through the teacher’s lectures in the classroom but also through the Internet. Teachers’ teaching methods are no longer single, and different teaching schemes can be designed by using the information-based teaching system [2]; Smart classroom, as a new form of deep integration of information technology and education, promotes the transformation of traditional classroom, develops in the direction of intelligence [3]. With the in-depth study of neural network and DL (deep learning) [4], the combination of intelligent classroom and AI (artificial intelligence) [5, 6] is bound to be the future development trend.

The arrival of the era of big data has driven the development of all walks of life, and the field of education is also being deeply affected. Infiltrating advanced big data technology into the field of education is the future development trend of educational informatization [7]. If the characteristics of learners and the learning process can be combined, effectively analyze learners’ characteristic information and learning data, identify and match learning resources that meet learners’ individual needs from massive online learning resources, and actively push them in the online learning platform, it not only improve the learning efficiency of learners but also obtain better learning effect, but also fully improve the utilization rate of the online learning platform [8, 9]. CF (collaborative filtering) is the most familiar, widely used, and mature recommendation technology in PR system at present. It uses the idea of similarity between users of computing system to filter information to solve the recommendation problem. By classifying system users, the interests and hobbies of one user in the class are extended to other users in the class [10]. Therefore, it is very important to study educational resources PR (personalized recommendation). In other fields, the research of PR technology was earlier, including e-commerce websites, video websites, and reading websites. CF technology is the most successful
technology used in PR in these fields. CF technology has many advantages such as simple calculation and high recommendation quality.

PR system can analyze the explicit and implicit information of users’ interests, preferences, ratings, and so on, using data mining in the original data source, and can quickly and effectively obtain the potential needs of users [11]. The key technology to consider in the design process of a PR system is the recommendation algorithm; which recommendation algorithm to use depends on the type of recommendation system [12]. This paper avoids the hassle of maintaining and upgrading C/S mode by recommending personalized learning resources in B/S mode. The system’s recommendation of personalized learning resources is based on the cognitive level, learning style, preferences, and learning process of the learners, among other factors. The CF algorithm’s optimization is also based on an analysis of learners’ characteristics and learning process data, giving learners the ability to push for the learning resources they really need.

2. Related Work

In people’s life and study, people rely more and more on the Internet to query knowledge and information. In recent years, network-based educational platforms have emerged continuously, and educational resources have also exploded. Literature [13], using CF and association rule mining algorithm, recommend information to teachers to improve teaching, and this recommendation method is also suitable for recommending educational resources to students. In literature [14], through learners’ rating of learning resources, CF algorithm and genetic algorithm are combined to recommend learning resources for students. This study proves that the recommendation system constructed by combining these two algorithms has better performance. In literature [15], learners’ interest model is obtained through analysis, and their learning progress is recorded, so as to recommend orderly learning resources for learners. This method uses neighborhood-based algorithm to recommend learning resources for learners. In literature [16], the Apriori algorithm is used to analyze students’ learning records, and the blind spots of students’ knowledge are obtained. Based on this, the corresponding question recommendation model is put forward. The model is based on users’ scores, and the K-means algorithm is used to improve it, and the design and development of the wrong question system are realized. The CF algorithm is applied to the form of group learning, and a group recommendation model is constructed according to the characteristics of group learning.

Most recommendation systems have high user participation, which requires users to make a lot of choices and inputs in the system. This is because most of these systems use content-based filtering and search engines, or they first classify their information, and then, users select categories to view it. In literature [17], establishing user’s interest model and making personalized recommendation constitute the whole personalized recommendation system of book learning materials. The user’s interest model is obtained by collecting and analyzing the user’s behavior information.

Literature [18] combines content-based recommendation with CF to achieve the effect of recommending various books and learning materials to users with various interests. Traditional CF recommendation technology is combined with content-based recommendation technology in literature [19]. Literature [20] proposes an intelligent agent method for mining the users’ demand information, creating a user model, and then generating recommendation results. To improve prediction accuracy, literature [21] uses the method of examining the weight of users and content items in the system. By analyzing users’ historical data on an e-commerce platform, literature [22] proposes a method of combining the K-means algorithm with a neural network to estimate users’ personal preferences for products. Simultaneously, other data mining-related technologies have been used to improve the accuracy of the recommendation system by analyzing weblogs.

Solved the bottleneck and scalability of recommendation algorithm, and achieved good results. Bayesian probability is used to transform the analyzed feature attributes into the nearest neighbor set of users, and the scoring matrix is constructed by Map Reduce model. Finally, the recommendation effect is achieved according to the association between the nearest neighbor set and users.

3. Research Method

3.1. Construction of Personalized Learning Resource Recommendation Model. In all recommendation systems, the premise of PR is user modeling [23], and the gap between the established user model and real users determines the final recommendation quality. The user models in teaching can be teacher models and student models. Obtain the behavior data of teachers and students through the teaching and learning platform, analyze and summarize this, and get the standard and calculable teacher-student models, and dynamically fine-tune the teacher-student models with the increasing amount of data.

Generally speaking, the principle of CF technology is to get the relevance between users or items based on the users’ favorite degree of content items, and then to carry out the corresponding PR based on these relevance relationships. Choose a method to calculate the similarity to calculate the K neighbors of a user, that is, the nearest users, which can be regarded as the most similar users with the user.

The basic principle of user-based CF algorithm can be seen from Figure 1, assuming that user $A$ is interested in $a$, $b$, $c$, $B$ is interested in $d$, and $C$ is interested in $b$, $c$; from these hypothetical user preference data, we can find that $A$, $C$ has a great correlation with the preference of items.

To calculate the current user’s interest in the $i$th item, the formula (1) can be used:

$$p(u, i) = \sum_{v \in S(u) \cap N(i)} w_{uv} r_{v,i}.$$  \hspace{1cm} (1)

Item-based recommendation algorithm is also suitable for personalized recommendation. If users buy a certain
product, they will wholeheartedly recommend similar products according to the attributes of the product.

The basic idea of the improved hybrid algorithm is to make full use of all users' historical preference information about the project and apply the results generated by the project-based algorithm to the input of the user-based algorithm. Finally, the user-based algorithm is used to predict the user’s preference for unrated content items, and the final prediction evaluation information and recommendations are generated.

It is found that the nearest neighbor set of the target user \( u \) is \( KNN(u) \), so the predicted score \( P_{u,j} \) of the unrated content item \( j \) by the user \( u \) is shown in formula (2):

\[
P_{u,j} = \bar{R}_u + \frac{\sum_{v \in KNN(u)} sim(u,v) \times (R_{v,j} - \bar{R}_v)}{\sum_{v \in KNN(u)} sim(u,v)},
\]

where \( \bar{R}_u, \bar{R}_v \) is the average of the historical evaluation information of the target user \( u \) and the user \( v \) in its nearest neighbor set, respectively; \( R_{v,j} \) is the historical evaluation of the target content item \( j \) by the user \( v \) in the nearest neighbor set of the target user \( u \); and \( sim(u,v) \) is the degree of association between user \( u \) and user \( v \) in the nearest neighbor set of target user \( v \).

All the unvaluated content items of user \( u \) are calculated by the above method, and the top \( N \) items with the highest estimated scores are taken as the Top-N recommendation set, and they are pushed to the target users.

User interest modeling is mainly used to predict the degree of users’ demand for unknown information and thus give personalized recommendations. When the user interest model is established, the recommendation system will recommend users according to relevant information, thus providing powerful help for users in the massive and complicated resources. In the evaluation matrix constructed in this paper, each row represents a student, each column represents an exercise, and each element represents the result of the student’s exercise. The result has two possibilities, namely, right and wrong, so it can be represented by 0 and 1, respectively, and a null value indicates that the student has not done the exercise. The representation method based on evaluation matrix is usually used in the recommendation system using CF technology, and the PR in this paper uses CF technology.

Whether it is explicit or implicit feedback, it is increasing over time, which means the user model will change, necessitating an update to the previous user interest model. The neural network updating technology begins by using neural networks to create a user interest model. The neural network will automatically adjust the weights of each edge in the network as user feedback increases, thus updating the user interest model. Incorporating information about teacher-student interactions both inside and outside the classroom intelligently pushing relevant learning resources based on students’ circumstances, completely subverting the traditional classroom mode, and allowing every student to participate in classroom learning and obtain learning materials and methods that are appropriate for him in a timely manner [24]. DL will be used in an intelligent classroom based on AI to accurately analyze each student’s learning situation and knowledge, push relevant learning materials and videos in a targeted manner, and push all materials in a differentiated, personalized, and intelligent manner, so that students of various levels can achieve the desired results.

The recommendation effect will get better and better, forming a virtuous circle. Prediction accuracy MAE (mean absolute error): The working principle of MAE is to first get the prediction score from the training set data and the average absolute difference between the calculated score and the real score of the test set. See (3) for MAE calculation formula.

\[
MAE_i = \sum_{n=1}^{u} |s_i - t_i|.
\]

The recall rate considers the ratio between the number of items in the recommendation list that meet the user’s interests and hobbies and the number of items that users like in the recommendation system [18].

\[
\text{Recall} = \frac{\sum_{u \in U} R(u) \cap T(u)}{\sum_{u \in U} |T(u)|}.
\]

The main task of online work is to collect students’ behavior data, which includes students’ record of doing exercises, preview, and class performance. Students’ behavior data provides data support for students’ modeling. The main task of off-line work is to analyze the students’ interest model and calculate the recommendation results by using the recommendation algorithm. The overall PR model is shown in Figure 2.

Students can learn in personalized learning platform, and the educational resource database provides data support for the learning platform, in which the educational resource model and knowledge point model can be constructed. In the process of students’ learning, the platform will record the data of students’ behaviors to build a student model. Then, these data will be used as the input data of the recommendation module.
3.2. Construction of Personalized Learning Platform. Under the background of big data, AI can use big data to help teachers quickly collect actionable opinions from students’ performance, find out students’ specific problems, and then help teachers formulate effective methods and collect materials to help students improve problems and make up for gaps. In addition, the development of intelligent empowerment education activities can provide endless impetus for students to learn information technology knowledge and skills.

The use of AI technology will subvert the previous educational knowledge system, allowing for no boundaries between disciplines, general and online education, and constant diversification of educational content. Furthermore, by utilizing the AI learning platform, the data of individual learning behavior, including everyone’s interests, preferences, and learning degree, will be fully recorded, and the system will devise learning contents, methods, and progress that are appropriate for everyone, resulting in truly personalized service. Full arousal of learners’ interest, cultivation of learners’ practical ability, ability to analyze and solve problems, creativity, and other skills will also liberate teachers’ productivity, stimulate teachers’ creativity, and achieve the goal of balanced education, allowing students in remote areas to receive a good education.

The algorithm principle of DL is to input the known data whose data model is not easy to find into the input layer and get the output data from the output layer through the function mapping of multiple hidden layers, so as to find the real relationship between variables. Massive research shows that DL can get regular information among data through repeated training and learning [25]. It is a supervised machine learning algorithm.

Let the number of hidden layers in DL model be 5, and the terminal output is

$$ y_b = f \left( \sum_{b=a}^{C} \mu_{bj} z_j \right) \quad b = a, \ldots, b, \ldots, c. \quad (5) $$

Then, it is the back propagation process of error. The error function is

$$ E = \frac{1}{2} \sum_{q=1}^{Q} \sum_{b=a}^{C} \left( t_{by} - y_{by} \right)^2 = \sum_{q=1}^{Q} E_Q. \quad (6) $$

Through the reverse transmission of the error, the error becomes smaller and smaller, close to the specification error value, until the condition of stopping the algorithm is met. Adjust the weights between input and hidden layer, hidden layer, and hidden layer and output according to formula (7):

$$ \Delta \omega_{jm} = -\eta \frac{\partial E}{\partial \omega_{jm}} = -\eta \frac{\partial}{\partial \omega_{jm}} \left( \sum_{q=1}^{Q} q \right) = \sum_{q=1}^{Q} \left( \eta \frac{E_q}{\partial \omega_{jm}} \right). \quad (7) $$

Among them, $0 < \eta < 1$.

Using technologies like deep learning, neural networks, and TensorFlow, a large amount of data from students’ learning processes, as well as all historical success data, should be used. Process data, such as teacher activities, student activities, classroom teaching situations, and practice test results, should be included in this data. Teachers can be aided in making teaching decisions by a thorough analysis of AI algorithms. This paper proposes to build a smart classroom platform based on AI using advanced, convenient, and practical technical tools, based on research and theoretical analysis of relevant smart classroom modes at home and abroad (see Figure 3).

AI integrated platform provides all the technologies of smart classroom informatization, real-time interaction with the whole process of smart classroom, smart learning space, and intelligent learning space, so that communication between teachers and students and between students can be unimpeded, and learning materials suitable for them can be selected in real time. The user layer provides many applications, so that teachers and students can form a dynamic “teaching” and “learning,” and monitors all the

![Figure 2: Educational PR model.](image_url)
information generated in this process to be stored as data, thus forming an all-round and full-coverage intelligent service.

Take the original course ID list recommended by the recommendation model and display it graphically on the platform page of the corresponding user. The intelligent classroom uses artificial intelligence to monitor and evaluate the entire teaching process in real time, as well as provide real-time feedback to teachers and students. Teachers can quickly grasp the entire learning process by using AI technology to analyze students’ basic data and status data and providing personalized assessment exercises. Learners can use mobile devices to take learning resources with them wherever they go to study. Learning resources, on the other hand, are presented to learners in digital form, so many can be stored on mobile devices offline or in the cloud online, and learners can learn at any time.

Figure 4 shows the specific process of the effect evaluation stage of personalized course recommendation system.

Extract new courses learned by active users within \( m \) days after the last recommendation system was implemented from the platform business database. Read the database of recommendation system configuration parameters, and extract the main parameters of CF algorithm, including the number of K neighbors and the number of recommended courses.

Take out the list of recommended courses from the database of recommended results, compare the recommended results with the courses actually studied by users, and calculate the effect evaluation index. The value of this effect evaluation index is stored, and the system counts whether the user has registered to study the courses recommended to him as the feedback of the recommendation effect.

4. Result Analysis and Discussion

Smart classroom breaks the communication and feedback limited to time and place in traditional classroom and enables teachers to know students’ needs in real time and adjust teaching strategies according to students’ needs in time.

Use the modified cosine similarity algorithm to get the similarity between items in the training set of Movie Lens 100K dataset, and some data are shown in Figure 5.

Education data can be collected in all directions and throughout the process in the era of big data and the Internet, thanks to data mining technology. Unstructured accurate data, such as personal emotional information, psychological tendency, and practical ability, can be obtained in addition to structured data, providing an objective basis for a comprehensive evaluation of the system. Students were previously evaluated through communication, investigation, classroom questioning, testing, and other methods, and it was difficult to obtain a timely and comprehensive understanding of each student’s unique situation. As evaluation indicators, use the network to gather information, share experiences, and so on. Students are evaluated primarily based on their active acceptance of teachers’ lectures and their level of interaction with teachers and classmates. The degree of information sharing and utilization of the network also requires comprehensive evaluation of students’ innovation and development potential.
When users operate in the interface through the browser, if they want to call codes and algorithms, they need the logic layer to solve the problem. The logic layer calls the background codes through the range technology of ADO.NET database and returns them to the presentation layer for users to use. The data access layer, as its name implies, is used to store all kinds of required data, mainly according to the knowledge tree structure, adding learning contents of chapters and sections, creating learning resources and forming a list of curriculum resources, which also includes the modification, deletion, and viewing of corresponding learning resources. A learning content or knowledge point is essentially a learning resource, and no other learning resources are required to be uploaded in the system.

Using the data obtained from the study, the MAE value of improved mixed similarity is obtained, thus obtaining Figure 7.

Because the improved hybrid similarity algorithm and the modified cosine similarity algorithm are both decentralized, that is, they consider the dimensional differences of each dimension, their MAE is lower than cosine similarity; MAE gradually decreases with the increase of the number of nearest neighbors, and eventually tends to be stable. The problem of sparsity of scoring information is partially solved, and the improved algorithm’s MAE value is lower than the three traditional similarity algorithms.

Figure 7 can intuitively show the relationship between RMSE (root-mean-square error) and the number of nearest neighbors of the improved algorithm and three traditional similarity algorithms.

RMSE gradually decreases with the increase of the number of nearest neighbors until it converges to a certain value, and finally tends to a stable state. Because the improved hybrid similarity algorithm fills the scoring vectors among users, it solves the problem of sparsity of scoring information to a certain extent and improves the stability of the algorithm to a certain extent. When the number of users’ nearest neighbors is small, the stability of the algorithm is very sensitive to the number of nearest neighbors, as shown in Figure 8.

In order to ensure the optimal performance of the designed machine learning algorithm, the parameters of CF algorithm are adjusted, focusing on the influence of different values of nearest neighbor parameter K on the performance of the algorithm. Select a subset of course selection data, in which the number of users is 4921 and the number of courses is 2034; select k values of 10, 50, 100, and 200 and evaluate the four indexes of the object-based CF algorithm and the calculation time (Figure 9).

By comparison, it is found that when the k value is 100, the values of the four evaluation indexes show the best effect and take the shortest time, so the optimal CF algorithm is deployed on the platform based on the k value of 100.

For implicit feedback data, the most important thing is to find clues that can reflect users’ preferences from the data and predict and evaluate the scoring range through a series of behavioral trajectories. Item-based collaborative filtering algorithm is significantly more time-consuming than user-based collaborative filtering algorithm, and the time-

---

**Figure 5:** Interitem modified cosine similarity value.

**Figure 6:** Comparison of similarity prediction scores.

Using the similarity data between users, the score prediction of user-item data in the training set u1.test of MovieLens 100 K dataset is carried out, and the score comparison based on the modified Pearson similarity prediction in the test set u1.test of Figure 6 is obtained.

Based on the convenience and simplicity of development and maintenance, and the lower management cost compared with other modes, the system adopts B/S mode for development. The system interface is designed with a unified master. Users do not need to install clients and software on the client and eliminate the trouble of upgrading. They can only use the system through the browser after networking. In order to facilitate the management and staff to update, maintain, and upgrade the system, the application program of the system runs on the server side, which not only reduces the dependence between the server and the client but also ensures the security of running and storing the program code.
consuming of user-based collaborative filtering algorithm is not optimistic. By introducing the idea of multiprocess, optimizing the single-machine single-process code into the form of single-machine multiprocess, it can be seen that the time is significantly reduced, and the multiprocess experiment has a significant effect.

With the training set and test set of a single machine and a single process as data input, the U_CF (user-based collaborative filtering algorithm) and O_CF (object-based collaborative filtering algorithm) of a single machine and a single process are run, respectively, and the calculation time, recall rate, accuracy rate, coverage rate, and average popularity are obtained, as shown in Figure 10.

Teachers should establish an effective teaching platform for the improvement of students’ practical ability, promote the flexible application of students’ knowledge, and strengthen students’ practical operation ability. Schools should increase capital investment and buy smart product components for teachers, encourage students to design and assemble independently, actively guide students to participate in AI competitions, stimulate students’ desire for independent learning in competition, and promote students to devote themselves to learning activities of information technology with full enthusiasm.

Deepen students’ understanding of information technology and AI technology, and allow students to work in groups to complete tasks assigned by teachers. Each group member has his or her own set of responsibilities, laying a solid foundation for improving students’ overall learning abilities. To perform on MoviesLens data, we used the U_CF algorithm, the O_CF algorithm, and an improved hybrid algorithm. Because MoviesLens’ data score range is 15, MAE was chosen as the evaluation index. The performance of these three algorithms is shown in Figure 11 for various \( k \) values.

As can be seen from the above experiments, using the CF algorithm in education data improves the accuracy of the recommendation result while also improving the performance of the improved algorithm. As a result, if the CF algorithm is implemented into the current teaching system, it will not only allow students to check for missing information and fill in gaps but it will also inform teachers about how well each student has mastered the course. Artificial intelligence is not unattainable. It is, in fact, intertwined with students’ academics and personal lives. Teachers can demonstrate AI application cases, teach students about AI image recognition and language recognition application scenarios, and help students understand the connection between AI and people’s lives in order to pique students’ interest in learning information technology [4].
The system will also recommend learning resources for students according to the learners’ mastery of test questions, so in this module, teachers need to add 2-3 social labels to each test question and classify each test question horizontally, and the test questions are divided into difficult, general, and easy. In the question bank part, you can query the summary of all test items, which are also arranged according to the knowledge tree structure and chapter contents. Teachers can also query, modify, and delete.

5. Conclusion

Individualization and intelligence are the key factors to develop online education. This paper studies the personalized learning platform based on CF thought and uses multi-process thought to optimize the training efficiency of the model. By building an AI-based smart classroom platform, using DL to analyze big data and monitor the whole teaching process in real time, the purpose of teaching students in accordance with their aptitude can be truly realized, and students’ learning passion can be greatly enhanced. Through the research of DL algorithm in big data, the mastery of each student to each exercise on the network teaching platform is collected, quantified as data expression form, input DL model, and get students’ understanding of all knowledge points. Finally, the feedback is given to teachers, so that teachers can repeat the explanation of knowledge points in a targeted way, and students can learn further. Taking the time decay factor as an input variable and considering the similarity measure, it reflects the timeliness of the input data of the system. Finally, the simulation experiment proves that the recommendation result of the improved algorithm can improve the performance compared with the recommendation result of the traditional algorithm.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author does not have any possible conflicts of interest.

References


