Research on Students' Course Selection Preference based on Collaborative Filtering Algorithm

Mustafa Man[1], Jianhui Xu[2], Ily Amalina Ahmad Sabri[3], Jiaxin Li[4]

Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Terengganu, Malaysia[1, 2, 3]

Software College, Liaoning Technical Universiti, Anshan, China[4]

Abstract—Due to the events caused by the COVID-19 pandemic, the education industry is no longer limited to offline, and online classroom education is widely used. The rapid development of online education provides users with more abundant educational course resources and flexible learning methods. Various online education platforms are also constantly improving their service models to give users a better learning experience. However, at present, there are few personalized information recommendation services in student course selection. Students receive the same course selection information and cannot be "tailored" according to their specific preferences. This paper focuses on the integration of collaborative filtering technology into a college course selection system to construct a rating matrix based on students' ratings of the courses they take through correlation between courses and correlation between students. Based on the collaborative filtering algorithm, a predictive rating matrix is generated to produce a recommendation list to achieve intelligent recommendation of suitable courses for students. The experimental results show that, based on the traditional collaborative filtering recommendation technique, the improved collaborative filtering algorithm based on both item and user weighting is used to achieve course recommendation with higher recommendation accuracy. The application of the improved collaborative filtering technique in the course selection recommendation system of colleges and universities is very good at recommending courses for students intelligently, and the recommended courses for students have good rationality and accuracy, and achieve more intelligent course selection for students, which has great practicality and practical significance.

Keywords—*Collaborative filtering; course selection system; recommendation; scoring matrix; weighting*

I. INTRODUCTION

With the rapid development of the Internet industry, the application of information technology is becoming more and more widespread in the management of academic affairs information, and the online course selection system has become a significance part of the management of academic affairs information. How to quickly find out the course you are interested in among the large amount of optional course information has become one of the research hot-spots of the course selection system. In the course selection system, most of them are based on the system search engine to query course information and take courses [1]. When facing a large number of available courses, students do not have the relevant knowledge as a basis to select courses because they do not know enough about the course information, which will lead to a waste of course resources. Currently, most of the course selection systems have no recommendation function or low quality of personalized recommendation, so it does not recommend the courses that students may be interested in. However, building a recommendation system based on feedback information such as students' major information and students' ratings of course information can solve these problems [2].

A. The Statement of the Research Problem

Elective courses for college students are courses that students can choose to study independently according to their preferences. Through elective courses, students can expand their knowledge. At present, the course selection system of most colleges and universities lists all the elective courses offered this semester in the system like commodities for students to choose [3]. The relevance and orientation between courses are poor. When college students choose courses, there are the following three problems:

1) Blindness: Students do not understand the relevance of courses and the direction of majors,Course selection is arbitrary.

2) Poor purpose, dealing with errands: Students only choose some courses that can easily pass in order to complete their credits, regardless of whether the courses they learn are helpful to their curriculum system. As a result, after the course is selected, the learning enthusiasm is poor and the learning effect is not good, which does not achieve the expected effect of elective courses(R. N. Behera and S. Dash,2016).

3) Instability and potential risks of course selection system: The traditional course selection mode has strong time constraint and does not take into account algorithmic fairness, which may often cause peak access and easily cause hidden danger to the security operation of the back-end system.

Therefore, personalized recommendation technology is applied to the course selection system to provide students with personalized elective course recommendations according to their needs and interest preferences, prevent students from choosing courses blindly, and greatly improve the utilization rate of university elective course resources and the operation efficiency of the course selection system.

B. Research Objectives

1) To analyzes and compares several recommendation algorithms, finds out their shortcomings and advantages, and determines the research idea of applying collaborative recommendation technology to course selection system.

2) By understanding the courses that students have taken and their evaluations, analyzing students' preferences, and pushing courses that target students may be willing to take.

3) To apply collaborative filtering technology to the field of students' course selection, and combine it with the basic course selection system to realize an efficient and convenient Personalized Course Selection recommendation system.

C. Research Question

The recommendation algorithm based on collaborative filtering is based on the similarity measure of individual behavioral characteristics. By calculating the similarity between the specified new sample (students) and the original sample in the database, the new sample is clustered, and the individuals with similar behavioral characteristics to the new sample are identified as the nearest neighbor samples (nearest students). After that, the selection set of the nearest neighbor sample is generated, and the selection set is sorted according to preference scores.

Ultimately, course selection recommendations are made to students based on the similarity of the new sample to the nearest neighbor students and the course selection preferences of the nearest neighbor students. There are two important issues that need to be addressed.

1) First, how to calculate the similarity between students through behavioral data, and require that the similarity reflect the interests and learning characteristics of students strengths.

2) Second, after identifying a new sample of near-neighbor students, how to determine the set of recommended courses to be selected based on the near-neighbor students' course selection records.

D. Rationale of the Study

In the field of teaching management practice of colleges and universities, in order to follow up the reform of higher education teaching and meet the demand of society for comprehensive and practical talents, the course selection system of China's colleges and universities has completed the conversion from academic year system to credit system, and schools provide students with a large number of elective courses, and students in colleges and universities are given more options to choose their favorite courses according to their interests. With the diversified development of the society, people are more and more concerned with a wide range of fields, and students' interests show a trend of divergence, so colleges and universities have also opened corresponding courses for students to choose in response to this phenomenon. However, in recent years, there are too many elective courses in colleges and universities, which lead to information overload when students choose courses, and it is difficult for students to choose courses that suit their personal development due to the structural deficiencies of course classification and specialization. This shows that the course selection process of students is not without research value, and there is a pattern of course selection behavior. Moreover, the traditional course selection mode has strong time constraint and does not take into account algorithmic fairness, which may often cause peak access and easily pose hidden risks to the security operation of

the back-end system. Therefore, personalized recommendation technology is applied to the course selection system to provide students with personalized elective course recommendations according to their needs and interest preferences, prevent students from choosing courses blindly, and greatly improve the utilization rate of university elective course resources and the operation efficiency of the course selection system [4].

The "department store" approach of simply improving the quality of teaching resources by simply listing them and letting students "pick and choose" is obviously no longer in line with the current requirements for "personalized learning".

Personalized learning requires adding a number of "shoppers" to the "department store" with a wide range of elective courses to help learners get the right course for them in a timely and accurate manner that is recognized by the learners. This "shopper" is the role that learning paths play in the learning process of learners, aiming to improve the precise guidance of learners, reduce the blindness of learners, and improve the efficiency of course selection [5].

E. Research Gap

At the same time, with the digital reform of higher education, some scholars have started to study the mining of student's one-card data to analyze student behavior. The students' one-card accumulates a large amount of student spending data and daily behavior data. However, there are still many shortcomings in the above-mentioned research on course selection recommendation systems. The recommendation systems based on students' course selection data use very limited course selection data of varying quality, which makes it difficult to accurately mine students' preferences, and they focus too much on the algorithm level, trying to copy the success of recommendation systems in e-commerce and entertainment fields to the education field, using various methods to improve the accuracy of the algorithm, while ignoring the characteristics of the education field itself and the limitations of the scoring matrix itself, resulting in no qualitative improvement in accuracy and hardly satisfactory recommendation results.

Web mining and bibliography mining are the theoretical basis of data mining technology in library user behavior analysis. Based on the research and analysis of library patron behavior composition and acquisition, library user behavior models are constructed by using machine learning and other algorithms. Through these models, we can understand the interest preferences of reader groups. However, there are still many shortcomings in the above-mentioned research on course selection recommendation systems. The recommendation systems based on students' course selection data use very limited course selection data of varying quality, which makes it difficult to accurately mine students' preferences, and they focus too much on the algorithm level, trying to copy the success of recommendation systems in e-commerce and entertainment fields to the education field, using various methods to improve the accuracy of the algorithm, while ignoring the characteristics of the education field itself and the limitations of the scoring matrix itself, resulting in no qualitative improvement in accuracy and hardly satisfactory recommendation results.

This paper proposes a collaborative filtering algorithm-based course selection recommendation system, which no longer pursues excessive algorithmic complexity, avoids the limitations of the scoring matrix itself, and realizes personalized course recommendation.

II. LITERATURE REVIEW

Collaborative filtering is a push technology that is often used to achieve the basic recommendation push function for the system. It is mainly to divide users into different sets by different tendencies, and to push items to target users according to the items that are closer to users' preferences, item's comment information, etc. as the basis for judging.

A. Theoretical Background

Collaborative filtering is to mine a small number of students with similar course preferences to the specified students in a large amount of data, and then designate these similar students as Then, we organize the course preferences of the near-neighbor students into a catalog sorted by preference, and finally recommend courses to the specified students based on the course preferences of similar students.

1) Similarity measure: The similarity measure between samples is the basis of cluster analysis. When doing classification, a sample is usually considered as 1 vector in an n-dimensional Euclidean space, so the similarity between 2 vectors in n-dimensional Euclidean space can be measured from the following 2 perspectives. One is from the fish degree of vector distance. Second, from the angle of vectors. In particular, since the Euclidean distance in high-dimensional space still satisfies the triangular inequality of distance, the Euclidean distance is the most common method to measure the vector distance in high-dimensional space [6].

a) Euclidean distance metric: In high-dimensional space, the Euclidean distance is a measure of the distance between points in vector space that is closest to the intuitive meaning of distance in three-dimensional space. By introducing the concept of Euclidean distance; the vector space has the concepts of length and angle. Suppose the samples x and y are 2 points in an n ($n \geq 1$) dimensional Euclidean space.

$$x = [x_1, x_2, x_3, \cdots, x_n]^T \tag{1}$$

$$y = [y_1, y_2, y_3, \cdots, y_n]^T \tag{2}$$

Then the Euclidean distance is calculated as follows.

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{3}$$

b) Manhattan Distance Metric: Manhattan distance is also called city block distance assuming that the sample x and y are 2 points in n ($n \geq 1$) dimensional space, we get equations (1) and (2), then Manhattan distance is calculated as follows.

$$d(x,y) = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n| = \sum_{i=1}^{n}|x_i - y_i| \tag{4}$$

c) Chebyshev distance metric: Chebyshev distance is also an important measure to define the distance of points in

vector space, which takes the maximum value of the distance in the component dimension as Chebyshev distance. Specifically, assuming that the samples x and y are two points in an n ($n \geq 1$) dimensional space, equations (1) and (2) are obtained, and the Chebyshev distance is calculated as follows

$$d(x,y) = lim_{p \to \infty} \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \tag{5}$$

d) Minkowski distance metric: The Min distance, sometimes referred to as the space-time interval, was first expressed by the Russian-German mathematician H. Minkowski (1864-1909). Assuming that the samples x and y are two points in an n ($n \geq 1$) dimensional space, equations (1) and (2) are obtained, and the Minkowski distance is calculated as follows.

$$d(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \tag{6}$$

From the calculation formula, Min's distance treats the components as obeying the same distribution, and also disregards the difference of the components in the magnitude.

e) Standardized Euclidean distance Metric: Similar to the Min distance, the simple Euclidean distance also suffers from the problem of treating the components as obeying the same distribution, and ignores the differences in the components in terms of mean and variance. By first standardizing the components and then calculating the Euclidean distance, an improved standardized Euclidean distance is obtained. Assuming that the samples x and y are 2 points in an n ($n \geq 1$) dimensional space, equations (1) and (2) are obtained, and the standardized Euclidean distance is calculated as follows.

$$d(x,y) = \sqrt{\sum_{i=1}^{n} \left(\frac{x_i - y_i}{s_i} \right)^2} \tag{7}$$

f) Angle cosine metric: The angle cosine is a measure of the similarity of sample points from the directional point of view, and is widely used in many fields. Suppose the samples x and y are two points in n ($n \geq 1$) dimensional space, and equations (1) and (2) are obtained, then the vector angle cosine is calculated as follows:

$$\cos(\theta) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \tag{8}$$

From the above equation, the absolute value of the cosine of the vector angle is less than or equal to 1. Its magnitude can reflect the similarity of the two vectors, and the larger the value, the higher the similarity of the two vectors [7]. Moreover, a positive value of the cosine of the vector angle indicates that the two vectors have an isotropic relationship, and vice versa, it indicates that the two vectors. The opposite indicates that the two vectors are negatively related [8]. Considering the similarity measures and the research needs of this paper, the similarity of students is calculated by using the similarity measure based on the cosine of the vector angle.

2) Student-based collaborative filtering: In the recommendation system, a sample of 343 senior students and their course selection records from the previous year at Liaoning National Normal College in China was selected as

the basis for this study and the similarity between classmates was calculated by referring to the angle cosine similarity measure. For the purpose of analysis, here is an example of seven students' collaborative over filtering process [9].

Suppose there are seven students, A, B, C, D, E, F and G, who choose the courses they want to take among five courses, a, b, c, d and e. Their course selections are as follows as shown in Fig. 1.

Based on the above students' course selection and the pinch cosine similarity measure, the similarity between students was calculated [10]. For example, the similarity between A and B is shown as below:

$$w_{AB} = \frac{|\{a,b,e\} \cap \{a,d\}|}{\sqrt{|\{a,b,e\}||\{a,d\}|}} = \frac{1}{\sqrt{6}} \quad (9)$$

The similarity between A and C is

$$w_{AC} = \frac{|\{a,b,e\} \cap \{c,d\}|}{\sqrt{|\{a,b,e\}||\{c,d\}|}} = 0 \quad (10)$$

Similarly, the similarity between A and D is $\frac{2}{\sqrt{6}}$, the similarity between A and E is $\frac{1}{3}$, the similarity between A and F is $\frac{1}{\sqrt{6}}$, and the similarity between A and G is $\frac{2}{3}$

Considering that the above algorithm needs to calculate the similarity between the specified sample (student) and any other sample, in order to improve the computational efficiency of the algorithm, the following improvement scheme is proposed [11].

Step 1: Create a course-to-student reverse lookup table.

Step 2: Build the student's congruence matrix based on the backwards checklist

Based on the selected status of each course in the sample, a backward checklist is created as shown in Fig. 2.
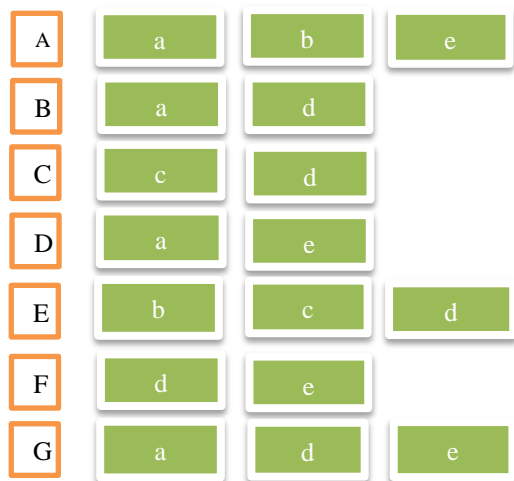

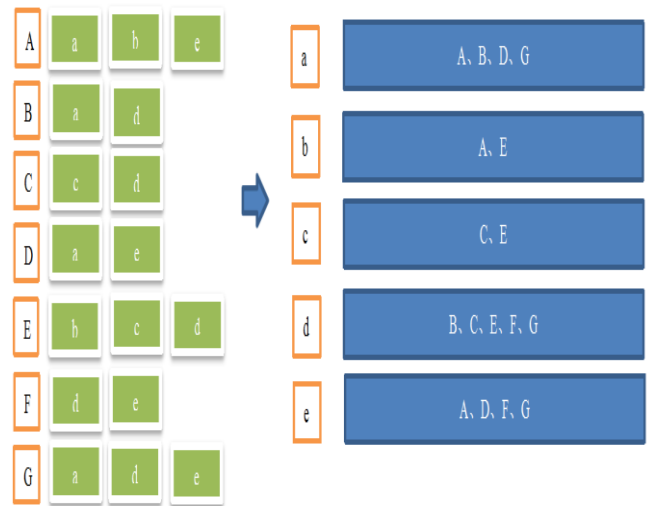Fig. 1.   Sample Student Course Selection List.

Fig. 2.   Sample Student Course Selection Backward Checklist.

Based on the inverted table in Fig. 2, the same present matrix was generated for different students, and the results are shown in Table I.

The co-occurrence table was normalized and the results are shown in Table II.

3) Collaborative course-based filtering: When designing a recommendation system based on collaborative filtering, it is necessary not only to find out the near-neighbor students of a given student based on behavioral data, but also to evaluate the similarity of courses In evaluating the similarity of courses, the algorithm uses students' ratings of courses to evaluate the similarity between courses [12]. In evaluating course similarity, the algorithm uses students' ratings of courses to evaluate the similarity between courses. To demonstrate the algorithm flow, the analysis is still based on the course selection records of seven sntudents, A, B, C, D, E, F, and G, in five courses, a, b, c, d, and e. The records are analyzed, and their course selections are shown in Fig. 3.

TABLE I.   CO-PRESENTATION MATRIX OF THE SAMPLE STUDENTS

	A	B	C	D	E	F	G
A		1		1	1	1	1
B	1		1	1	1	1	1
C		1			1	1	1
D	1	1				1	1
E	1	1					
F	1	1	1	1	1		1
G	1	1	1	1	1	1	

TABLE II. NORMALIZED MATRIX OF THE SAME-PRESENT MATRIX OF THE SAMPLE STUDENTS

	A	B	C	D	E	F	G
A		0.2		0.2	0.2	0.2	0.2
B	0.167		0.167	0.167	0.167	0.167	0.167
C		1			0.25	0.25	0.25
D	0.25	0.25				0.25	0.25
E	0.167	0.167	0.167	0.167		0.167	0.167
FF	0.167	0.167	0.167	0.167	0.167		0.167
GG	0.167	0.167	0.167	0.167	0.167	0.167	

Fig. 3. Sample Student Course Selection Records.

Then, the co-occurrence matrix of the course is normalized. The formula is as follows:

$$w_{ij} = \frac{|N(i \cap j)|}{|N(i)|} \tag{13}$$

The results of the treatment are shown in Table III.

TABLE III. NORMALIZED MATRIX OF COURSE CO-OCCURRENCE MATRIX

	a	b	c	d	e
a		0.33		0.33	0.33
b	0.25		0.25	0.25	0.25
c		0.5		0.5	
d	0.25	0.25	0.25		0.25
e	0.25	0.25		0.25	

After obtaining the students' ratings of the courses, and the similarity between the courses, estimating the preference of students u for the alternative course o [13].

The formula is as follows.

$$p_{u,0} = \sum_{i \in C(0,N)} w_{u,i} p_{u,i} \tag{14}$$

In Equation (14), the measure is the degree of preference of student u for course o. Calculating the degree of preference of a given student for all alternative courses is the basis for constructing the student course selection recommendation system [14].

In addition, C(0, N) in equation denotes the set of courses similar to the alternative course o, $w_{u,i}$ represents the similarity between the alternative course o and course i, and $p_{u,i}$ is the degree of preference of student u for course i calculated based on the nearest neighbor students [15].

B. Overview of Recommendation Systems

Currently, general education platforms use the course recommendation method based on data statistics, ranking the platform courses based on the number of course selections, and recommending the courses that most platform users are interested in, i.e. the popular course selection list. Among the current recommendation systems, collaborative filtering-based recommendation algorithms are the most widely used[16].The Recommended methods have their own advantages and disadvantages, in order to better understand their interrelationship and their respective advantages, the two types of algorithms are compared as shown in Table IV.

There are two types of collaborative filtering methods commonly used.

1) user-based collaborative filtering (User-based CF) algorithm, which judges the user's favorite degree of the project through the user's historical behavior, calculates the relationship between users according to different users' preferences for the same project, and recommends the project among users with the same preferences [17].

2) Item-based algorithm focuses on Item, whose similarity is mainly based on its inherent feature values, so it can be classified according to its feature values, calculate the proximity between them, and give suggested results. Since the classification of item is more stable, it can be pushed offline [18].

The comparison of user-based and item-based advantages and disadvantages, and the scope of application have been summarized as shown in Table V.

TABLE IV. COMPARISON OF CLASSIC RECOMMENDATION METHODS

Comparison of Classic Recommendation Methods		
Recommended method	*Advantages*	*Disadvantages*
Collaborative filtering recommendations	The ideas are relatively simple and easy to understand; No domain knowledge required; Good comprehension; Can handle unstructured data	Cold start problems; Data sparsity ;
Content-based recommendations	Avoid cold starts and data sparsity;	Difficulty in processing unstructured data Complex feature extraction Recommendation information is prone to over-fitting;

TABLE V.   COMPARISON OF USER-BASED AND ITEM-BASED ALGORITHM

Comparison of User-Based And Item-Based Algorithm		
Recommended method	**Advantages**	**Disadvantages**
User-based	Push for more socialization. Simply maintain user resemblance Table	Less pronounced specialization. Difficult to provide a push explanation
Item-based	Push specialization. Long-tail item enrichment. Convincing push explanations are available	Unable to adapt to Item updates fast speed

C. The limitation of Previous Studies

The current course selection recommendation system in universities is separate from the online course system. The vast majority of scholars have done a lot of complex work and research in the direction of personalized social network recommendations, e-commerce product recommendations, user emergence mining models, personalized recommendation videos, learning resources, and student user preferences, user profiles, deep recommendation algorithm models, etc.

However, there is very little content related to students' real course selection recommendations, as each school has a different course selection platform. Most schools are concerned about how to complete the task of course selection more easily and quickly, and guarantee the stable operation of the system's course selection platform by randomly drawing lots, or by centrally opening the platform for students to grab courses within a fixed period of time, and do not consider the satisfaction of students' diverse needs through a simple and brutal way.

Algorithms are at the core of personalized recommendation and collaborative filtering techniques. However, how to develop effective and accurate evaluation criteria for the recommendation results of algorithms is an issue that deserves constant attention both for academia and industry. Different evaluation criteria have different focuses, and a single evaluation criterion generally only evaluates a certain aspect of the algorithm, which is more or less deficient. Therefore, how to choose appropriate evaluation metrics to evaluate the recommendation results has a crucial impact on the development of the whole personalized service field [19].

III. METHODOLOGY

Combine course selection function and collaborative filtering and recommendation technology to realize intelligent course selection function. To analyze the trend of students' interest through their information in the system, such as courses taken, grades, ratings, comments, etc., and give them a list of courses. Focusing on the application of pushing process, user-based and item-base are applied to the selection of courses with student and course as the main objects of study. Since the traditional algorithms have some shortcomings, this paper applies item-based weighted and user-based weighted collaborative filtering algorithms to the course selection push system to improve the accuracy. In the design of the course selection push system, consider the major attributes of students, the relevant attributes of courses and the ratio of students to courses to ensure that the recommendation system can push the set of courses that students are really interested in [20].

A. Data Collection

1) Selection of data set: In this paper, we use the data set of the academic system of Liaoning National Normal University to implement the traditional pushing process. A part of the information is extracted as the initial data, and 238 students' evaluation information is obtained. The records were tabulated into students, courses, and ratings tables. Each data file contains the following details.

- Students：Stu-name、Stu-number、Sex、Grade、Score、Zx/Gx (Professional Elective Courses/Public Elective Courses)

- Ratings：Stu-number、Course ID, Rating,and Times.

- Courses: Course-ID、Course-name、Course-category.

The data set is the basis for implementing the course selection push function. Based on the students' course selection and course rating over a period of time, the students' interest level in various elective courses is analyzed and expressed by rank. In this paper, we use a two-dimensional matrix to represent the student's interest in a course, i.e., a $v \times w$ student-course favorite table vw, where v represents the number of students and w represents the number of selected courses. vw value represents the vth student's interest in the wth course. This matrix can be explained by Table VI.

TABLE VI.   STUDENT-COURSE GRADING SCALE

	Java	Python	Basketball	H5	Android	PS	Database
student1	5	1	2	1	0	1	4
Student2	2	3	1	2	2	2	5
Student3	2	0	0	4	5	5	2
Student4	3	3	2	3	3	1	3
…	…	…	…	…	…	…	…
Student N	2	5	0	4	4	4	3

B. Data Description

- Extraction of student attributes and student behavior characteristics:

Define student attributes and behaviors:

student（name、id、sex、grade、Course-name、score、Professional Electives/Public Electives）, For example:Students (Bao Fuyu, 201503, difficult, 21 Computer Applications, Computer Composition, 85, major elective).

- Course Properties

Define course attributes:

Course (course-name, course-id, course-time, score, grade, course-type), such as course (Java programming core technology, F1025, 36 hours 90, 19computer, elective).

C. Research Procedures

The basic framework of intelligent course selection push system is shown in Fig. 4.
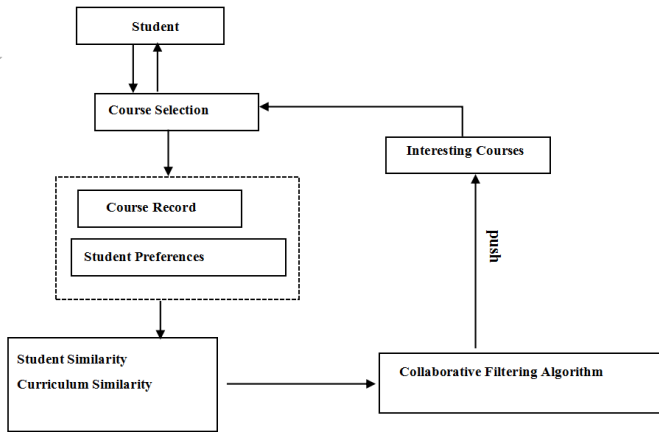

Fig. 4.   The Framework of the Push System.

The process of Item-based weighting and User-based weighting push is shown in the following Fig. 5.
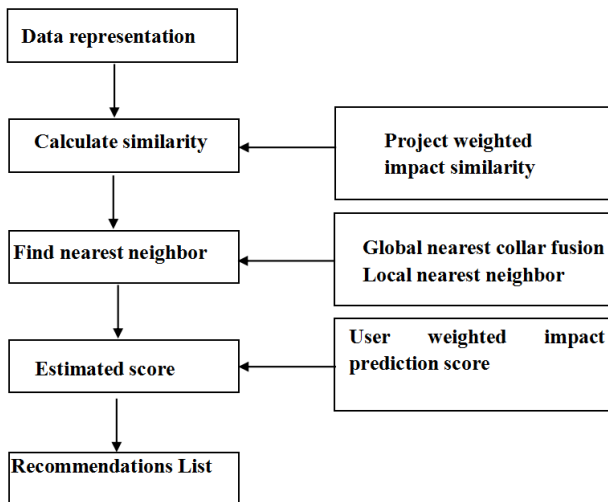

Fig. 5.   Item based Weighted and user based Weighted Recommendation Process.

1) Basic idea: Consider both course-weighted and student-weighted, and use item-based weighted user cosine similarity and user-based weighted user collaborative filtering to predict the evaluation and improve the accuracy of recommendation [21].

2) Description of the algorithm process: Input: target student T, course evaluation form R, course characteristics form A, number of neighbors k;

Output: Top-N recommendation set of target student T;

Step1: Find the set of graded courses of all students, courses and target students T in the system from the course grading table R, and denote them as U_m, I_n, I_T [22].

Step2： The top k students with the highest median value are selected as the closest set of students to student T according to the influence of different course weights to obtain the weighted cosine similarity N_T={ $j_1,j_2....,j_k$}.

Step3： For any ungraded course i of the target student T, a weighted approach is used to combine the predicted evaluation of student impact with the predicted evaluation of the student's historical evaluation.

Step4： Select N courses with higher predicted scores as the push result set for target student T.

IV.   ANALYSIS AND RESULTS

This chapter is the realization of the system function based on the description and design of the course selection system, then expounds the implementation effect of the recommendation function of the system, and effectively evaluates the recommendation function of the main program of the system according to the evaluation index.

A. Experimental Information

The computer test environment of this machine is shown in Table VII.

TABLE VII.   EXPERIMENTAL ENVIRONMENT AND INFORMATION SOURCES

Central processing unit model	Internal Memory (GB)	Operating System	Hard Disk (GB)	Database
Intel(R) Xeon(R) CPU X5650 @ 2.67GHz (12 CPUs)	16GB	Windows11	1TB	Sql Server 2018 R2

B. Data Conversion

The item based weighting and user based weighting recommendation methods are adopted. The item based weighting will affect the proximity value and nearest neighbor selection, and the user based weighting will affect the prediction and evaluation. It is applied to the university educational administration course selection push system to realize the purpose of intelligently pushing elective courses for students. First, get the original data from the database, sort out the original data, retain useful data, delete irrelevant records, and improve the efficiency and accuracy of the algorithm. The data involved in the algorithm studied in this paper are from the educational administration system of Liaoning National Normal University. The data will be processed separately to meet the requirements of the algorithm [23].

From the database, 7 data tables related to students' previous course selection and evaluation data of students' courses are selected. There are 8 tables, from which students' attributes, course selection attributes and students' teaching evaluation information can be obtained. From the interview table collected from students, it can intuitively obtain the students' interest in course in some directions [24]. Obtain the student number, name, course name, major, score, evaluation and other records useful for the algorithm. After re integrating the records, establish the correlation between tables and rewrite them into the database. See Fig. 6 as follows.
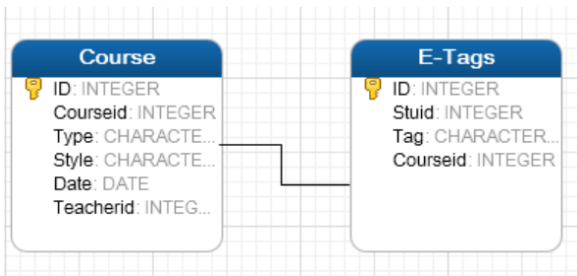
Fig. 6.   Association Table after Data Conversion.

C. Accuracy Comparison Results

Experiments show that the improved algorithm can improve the push accuracy. During the experiment, the data comes from.

The data of Liaoning National Normal University educational administration system and the information is collected by questionnaire survey. The test set takes three tenths of the information, the training set takes seven tenths of the information and arbitrarily turns it into five parts, which are expressed as data set 1, data set 2, data set 3, data set 4 and data set 5, respectively. The accuracy flow of push algorithm is shown in Fig. 7.
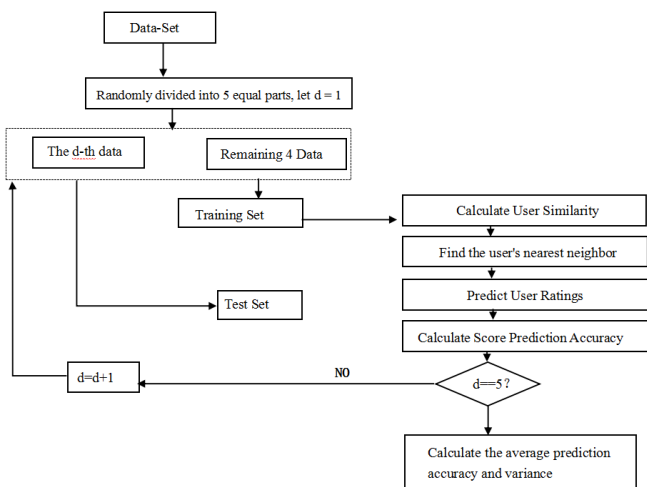

Fig. 7.   Accuracy Flow of Collaborative Filtering Algorithm.

V.   DISCUSSION

In order to test the prediction accuracy of the recommendation algorithm, the course data set is used for off-line calculation, the student behavior model is established on the training set to predict the student behavior on the test set, and the score prediction accuracy is calculated through the root mean square error RMSE. The detailed experiments are as follows:

- The experimental design uses different data to test the influence of algorithm convergence, and takes the course data set to test, including 343 students and 40 course scores. The data set is divided into two parts: 80% training data and 20% test data.

- Experimental data: under the condition of comprehensive prediction accuracy and efficiency, the experiment carries out three iterative tests of 10K, 100k and 1m data: 5, and 10. The RMSE data is shown in Table VIII.

A. Experimental Analysis

Through the experimental analysis, the following results can be obtained. When the number of iterations increases, the prediction accuracy of RMSE will decrease, and each doubling will decrease by 0.04, indicating that the increase of the amount of data will not significantly reduce the performance of the recommended algorithm, and the increase of the amount of data will make the convergence effect of the algorithm better [25]. At the same time, it can be obtained from the analysis that when the amount of data is the same, the choice of K value affects the prediction accuracy, that is, the smaller the K value, the higher the accuracy, and the larger the K value, the lower the accuracy. Therefore, K is selected as 5 as the number of iterations. The collaborative recommendation algorithm combines the characteristics of User-CF and Item-CF algorithms, and filters the students' information with tags at the initial stage of recommendation, which reduces the search scope to a certain extent. When recommending, filter the initial results of User-CF recommendation, and then make the final recommendation according to the score, so as to make the recommendation effect better, as Fig. 8 and 9.

After many times of verification and improvement, on the whole, the system meets the design requirements. The accuracy of the core algorithm in the recommendation function is evaluated. The experimental results show that the algorithm can produce.

TABLE VIII.   RMSE IN DIFFERENT ITERATION DATA

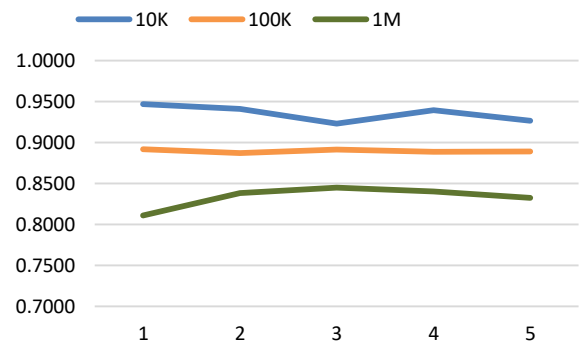| Data-Set	1	2	3	4	5	K
10K	0.9468	0.9411	0.9228	0.9393	0.9265	5
10K	0.9221	0.9254	0.9265	0.9162	0.9232	10
100K	0.8916	0.8870	0.8915	0.8886	0.8891	5
100K	0.8559	0.8671	0.8673	0.8672	0.8655	10
1M	0.8107	0.8382	0.8449	0.8402	0.8322	5
1M	0.8182	0.8182	0.8184	0.8163	0.8153	10

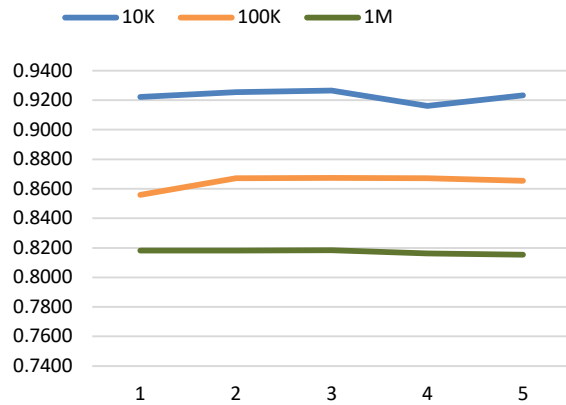Fig. 8.   Association Table after Data Conversion when k=5.

Fig. 9. Association Table after Data Conversion when k=10.

VI. CONCLUSION

In this paper, is used the collaborative filtering algorithm to study students' course selection preferences based on their course selection history data, and build a university course selection recommendation system based on this algorithm. A university course selection recommendation system is built based on this algorithm. The experimental results show that the collaborative filtering algorithm is able to mine and extract the attributes, behavioral characteristics and preferences of students, which can help solve the problems of students' course selection. It can help solve the problem of students' randomness and blindness in course selection, and improve the management efficiency of university education and teaching. In particular, this paper shows that the collaborative filtering algorithm-based student course selection recommendation system can better balance the relationship between student characteristics and course characteristics, and can help achieve the optimal matching between student learning ability and course requirements, student course selection preference and course characteristics, and student ability development and career requirements.

A. Innovation Points

Through the shortcomings of the existing course selection system and the urgent need for an intelligent course selection system, we select the existing course selection system of Liaoning National Normal University and analyze it. Then, introduce collected students' course selection data from Liaoning National Normal University's academic affairs system, pre-processed the data, and compared the user collaborative filtering algorithm based on user weighting with the traditional filtering algorithm through experiments, and the improved way considered the weights of student and course, verified that it can improve the recommendation pushing accuracy in a certain range, and the better pushing result is achieved.

B. Future Work

When looking for nearest neighbors, it can consider the fusion of local nearest neighbors and global nearest neighbors. Global nearest neighbors, local interests are not similar, local nearest neighbors, and global interests are not similar. Through

the optimization algorithm of similarity, the recommendation accuracy can be improved, the data can be used to a greater extent, and the error problem caused by sparse data can be improved. Following problems need to be further studied in the future to supplement and improve the paper: How to make the intelligent course selection recommendation system more accurate and more real-time, and consider the time complexity and space complexity of the algorithm to make the efficiency better?

REFERENCES

[1] J. Xiao, M. Wang, B. Jiang, and J. Li, "A personalized recommendation system with combinational algorithm for online learning," Journal of Ambient Intelligence and Humanized Computing, vol. 9, no. 3, pp. 667–677, 2018.

[2] R. N. Behera and S. Dash, "A particle swarm optimization based hybrid recommendation system," International Journal of Knowledge Discovery in Bioinformatics, vol. 6, no. 2, pp. 1–10, 2016.

[3] M. K. Najafabadi, A. Mohamed, and C. W. Onn, "An impact of time and item influencer in collaborative filtering recommendations using graph-based model," Information Processing and Management, vol. 56, no. 3, pp. 526–540, 2019.

[4] G. Geetha, M. Safa, C. Fancy, and D. Saranya, "A hybrid approach using collaborative fifiltering and content based fifiltering for recommender system," Journal of Physics: Conference Series, vol. 1000, pp. 012101–012123, 2018.

[5] L. Guo, J. LiangYing, Z.Y. Luo, and L. Sun, "Collaborative fifiltering recommendation based on trust and emotion," Journal of Intelligent Information Systems, pp. 1–23, 2018.

[6] H. Zarzour, F. M. Mohamed, S. Chaouki, and C. Chemam, "An improved collaborative fifiltering recommendation algorithm for big data," Computational Intelligence and Its Applications, pp. 660–668, 2018.

[7] H. Zhang, T. Huang, Z. Lv, S. Liu, and Z. Zhou, "MCRS: a course recommendation system for MOOCs," Multimedia Tools and Applications, vol. 77, no. 6, pp. 7051–7069, 2018.

[8] R. Logesh, V. Subramaniyaswamy, D. Malathi, N. Sivaramakrishnan, and V. Vijayakumar, "Enhancing recommendation stability of collaborative fifiltering recommender system through bio-inspired clustering ensemble method," Neural Computing and Applications, vol. 32, no. 7, pp. 2141–2164, 2020.

[9] Y. Yang, Y. Xu, E. Wang et al., "Improving existing collaborative fifiltering recommendations via serendipity-based algorithm," IEEE Transactions on Multimedia, vol. 20, no. 7, pp. 1888–1900, 2017.

[10] H. Li, H. Li, S. Zhang, Z. Zhong, and J. Cheng, "Intelligent learning system based on personalized recommendation technology," Neural Computing and Applications, vol. 31, no. 9, pp. 4455–4462, 2019.

[11] A. Klasnja-Milicevic, M. Ivanovic, B. Vesin et al., "Enhancing e-learning systems with personalized recommendation based on collaborative tagging techniques," Applied Intelligence, vol. 48, no. 6, pp. 1519–1535, 2018.

[12] J. Chen, C. Zhao, L. Uliji, and L. Chen, "Collaborative filtering recommendation algorithm based on user correlation and evolutionary clustering," Complex and Intelligent Systems, vol. 6, no. 1, pp. 147–156, 2020.

[13] L. Jiang, Y. Cheng, L. Yang, J. Li, H. Yan, and X. Wang, "A trust-based collaborative fifiltering algorithm for E-commerce recommendation system," Journal of Ambient Intelligence and Humanized Computing, vol. 10, no. 8, pp. 3023–3034, 2019.

[14] X. Liu, "A collaborative filtering recommendation algorithm based on the inflfluence sets of e-learning group's behavior," Cluster Computing, vol. 22, no. 2, pp. 2823–2833, 2019.

[15] J. Chen, C. Zhao, L. Uliji, and L. Chen, "Collaborative fifiltering recommendation algorithm based on user correlation and evolutionary clustering," Complex and Intelligent Systems, vol. 6, no. 1, pp. 147–156, 2020.

[16] Karalis, T., & Raikou, N. (2020). Teaching at the times of COVID19: Inferences and implications for higher education pedagogy.International Journal of Academic Research in Business and Social Sciences,10(5),pp 479-493.

[17] Wang, X., Zhu, Z., Yu, J., Zhu, R., Li, D., Guo, Q. (2019). A learning resource recommendation algorithm based on online learning sequential behavior. International Journal of Wavelets, Multiresolution and Information Processing, 17(2).

[18] Almusharraf, N., Khahro, S. (2020). Students Satisfaction with Online Learning Experiences during the COVID-19 Pandemic, International Journal of Emerging Technologies in Learning, 15(21):pp 246-267.

[19] Diao, X., Zeng, Q., Duan, H., Song, Z., Zhou, C. (2021). Personalized learning resource recommendation based on course ontology and cognitive ability. Journal of Computers, 32(2): pp149-163.

[20] Liu, Z.Y., Lomovtseva, N., Korobeynikova, E. (2020). Online Learning Platforms: Reconstructing Modern Higher Education, International Journal of Emerging Technologies in Learning, 15(13): 4-21.

[21] Chung, E., Subramaniam, G., & Dass, L. C. (2020). Online Learning Readiness among University Students in Malaysia amidst COVID-19. Asian Journal of University Education, 16(2),pp 46–58.

[22] Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. Human Behavior and Emerging Technologies, 2(2), pp113–115.

[23] Adnan, M., & Anwar, K. (2020). Online Learning amid the COVID-19 Pandemic: Students' Perspectives. Online Submission, 2(1), 45-51.

[24] Sahbaz, A. (2020). Views and evaluations of university students about distance education during the COVID-19 pandemic. Educational Process: International Journal (EDUPIJ), 9(3), pp184-198.

[25] Al-Salman, S., & Haider, A. S. (2021). Jordanian University Students' Views on Emergency Online Learning during COVID-19. Online Learning, 25(1), pp286-302.