Design of online course knowledge recommendation system based on improved learning diagnosis model

Baiqiang Gan1*, Chi Zhang2, Qiuping Dong3, Wei Sun4

1,4 Guangzhou Nanyang Polytechnic, Guangzhou, Guangdong, 510900, China
2 Nanfang College of Sun Yat-sen University, Guangzhou, Guangdong, 510900, China
3 Guangzhou College of Technology and Business, Guangdong, 528100, China
*jackian@163.com

Abstract. In order to better manage curriculum resources, give full play to the practical value and efficiency of university curriculum, provide an efficient and convenient interface for the majority of students and curriculum resources, and meet the different needs of different students in learning, this paper aims at the current content-based recommendation algorithm cannot solve the problem of user interest drift, online curriculum resource sharing platform cannot directly communicate with students this paper proposes a learning diagnosis method which combines the similarity of courses and user interest model to calculate the similarity. According to the actual characteristics and attributes of online course resources, and using the improved algorithm, this paper designs and implements a set of online course knowledge recommendation system based on the improved learning diagnosis model, to effectively alleviate the problems of low user interest and direct communication between online and offline classroom users in the current online course learning process.

1. Introduction
With the gradual integration of education industry and information technology, the core goal of internet education is to maximize the coverage of high-quality education resources, gradually realize a new educational thinking and practice paradigm, and connect all elements in the field of education. Online education not only brings us convenience, but also brings us many new problems, which need us to study and solve. With the increasing demand of online education market, online learning resources on the internet are growing geometrically. On the one hand, it provides users with convenient learning resources, on the other hand, it also causes users to find and match their own learning resources in the complex resources. In the face of information overload, users are easy to get lost, so they cannot choose the original target. Therefore, the recommendation system can better solve the problem of information overload. If we can use the recommendation system to recommend the courses that users may be interested in to users according to the characteristics of specific users' personal information, interests, selected courses, and evaluation of courses, we can make personalized recommendation for users, help users choose, and better serve users. In the personalized curriculum resources recommendation system, students can get more curriculum resources information, and the system can also dynamically obtain students' preferences with the passage of time. The characteristics of each student user are analysed, and the curriculum resources with high matching degree are spontaneously and actively recommended to students according to certain standards, to realize the personalized recommendation of the curriculum more accurately. The personalized recommendation
system constructed in this paper is a learning diagnosis recommendation method which combines online course resources with user interest model. Based on the user's learning diagnosis status and online course content introduction, combined with user browsing history, the system dynamically analyses the user's interest preference, updates the personalized recommendation for each user in time, and finds the most suitable course for users in line with their own learning direction. At the same time, the system combines offline classroom with course resource information, which is more conducive to the improvement of users' learning ability.

2. Related theories and technologies

2.1. Learning the theory of diagnosis model
At the beginning, students do not have a complete set of learning diagnosis model, so a set of authoritative knowledge in related fields is very important, which can bring some reference to students' learning direction, and will not let students have certain blindness in the learning process. His main function is to have a common understanding of his domain knowledge, and have a complete set of relationship models between knowledge and knowledge, knowledge, and learners, to achieve the integration and reuse of different knowledge, to guide the learning purpose of basic learners. The construction of learning diagnosis model is still in the initial stage of exploration, without a complete standard. Batarseh et al. [1] focus on the representation of domain knowledge based on the reasoning of students' ability and teachers' teaching strategies. Umek [2] uses knowledge management and tree document directory to build a simple domain learning diagnosis model. KN-AHS is an adaptive hypertext client. It divides the domain knowledge level based on learners' cognitive state through learners' information and other concepts. It adopts a tree structure, and the leaf nodes of the tree are used as different knowledge points. It is also the most common way to build a domain learning diagnosis model. In the process of personalized recommendation of learning diagnosis model, user behaviour information is the basis of the whole recommendation process, and one of the key links in the whole recommendation process. The user's behaviour information will eventually be expressed in a mathematical way to form a model for individual users. This model will be updated iteratively with the user's behaviour information and time to ensure the real and real-time embodiment of the user's interest. Therefore, if the system wants to form a user's real model, it needs to obtain the user's real needs and hobbies. There are two ways to obtain them. One is that the system presents the information selectively or according to the public user characteristics and product characteristics, and the user takes the initiative to choose, to inform the system. The other is to collect user's behaviour information by burying different functions. Therefore, this paper uses a reasonable data structure to express the user's interest model by combining the attribute characteristics of the system, which not only requires the accurate expression of the user's interest model, but also has a certain practicality in the calculation.

2.2. Online course knowledge point recommendation algorithm based on user interest points
Aiming at the problem that users' interests drift with time in content-based recommendation algorithm, this paper proposes a learning diagnosis method combining course similarity and user interest model to calculate the similarity, and designs a specific online course resource recommendation algorithm. Firstly, the TF-IDF technology [3] is used to extract the keywords of the course content, and the similarity of the course content is calculated according to the user interest information and online course information content. Then, the improved learning diagnosis model is used to calculate the weight preference of users for different course attributes, and the similarity matrix of course attributes is dynamically calculated according to the time control window to solve the problem of user interest drift. Finally, the two parts are weighted to get the final course similarity and results, to obtain the recommendation list. The implementation process of the improved model and algorithm is divided into three parts.

Step 1: key points extraction of online course knowledge. The key point of online course knowledge is the words that describe the main characteristics and content of the course. This system
uses TF-IDF algorithm to obtain the set of keywords that can describe the course content from the course information. Among them, TF-IDF algorithm is mainly combined with word frequency and reverse file frequency. When the TF-IDF value of a word is higher, it means that the word is more important to the whole article. Suppose that the course set is \( M = (m_1, m_2, m_3, \ldots, m_n) \), the word set in the course information content is \( N = (n_1, n_2, n_3, \ldots, n_n) \), and \( M_j = (m_1^j, m_2^j, m_3^j, \ldots, m_n^j) \) represents the information content of the \( j \) course, where \( w_n^j \) represents the weight of the \( n \) word \( n_n \) in the course information, and the weight value is calculated according to the TF-IDF calculation formula (1), and formula (2) is the weight calculation formula.

\[
\text{TF-IDF}(N_j, M_j) = \frac{\text{TF}(N_j, M_j) \log \frac{N_j}{N}}{\sqrt{\sum_{i=1}^{N} \text{TF-IDF}(N_i, M_i)^2}}
\]

Step 2: calculate the online course attribute weight. Due to the continuous changes of user interests, long-term and lasting interests are called long-term interests, and those that change greatly in a short period of time are called short-term interests. Therefore, the long-term and short-term interest model is helpful to optimize the personalized recommendation service. The recommendation system model constructed in this paper uses the window control method to control and classify the user interests, and sorts the user interests according to the total number of visits to the attributes from more to less, to calculate the long-term interest attribute weight \( (w_{ij}) \) and short-term interest attribute weight \( (w_{sij}) \) of the user. When the number of visits to the \( j \) attribute is higher than the average number of visits to all attribute tags, the attribute tag is added to the user's long-term interest tag set \( L \). The number of long-term interest visits is \( q_{ij}^l \), and the number of visits to the most attribute is \( Q_{max} \). Then the formula for calculating the weight of long-term interest attribute is \( w_{ij} = \frac{q_{ij}^l}{Q_{max}} \). When the user's access frequency to the \( j \) attribute is greater than the threshold \( \theta \), we add this attribute to the user's short-term interest set \( S \). The number of short-term interest visits is \( q_{ij}^s \), and the total number of short-term visits is \( Q_{sum} \), so the short-term interest attribute weight calculation is \( w_{sij} = \frac{q_{ij}^s}{Q_{sum}} \).

Step 3: get the online course recommendation results. According to the key points of online course knowledge and the weight of course attributes, the model and algorithm proposed in this paper use cosine similarity calculation method and user interest model to calculate the similarity of course attributes, and on this basis, calculate the similarity of courses by weighting, finally, get the similarity in the order of value from high to low recommendation list \( \text{Top} - N \), to realize the generation of recommended courses. Formula (3) is the cosine similarity calculation method, where \( \vec{a} \) and \( \vec{b} \) are two \( n \) dimensional vectors and \( \alpha \) is the angle between the two vectors.

\[
\cos \alpha = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} \sqrt{\sum_{i=1}^{n} b_i^2}}
\]

3. Framework design of recommendation system

The improved learning diagnosis model online course knowledge recommendation system uses JavaEE three-tier architecture [4], adding a component layer between the client and the database. Therefore, the three-tier structure of the application is divided into presentation layer, business logic layer and data layer. Based on TF-IDF algorithm and learning diagnosis model, this paper proposes a learning diagnosis method which combines course similarity and user interest model to calculate the
similarity, to solve the problem of user interest drift and recommend appropriate personalized course resources to users. Users include two groups: online users and offline users. The recommendation mechanism of the two groups is based on the improved learning diagnosis model. For the recommendation mechanism of online users, the system will view the record data of the target user's system interaction line in the system database according to the user's unique identity. If there are records in the database, there will be two processes. One process is special feature matching [5], the system will match the data of learning and interest features and hobbies filled in by the target user on the first landing page recorded in the database to form a data set, and then use the cosine similarity formula to calculate the data in the data set, and screen out the Top $N$ list with high similarity. The second process is offline user recommendation, the system will search the user's historical interaction record data, learning data and user comment data, and calculate the data through intelligent algorithm to obtain the appropriate online course resources. Figure 1 is the framework flow of recommendation system.

![Figure 1. Recommender system framework process.](image)

#### 4. Module design of recommendation system

**4.1. Function design of recommendation system**

The function of the improved learning diagnosis model online course knowledge recommendation system is divided into two parts: front-end design and back-end design. The front-end design includes user module, course module and learning material module [6]. User information module includes: user registration, user login, user information modification. The course module includes course retrieval module, view course details module, view recommended course module, browse history module, and course evaluation module. The learning materials module includes uploading files, downloading learning materials and viewing learning materials. The background of the recommender system includes six modules, which are user information management module, course information management module, learning materials management module, learning information management module, location information management module and course review management module.

**4.2. Implementation of recommendation system core module**

The recommendation system module can be divided into user information management module, recommendation system page display module, user history module, course retrieval module, course interface display module, learning resource module, user center module, course details module, administrator module and course management module, among them, user information management module and course recommendation module are the core of the whole recommendation system. The user information management module is the entrance to the recommendation system. After the user
enters the system, the login interface is displayed. The system makes a server request for the user's account and password. The result of the request is that the user exists or does not exist. If the user does not exist, it enters the user registration interface. If the user exists, it checks whether the user remembers the password function. If it is selected, the account and password will be saved in the local cookie, this interface and the next preference feature interface help to solve the cold start problem of the system and improve the basic information of users. On the one hand, it can help users enter the system smoothly, on the other hand, it can accumulate data for users’ recommendation. Figure 2 is the sequence flow of user information module.

The course recommendation module is one of the core functions of the online course knowledge recommendation system based on the improved learning diagnosis model. The module uses the learning diagnosis method which combines the course similarity and user interest model to calculate the similarity, solves the problem of user interest drift, and provides users with more accurate course recommendation. Figure 3 shows the process of course recommendation.

---

**Figure 2. Sequence flow of user information module.**

**Figure 3. Course recommendation process.**

---

Core code of course recommendation module

```java
Coursera=MapSort.sortMapByValue(Coursera);
String[] id_array = new String[6];
Int n = 0;
for (Map.entry<String,Float> entry: Coursera.entrySet())
{ System.out.println(entry.getKey()+entry.getValue());
  if(n<=6)
```
try{
    key = entry.getKey();
    Key = key.split(",")[0];
    if(LoopArray(id_array,key))
    {id_array[n]=key;}
}

5. Recommendation system test and results

In order to verify the effectiveness of the recommended system, the hardware environment of the test is Intel(R) Xeon(R) CPU E5-2689 @ 2.6Ghz, memory is 32G, and the operating system is windows 10 professional version. The system test is divided into two parts. The first part is function test [7], which is mainly used to check whether the system can work normally for users. In the test process, the appropriate platform or browser is used to ensure the user's good experience. Starting from the interface of the system, regardless of the internal architecture and source code of the system, through the functional test of each function, to verify whether the function still maintains the correctness of the system and the fluency of operation under various operations of users. Through the detailed test of the user information management module and the course recommendation function module in the core module of the recommendation system, the expected results are consistent with the final actual results, and the target requirements are achieved. The test results show that the function of the recommendation system proposed in this paper is more stable and the efficiency is more efficient. Table 1 shows the functional test results.

| Test items          | User information management module | Course recommendation function module |
|---------------------|-------------------------------------|---------------------------------------|
| Test method         | Black box test                      | Black box test                        |
| Test content        | Registered users input correct or wrong user name and password; Administrator input correct or wrong user name and password; Unregistered users enter information to register; | Log in and out of the system to view the course information; Log in to the system to view the recommended courses; The user can view more courses in different directions; Different users log in to view the recommended courses; |
| Expected results    | Enter the user interface correctly; Promt user name or password error, login failure; Enter the management interface correctly; Prompt user name or password error, login failure; Successful registration; | Users can view the recommended courses; The results of the display recommendation are different; Different users display different recommended courses; |
| Actual results      | The actual operation is consistent with the expected results | The actual operation is consistent with the expected results |

The second part is the compatibility test, which mainly tests whether the display effect of the recommender system proposed in this paper is correct in different browser interfaces. The recommender system is tested with 360 browser, Google browser and QQ browser respectively. The test results show that the recommender system displays normally and meets the compatibility test standard. Table 2 shows the results of the recommended system compatibility test.

| Browser name         | 360 browser, Google browser, QQ browser |
|----------------------|----------------------------------------|

Table 1. Function test results of recommender system.

Table 2. Recommender system compatibility test results.
| Number | Problem description                                                                                                                                  | Test result |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 1      | After successful login, whether the account number and password will be filled in automatically by clicking                                           | Yes         |
| 2      | Is the text box of the registration interface normally displayed without focus                                                                       | Yes         |
| 3      | Does the text box of login and registration change the border color after getting the focus                                                         | Yes         |
| 4      | After entering the main interface, whether the matching results are displayed normally                                                                | Yes         |
| 5      | Can the tabs at the top of the page switch normally and smoothly                                                                                    | Yes         |
| 6      | In the recommendation interface, whether the graphic arrangement of the recommendation results is displayed normally                                  | Yes         |
| 7      | In the view details interface, is the comment content displayed in secondary form                                                                   | Yes         |
| 8      | In each interface, whether the text arrangement is beautiful and neat                                                                               | Yes         |
| 9      | In each interface, is the layout reasonable                                                                                                | Yes         |

6. Conclusion
In this paper, aiming at the problems of low online learning interest of current users, online course resource sharing platform cannot directly contact with offline classroom, a learning diagnosis method combining course similarity and user interest model to calculate the similarity is proposed, and according to the proposed method, an improved learning diagnosis model online course knowledge recommendation system is designed to help users solve the problem of interest drift question. Finally, the stability and reliability of the recommender system proposed in this paper are verified by function test and compatibility test, which can provide reference for researchers engaged in personalized recommender system.

Acknowledgments
Research and implementation of online course knowledge recommendation system based on learning diagnosis model (no. 2020KTSCX378). Online open course project of film and television synthesis. Provincial excellent online open course project of Guangdong Provincial Department of education in 2018. Research on the third language teaching quality monitoring mechanism based on PDCA Cycle Theory (no.2020WQNCX109).

References
[1] Batarseh, F., Gonzalez, A. (2015) Validation of knowledge-based systems: a reassessment of the field. Artificial Intelligence Review.43(4):485-500.
[2] Umek, A., Zhang, Y., Tomažič, S., et al. (2017) Suitability of strain gage sensors for integration into smart sport equipment: A golf club example. Sensors.17(4): 916.
[3] Kos, A., Umek, A. (2018) Smart sport equipment: SmartSki prototype for biofeedback applications in skiing. Personal and Ubiquitous Computing.22(3) : 535-544.
[4] Palumbo, E., Rizzo, R. (2017) Entity2rec: Learning User-Item Relatedness from Knowledge Graphs for Top-N Item Recommendation. In: Proceedings of the Eleventh ACM Conference on Recommender Systems. Hainan. pp. 32-36.
[5] Wang, Q., Mao, M., Wang, B. (2017) Graph Embedding: A Survey of Approaches and Applications. IEEE Transactions on Knowledge and Data Engineering.29(12):2724-2743.
[6] Wang, H., Wang, J., Pi, L. (2019) Recommendation Algorithm based on User Reviews and Ratings. International Journal of Performability Engineering.15(3):803-812.
[7] Hao, L., Zhang, Y., Kang, Y., Yang. (2020) Movie recommendation algorithm based on circular knowledge mapping and collaborative filtering. Computer engineering and applications.6(2):106-114.