International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 1031-1040

https://doi.org/10.70517/ijhsa46379

Optimization of Knowledge Point Association Structure and Learning Path Planning Based on Partition Algorithm in Massive Online Civic and Political Education Course System

Mingcheng Wang[1] and Xingmin Qi[2,*]
[1] School of Marxism, Nanning College of Technology, Nanning, Guangxi, 530006, China
[2] Hubei Institute of Logistics Technology, Xiangyang, Hubei, 441100, China
Corresponding authors: (e-mail: qxm_2201@163.com).

Abstract This paper proposes an efficient attribute approximation algorithm based on partition method. Through the predefined attribute order relationship, the positive region of decision table is decomposed into multiple equivalence classes, combined with the fast sorting method to reduce the computational complexity and realize the efficient attribute approximation. The knowledge structure tree model is constructed, and the core keyword similarity calculation and subtree clustering methods are integrated to obtain the knowledge point aggregation relevance and dynamically aggregate the cross-curricular knowledge points of Civic and Political Education. A collaborative filtering recommendation model based on knowledge association is introduced, combining the TF-IDF algorithm and user similarity measure to optimize the prediction accuracy of resource scores and realize personalized recommendation of resources. The method of this paper is applied to the knowledge association mining of Civics and Political Science courses in colleges and universities, and the optimization of resource recommendation realizes the planning and guidance of students' learning path. The study shows that the method of this paper can effectively mine the knowledge point association and calculate the aggregated correlation degree, and the aggregated correlation degree of the three main Civic and Political Science courses is more than 0.75, 0.85, and 0.60, respectively. The maximum RMSE value of the recommendation model is 2.03563, the maximum MAE result is 1.50122, and the accuracy of the score prediction is better than the comparison method.

Index Terms partitioning algorithm, attribute approximation, knowledge structure tree, aggregated association, collaborative filtering recommendation

I. Introduction

Entering the new era, higher education has ushered in a new development opportunity. In line with the development of information technology, the establishment of a large-scale online ideological and political education program system is not only the requirement of the times, but also the responsibility that colleges and universities should undertake [1]. Ideological and political theory course, as a key course to implement the fundamental task of establishing moral character, is the main channel and main position to cultivate the soul with innovative theories, enlightenment and moistening the heart, and has an irreplaceable role in cultivating socialist builders and successors [2], [3]. In the final analysis, it is a course to transform theoretical knowledge into the ideological and political quality of the educated, and the starting point is in the education of people [4]. Therefore, firmly grasp the focus of online teaching of ideological and political theory courses, strengthen the construction of online resources, build the cornerstone of the development of online teaching of ideological and political theory courses, and help to enhance the students' sense of independent learning and interest [5]-[7].

In order to fully utilize the massive data resources in online ideological and political education, and further assist the decision-making of the administrators, teachers' lectures, and guidance of students' learning, educational data mining methods have emerged [8]. Among them, grades and knowledge scores reflect the performance of students' academics, and data mining of related content, linking the results to the state of students, the difficulty of knowledge topics, assisting in understanding students' learning, and providing help to students in resource recommendation [9]-[11]. Combined with intelligent algorithms for accurate chemistry guidance helps students deepen their knowledge understanding, form a personalized knowledge system, and consolidate the connection between what they have learned and what they are learning now, and can also be based on the students' own learning level and characteristics, planning a learning path for them, and ultimately achieving the win-win goal of a more accurate and scientific teaching and learning process [12]-[14].

The expansion of the scale of online Civics education puts forward higher requirements on the tightness of the structure of knowledge points and the accuracy of resource recommendation. In this paper, we reduce the complexity of attribute approximation by partition method. Combined with the semantic similarity calculation of key knowledge points, the aggregated associations of different course knowledge points are mined to construct the course knowledge structure tree. The improved collaborative filtering recommendation model and TF-IDF algorithm are utilized to carry out accurate recommendation of Civics learning resources and intelligent planning of learning paths. The application effectiveness of the method of this paper is verified through experiments.

II. Technical support for knowledge linkage and learning pathway planning

This chapter describes the technical support related to the optimization of the knowledge-point association structure and learning path planning of the online Civic and Political Education course system from several dimensions, such as algorithm design, knowledge-point association analysis, and resource recommendation implementation method.

II. A. Efficient Attribute Approximation Algorithm Based on Partitioning Approach

II. A. 1) Concepts related to the order of attributes

On the attribute space, a fast decomposition of the domain objects can be performed using the partition method. However, a premise is needed here: the decomposition order of attributes needs to be given in advance. The study of attribute order is of great significance in data mining for domain users. In this chapter, the idea of the partition method is integrated into the positive region computation and attribute approximation process of the decision table under a given attribute order, and a new attribute approximation algorithm is proposed, and the basic concept of the attribute order will be given below.

Theorem 1 (Skowron's approximation rule based on Skowron's discrimination matrix) Let M be the discrimination matrix of the decision table $S=\langle U, A=C\cup D, V, f\rangle$. $R\left(R\subseteq A\wedge R\neq\varnothing\right)$ is an approximation if and only if:

$$\forall_{\alpha\in M}\left(\alpha\neq\varnothing\rightarrow\alpha\cap R\neq\varnothing\right) \tag{1}$$

We define a complete ordinal relation for the entire set of conditional attributes. Based on this ordinal relation, we can give an equivalence relation for the Skowron discrimination matrix. For the convenience of the subsequent presentation, the basic concepts about attribute order are first briefly introduced here.

Let $S=\langle U, A=C\cup D, V, f\rangle$ be a decision table. We define a complete ordinal relation "$\prec$" on C, and at the same time, we label all attributes in C with 1 to $|C|$ respectively. In this way, we have a sequence of attributes on C, which in this paper is called the "attribute sequence" $SO: c_1\prec c_2\prec\cdots\prec c_{|C|}$.

Let M be the Skowron discriminant matrix of $S=\langle U, A=C\cup D, V, f\rangle$. $\forall\delta\in M$, the attributes in δ inherit the sequence SO from left to right, e.g., $\delta=c_jB$, $c_j\in C$, B is a subset of the attributes in C, and c_j is the first attribute of δ under the sequence SO, here we call c_j called the labeling attribute of δ.

For c_j, we define the set: $L(SO)=\left\{\delta\mid\delta=c_jB,\delta\text{ inherits the sequence }SO\text{ from left to right }\delta\in M\right\}$. It is easy to verify that: $L(SO)$ is an equivalence relation on M and divides M into multiple equivalence classes. One can express its division in terms of the quotient set: $M/L(SO)=\left\{[c_1],\cdots,[c_{|C|}]\right\}$. since $[c_i]\cap[c_j]=\varnothing$ (when $i\neq j$) and hence this division is unique. Therefore, each element of the Skowron discrimination matrix belongs to only one equivalence class and that equivalence class is determined by the labeling attribute of the discrimination matrix element. This division can be expressed in terms of the subscripts of the attributes: $M/L(SO)=\left\{[1],\cdots,[|C|]\right\}$.

In the equivalence class obtained by dividing $L(SO)$, the labeled attribute with the largest subscript plays a very important role in attribute approximation. Let

$$N=\max\left\{\left[M/L(SO)\right]\right\}, 1\leq N\leq|C| \tag{2}$$

Its labeling attribute is a_N. R is an approximation, assuming $R=\varnothing$.

II. A. 2) Fast Approximation Algorithm Combining Partitioning Methods under Attribute Ordering

In attribute approximation based on the invariant positive region of the decision table, computing the positive region of the decision table is one of the most important computations. For the computation of the positive region the partition method can be used (the method of fast sorting belongs to the partition method). The average time complexity of fast sorting for multidimensional tables results in $O\left(|U|\times\left(|C|+\log|U|\right)\right)$, and the space complexity is $O(|U|)$. Considering the space complexity, this chapter adopts the fast sorting method to compute the positive region of the decision table, and the attribute approximation algorithm under the attribute order is given below.

Theorem 2 A sufficient condition that the discriminant matrix element B_{xy}^{s} in M is not empty is:

$$\begin{aligned}&\left(x\in Pos_C(D)\wedge y\in Pos_C(D)\wedge(d(x)\neq d(y)\right)\\&\vee(x\in Pos_C(D)\wedge y\notin Pos_C(D))\\&\vee(x\notin Pos_C(D)\wedge y\in Pos_C(D))\end{aligned}\tag{3}$$

Theorem 3 Let $[k]\in M/L(S)$, then $[k]\neq\varnothing$ is sufficient if $\exists x,y\in U$, satisfying the following two conditions:

(1) The discriminant matrix element B_{xy}^{s} in M is not empty;

(2) $\forall_{k_1(1\leq k_1<k)}\left(c_{k_1}(x)=c_{k_1}(y)\right)\wedge\left(c_k(x)\neq c_k(y)\right)$.

Theorem 4 Given Decision Table $S=\langle U,A=C\cup D,V,f\rangle$ and given Attribute Sequence $SO:c_1\prec c_2\prec\cdots\prec c_{|C|}$, a sufficient condition for Attribute $c_k\left(c_k\in C\wedge 2\leq k\leq|C|\right)$ to be a non-null-labeled attribute is that $\exists B_{xy}^{s}\in M$ satisfies $c_k\in B_{xy}^{s}$ after sequentially removing the elements of the discrimination matrix containing Attribute $c_1,c_2,\cdots,c_{k-1}$ from the Skowron discrimination matrix.

II. B.Logical structures and formal definitions

In order to describe the hierarchical relationships, coupling associations, aggregation associations between nodes, and quantitative associations between different categories of nodes, the whole model knowledge structure tree and associations are composed of two basic elements.

Definition 1 The course knowledge structure tree is a multi-branching tree with a depth of 4, consisting of branching nodes (nodes at levels 1 to 3) and leaf nodes (nodes at level 4). In this case, the branching nodes on the root node of the tree (i.e., layer 1) to layer 3 are uniformly mapped as: course - base content branching - knowledge point distribution, and the leaf nodes (layer 4 nodes) are resource nodes. T_c - the course knowledge structure tree:

$$T_c=\left\langle t_{sub},t_{source},R_t\right\rangle\tag{4}$$

where: the branch node t_{sub} is a quaternion: $t_{sub}=\left(t_{id},f,key,d\right)$; t_{id} is the number id that uniquely identifies the tree node; f is the set of subject domains; key is the set of keywords of the knowledge point; and d is the definition of the knowledge point or textual description.

The leaf node (resource) t_{source} is a quintuple:

$$t_{source}=\left(S_{id},S_k,S_d,S_{ty},S_{URL}\right)\tag{5}$$

where S_{id} is the id encoding that uniquely identifies the resource label; S_k is the set of knowledge point keywords; S_d is the resource description; S_{ty} is the resource category description; and S_{URL} is the resource storage link address.

The intra-tree node association R_t describes the association property between two nodes within T_c, represented as a quaternion:

$$R_t=\left(r_m,r_s,w,t\right)\tag{6}$$

where r_m is the number id of the master node (parent node); r_s is the number id of the slave node (child node); w is the weight of the degree of association between the two; and t is the class of the association (in both cases, the branch node association or the branch node and the leaf node association).

Definition 2 User knowledge structure tree

$$U_t=\left(u_{ifo},<u_f,T_c>,R_{u_t}\right)\tag{7}$$

where: u_{ifo} is the user information;

$$u_{ifo} = \left\{ u_i \mid (u_{id}, u_{age}, u_{deg}, u_{dept}, \cdots) \right\}, i = 1, 2, \cdots, n \tag{8}$$

u_{id} member id; u_{age} age; u_{deg} education; u_{dept} unit and department, etc;

$\langle u_f, T_c \rangle$ is a tree describing the user's area of specialization, the core knowledge it possesses and its counterpart; R_{u_t} is a user-skill association describing the user's competence in each core knowledge and its corresponding weight coefficient.

Definition 3 The user-resource association $R_{us} = (u_{id}, S_{id}, C, D)$, denotes the association information between users and resources.

Where u_{id} is the id encoding of the user; S_{id} is the resource node encoding; C is the user's evaluation, rating, etc., of the resource; and D is the preference weight accounted by the user's evaluation, rating, etc., of the resource.

where T_p is the domain core knowledge tree, describing the main core knowledge structure of each specialized domain; M_F is a finite set, $M_F = \left\{ T_{c\_i} \mid T_{c\_1}, T_{c\_2}, \cdots, T_{c\_m} \right\}$ is the set of core knowledge trees covered by the sub-corresponding domain after being split; m is the number of courses after being split; R_{PF} is the core course association within the domain, which describes the composition of the core knowledge within the corresponding domain and the corresponding weight coefficients (i.e., a description of the importance of the corresponding knowledge points).

Definition 5 Knowledge point aggregation association $R_{kc} = (K_{id}, t_{id}, C_e, F)$, which denotes the knowledge points that are similar, analogous, or the same among the courses, and it can effectively describe the knowledge points that are repeated in the course clusters, and it can synergize the allocation of learning time or mastery of the knowledge points in the study of the various courses. Degree.

Where, K_{id} is the aggregation association id; t_{id} is the set of aggregatable knowledge points id; C_e is the text describing the aggregation association or the recognized definition of the knowledge point (valid definition); F is the degree of association or contribution of the knowledge point to the aggregation of the weighting factor.

Example of "User-Course Knowledge Structure Tree-Resource-Association", e.g., describes the user (U202101**, ***, ...) , area of specialization (computing, automation), courses (data structures T_{c_DS}, databases T_{c_DB}, operating systems T_{c_OS}, signal processing T_{c_SP}), aggregated associations of the knowledge points in the courses R_{kc} (Queue, B-tree/B+-tree), and interconnections.

II. C.Discovering Knowledge Point Aggregation Associations Rkc

According to the open-source tokenizer-Jieba, a custom dictionary of core keywords is defined based on the National Standard Discipline Classification and Code and the recognized entry: "HowNet Entry", and loaded into the knowledge point aggregation and association operation in the format of "Word_Part of Speech_Word Frequency".

In this paper, we propose the knowledge point aggregation association R_{kc} based on the similarity of labeled keywords. Firstly, the similarity of the nodes in the knowledge structure tree in terms of the core keywords is calculated; secondly, the similarity value is used for clustering; finally, the threshold value is used for filtering to complete the knowledge point aggregation association.

II. C. 1) Calculating the similarity of core words

The similarity of the two keywords w_1, w_2 is calculated using the earlier research results for the data field "key" in the third level node of the knowledge structure tree: "core words of the research field". The keyword similarity function formula $sim(w_1, w_2)$ is defined as formula (9):

$$sim(w_1, w_2) = \frac{p(a) * (s(w_1) + s(w_2))}{\left\{ p(a) * (s(w_1) + s(w_2)) + p(aw_1) * s(w_1) + p(aw_2) * s(w_2) \right\}} \tag{9}$$
$$\log(10 * \max\_ p(aw))$$

where $p(a)$ is the position where the two keywords w_1, w_2 common ancestor nodes are located; $s(w_1)$ and $s(w_2)$ denote the positions of the two keywords w_1, w_2 in the tree; $p(aw_1)$ and $p(aw_2)$ denote the positional differences with the two keywords w_1, w_2 common ancestor nodes; and $\max\_p(aw) = \max\left(p(aw_1), p(aw_2)\right)$, i.e., the maximum of $p(aw_1)$ and $p(aw_2)$.

II. C. 2) Clustering by similarity values

On the basis of word similarity, we realize knowledge point aggregation association calculation. The similarity of the aggregated associated knowledge point structure and key description matching is transformed into the similarity calculation of two subtrees, which is then transformed into the similarity calculation of the data domain "key" of each node in the structure tree. The knowledge point aggregation correlation is calculated as formula (10):

Let t_{sub_i} and t_{sub_j} be the two knowledge nodes (branch nodes on layer 3 of T_c) to be verified and aggregated, k_i and k_j are the number of keywords in the node data field "key", and F_T_n represent the number of branches contained in the t_{sub_i} and t_{sub_j} subtrees.

$$sim\left(t_{sub\_i}, t_{sub\_j}\right) = \frac{1}{F\_T\_n} \sum_{K=1}^{F\_T\_n} \max_{1 \le P \le k_i, 1 \le q \le k_j} sim\left(w_{p\_i}, w_{q\_j}\right) \tag{10}$$

where w_{p_i} and w_{q_j} are the p, q th keywords in the set of "key" descriptors in the data domain of the node, and the computation of the sub-equation $\max_{1 \le q \le k_j, 1 \le P \le k_i} sim\left(w_{p\_i}, w_{q\_j}\right)$ denotes that the similarity is taken to be the maximum of the word similarity $sim\left(w_1, w_2\right)$ within the set of descriptive keywords.

II. D.Resource Recommendation Implementation

II. D. 1) TF-IDF similarity calculation based on knowledge association

Traditional TF-IDF methods mainly focus on word frequency and document frequency. When using the TF-IDF algorithm for association analysis between learning resources and knowledge, the deep semantic relationship between words and knowledge needs to be considered. In order to extract keywords that are more relevant to the knowledge points, the semantic distance between the vocabulary and the knowledge points needs to be incorporated when calculating the inverse document frequency.

Firstly, based on the semantic similarity between the vocabulary w_i and the knowledge points, screening is performed by setting a similarity threshold u, and when the similarity of a keyword reaches or exceeds this threshold, the corresponding knowledge point z will be selected into the set Z_i. Subsequently, the knowledge points that are semantically closest to w_i are selected from the set Z_i, and their highest similarity values are used as weights to adjust the inverse document frequency. Meanwhile, in order to comprehensively measure the relationship between vocabulary and knowledge points, the inverse query strategy is further introduced to calculate the proportion of the sum of the TF-IDF values of all keyword items t with similarity to w_i exceeding another threshold value λ. This statistical value reflects the strength of interrelationships between keywords and knowledge points, and reveals the degree of prominence of these keywords in the document. By fusing this semantic and statistical information, the association between learning resources and knowledge points can be analyzed more precisely. The improved calculation formula is shown below.

$$tf_i = \max\left(sim\left(w_i, z_i\right)\right) \times \sum_{t \in \{sim(i,t) > \lambda\}} sim(i,t) \times tf_t \tag{11}$$

$$Z_i = \left\{z \in Z, sim(w_i, z) \ge u\right\} \tag{12}$$

II. D. 2) Specific recommendation process

Figure 1 shows the flow of collaborative filtering recommendation algorithm based on knowledge association. The algorithm in this section aims to recommend a Top-N collection of resources for the target learner u. The specific steps of the collaborative filtering recommendation algorithm based on knowledge association are as follows:

(1) Calculate the similarity between learners based on the cosine similarity measure and determine the set of learners S_u that are most similar to the target learner u;

(2) Calculate the similarity between knowledge resources using the TF-IDF algorithm based on knowledge correlation, and construct a prediction model to predict and fill in the missing scores in the learner resource scoring matrix by combining the known scoring data;

(3) Utilizing the set of learners S_u to which the target learner u is most similar, predict learner u's ratings $P_{u,i}$ of unassociated learning resources based on the learner resource ratings matrix with the following formula;

$$P_{u,i} = \bar{R}_u + \frac{\sum_{i \in NN} sim(u,n) * \left(R_{n,i} - \bar{R}_n\right)}{\sum_{i \in NN} \left(|sim(u,n)|\right)} \tag{13}$$

where $sim(u,n)$ represents the cosine similarity between learner u and learner n, NN represents the scoring matrix of "learner-learning resource", $R_{n,i}$ is learner n's evaluation of learning resource i, and $\bar{R}_n$ is the average overall evaluation score of learner n. The predicted score $P_{u,i}$ of learner u on unassociated learning resources i is obtained through computation. The higher the predicted score, the greater the interest of learner u in learning resource i, which makes it more suitable to recommend to learner u;

4) Perform descending sorting based on the predicted scores, and select the top N learning resources with high scores as the recommended results to the target learner u.

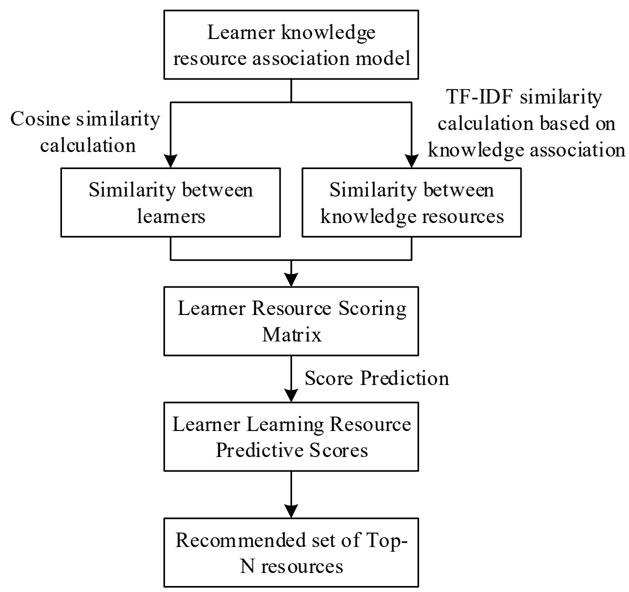


Figure 1: Collaborative filtering recommendation based on knowledge association

III. Practice and analysis of knowledge linkage and learning path planning

Analyzing the specific contents of the large-scale online Civics education course system and the actual Civics course system in colleges and universities, it is found that there exist 3 major compulsory Civics courses for students, which are Thought Ethics and the Rule of Law, Basic Principles of Marxism, and Situation and Policy. This chapter takes these 3 compulsory Civics and Politics courses as an example to verify the effectiveness of the knowledge point association structure optimization and learning path planning of this paper's method.

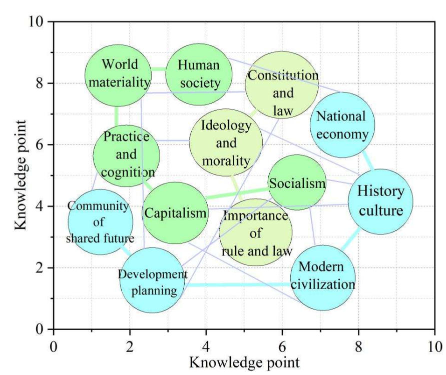


Figure 2: Knowledge network of 3 main ideological and political courses(Top 13)

III. A. Analysis of the knowledge linkage of the Civic Education program

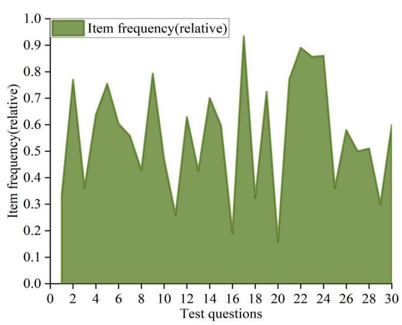
Figure 2 shows the knowledge point network of the 3 key Civics courses in the large-scale online Civics education course system. Through the network of knowledge points in Figure 2, it can be clearly seen that there is a correlation between the 13 key knowledge points of the 3 courses in the massively online Civic and Political Education course system. For example, there is a correlation between development planning in Situation and Policy and the materiality of the world in Fundamental Principles of Marxism, and these 2 knowledge points are in turn correlated with the Constitution and Law in Ideology, Ethics, and the Rule of Law. Therefore, specific association mining of user-course-basic content-knowledge points, etc. based on keyword similarity can be further carried out to optimize the course knowledge structure tree.

III. B. Course Knowledge Point Association Mining and Visualization

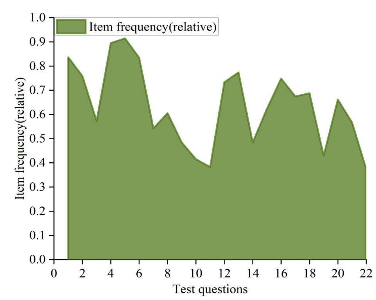
III. B. 1) Association Rule Frequency Mining for Knowledge Acquisition

Based on students' Civics test paper answer data, the applicability of the partitioning algorithm to different Civics courses possessing different knowledge point association rules is compared. Specifically, the algorithm is used to analyze the test paper for knowledge association; in order to judge the accuracy of knowledge association, the hierarchical structure of the test questions and the attributes of the knowledge points are calibrated after discussion with experts in Civic and Political Education and frontline teachers. The data of Civic and Political Education final test papers answered by a total of more than 10,000 students in the sophomore year of a university were studied for the knowledge association relationship. It included three tests on Ideological Ethics and Rule of Law, Basic Principles of Marxism, and Situation and Policy. Among the tests, only objective questions were retained and subjective questions were deleted. And the scores of the objective questions were processed with secondary scoring conversion to ensure the implementability of the algorithm and improve the accuracy of the results.

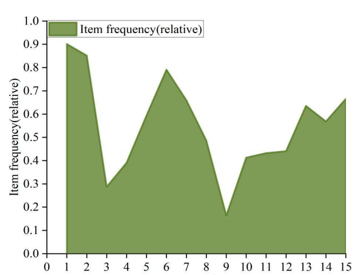
Figure 3 shows the association rule frequency mining of students' mastery of each course knowledge point, i.e., the probability of correctly answering each question. Combined with Figure 3, it can be found that in the test questions in the three courses, those with high scores (>0.6) are generally basic knowledge points, such as the connotation of the basic concepts of ideology and morality, etc., with a low knowledge point attribute hierarchy and low difficulty; those with medium or low scores (<0.6) are complex knowledge points belonging to the same hierarchical type, for example, the impact of rule of law construction on the development planning of the country, with a high knowledge point attribute hierarchy and difficulty. Using the model of this paper can effectively analyze the students' answer situation and explore the correlation between students' knowledge mastery and the knowledge points of related courses.



(a) Ideology, morality and rule of law



(b) The basic principles of Marxism



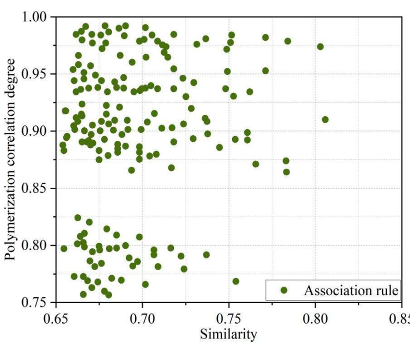
(c) Situation and policy

Figure 3: Frequency mining of association rules for knowledge mastery

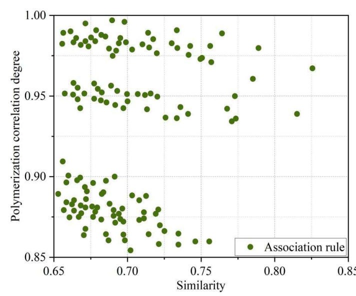
III. B. 2) Visualization of the association of knowledge points in different courses

Based on the mined association rules of different course knowledge points, the specific relationships between different course knowledge points are obtained. It proves the value of the logical structure of user-course-basic content-knowledge points, etc. of the model in this paper, as well as the accuracy of the calculation of knowledge point similarity and aggregated correlation.

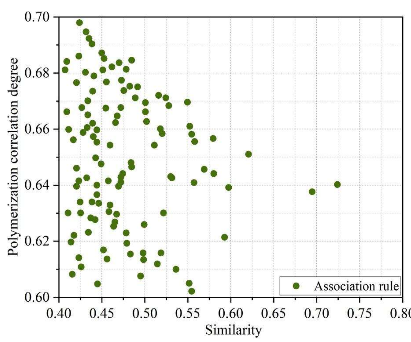
Figure 4 shows the visualization of the association of different course knowledge points. The scatter plot in Figure 4 shows that there are differences in the knowledge associations of different courses. The knowledge points of Ideology, Ethics and the Rule of Law and Fundamental Principles of Marxism are relatively close, while Situation and Policy is relatively scattered. The similarity of the knowledge points of Ideology, Ethics and the Rule of Law is above 0.65, and the aggregated correlation is above 0.75; the similarity of most of the knowledge points of Fundamental Principles of Marxism is above 0.65, and the aggregated correlation is above 0.85; and the similarity of most of the knowledge points of Situation and Policy is above 0.4, and the aggregated correlation is above 0.6. The correlation between the knowledge points of Basic Principles of Marxism is the strongest, followed by Ideology, Morality and Rule of Law, and Situation and Policy is the lowest. Combined with the actual course content, Situation and Policy focuses on the more macro domestic and international related content, while Basic Principles of Marxism focuses on the core content of Marxism, and Ideology, Ethics and the Rule of Law is divided into two key sections, Ideology, Ethics and the Rule of Law. Therefore, using the model of this paper for the analysis of students' answers between different courses, as well as the calculation of the related knowledge point association situation, we can get the actual association in line with the course content-knowledge points.



(a) Ideology, morality and rule of law



(b) The basic principles of Marxism



(c) Situation and policy

Figure 4: Visualization of association rules of different course knowledge points

III. C. Validation of Resource Recommendation Effect

III. C. 1) Experiments on resource recommendation based on user similarity

In order to improve the accuracy of resource recommendation, the cosine similarity measure is utilized to calculate the similarity between a total of more than 10,000 students in the sophomore year of this university. According to the calculation results, this paper categorizes these students into 4 categories according to the similarity of knowledge mastery, A, B, C, D, with about 2500 students in each category. Class A is the student with the best mastery of the knowledge points of the Civics course, class B is the second, class C is the second, and class D has the lowest mastery. On the basis of classification, the method of this paper is used to realize the recommendation of corresponding Civics resources. Figure 5 shows the results of resource recommendation based on knowledge association. Observing Figure 5, it can be seen that the resources recommended by the model to the students of category A are mainly advanced resources, and the amount of recommended resources reaches about 0.97. The students of this category have a higher degree of mastery of the knowledge points, and they are more familiar with the knowledge points between different Civics courses, so recommending the advanced resources to them will help them to improve the level of Civics in this category of students. Accordingly, Category D students have the lowest level of mastery of the knowledge points related to the Civics courses, so more review resources need to be recommended, accounting for about 1/3 of the total resources. and if only review resources are recommended, the curiosity of this category of students about the related Civics courses may be reduced, so a suitable collection of basic and advanced resources need to be recommended, so as to anticipate and correct the students' learning paths, and to guide the students of Category D in the continue learning at the stage where the knowledge points are not mastered to a high degree. Therefore, it is initially judged that the method in this paper can achieve the desired effect of resource recommendation.

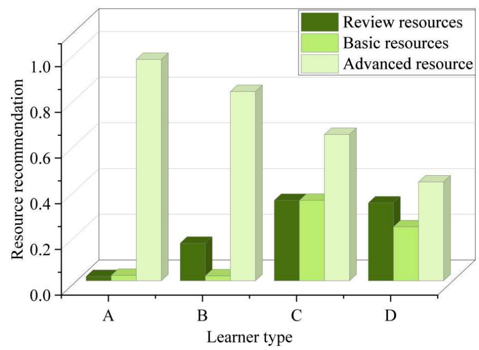
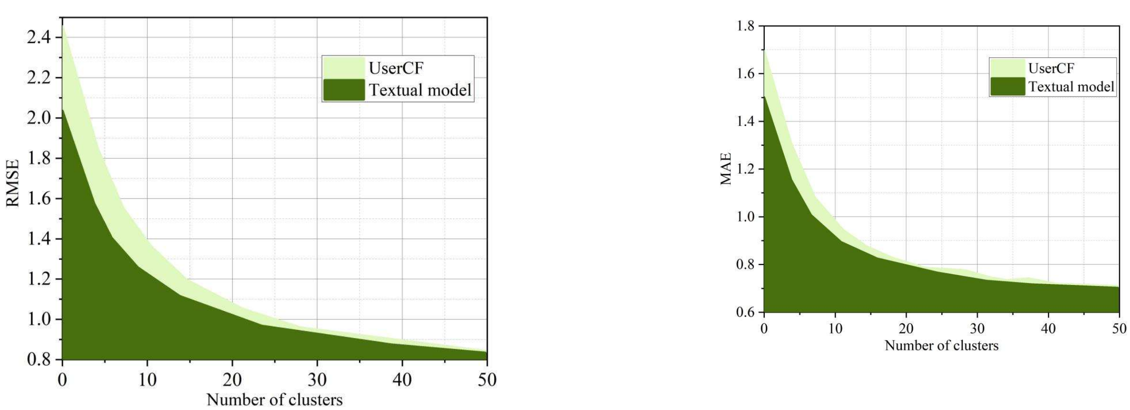


Figure 5: Resource recommendation results based on knowledge association

III. C. 2) Comparative Experiments on the Effectiveness of Resource Recommendation

The recommendation performance of this paper's collaborative filtering recommendation algorithm based on knowledge association is compared with the traditional collaborative filtering recommendation algorithm (UserCF) to understand the rating prediction accuracy of this paper's method and to determine the algorithm's resource recommendation performance advantage. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are used to measure the rating prediction accuracy of the algorithm. Figure 6 shows the experimental results of RMSE and MAE of the 2 methods with different number of clusters. According to the experimental results in Fig. 6, it can be seen that the range of RMSE values calculated by this paper's method is [0.83371,2.03563] and the range of MAE results is [0.70235,1.50122]; while the range of RMSE values for UserCF is [0.84219,2.45588] and the range of MAE results is [0.70981,1.69533]. The RMSE and MAE results calculated by the method in this paper are smaller than those of UserCF, indicating that the prediction accuracy has been improved to some extent. Then in the recommendation of learning resources with close association of Civics knowledge, the quality of Civics resource recommendation of this paper's method is higher than the traditional collaborative filtering recommendation algorithm.



(a) RMSE (b) MAE

Figure 6: Comparison of RMSE and MAE under different cluster numbers

IV. Conclusion

In this paper, we optimize the association structure of knowledge points through the partition algorithm and combine it with the improved collaborative filtering model to realize the personalized planning of learning paths in Civics courses. The method proposed in this paper can effectively calculate the similarity and aggregation relevance between the knowledge points of different Civics courses. The aggregation relevance of Basic Principles of Marxism is over 0.85, the aggregation relevance of Ideology, Morality and Rule of Law is over 0.75, and the aggregation relevance of Situation and Policy is more than 0.6. Based on the degree of mastery of the students' knowledge points, these courses are divided into four categories, and personalized recommendation of learning resources is realized. The RMSE value of this paper's method under different number of clusters ranges from [0.83371,2.03563], and the MAE ranges from [0.70235,1.50122], and the quality of resource recommendation is higher than that of the comparison algorithm. In the future, deep learning can be further introduced to enhance the semantic association analysis of key knowledge points, in order to improve the quality and efficiency of large-scale Civics course knowledge point association mining, and optimize the students' learning path.

Funding

This work was supported by the National Social Science Foundation of China (Grant No.21&ZD013).

References

[1] Gao, H. W. (2023). Innovation and development of ideological and political education in colleges and universities in the network era. International Journal of Electrical Engineering & Education, 60(2_suppl), 489-499.
[2] Fan, Y. (2025). Research on the Innovation Path of Network Ideological and Political Education in Colleges and Universities. Curriculum Learning and Exploration, 2(4).
[3] Yang, G. (2025). Constructing the Ideological and Political Culture Education System of Career Planning Class in Higher Vocational Colleges and Universities from a Strategic Perspective in the Context of Information Technology. Cultura: International Journal of Philosophy of Culture and Axiology, 22(2), 274-290.
[4] Li, H. (2025). Multicultural data assistance mining analysis for ideological and political education in smart education platforms using artificial intelligence. Wireless Networks, 31(1), 567-581.
[5] Gao, H., & Yasin, M. H. M. (2023). Challenges and development strategies in ideological and political education in the internet era in Higher Education. The Educational Review, USA, 7(11), 1661-1665.
[6] Yu, Y. (2022). On the ideological and political education of college students in the new media era. Open Journal of Social Sciences, 10(1), 1-14.
[7] Yang, H. (2024). E-learning platforms in ideological and political education at universities: students' motivation and learning performance. BMC Medical Education, 24(1), 628.
[8] Wang, Z., & Yu, N. (2021). Education Data-Driven Online Course Optimization Mechanism for College Student. Mobile Information Systems, 2021(1), 5545621.
[9] Liu, J., & Wen, B. (2020). Construction and optimization of educational technology course knowledge network. International Journal of Information and Education Technology, 10(9), 694-703.
[10] Liang, Y. (2023, December). Analysis and Optimization of Teaching and Learning Paths in Universities Based on Association Data Mining. In International Conference on Educational Technology and Administration (pp. 257-267). Cham: Springer Nature Switzerland.
[11] Niu, Q. (2021). Optimization of teaching management system based on association rules algorithm. Complexity, 2021(1), 6688463.
[12] Rafiq, M. S., Jianshe, X., Arif, M., & Barra, P. (2021). Intelligent query optimization and course recommendation during online lectures in E-learning system. Journal of Ambient Intelligence and Humanized Computing, 12(11), 10375-10394.
[13] Sato, Y., Izui, K., Yamada, T., & Nishiwaki, S. (2019). Data mining based on clustering and association rule analysis for knowledge discovery in multiobjective topology optimization. Expert Systems with Applications, 119, 247-261.
[14] Dwivedi, P., Kant, V., & Bharadwaj, K. K. (2018). Learning path recommendation based on modified variable length genetic algorithm. Education and information technologies, 23, 819-836.