International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 348-365

https://doi.org/10.70517/ijhsa46325

Research on Intelligent Optimization Strategies for English Vocabulary Learning Based on Group Intelligence Algorithm in Big Data Era

Xiaochao Yao[1,*]
[1] Foreign Language Teaching Department, Hainan Vocational University of Science and Technology, Haikou, Hainan, 571126, China
Corresponding authors: (e-mail: mamalin8483@163.com).

Abstract In this paper, based on the characteristics of English vocabulary and students' behavioral data, a cognitive level test learning model based on IRT theory is constructed, and for the defects existing in K-means clustering algorithms, a user clustering recommendation algorithm based on the minimum variance is obtained by using the minimum variance optimization initial cluster heart method. On this basis, a personalized recommendation platform for English vocabulary learning based on students' vocabulary level with dichotomous K-means clustering is designed and implemented, and the effectiveness of the platform is verified. The experimental results show that the method proposed in this paper can very directly observe that the topic parameters and students' learning ability values match the information reflected in the actual data. In addition, the accuracy of this paper's model in successfully recommending learning resources can be improved by up to 58.23% compared with the traditional model, and relevant extended knowledge of non-vocabulary subjects such as oral expressions is given in the recommendation results, which alleviates the problem of increasingly narrow vision of students caused by the cocoon effect. Teaching experiments show that this strategy can significantly improve students' learning of English vocabulary and increase the mean value of their English scores by 7.4628 points. Obviously, the model in this paper solves the defects of the existing English vocabulary learning software that does not meet students' individual needs.

Index Terms IRT theory, K-means, minimum variance optimization, English vocabulary learning, personalized recommendation

I. Introduction

With the deep development of globalization, English, as the main language of international communication, occupies a pivotal position in China's education system. As the cornerstone of English education, English vocabulary teaching has a direct and far-reaching impact on the improvement of students' language proficiency, communicative competence and comprehensive quality [1], [2]. However, traditional English vocabulary teaching mainly relies on classroom explanations, demonstrations and exercises, while students learn English vocabulary by listening, memorizing and imitating [3]. This teaching mode can help students master a certain amount of vocabulary to a certain extent, but the efficiency is low and the results are not obvious, which makes it difficult for students to master and apply English vocabulary [4], [5]. With the development of big data, the application of population intelligence algorithm in English vocabulary learning has gradually been paid attention to, which can achieve intelligent optimization strategies for vocabulary learning, thus improving the effect of vocabulary learning [6], [7].

Population intelligence optimization algorithm refers to a kind of optimization method based on multi-intelligence body collaboration, self-organization, learning and evolution developed by integrating cross-disciplinary knowledge of computer science, artificial intelligence, mathematics and other disciplines with reference to the behavioral pattern of group intelligence in nature, which is mainly divided into particle swarm algorithm, artificial fish swarm algorithm, ant colony algorithm, immune algorithm and so on [8]-[11]. By modeling the excellent problem-solving ability of population intelligence in nature, it enables computer systems to find the best solution to a problem by distributed algorithms in a process similar to evolution in nature [12]-[14]. In optimization problems, the application of population intelligence algorithms is becoming more and more widespread, not only to improve the solution efficiency in complex optimization problems, but also to provide a basis in the fields of social network analysis, intelligent manufacturing, intelligent transportation, and learning intelligent optimization strategies [15]-[18].

Literature [19] proposes a hybrid model of deep learning aimed at facilitating English vocabulary learning, highlighting the fact that advanced Natural Language Processing (NLP) techniques can be applied to the development of smart educational technologies to help non-native English-speaking students learn vocabulary at speed, providing a personalized, adaptive and immersive learning experience. Literature [20] proposes student mental optimization of intelligent deep neural networks for personalized English vocabulary learning recommendations, which focuses on English vocabulary learning performance metrics and excels in terms of recall and correctness, which helps in language comprehension and application. Literature [21] discusses the integration of Artificial Intelligence (AI) in English vocabulary acquisition, emphasizing that AI provides effective instructional solutions through adaptive learning platforms, intelligent tutoring systems, and natural language processing, which focuses more on student engagement, retention, and overall proficiency improvement than traditional methods. Literature [22] examined the application of adaptive algorithms in English vocabulary learning, which can adjust the learning content and difficulty in real time according to the learning progress and performance, ensuring that students learn in the most suitable learning environment, and verified that the algorithms effectively improve students' vocabulary learning. Literature [23] examined the application and effectiveness of AI technology in college English vocabulary teaching, and based on the limitations of vocabulary teaching, proposed the application of AI technology in personalized learning, adaptive technology, etc., which showed that AI can effectively improve the effect of vocabulary teaching. Literature [24] introduces NLP and proposes an English vocabulary assisted learning system based on digital twin Wasserstein generative adversarial network, which reveals the effectiveness of the method in English vocabulary learning by comparing it with other methods, with a significant improvement in its accuracy. Literature [25] proposed an innovative adaptive English vocabulary recommendation system based on deep reinforcement learning, which models vocabulary learning as a continuous decision-making process so as to dynamically adjust the recommendations based on user performance and contextual factors, providing a powerful solution strategy for personalized learning. Literature [26] developed a vocabulary learning model for human-computer interaction based on generative artificial intelligence, which improved learners' vocabulary richness and comprehension by analyzing their learning outcomes and experiences, providing insights for in-depth exploration of human-computer dialogue learning models. The above study outlines the application of artificial intelligence methods such as natural language technology, digital twin technology, and deep reinforcement learning to optimize English vocabulary learning, which provide students with personalized learning approaches to facilitate English vocabulary learning.

In this paper, we first design the construction process, data items, and storage method of the student model, and then use item response theory as a test method for students' cognitive level to realize the dynamic update of the student model. After that, the minimum variance-based K-means algorithm is used to cluster the data, reduce the nearest neighbor search space, and improve the scalability of the algorithm. According to the idea based on waterfall hybrid technology, the bifurcated K-means clustering algorithm based on students' abilities is combined with the collaborative filtering recommendation algorithm based on students' dynamic interests to compute the similar user set and recommend high-frequency vocabulary resources in the similar user set for students. Finally, the parameters and recommendation performance of the model are determined, and the application effect of the model is verified through real cases.

II. IRT-based personalized recommendation algorithm for English vocabulary learning resources

II. A. Learning system modeling based on IRT theory

II. A. 1) Student modeling methodology

The student model is a data structure used to characterize the current knowledge state of students, which reflects their individual characteristics, knowledge state, and cognitive abilities. The student data in the student model is an important basis for the system to make content recommendations. The specific construction process is as follows:

Step1: Basic information is collected after the first registration, which is used to build the static model of students on the one hand, and to generate the credentials for logging into the system entrance on the other hand;

Step2: After logging into the system, students can choose content learning and testing, and student behavior data is collected during the learning process, which is used to record the current learning progress of students and update the knowledge status data, so as to facilitate the next step of learning when students log into the system again;

Step3: When students finish learning the current content, they need to answer the corresponding test questions. The system utilizes the Item Response Theory testing model to test the cognitive level as an important influencing factor for updating the knowledge status data;

Step4: If students have already learned certain content, they can directly enter the test session to avoid wasting time;

Step5: After students complete the learning or testing session, the model is dynamically updated by combining the learning behavior and cognitive level data;

Step6: Finally, the updated model data will be stored in the student database, which is convenient for the system to call subsequently.

II. A. 2) Cognitive level testing method based on IRT theory

Item response theory allows for the effective use of data from testing sessions by examining the relationship between ability level, item difficulty, item discrimination, and guessing coefficients. The basic principle is to present the test questions adaptively according to the cognitive level of each tester and the response situation during the test.

(1) Test question design rules

In order to ensure that the test session can generate more valuable learning data, we need to develop more effective test items. In the IRT model [27], an information function is used to characterize the validity of the test questions, and the larger value of the function represents the more accurate estimation of the ability level. Its mathematical expression is: $I_i(\theta)=\frac{(P_i)^2}{P_i(1-P_i)}$. For the three-parameter logistic model the information function can again be expressed as:

$$I_i(\theta)=\frac{1.7^2(a_i)^2(1-c_i)}{\left[c_i+e^{1.7a_i(\theta-b_i)}\right]\left[1+e^{-1.7a_i(\theta-b_i)}\right]^2} \tag{1}$$

where a_i, b_i, c_i, and θ have the same meaning as the parameters in the Logistic model. Item response theory holds under the assumption of locality, so the test items are designed to be non-interfering and independent of each other. In this case, the test information function is the sum of all information functions, i.e., $I(\theta)=\sum_{i=1}^{m}I_i(\theta)$, m is the total number of test items, which is an important parameter used to identify the overall performance of the test items.

Here we use the standard error of estimate of the parameter θ that characterizes the ability level of the subjects as a measure of test accuracy. The estimated standard error is denoted as $SE(\theta)$, and the mathematical expression is: $SE(\theta)=\frac{1}{\sqrt{\sum I_i(\theta)}}$, which in turn can be used to obtain the relationship between the manometric information function and the estimated standard error.

(2) Cognitive level testing process

Step1: During the construction of the test question bank, try to follow the design rules in the previous section to select representative test questions, then let students answer them, and calculate the difficulty coefficient and differentiation degree of the test questions according to the answer situation.

Step2: Before conducting the cognitive level test, the initial level is assigned to the students, who are first allowed to answer a certain number of test questions, and then the natural logarithm of the ratio of the scores of the subjects in the test to the lost scores is calculated.

Step3: During the testing process, test questions comparable to the current cognitive level of the subjects are presented first, the level value of the subjects is reassessed according to the answering situation, and the termination condition is determined, and the relevant information is outputted when the termination condition is satisfied.

Step4: When the termination conditions are not met, the test questions are pushed according to the answering situation. A slightly easier test question is pushed when the previous question is answered incorrectly, otherwise a more difficult test question is pushed.

(3) Cognitive level estimation method

By analyzing the calculation method of each ability value in item response theory, the cognitive level of the tester is estimated under the assumption that each parameter of the test question is known. The specific realization method is as follows:

Assuming that there are N students answering m test items, the cognitive level of the α student is denoted as $\theta_\alpha\left(1\le\alpha\le N\right)$, $a_i,b_i,c_i\left(1\le j\le m\right)$, the differentiation, difficulty and guessing coefficients of the jth test item, respectively, and let all the test questions be rated 0-1, denoted as:

$$U_{\alpha j}=\begin{cases}1 & \text{Student with proficiency value } \theta_\alpha \\ & \text{answered the } j \text{ test question correctly} \\ 0 & \text{Student with proficiency value } \theta_\alpha \\ & \text{answered the } j \text{ test question incorrectly}\end{cases} \tag{2}$$

The responses of students with ability θ_α on m items are represented below:

$$U_\alpha=(U_{\alpha 1},U_{\alpha 2},\cdots,U_{\alpha m})(1\le\alpha\le N) \tag{3}$$

Then the responses of N students to m test questions can be represented by the matrix U :

$$U=(U_{\alpha j})_{N\times M}=\begin{bmatrix}U_{11} & U_{12} & \cdots & U_{1m}\\ U_{21} & U_{22} & \cdots & U_{2m}\\ U_{31} & U_{32} & \cdots & U_{3m}\\ \cdots & \cdots & \cdots & \cdots\\ U_{N1} & U_{N2} & \cdots & U_{Nm}\end{bmatrix} \tag{4}$$

where $U_{\alpha j}$ takes the value of 0 or 1 of the random variable whose observation is $u_{\alpha j}$, and let $P_{\alpha j}$ be the probability that a subject with an ability of θ_α will answer the jth test question correctly, i. e:

$$P_{\alpha j}=P(u_{\alpha j}=1) \tag{5}$$

Assuming that the test questions satisfy the local independence assumption condition, the probability of various response scenarios of students with cognitive level θ_α respectively in the test with test question parameters $a_i,b_i,c_i\,(1\le j\le m)$ can be expressed as L, i.e.:

$$L=\prod_{\alpha=1}^{N}\prod_{j=1}^{m}P_{\alpha j}^{u_{\alpha j}}(1-P_{\alpha j})^{1-u_{\alpha j}} \tag{6}$$

The value of the independent variable for which the likelihood function L takes a great value is taken as the value of the cognitive level to be estimated. Since L and $\ln L$ have the same point of greatness, taking the logarithm of L yields:

$$\ln L=\sum_{\alpha=1}^{N}\sum_{j=1}^{m}(U_{\alpha j}\ln P_{\alpha j}+(1-U_{\alpha j})\ln(1-P_{\alpha j})) \tag{7}$$

Let the partial derivative of $\ln L$ with respect to each parameter be equal to zero, i.e:

$$\begin{aligned}&\frac{\partial\ln L}{\partial\theta_\alpha}=0\left(1\le\alpha\le N\right)\\ &\frac{\partial\ln L}{\partial a_j}=0\left(1\le j\le m\right)\\ &\frac{\partial\ln L}{\partial b_j}=0\left(1\le j\le m\right)\\ &\frac{\partial\ln L}{\partial c_j}=0\left(1\le j\le m\right)\end{aligned} \tag{8}$$

Bringing Eq. (7) into Eq. (8) can be further organized to obtain:

$$\begin{aligned}&\sum_{j=1}^{m}\frac{\left(U_{\alpha j}-P_{\alpha j}\right)}{P_{\alpha j}\left(1-P_{\alpha j}\right)}\frac{\partial P_{\alpha j}}{\partial\theta_\alpha}=0\\ &\sum_{\alpha=1}^{N}\frac{\left(U_{\alpha j}-P_{\alpha j}\right)}{P_{\alpha j}\left(1-P_{\alpha j}\right)}\frac{\partial P_{\alpha j}}{\partial a_j}=0\\ &\sum_{\alpha=1}^{N}\frac{\left(U_{\alpha j}-P_{\alpha j}\right)}{P_{\alpha j}\left(1-P_{\alpha j}\right)}\frac{\partial P_{\alpha j}}{\partial b_j}=0\\ &\sum_{\alpha=1}^{N}\frac{\left(U_{\alpha j}-P_{\alpha j}\right)}{P_{\alpha j}\left(1-P_{\alpha j}\right)}\frac{\partial P_{\alpha j}}{\partial c_j}=0\end{aligned} \tag{9}$$

The above equation is a nonlinear equation about cognitive level θ_α, and the Newton-Raphson iterative method is utilized to find an estimate of θ_α using the Newton-Raphson iterative method with the parameters of each item known, then $g(\theta_\alpha)$ is:

$$g(\theta_\alpha)=\frac{\partial \ln L}{\partial \theta_\alpha}=\frac{\partial \ln L}{\partial P_{\alpha j}}\frac{\partial P_{\alpha j}}{\partial \theta_\alpha}=\sum_{j=1}^{m}\frac{(U_{\alpha j}-P_{\alpha j})}{P_{\alpha j}(1-P_{\alpha j})}\frac{\partial P_{\alpha j}}{\partial \theta_\alpha}=0 \tag{10}$$

In the three-parameter logistic model [28] $g(\theta_\alpha)$ can again be expressed in the following equation:

$$g(\theta_\alpha)=\sum_{j=1}^{m}Da_j\frac{1-c_j}{P_{\alpha j}-c_j}\frac{(u_{\alpha j}-P_{\alpha j})}{P_{\alpha j}}=0 \tag{11}$$

In this paper the iterative formulation of the problem for cognitive level solving is:

$$\theta_{\alpha(k+1)}=\theta_{\alpha k}-g(\theta_{\alpha k})/g'(\theta_{\alpha k}) \tag{12}$$

An expression for $g'(\theta_k)$ is obtained by derivation of Eq. (11):

$$g'(\theta_{\alpha k})=\sum_{j=1}^{m}D^2a_j^2(1-c_j)(P_{\alpha j}-c_j)(1-P_{\alpha j}) \tag{13}$$

The initial value of the iteration is the first step in the solution of the equation, where the initial value of the cognitive level is $\theta_{\alpha 0}$, which is found by the natural logarithm of the ratio of the subject's score to the missing score on the test, i.e.:

$$\theta_{\alpha 0}=\ln\frac{\sum_{j=1}^{m}u_{\alpha j}}{m-\sum_{j=1}^{m}u_{\alpha j}} \tag{14}$$

The iterative termination rule is shown in equation (15): in practice ε is taken to be 0.01 or 0.001. i.e.:

$$\left|\theta_{\alpha(k+1)}-\theta_{\alpha k}\right|<\varepsilon \tag{15}$$

Above is the solution process and working principle of IRT cognitive level prediction model.

II. B.K-means based clustering recommendation algorithm

II. B. 1) K-means algorithm

The K-Means algorithm [29] initial cluster centers are chosen randomly, the data samples in the dataset are grouped together with the nearest initial cluster centers by similarity calculation, and the process is repeated until the initial cluster centers do not change within a certain accuracy range. The basic idea is to find a clustering result that minimizes the error criterion formula by iterating through a loop.

Let the data set to be clustered be: $X=\left\{x_i \mid x_i\in \mathrm{R}^P, i=1,2,\cdots,n\right\}$. The K clustering centers are $M_1,M_2,\cdots,M_k$. Denote the k categories of clustering by $W_j\,(j=1,2,\cdots,k)$. The specific steps of the algorithm are as follows:

Definition 1: Euclidean distance between two data objects:

$$d(x_i,x_j)=\sqrt{(x_i-x_j)^T(x_i-x_j)} \tag{16}$$

Definition 2: The arithmetic average of data objects in the same category:

$$M_j=\frac{1}{N_j}\sum_{x\in W_j}X \tag{17}$$

Definition 3: Clustering criterion function:

$$E=\sum_{i=1}^{k}\sum_{j=1}^{n_j}d(X_j,Z_i) \tag{18}$$

1) Randomly select k sample data in X containing N samples as the initial cluster center $M_i\,(i=1,2,\cdots,k)$;
2) Using Equation (16), calculate the distance $d(p,M_i)$ from each sample data p to M_i in X;
3) Find the smallest $d(p,M_i)$ of each sample data p and add p to the same cluster as M_i;
4) After completing the traversal of all the samples, recalculate the value of M_i as the new cluster center by Eq. (17);
5) Repeat steps (2)-(4) until the value of the objective criterion function E no longer changes.

II. B. 2) Collaborative Filtering Recommendation Algorithm Based on K-means Clustering

Each piece of data used in the algorithm needs to have 3 components: users, items, and ratings. Let the set of users be $U=\{u_1,u_2,\cdots,u_m\}$, and the set of users generated based on the K-means algorithm is denoted as $U^a=\left\{C_1^a,C_2^a,\cdots,C_k^a\right\}$. where k is the number of generated cluster classes and C_k^a is the kth cluster class. The collaborative filtering recommendation algorithm based on K-means clustering describes the steps as follows:

Input: user u , matrix R_{m*n} , number of k cluster classes

Output: N recommended items

Step1: Eliminate the 0 elements in R_{m*n} to obtain the matrix R'_{m*n} ;

Step2: Take $M_i\left(i=1,2,\cdots,k\right)$ as the initial cluster center, and classify the data in R'_{m*n} into k classes by K-means algorithm;

Step3: Calculate the similarity between u and k cluster centers, and then add u to the class that is most similar to it;

Step4: Calculate the similarity between u and other users in the same class to get its nearest neighbor set $N_{uj}^a\left(j=1,2,\cdots,m\right)$;

Step5: After getting the similar nearest neighbors, the prediction score of u for the items to be recommended can be obtained based on their ratings of the items, and the top N items are recommended to u after sorting them from high to low.

II. C. User clustering recommendation algorithm with optimized clustering center

II. C. 1) Optimizing initial cluster centers based on minimum variance

Among many clustering algorithms, K-means algorithm is very typical, although it is simple and convenient to implement, but it also has some disadvantages: firstly, the K value is randomly determined according to human experience, which has a certain degree of blindness, and if you don't know the data to be clustered, then it will be very difficult to give a reasonable K value; secondly, the selection of the initial cluster centers is random, and the different cluster centers will lead to different clustering effects, and if isolated points are selected, the clustering results will be very different.

For the defects of K-means clustering algorithm, many researchers have improved it. In this paper, the initial clustering center is optimized based on the minimum variance, and the K samples with the smallest variance in different ranges are selected as the initial cluster center. According to the definition of variance, the smaller the variance of a sample is, the denser the data distribution in its neighborhood will be, making the selection of cluster centers more objective and the clustering results more accurate. The specific steps for cluster center selection are as follows:

Definition 4: The average of the distance from sample x_i to all samples:

$$m_i=\frac{1}{n}\sum_{j=1}^{n}d(x_i,x_j) \tag{19}$$

Definition 5: The variance of a sample x_i :

$$\mathrm{var}_i=\frac{1}{n-1}\sum(d(x_i,x_j)-m_i)^2 \tag{20}$$

Definition 6: The average distance of the samples of a data set:

$$cmean=\frac{2}{n(n+1)}\sum_{i=1}^{n}\sum_{j=1}^{i}d(x_i,x_j) \tag{21}$$

1) Calculate the variance of each sample in the data set X , and then find the sample x_i' with the smallest variance in X and add it to the set C' as the first cluster center M_1' ;

2) Calculate the mean cmean of the distances between individual samples in X ;

3) Find another sample with the smallest variance outside the circle with $cmean$ as the radius and add it to the set C' as the second cluster center;

4) Repeat the previous step and keep searching among the remaining samples, the algorithm ends after finding K cluster centers.

II. C. 2) Minimum variance based user clustering recommendation algorithm

Similarly, each piece of data used by the algorithm needs to have 3 components: users, items, and ratings. Let the set of users be $U=\{u_1,u_2,\cdots,u_m\}$, and the set of users generated by the K-means algorithm based on the optimization of the minimum variance is denoted as $U^b = \left\{C_1^b, C_2^b, \cdots, C_k^b\right\}$. where k is the number of generated cluster classes and C_k^b is the kth cluster class. The process of the minimum variance based user clustering recommendation algorithm is as follows:

Input: user u, matrix R'_{m*n}, number of cluster classes k.

Output: N recommended items.

Step1: Take the element $M'_i(i = 1, 2, \cdots, k)$ in the set C' as the initial cluster center, and classify the data in R'_{m*n} into k classes by K-means algorithm;

Step2: Calculate the similarity between u and k cluster centers, and then add u to the class that is most similar to it;

Step3: Calculate the similarity between u and other users in the same class to get its nearest neighbor set $N_{uj}^b\ (j = 1, 2, \cdots, m)$;

Step4: After obtaining similar nearest neighbors, the predicted score of u for the items to be recommended can be obtained based on their ratings of the items, and the top N items are recommended to u after sorting them from high to low.

II. D.Personalized Recommendation Algorithm for English Vocabulary Learning Resources

The overall recommendation strategy of the English vocabulary learning recommendation system constructed in this study is: when a student enters the system, the system first detects whether the existing database contains the vocabulary learning behavior data of the current student. If there is no vocabulary learning behavior data of the current student in the system or the existing vocabulary learning behavior data in the system is too little, the system cannot determine the vocabulary learning needs of the student, and the system will recommend vocabulary resources to the student based on the student's initial model. If the vocabulary learning behavior data of the current student exists in the system, which can support the system to analyze the vocabulary learning needs of the student, then the system will initially analyze the vocabulary level and learning interest of the student based on these data, and then recommend vocabulary resources to the student based on the student's vocabulary level and dynamic interest model. The system firstly finds the set of users with similar vocabulary ability level as the student through bisection K-means clustering algorithm, which reduces the complexity of similarity calculation, and then based on waterfall mixing technology, in the set of users with similar vocabulary ability level as the student, the system further finds the set of users with similar interest in learning through the collaborative filtering recommendation strategy based on the dynamic interest of the student, and then counts the frequency of occurrence of all vocabulary resources among the users of similar level. The frequency of all vocabulary resources in similar users is counted, and the vocabulary learning resources are sorted according to the frequency, and finally the vocabulary resources are recommended to students according to the sorting. The overall strategy of vocabulary resource recommendation is shown in Figure 1.

II. D. 1) Recommendations based on initial student modeling

In this study, vocabulary recommendation based on students' initial model, i.e., recommending vocabulary resources to students based on students' registration information. The student's registration information includes the student's major and the student's initial learning interest information, and the construction of the student's initial model is completed based on this basic information. The recommendation process based on the initial model of students is as follows: the vocabulary categories that students want to learn are initially analyzed by obtaining their basic information, including their majors, and their initial learning interests, and then the vocabulary resources are recommended to the students through the collaborative filtering recommendation algorithm based on the users.

II. D. 2) Dichotomous K-means clustering based on students' vocabulary level

Recommendation based on students' vocabulary level is to recommend vocabulary resources that meet students' current vocabulary level according to their mastery of vocabulary, and to help students actively adjust their learning behaviors according to their own vocabulary mastery. In this paper, we use the bipartite K-means clustering algorithm [30] to find the set of users with similar vocabulary levels to recommend vocabulary resources for them.

(1) Constructing the student-vocabulary level scoring matrix

Based on the recommendation of students' vocabulary level, the student-vocabulary level scoring matrix N, which represents the students' vocabulary level model, is obtained based on the analysis and construction of the student-vocabulary level model. The student-vocabulary level scoring matrix is shown in Table 1, where S denotes students, k denotes vocabulary resources, and $N_{m,n}$ denotes the learning mastery of the nth vocabulary by the mth student.

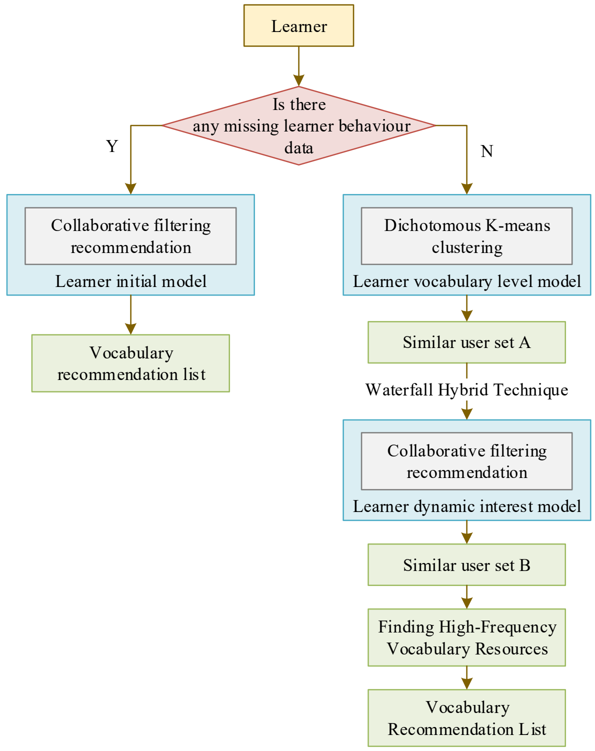


Figure 1: The overall strategy recommended by vocabulary resources

Table 1: Learner - vocabulary level matrix

S/K	k_1	k_2	k_3	…	k_n
S_1	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	…	$N_{1,n}$
S_2	$N_{2,1}$	$N_{2,2}$	$N_{2,3}$	…	$N_{2,n}$
S_3	$N_{3,1}$	$N_{3,2}$	$N_{3,3}$	…	$N_{3,n}$
…	…	…	…	…	…
S_m	$N_{m,1}$	$N_{m,2}$	$N_{m,3}$	…	$N_{m,n}$

(2) Dichotomous K-means clustering

Bisection K-means clustering is computed on the student-vocabulary level scoring matrix N, by which all students with similar vocabulary levels are gathered into the same cluster. The detailed algorithmic steps to perform bisection K-means clustering are:

Input data: student-vocabulary level scoring matrix N.

Output data: set of clustered clusters $C=\{C_1,C_2,\cdots,C_n\}$ and n cluster centers.

Step1: Add all data as a cluster to the set of clustered clusters.

Step2: Repeat.

Step3: Select the cluster with larger sum of squares of errors from the set of clustered clusters to be added to the set of clustered clusters.

Step4: for $i=1$ to the initially set number of cycles.

Step5: Divide the selected clusters into two subsets of clustering clusters by k-means algorithm.

Step6: Sum the error squares of the two cluster cluster subsets.

Step7: end for.

Step8: Add the two subsets with the smallest sum of error squares in the for loop to the set of clustered clusters.

Step9: Until the set of clustering clusters has n cluster centers.

II. D. 3) Collaborative Filtering Recommendations Based on Students' Dynamic Interest Models

According to the waterfall mixing technology obtained with the students' vocabulary level similar to the user set A, through the recommendation strategy of collaborative filtering based on the students' dynamic interests, to find similar to the students' learning interests of the user set B. Finally, statistics on the frequency of all vocabulary resources in the similar user set B, vocabulary learning resources in accordance with the frequency of the high and low sorting, the high-frequency vocabulary resources will be recommended to the students, based on the student's initial model of the The recommendation results are shown in Fig. 2.

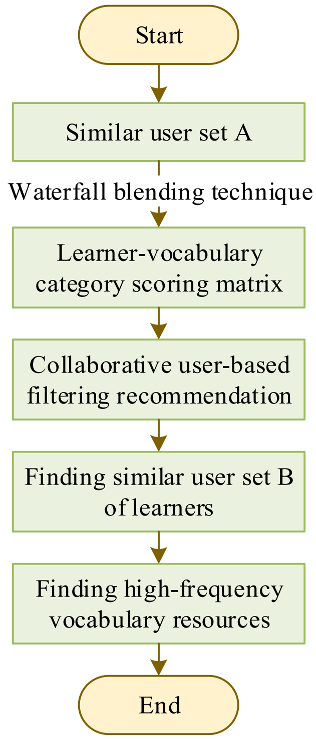


Figure 2: Recommendations based on the learner's initial model

(1) Establishing the student-vocabulary category scoring matrix

Student dynamic interest model, calculate the weight value of students' vocabulary categories under the same time window, and obtain the student-vocabulary category scoring matrix M, S denotes students, C denotes vocabulary categories, there are thirteen categories in total, and $M_{m,13}$ denotes the m th student's interest in the thirteenth category of vocabulary.

(2) Calculate similarity and find similar user sets.

In this paper, we calculate the similarity $sim(a,b)$ between students and students in the student-vocabulary category scoring matrix M according to the cosine similarity formula, and select the top N set of nearest-neighbor users $S(a,N)$ based on the size of the similarity.

$$sim(a,b)=\frac{\sum_i r_{a,i}r_{b,i}}{\sqrt{\sum_i r_{a,i}^2\sum_i r_{b,i}^2}} \tag{22}$$

(3) Finding high-frequency resources

Suppose the set $a=\{U_1,U_2,U_3,U_4,\cdots\}$ is the nearest-neighbor set of the students, in which the vocabulary learning records of each student are shown in Table 2. From the table, it can be seen how often each vocabulary learning resource appears in the vocabulary learning records of all similar students, in which the vocabulary resource with the highest frequency is k_2, followed by k_1 and k_3. The frequency of occurrence of all vocabulary resources among similar students is counted, the vocabulary learning resources are sorted according to the frequency, and finally the vocabulary resources are recommended to students according to the sorting.

Table 2: Similar collection of learners' learning records

Similar learner	History of vocabulary resources
U_1	k_1, k_2, k_3, k_5, k_8
U_2	k_1, k_2, k_3, k_5
U_3	k_2, k_3, k_4, k_6
U_4	k_1, k_2, k_7, k_9
…	…

III. Analysis of the effect of personalized recommendation of English vocabulary learning resources

III. A. Determination of model parameters based on learning ability

The data source currently used in this paper contains a total of six English courses, with C, D, E, and F being economics courses and A and B being social science courses, each containing data from four semesters. There are a total of 31505 students and 153428 quiz data, and after preprocessing the data there are actually 27531 students and 143274 quiz data.

The statistical results of the student grade data are shown in Table 3. It can be seen that the three courses B, D, and F have all four semesters of data, and the amount of data in each semester is relatively large, so this section mainly selects the course data of these three courses for the experiment.

Table 3: Student performance data statistics

	2023B	2023J	2024B	2024J
A	--	290	--	276
B	627	883	550	1244
C	--	--	503	811
D	443	769	389	888
E	--	613	357	682
F	821	1119	696	1107

The 0-1 scoring questions were first analyzed, and the distribution of scores for the 0-1 scoring questions is shown in Table 4. The first column represents the question number, the second column represents the percentage of students scoring 0 points for each question, the third column represents the percentage of students scoring 1 point for each question, and the fourth column represents the percentage of missing data in the data set.

Table 4: 0-1 scoring distribution

Topic number	Zero fraction	1 account ratio	Missing value ratio
Q1A	0.3100	0.6934	0
Q1B	0.2470	0.7413	0
Q1C	0.3829	0.6252	0
Q1D	0.4104	0.5887	0
Q1E	0.1565	0.8360	0
Q2A	0.2528	0.7480	0
Q2B	0.2685	0.7357	0
Q2C	0.3383	0.6495	0
Q2D	0.4701	0.5239	0
Q2E	0.4867	0.5207	0

A sample of student responses to the 0-1 scoring questions is shown in Table 5. Each row represents a student, and each column represents a question except for the first column, which represents the student's number. The values in the table represent the student's score on the corresponding question, with 0 representing a student's score of zero and 1 representing a student's score of one.

Table 5: 0-1 sample of the student answer problem

Student number	1	2	3
Q1A	1	0	1
Q1B	0	0	1
Q1C	1	0	0
Q1D	0	0	1
Q1E	1	1	0
Q2A	0	0	1
Q2B	1	0	0
Q2C	0	0	1
Q2D	1	1	0
Q2E	0	1	0

The results of the parameter estimation for the 0-1 scoring question topics are shown in Table 6. Each row represents a question, the first column indicates the number of the question, the second column represents the differentiation of the question, the third column represents the difficulty of the question, and the fourth column represents the guessing coefficient of the question.

Table 6: The parameter estimation result of the 0-1 scoring problem

Student number	Topic number	Differentiating	Difficulty
Q1A	0.6538	-1.3506	-0.0080
Q1B	0.2493	-4.1968	0.0101
Q1C	2.0973	-0.4058	0.0080
Q1D	0.8047	-0.5300	-0.0127
Q1E	0.4282	-3.9678	0.0044
Q2A	0.4340	-2.6420	0.0002
Q2B	0.5198	-2.0499	-0.0054
Q2C	0.6612	-1.0643	0.0035
Q2D	0.8072	-0.1431	0.0076
Q2E	0.9216	-0.0950	0.0002

The 0-1 scoring question characteristic curve is shown in Figure 3. The horizontal axis represents the student learning ability θ and the vertical axis represents the probability value of answering the question correctly for the corresponding ability value. Item response theory regards questions as the basic unit for measuring students' learning ability, and the attributes of questions are expressed through question parameters. Topics generally contain three parameters, all of which can be responded to in the characteristic curve. The difficulty parameter b is equal to the value of θ at the inflection point of the curve, the magnitude of the differentiation parameter a is equal to the value of the slope of the curve at the inflection point, and the guessing parameter c is equal to the value of the asymptote converging to the lower left of the curve. Obviously, the greater the differentiation of a question, the better, the smaller the guessing parameter, the better, so, from the figure, we can see that the differentiation of the question Q1C is larger, the difficulty is in the middle of a slightly simpler level, and it is difficult for the students to guess the right question, which indicates that Q1C can play a better role in the testing of the majority of the students, and the opposite of the question Q1E, the guessing parameter is larger and the differentiation is very small, which indicates that the design of the question Q1E is not too reasonable.

The learning ability of the students of the 0-1 scoring quiz is shown in Table 7. It is the value of the learning ability of the students assessed by the algorithm, the learning ability of the students usually we take the value of the range of [-2,0], the learning ability of the value corresponds to the greater the value indicates that the student's learning ability is stronger, and vice versa indicates that the student's learning ability is weaker. Many of the graphs in this paper have learning ability values beyond this range in order to make the topic curves more complete and make the intuitive analysis more convincing. It can be visualized that the assessed learning ability values are reasonable. Because the data is not labeled, so this paper uses the professional item response theory data generation software WinGen to generate data, using the parameter estimation method in this section to evaluate the parameters, the final evaluation of the learning ability value, the title of each parameter is consistent with the parameters of the Win Gen generated data.

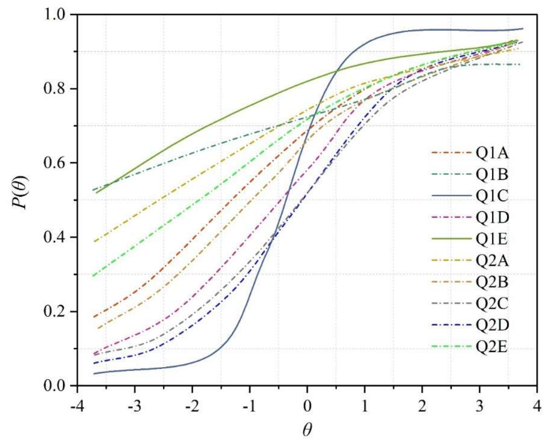


Figure 3: 0-1 scoring feature curve

Table 7: 0-1 scoring test learners' learning ability

Student number	1	2	3
Learning ability	-1.2087	-0.5745	-0.8822

A sample of student responses to the multilevel scoring questions is shown in Table 8. Each row represents a student, and each column represents a question except for the first column, which represents the student number, and each question has a full score of 4. The numerical values in the table represent the students' scores on the corresponding questions.

Table 8: Sample of students answering questions on multistage scoring

Student number	Item1	Itme2	Item3	Item4
1	4	4	2	2
2	3	3	2	4
3	4	2	3	3

The results of the data distribution for the multi-level scoring questions are shown in Table 9. The first column represents the question number, the second column represents the percentage of students scoring 1 point for each question, the third column represents the percentage of students scoring 2 points for each question, the fourth column represents the percentage of students scoring 3 points for each question, the fifth column represents the percentage of students scoring 4 points for each question, and the fifth column represents the percentage of missing data in the data set. From the table we see that this dataset is special because there is no score of 0 for each question and the lowest score the students scored on each question was 1.

Table 9: Results of multistage scoring data distribution

Topic number	1 account ratio	2 fraction ratio	3 fraction ratio	4 fraction ratio	Missing value ratio
Item1	0.02056	0.08343	0.68477	0.22714	0
Itme2	0.08078	0.2514	0.53231	0.13461	0
Item3	0.03691	0.17927	0.53921	0.2441	0
Item4	0.06757	0.2572	0.50525	0.18548	0

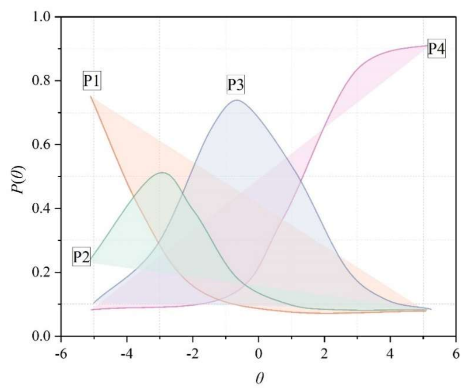
The parameters characterizing the multilevel scoring questions are shown in Table 10, where each row represents a question, the first column indicates the number of the question, and the second column represents the differentiation level of the question.

All students can score a minimum of 1 point for each question, which means that there is no difficulty in scoring 1 point for each question, so the question difficulty parameter is calculated from scoring 2 points. The third column represents the difficulty of the corresponding question with 2 points, the fourth column represents the difficulty of the corresponding question with 3 points, and the fifth column represents the difficulty of the corresponding question with 4 points.

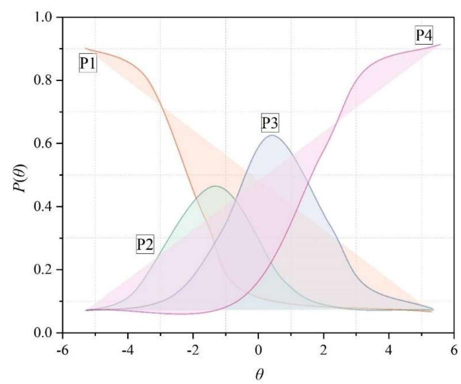
Table 10: Multistage scoring feature parameters

Topic number	Differentiating	2 points is difficult	3 points is difficult	4 points is difficult
Item1	1.0413	-4.6703	-2.5412	1.4161
Item2	1.2326	-2.3853	-0.7217	1.8573
Item3	2.2898	-2.2802	-0.9664	1.8726
Item4	1.0963	-3.0567	-0.8995	1.5396

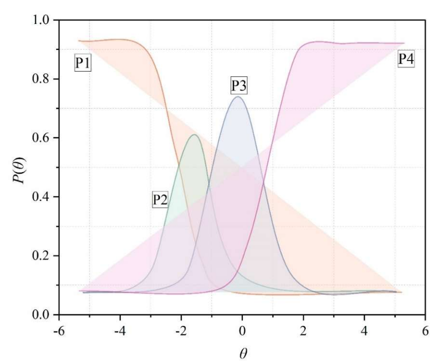
The arithmetic characteristic intervals of multilevel scoring questions are shown in Fig. 4, (a)~(d) represent item1~ item4 respectively. Taking the first question item1 as an example, P1~P4 represent the probability of scoring 1~4 points for students with different learning ability in turn, it can be seen from Fig. (a) that the probability of scoring 1 point (corresponding to the interval of P1) gradually decreases as the value of the learning ability increases, and the probability of scoring 2~4 points (corresponding to the interval of P2~P4) gradually increases when the value of learning ability is about -3.5, and the probability gradually decreases thereafter. The probability of scoring 4 points (corresponding to the interval P2~P4) gradually increases, when the value of learning ability is about -3.5 the probability of students scoring 2 points is the largest, and after that the probability gradually decreases, when the value of learning ability is about 0, the probability of students obtaining 3 points in the first question reaches the maximum, and after that the probability gradually decreases, and the probability of scoring 4 points gradually increases and gradually converges to the probability 1.



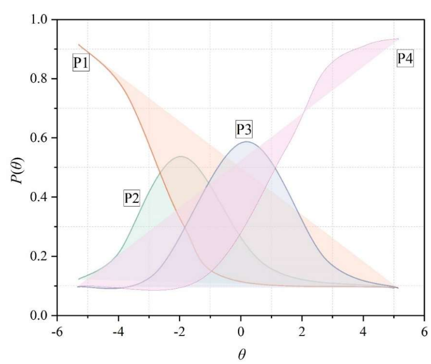
(a) item1



(b) item2



(c) item3



(d) item4

Figure 4: The operation feature interval of multistage scoring problem type

The final aptitude values of students on the multilevel scoring test are shown in Table 11. The range of students' learning ability is usually [-3,+3], and the larger the value corresponding to the learning ability value indicates the stronger the students' learning ability, and vice versa indicates the weaker the students' learning ability. Since the data are unlabeled, this paper utilizes the professional item response theory data generation software WinGen to generate the data, and evaluates the parameters using the parameter estimation method in this section, and the final evaluated learning ability value, each parameter of the topic is consistent and reasonable with the parameters of the data generated by WinGen and the evaluated learning ability value.

Table 11: Multistage scoring test students' ultimate ability

Student number	1	2	3
Learning ability	0.4036	0.0527	-0.8903

III. B. Model Performance Comparison

The dataset MOOCCube5 chosen for the experiments in this paper is the online learning data collected by the MOOC platform's academy online.5 The dataset contains the records of 200,000 students' selections and video viewings in 710 real English online courses, which involve 38,542 teaching videos and 121,543 knowledge concepts. The experiment divides 80% of the data in the dataset into training and validation sets, and 20% into testing sets.

In order to verify the effectiveness of the proposed model in this paper, the following 10 typical benchmark methods are selected to compare the experimental results:

MLP: A collaborative filtering method utilizing multilayer perceptron to learn user-item interactions.

FISM: a content-based approach to generate top-N recommendations, which learns the item similarity matrix as the product of two low-dimensional latent factor matrices, alleviating the problem that the performance of the model decreases due to the increase of data sparsity.

NAIS:A neural attentional item similarity model based on collaborative content filtering, which distinguishes which historical items in the user interaction record are more important for prediction through an attention network.

NARM:A neural attention recommender machine with a coding and decoding structure that captures the successive behaviors and the main purpose of the user in the current session by incorporating an attention mechanism into the RNN.

metapath2vec: a metapath-guided random walk strategy in heterogeneous networks that captures structural and semantic associations of different types of nodes and relations.

ACKRec: an attentional convolutional network knowledge recommender based on graph neural networks that builds a heterogeneous information network to capture valid semantic relationships between different types of entities and incorporates them into the representation learning process.

Referring to most of the research works on MOOC recommendation models, this experiment adopts Hit Rate (HR) and Normalized Discounted Cumulative Gain (NDCG) as the evaluation metrics.HR is a metric for evaluating the recall accuracy, which is used to measure the percentage of successful recommendations to the students, and is computed by the following formula:

$$HR@K = \frac{\sum_{u=1}^{U} Hits_u @K}{|GT|} \tag{23}$$

where GT denotes the sum of the length of the test set of all students, and $Hits_u @K$ denotes the number of items in the top-K recommendation list of the u th student in the test set. NDCG is a measure of accuracy, and the larger the value, the higher the ranking of the results for which the recommendation is accurate, and is calculated by the following formula:

$$DCG_u @ K = \sum_{i=1}^{K} \frac{2^{rel_u^i} - 1}{\log_2(i+1)} \tag{24}$$

$$NDCG @ K = \frac{1}{U} \sum_{u=1}^{U} \frac{DCG_u @ K}{IDCG_u @ K} \tag{25}$$

where rel_u^i is the match between the recommended result ranked at the i th position and the u th user, $rel_u^i = 1$ if it hits the courses in the test set, and 0 otherwise, and $IDCG_u @K$ is the ideal value of $DCG_u @K$, i.e., the maximum value that may be achieved.

III. B. 1) Experimental results and model performance comparison

The experimental results and model performance comparison results are shown in Table 12. From the experimental results, it can be seen that the proposed model in this paper achieves the best performance in HR@5, HR@10, NDCG@5, NDCG@10 and NDCG@20 metrics, which are improved by 3.34%, 10.07%, 17.58%, 12.94% and 23.66% respectively compared with the suboptimal model, and the experimental results prove the effectiveness of the model in this paper.

Table 12: Model performance comparison analysis results

Models	HR@1	HR@5	HR@10	HR@20	NDCG@5	NDCG@10	NDCG@20
MLP	0.0593	0.373	0.5898	0.7154	0.2124	0.2813	0.3387
FISM	0.2357	0.4274	0.5614	0.7432	0.3766	0.4068	0.3736
NAIS	0.155	0.5844	0.7139	0.7026	0.3751	0.4124	0.4053
NARM	0.0605	0.4375	0.6581	0.9139	0.2536	0.3089	0.3689
metapath2vec	0.1572	0.4442	0.6411	0.7332	0.2226	0.3198	0.3862
ACKRec	0.2278	0.6308	0.7562	0.8441	0.4191	0.4718	0.4391
Ours	0.2855	0.6526	0.8409	0.9372	0.5085	0.5419	0.5752

III. B. 2) Visualization of knowledge preference states

The model in this paper can dynamically track students' changing learning preferences and needs with learning. In order to prove the reasonableness and interpretability of the model in tracking students' knowledge preference status, we visualize and track the same student's preference status in the six knowledge points: synonyms-antonyms, root words-affixes, fixed collocations, multiple meanings of words, compound structures, and lexical transformations. The student's preference status is visualized and tracked, and the visualization of the student's knowledge preference status with learning changes is shown in Figure 5, where darker colors indicate a higher preference status for the corresponding knowledge points. It can be seen that after the student completes the 5th study, the student has higher relevance to the knowledge points of “root-affixes, fixed collocations and compound structures”, thus capturing the student's possible tendency to learn the corresponding knowledge points, and by continuously tracking the student's changing learning preferences and needs, it can enable the model to recommend relevant MOOC resources more accurately, and make the model more accurate. By tracking students' changing learning preferences and needs, the model can recommend relevant MOOC resources more accurately and provide explanations for students' knowledge preferences.

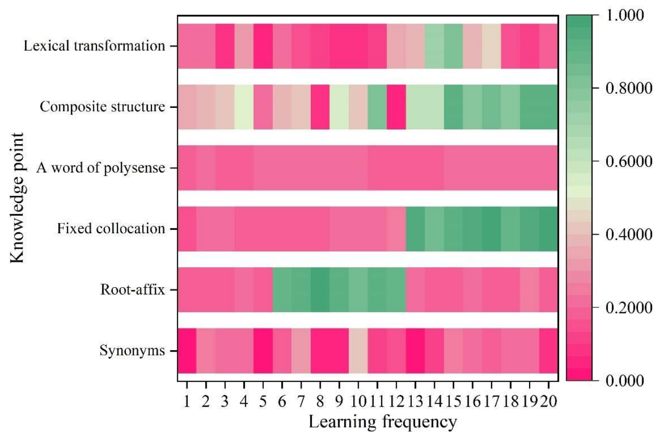


Figure 5: Student knowledge preference status visualization results

III. B. 3) Recommendation effect analysis of the model in this paper

The model obtains the correlation of MOOC content at the semantic level by calculating the feature similarity between MOOC subtitle texts, and this correlation can be used as the relying information for recommending MOOCs to students. Figure 6 shows the results of MOOC video recommendation and its correlation analysis, the vertical axis is the MOOC videos that a student has studied, and the horizontal axis is the six MOOC videos recommended to the student for “pronunciation rules, oral expression, contextual usage, etymological evolution, abbreviated forms, and synonym conversion”, where the larger value indicates a high content-level correlation between the courses. The larger the value, the higher the correlation between the courses. For example, pronunciation rules and “compound structure, root words, multiple meanings” have a high correlation in content, and this correlation may be reflected in the knowledge points involved or common features at the semantic level, and these automatically learned results can be used as a supplement to the data in the field of education.

In addition, the recommendation results show that the model has a good ability to expand students' interests, from the viewpoint of the student's historical learning records, the student may be interested in compound structures and synonyms-antonyms and other learning content, and in the recommendation results are given in the non-vocabulary subjects such as oral expression and other related extended knowledge, which can alleviate to a certain extent the cocoon effect caused by the increasing narrowness of students' horizons and other problems.

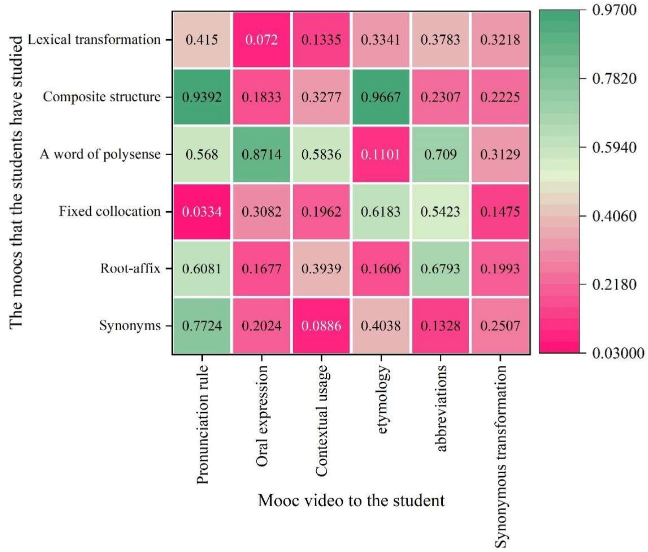


Figure 6: The selection results of mooc video and its correlation analysis

III. C. Practical case studies

In order to assess the learning effectiveness of the personalized mobile English vocabulary learning system, 20 students were recruited as volunteers to participate in the experiment. In order to ensure the smooth running of the experiment, it was first assumed that the participants had been educated in basic teaching skills in English listening, speaking, reading, and writing for at least five years on a continuous basis.

III. C. 1) Results of comparison of students' pre-test-post-test scores

Figure 7 shows the results of comparing the learning performance of the pre-test and post-test. The pre-test (Pre) and post-test (Pro) consisted of 50 multiple-choice (MC) and 50 perfect-fill-in-the-blank (FF) question papers out of 100 points. The difficulty of the 2 tests was kept within a similar range to ensure accuracy. There was a significant improvement in students' learning performance in both multiple-choice and fill-in-the-blank questions, which shows that the method in this paper has obvious advantages in enhancing students' English vocabulary learning.

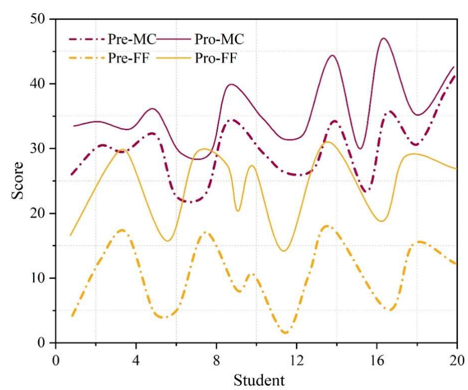


Figure 7: Comparison of the learning ability of the previous and post-test

III. C. 2) Differences between students' vocabulary scores on the pre-test and post-test

In order to compare the differences in students' vocabulary skills before and after using the system, the results of the pre-test and post-test were analyzed. According to the statistical results, the mean scores of the pre-test and post-test of the 20 students were 23.17, respectively. In addition, Table 13 shows the results of the comparison of pretest and posttest scores for the paired samples t-test. The difference between the mean scores of the pre-test and post-test is 7.4628, and the result reaches the significant level under the degree of freedom 19 (t=-7.6528, p=0.0014). In other words, the use of the learning system had a significant contribution to the students' academic performance, with an increase of 7.4628 points in the average test score.

Table 13: Comparison of test results of sample t test and post-test score results

Type	The mean difference is before and after	Standard deviation	Difference (95% confidence interval)		t	df	Significance
			Lower bound	Upper bound			
T test	7.4628	2.3466	-3.0157	-1.2013	-7.6528	19	0.0014

IV. Conclusion

In this paper, a dichotomous K-means clustering algorithm based on students' vocabulary level is proposed in the context of the big data era, and a personalized vocabulary is recommended for students through the English vocabulary learning recommendation platform.

In this paper, a learning ability assessment model is constructed based on the item response theory, and the parameters of the topics are obtained while the learning ability values of the existing students are assessed, and then these parameters are used to assess the ability of the new students. The experimental results show that this paper's model has an overall improvement of 3.34%-58.23% in HR@5, HR@10, NDCG@5, NDCG@10 and NDCG@20 evaluation indexes compared with the traditional comparative methods, which is very significant. The results of the actual case study show that the English vocabulary learning system proposed in this paper helps students memorize English vocabulary effectively, and the use of this learning system has a significant contribution to the students' academic performance ($P < 0.005$), and the average score of the pre and post-tests has been increased by 7.4628 points. It can be seen that the method of this paper can make the learned vocabulary transformed into long-term memory through an effective review process and no longer forgotten easily.

Funding

This work was supported by 2024 Key Projects of Education and Teaching Reform of Hainan Vocational University of Science and Technology (HKJGZD2024-04): "Research on the Optimization of College English Classroom Teaching Methods Based on Learning Data".

References

[1] Dakhi, S., & Fitria, T. N. (2019). The Principles and the Teaching of English Vocabulary: A Review. Online Submission, 5(1), 15-25.
[2] Syafrizal, S., & Haerudin, H. (2018). The implementation of vocabulary building strategy in teaching English vocabulary to young learners. Journal of English language teaching, 5(1), 40-48.
[3] Yunus, A., Seftina, A., Fatimah, D., Syaharani, N., & Simanjuntak, K. (2025). Improving Students' Vocabulary Ability: Teacher Planning Strategies in English Language Teaching. Fonologi: Jurnal Ilmuan Bahasa dan Sastra Inggris, 3(1), 222-229.
[4] Cui, H., & Kaur, N. (2025). Challenges in Teaching Medical English Vocabulary to Tertiary Students in China: A Systematic Literature Review. Theory & Practice in Language Studies (TPLS), 15(2).
[5] Wulandari, E. M. (2025). The effectiveness of using flashcards on teaching English vocabulary. JEET, Journal of English Education and Technology, 5(04), 324-338.
[6] Hussein, A. A., & Abd Khalid, R. (2019). A comparative study of swarm intelligence-based optimization algorithms in WSN. Asian Journal of Engineering and Applied Technology, 8(3), 1-7.
[7] Barabadi, E., & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English vocabulary. Cogent Education, 4(1), 1283876.
[8] Ma, H., Ye, S., Simon, D., & Fei, M. (2017). Conceptual and numerical comparisons of swarm intelligence optimization algorithms. Soft Computing, 21, 3081-3100.
[9] Zhiheng, W., & Jianhua, L. (2021). Flamingo search algorithm: a new swarm intelligence optimization algorithm. IEEE Access, 9, 88564-88582.
[10] Chakraborty, A., & Kar, A. K. (2017). Swarm intelligence: A review of algorithms. Nature-inspired computing and optimization: Theory and applications, 475-494.
[11] Hu, G., Huang, F., Chen, K., & Wei, G. (2024). MNEARO: A meta swarm intelligence optimization algorithm for engineering applications. Computer Methods in Applied Mechanics and Engineering, 419, 116664.
[12] Yasear, S. A., & Ku-Mahamud, K. R. (2021). Review of the multi-objective swarm intelligence optimization algorithms. Journal of Information and Communication Technology, 20(2), 171-211.
[13] Wang, X., Hu, H., Liang, Y., & Zhou, L. (2022). On the mathematical models and applications of swarm intelligent optimization algorithms. Archives of Computational Methods in Engineering, 29(6), 3815-3842.
[14] Cai, Y., & Sharma, A. (2021). Swarm intelligence optimization: an exploration and application of machine learning technology. Journal of Intelligent Systems, 30(1), 460-469.
[15] Brezočnik, L., Fister Jr, I., & Podgorelec, V. (2018). Swarm intelligence algorithms for feature selection: a review. Applied Sciences, 8(9), 1521.

[16] Tang, J., Liu, G., & Pan, Q. (2021). A review on representative swarm intelligence algorithms for solving optimization problems: Applications and trends. IEEE/CAA Journal of Automatica Sinica, 8(10), 1627-1643.
[17] Liu, R., Mo, Y., Lu, Y., Lyu, Y., Zhang, Y., & Guo, H. (2022). Swarm-intelligence optimization method for dynamic optimization problem. Mathematics, 10(11), 1803.
[18] Gad, A. G. (2022). Particle swarm optimization algorithm and its applications: a systematic review. Archives of computational methods in engineering, 29(5), 2531-2561.
[19] Zheng, F. (2025). Improving english vocabulary learning with a hybrid deep learning model optimized by enhanced search algorithm. Egyptian Informatics Journal, 29, 100619.
[20] Qi, D. (2025). Personalized Recommendation Algorithm for Optimizing English Vocabulary Learning Using Neural Networks. International Journal of High Speed Electronics and Systems, 2540227.
[21] Nykyporets, S. S., Pradivlyanny, M. H., Boiko, Y. V., Chopliak, V. V., & Kukharchuk, H. V. (2024). Innovative techniques in vocabulary acquisition for foreign language learning: the impact of artificial intelligence. Суспільство та національні інтереси№ 5 (5): 113-127.
[22] Chang, L. (2024, June). Improvement of Vocational English Vocabulary Learning Based on Adaptive Algorithms. In Proceedings of the 2024 International Conference on Intelligent Education and Computer Technology (p. 1).
[23] Cui, Y. (2024). Application of Artificial Intelligence Technology in College English Vocabulary Teaching. Development, 6(7), 196-200.
[24] Wu, F. (2024). English Vocabulary Learning Aid System Using Digital Twin Wasserstein Generative Adversarial Network Optimized With Jelly Fish Optimization Algorithm. Applied Artificial Intelligence, 38(1), 2327908.
[25] Zhang, T., & Li, C. (2025, March). Adaptive English vocabulary recommendation systems: a computational intelligence approach using deep reinforcement learning. In Second International Conference on Big Data, Computational Intelligence, and Applications (BDCIA 2024) (Vol. 13550, pp. 1231-1240). SPIE.
[26] Wang, Y., Liu, M., & Zhou, Z. (2024, March). Enhancing ESP Vocabulary Learning through ChatGPT: A Case Study. In Society for Information Technology & Teacher Education International Conference (pp. 907-913). Association for the Advancement of Computing in Education (AACE).
[27] Leonie V D E Vogelsmeier,Irina Uglanova,Manuel T Rein & Esther Ulitzsch. (2024). Investigating dynamics in attentive and inattentive responding together with their contextual correlates using a novel mixture IRT model for intensive longitudinal data.. The British journal of mathematical and statistical psychology,
[28] Zheng Chanjin,Guo Shaoyang & Kern Justin L. (2021). Fast Bayesian Estimation for the Four-Parameter Logistic Model (4PLM). SAGE Open,11(4),
[29] Xingzhen Li,Yiwei Ma,Hao Zhong & Miao Huang. (2025). A Novel Clustering Method for PV Power Curve Patterns based on Multidimensional Feature, Entropy Weight, and K-means. Engineering Letters,33(4),
[30] Jian Di & Xinyue Gou. (2018). Bisecting K-means Algorithm Based on K-valued Selfdetermining and Clustering Center Optimization.. JCP,13(6),588-595.