

A Resource Recommendation Algorithm for Online English Learning Systems Based on Learning Ability Evaluation

<https://doi.org/10.3991/ijet.v16i09.22745>

Zhenfang Zhou

Hunan University of Medicine, Huaihua, China

330463327@qq.com

Abstract—Mature online English learning platforms should provide students with necessary learning resources, ensure efficient access to learning projects, and offer the optimal learning experience. However, the traditional recommendation methods for English learning resources cannot satisfy the in-depth learning demand of students. To solve the problem, this paper designs a resource recommendation algorithm for online English learning systems based on learning ability evaluation. Firstly, the workflow of the designed algorithm was introduced, and a four-layer test system was developed for students' English learning ability evaluation. Next, an English learning ability evaluation method was proposed based on the maximum expectation algorithm, as well as the estimation methods for parameters like learning ability, degree of discrimination, difficulty, and guess coefficient. Experimental results demonstrated the good effect of the proposed resource recommendation algorithm. The research findings provide a reference for resource recommendation of other online learning systems.

Keywords—Learning ability evaluation, resource recommendation, online English learning

1 Introduction

As the information technology is developing at an astonishing pace, students' learning habits and behavior patterns have undergone major changes along with the wide application of handheld smart mobile terminals such as mobile phones and PDAs. Due to its unique advantages of convenient, fast, and can support fragmented learning, online learning has now been recognized by student groups at all levels, and its application advantages are even more obvious in terms of language disciplines such as English. The evaluation of English learning ability includes various aspects such as English listening, English grammar, English speaking, English reading, English writing, and English translation, etc. [1-2]. Each student's basic English knowledge and learning ability are different, and it is difficult to accurately find out the content suitable for their own learning conditions from the massive English learning resources [3]. A mature English online learning platform should be able to meet students' demand

for learning resources, improve their access efficiency to each learning project, and optimize their learning experience.

Now a few resource recommendation systems that can effectively alleviate the problems of large resource volume and slow screening speed have been applied to various Internet platforms [4-8]. The basic idea of these resource recommendation systems is to mine the data of students' preferences or abilities based on their operation history or evaluation information to further realize the prediction and recommendation services of their requirements and interests [9-11]. For an online learning platform with a complex knowledge system and diverse resource categories, the recommendation of learning resources can be considered from two aspects: the organization and processing of heterogeneous data on the platform side, and the knowledge level differences on the student side.

The traditional resource recommendation algorithm models established based on RNN sequence have limited data mining depth of students' learning requirements [12-16]. Scholars Sarra and Inès [17] combined edge information learning with user behavior learning, and introduced a collaborative attention mechanism to build a short-term conversation recommendation model that can analyze the potential preference of students' learning project visit history, and the proposed model had achieved good recommendation effect. Scholars Mehdi and David [18] took into account factors such as the different categories in the resource systems and the educational process of the resource subjects and proposed a recommendation model that is sensitive to students' learning ability level, and the proposed model contains three main function modules, namely the student ability level characterization module, the candidate learning resource characterization module, and the resource recommendation module. In the current knowledge push or resource recommendation services, the main reason for the inability to fully explore the information of the potential relationships between students, learning projects, and learning resources is the sparseness of data [19-22]. In order to improve the accuracy of knowledge push or resource recommendation, Yang et al. [23] applied a resource content feature extraction optimization algorithm based on multi-source data fusion to the dynamic tracking of students' learning preferences; in order to avoid the problem of undiversified content that is commonly seen during resource recommendation, they also proposed a multi-user joint-comparison algorithm to evaluate the candidate learning resources. Yang et al. [24] designed and constructed a personalized knowledge recommendation system suitable for networked education mode based on a repeated prediction mechanism and the resource collaborative filtering method, and detailed the system's function objectives, design mode, software and hardware structure, as well as debugging and solutions. As the application environment of online learning systems is becoming more and more complex, current resource recommendation models generally have problems such as data sparseness, user interest transfer, and cold start, etc., and these models need to be further optimized [25]. Aiming at the problem of data sparseness caused by the asymmetry of students' abilities and interests, Jonathan et al. [26] evaluated the asymmetric similarity between students and learning projects, and introduced a weighted neural network that integrated a recurring mechanism to construct a matrix factorization recommendation algorithm model with higher ranking accuracy. At

current stage, online education models with more obvious periodic resource update features generally have more complicated data storage structures and forms. Instead of the traditional clustering algorithms, Bahman et al. [27] used the association rule retrieval method of knowledge maps to perform preliminary screening on candidate recommendation resources, and built a crowdfunding and crowd-innovation education resource recommendation platform that integrated resource integration, knowledge completion, and personalized recommendation.

In the process of online English learning resource recommendation, it is necessary to comprehensively analyze the relationship between the independent learning resources and knowledge points of different English learning projects; then, combining with the advantages of online teaching, staged learning objectives and learning path should be planned based on students' English learning ability differences; further, the relevant learning resource choice and recommendation should be given accordingly. Different from traditional algorithms, this paper proposes a resource recommendation algorithm for English online learning systems based on student learning ability evaluation. The second chapter of this paper gives the flow of the algorithm, since the algorithm is based on student learning ability evaluation, a four-layer test system is constructed. The third chapter proposes an English learning ability evaluation method based on the maximum expectation algorithm, and gives the detailed estimation methods for parameters like learning ability, degree of discrimination, difficulty, and guess coefficient. The fifth chapter uses experimental results to verify the recommendation effect of the proposed resource recommendation algorithm.

2 The English Learning Resource Recommendation Algorithm Based on Learning Ability Evaluation

In order to improve the learning efficiency and learning effect of students' online English learning, combining with the massive student behavior data and performance data stored in the online English learning systems, an English learning resource recommendation mechanism based on student learning ability feature mining and group analysis is proposed, as shown in Figure 1. In this mechanism, at first, the performance fluctuation is analyzed and the trend is predicted based on students' English learning ability evaluation results and their online-learning process test results; then, combining with the evaluation results of students' English learning ability, it clusters the students based on their ability level; at last, according to the analysis on the student groups' cognition level, knowledge acceptance level, and individual learning preferences, it achieves the planning of students' online English learning paths and the recommendation of personalized learning resources.

The English learning resource recommendation algorithm proposed in this paper, it includes 5 steps:

Step 1: Pre-process the massive student learning process behavior data, English ability evaluation result data, and the process test result data stored in online English learning systems, eliminate invalid data, and supplement vacant and missing items.

Step 2: Based on the above data, evaluate students' English learning ability, and extract the features; then, with the help of the feature preference and ability-level division clustering algorithm, find out student groups of resource recommendation.

Step 3: For students divided into a same group, their learning preference, cognition development level, knowledge acceptance level, and other factors that can affect the recommendation of English learning resources are analyzed to obtain differences in their English learning status.

Step 4: According to the above analysis results, English learning resources are recommended to students from two aspects, mainly including the semantic association analysis between English learning resources, and the relationships among students, learning projects, and learning resources. The learning projects mainly include individual ability improvement projects such as listening ability training, speaking ability training, writing ability training, and reading ability training, etc.

Step 5: To avoid repeated recommendation resources, the similarity of learning resources is measured based on learning projects; according to the measurement results, the resources are screened before recommendation.

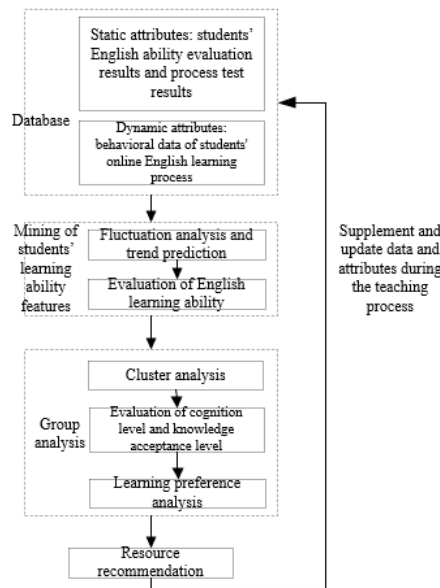


Fig. 1. Resource recommendation mechanism based on learning ability feature mining and group analysis

3 Evaluation Method of English Learning Ability

Since the basis for the proposed resource recommendation algorithm is to accurately measure the English learning ability of students, this paper also proposed the corresponding English learning ability evaluation model, the overall framework of the model is shown in Figure 2. The weighted sum of test scores of questions of each learning project is the evaluation result of a student's English learning ability. The question scores of students vary greatly with the question type, difficulty level, and degree of subjectivity. In view of these influencing factors, this paper proposed a maximum expectation algorithm for English learning ability evaluation, and the results obtained can be used to determine the content ratio of the learning resources to be recommended.

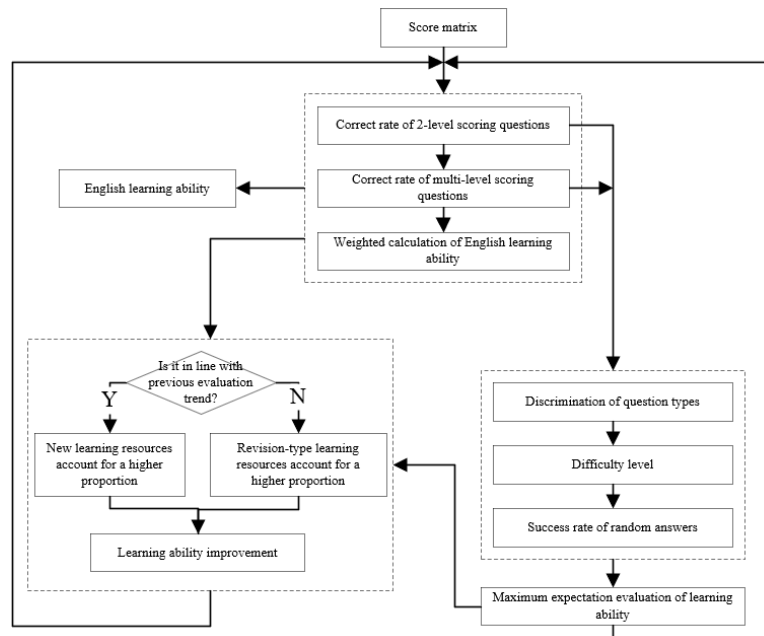


Fig. 2. Framework of the English learning ability evaluation model

3.1 Generation of learning ability evaluation functions

In the English learning ability evaluation process, assume the correct rate of the test questions conform to the shoulder-type normal distribution, Formula 1 is the function of correct rate of 2-level scoring questions (including one-choice questions, blank-filling questions, and true-or-false questions):

$$CR_i(a) = \int_{-\infty}^{d_i(a-h_i)} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \tag{1}$$

where, a is the student's English learning ability; d_i and h_i are the discrimination and difficulty parameters of the i -th 2-level scoring question. Since the 2-level scoring questions accounted for 80%-90% of all questions in the English learning ability evaluation, there is a certain probability that a random answer might be the correct answer. Therefore, by introducing a guess coefficient ε_i that describes the probability of random correctness, Formula 1 can be updated to Formula 2, a three-parameter correct rate function conforming to shoulder-type normal distribution:

$$CR_i(a) = \varepsilon_i + (1 - \varepsilon_i) \int_{-\infty}^{d_i(a-h_i)} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \quad (2)$$

Formulas 1 and 2 both include integral operations, which are not conducive to parameter estimation or practical application. By further optimizing Formula 2, a three-parameter logistic correct rate function could be obtained, as shown in Formula 3:

$$CR_i(a) = \varepsilon_i + \frac{1-\varepsilon_i}{1+e^{-d_i(a-h_i)}} \quad (3)$$

In the process of learning ability evaluation, for multi-level scoring questions such as English translation, oral interpretation, and English writing, the test system or the grading teacher will give scores according to the actual answering situations of the students. Compared with the 2-level scoring questions, the multi-level scoring questions cannot be answered at random, therefore they don't have the guess coefficient, and they can measure the learning ability of students more objectively. This paper subdivided the difficulty parameters of the multi-level scoring questions, then the corresponding three-parameter logistic correct rate function when the score of multi-level scoring questions is s could be expressed by Formula 4:

$$SR_i(a) = \frac{1}{1+e^{-1.7d_i(a-h_{is})}} \quad (4)$$

where, d_i and h_{is} are respectively the degree of discrimination of the i -th multi-level scoring question and the difficulty parameter when the score of the question is s .

3.2 Parameter estimation method

This paper evaluated students' English learning ability using Formula 3 (the correct rate function of 2-level scoring questions) and Formula 4 (the probability function of the multi-level scoring questions). The parameters in the functions need to be estimated using the score matrix constructed based on information data of questions and students' answers.

Suppose the quantized value of a student's English learning ability is a , and his/her scores on the N test questions of the English learning ability evaluation are $k_1, k_2, k_3, \dots, k_N$, respectively; then, the probability of this student's score in this evaluation is the product of the probabilities of scores of all questions:

$$\prod_{i=1}^N SP(K_i = k_i; a) \quad (5)$$

According to above formula, if the value of a changes, the value of the product of probabilities will change accordingly. Since the quantized value of a student's English learning ability actually exists, there's at least one value of a that can maximize the above formula, to write the value of the product of probabilities as a likelihood function of a , then, there is:

$$LF(a) = LF(k_1, k_2, k_3, \dots, k_N; a) \tag{6}$$

In this way, the problem of solving the value of a is transformed into finding the maximum value of $LF(a)$, that is, to find out the estimated value a^* that can maximize $LF(a)$ within the value range of a :

$$LF(a) = LF(k_1, k_2, k_3, \dots, k_N; a) = \max_{a \in A} LF(k_1, k_2, k_3, \dots, k_N; a) \tag{7}$$

Maximum expectation algorithm: Due to the uncertainty of the question parameters of the English learning ability evaluation and the learning ability parameters of students, the actual operation effect of the traditional maximum likelihood estimation algorithm is not ideal. Therefore, in this paper, the maximum expectation algorithm, which is not particularly sensitive to the completion of the data set, was adopted to estimate the quantized value of a . To facilitate the calculation, a_i was written in discrete form as $\delta_j, j=1,2,\dots,J$, and the corresponding score probability was represented by $DP(a_j)$. Suppose the score matrix of the English learning ability evaluation of a student can be described by an $M \times N$ matrix, $U = (u_1, u_2, u_3, \dots, u_M)$, then, for the i -th student, his/her score of the l -th question is u_{il} ; if the parameter of the l -th question is set as c_l , then, the test paper parameters are $C = (c_1, c_2, c_3, \dots, c_N)$, and the conditional distribution of the score matrix can be expressed by Formula 8:

$$f(u|C, DP(a)) = \sum_{j=1}^J f(u, \delta_j | C, DP(a_j)) = \sum_{j=1}^J f(u, \delta_j | C) DP(a_j) \tag{8}$$

Assume the correct rate of a student with a learning ability of δ_j in answering the l -th question is $CR(\delta_j | c_l)$, then, based on the local independence hypothesis theory, the above formula can be converted into the following formula:

$$f(u|C, DP(a)) = \prod_{l=1}^N CR(\delta_j | c_l)^{u_{il}} [1 - CR(\delta_j | c_l)]^{1-u_{il}} \tag{9}$$

Assume the number of students whose English learning ability value reaches δ_j is g_j , among the g_j students, the number of students who had correctly answered the l -th question is v_{jl} ; then, the likelihood function of the score matrix shown as Formula 10 could be obtained:

$$\begin{aligned} LF(u, a | C, DP(a)) &= \prod_{i=1}^M \prod_{l=1}^N CR(a_i | c_l)^{u_{il}} [1 - CR(a_i | c_l)]^{1-u_{il}} f(a_i | DP(a)) \\ &= \prod_{i=1}^M \prod_{l=1}^N CR(a_i | c_l)^{u_{il}} [1 - CR(a_i | c_l)]^{1-u_{il}} f(a_i | DP(a)) \\ &= \prod_{l=1}^N \prod_{i=1}^M CR(a_j | c_l)^{v_{jl}} [1 - CR(\delta_j | c_l)]^{g_j - v_{jl}} DP(a_j)^{g_j} \end{aligned} \tag{10}$$

Assume TMI is the number of students who answered the l-th question correctly among all the students participating in the test, then the above formula can be written in the logarithmic form shown as Formula 11 below:

$$\log[L(V, G|C, DP(a))] = \prod_{l=1}^N \prod_{i=1}^M v_{il} \log[CR(a_j|c_l)] + (g_j - v_{jl}) \log[1 - CR(\delta_j|c_l)] + g_j \log(DP(a_j)) \quad (11)$$

In the iteration process of the maximum expectation algorithm, the conditional probability of δ_j in the current iteration needs to be calculated based on the δ_j obtained in the previous iteration:

$$f(\delta_j|u_i, C^t, DP(a)) = \frac{f(u_i|\delta_j, C^t) DP(a_j)}{\sum_{j'=1}^J f(u_i|\delta_{j'}, C^t) DP(a_{j'})} \\ = \frac{DP(a_j) \prod_{l=1}^N CR(\delta_j|c_l^t)^{u_{il}} [1 - CR(\delta_j|c_l^t)]^{1-u_{il}}}{\sum_{j'=1}^J DP(a_{j'}) \prod_{l=1}^N CR(\delta_{j'}|c_l^t)^{u_{il}} [1 - CR(\delta_{j'}|c_l^t)]^{1-u_{il}}} \quad (12)$$

where, v_{jlt} and g_{jt} are respectively the conditional expectation values of the t-th iteration of the maximum expectation algorithm. g_{jt} can be calculated by Formula 13:

$$g_j^t = E(g_j|U, C^t, DP(a)) = \sum_{i=1}^M f(u_i|\delta_j, C^t, DP(a)) \\ = \sum_{i=1}^M \frac{f(u_i|\delta_j, C^t, DP(a_j))}{\sum_{j'=1}^J f(u_i|\delta_{j'}, C^t, DP(a_{j'}))} \\ = \sum_{i=1}^M \frac{DP(a_j) \prod_{l=1}^N CR(\delta_j|c_l^t)^{u_{il}} [1 - CR(\delta_j|c_l^t)]^{1-u_{il}}}{\sum_{j'=1}^J DP(a_{j'}) \prod_{l=1}^N CR(\delta_{j'}|c_l^t)^{u_{il}} [1 - CR(\delta_{j'}|c_l^t)]^{1-u_{il}}} \quad (13)$$

v_{jlt} can be calculated by Formula 14:

$$v_{jl}^t = u_{il} g_j^t = \sum_{i=1}^M \frac{u_{il} DP(a_j) \prod_{l=1}^N CR(\delta_j|c_l^t)^{u_{il}} [1 - CR(\delta_j|c_l^t)]^{1-u_{il}}}{\sum_{j'=1}^J DP(a_{j'}) \prod_{l=1}^N CR(\delta_{j'}|c_l^t)^{u_{il}} [1 - CR(\delta_{j'}|c_l^t)]^{1-u_{il}}} \quad (14)$$

Substitute Formulas 13 and 14 into Formula 12 to realize the maximization of Formula 12:

$$\log[LF(V^t, G^t|C, DP(a))] = \sum_{l=1}^N [z(c_l) + z(DP(a))] = \\ \sum_{l=1}^N \left[\sum_{j=1}^J \log(CR(\delta_j|c_l)) + (g_j - v_{jl}) \log(1 - CR(\delta_j|c_l)) \right] \\ + \sum_{j=1}^J g_j \log(DP(a_j)) \quad (15)$$

Based on Formula 15, the conditional expected value clt of the question parameter can be obtained. After completing the calculation of all the parameters of the t-th iteration, all these parameters could be brought into the t+1-th iteration to continue the calculation.

Since only the $z(c_l)$ in the log-likelihood function is related to the question parameter, finding the partial derivative of the log-likelihood function can be equivalent to finding the derivative of $z(c_l)$ with respect to c_l :

$$\sum_{j=1}^J \frac{v_{jl}^t - g_j^t CR(\delta_j|c_l)}{[1 - CR(\delta_j|c_l)]CR(\delta_j|c_l)} \frac{\partial CR(\delta_j|c_l)}{\partial c_l^t} = 0 \quad (16)$$

Since the maximum expectation algorithm can only estimate the parameters of the test questions, the learning ability parameters of students need to be estimated based on the maximum likelihood algorithm and combined with the solved question parameter C . The maximum likelihood function of a can be expressed by Formula 17:

$$LF(u_1, u_2, \dots, u_M|a) = \prod_{l=1}^N SP(U_l = u_l) \quad (17)$$

In the same way, by finding the derivative of the likelihood function, the desired a value could be obtained:

$$\frac{\partial LF(u_1, u_2, \dots, u_M|a)}{\partial a} = 0 \quad (18)$$

Estimation of learning ability: Formula 18 has a unique solution and it is the maximum value point, then this solution must be the desired maximum likelihood estimate of student's learning ability. Because the derived function of unknown students' learning ability in Formula 18 has a certain degree of nonlinearity, in order to solve the above-mentioned nonlinear optimization problem, this paper selected the Newton's method which is often used to solve approximate equations in real and complex number fields instead of the traditional direct equation method to solve the equations. The problem of solving the extreme value of the objective function $\phi(x)$ to-be-optimized can be transformed into the problem of solving the equation $\phi'(x)=0$. By expanding $\phi(x)$ according to the Taylor series expansion, we can get Formula 19 as:

$$\phi(x + \Delta x) = \phi(x) + \phi'(x)\Delta x + \frac{1}{2}\phi''(x)\Delta x^2 \quad (19)$$

The only condition for the above formula to hold is that Δx infinitely approaches 0, and it is equivalent to Formula 20:

$$\phi'(x) + \phi''(x)\Delta x = 0 \quad (20)$$

By solving Formula 20, we can get:

$$\Delta x = -\frac{\phi'(x_n)}{\phi''(x_n)} \quad (21)$$

The update iteration of x can be completed by Formula 22:

$$x_{n+1} = x_n - \frac{\phi'(x_n)}{\phi''(x_n)} \quad (22)$$

Based on the expectation posterior estimation algorithm that is not easily affected by iteration problems, the prior distribution was adopted to estimate the students'

English learning ability. The expected value of a can be calculated by Formula 23 when the score matrix is known:

$$E(a|U) = \frac{\int_{-\infty}^{\infty} aLF(U|a) \cdot \varphi(a) da}{\int_{-\infty}^{\infty} LF(U|a) \cdot \varphi(a) da} \quad (23)$$

Suppose IP_j and NIP are respectively the integration points of the expectation posterior estimation algorithm and the number of the integration points. LF(IP_j) represents the likelihood function value at the integration point, namely the probability for each test question to get the score of the score matrix when the student's English learning ability value is IP_j. The integration point value of the prior distribution function is represented by φ(IP_j), that is, the probability that the student's English learning ability value is IP_j. The Gaussian integration method was applied to calculate the estimated value a* of the expectation posterior estimation algorithm, as shown in Formula 24:

$$a^* = \frac{\sum_{j=1}^{NIP} IP_j \cdot LF(IP_j) \cdot \varphi(IP_j)}{\sum_{j=1}^{NIP} LF(IP_j) \cdot \varphi(IP_j)} \quad (24)$$

In the case there are N test questions in the English learning ability evaluation, the likelihood function can be described by Formula 25:

$$LF(U|a) = \prod_{l=1}^N CR_l(a)^{u_l} \cdot (1 - CR_l(a))^{1-u_l} \quad (25)$$

4 Experimental Results and Analysis

Table 1. Score probabilities of 2-level scoring questions

Test question type	Test question number	Error rate	Correct rate	Is there missing item?
One-choice question	q ₁₁	0.214	0.786	N
	q ₁₂	0.336	0.664	N
	q ₁₃	0.364	0.636	N
	q ₁₄	0.483	0.517	N
	q ₁₅	0.126	0.874	N
Cloze	q ₂₁	0.289	0.711	N
	q ₂₂	0.454	0.546	N
	q ₂₃	0.528	0.472	N
	q ₂₄	0.415	0.585	N
	q ₂₅	0.381	0.619	N

The statistical information of the data source used in the experiment in this study contained the data of the entire learning cycle of 6 learning projects of the English online learning platform, including the training of English listening, speaking, grammar, reading, translation, and writing. Question types included objective questions such as one-choice questions and cloze questions, and subjective questions such as translation, oral interpretation, and composition. A total of 156,332 pieces of learning behavior data and performance data of 500 students were included in the experiment,

after data preprocessing, 485 students and 148421 pieces of experimental data were retained.

Table 1 lists the score probabilities of some 2-level scoring questions. The first column is the test question type, the second column is the test question number, and the third and fourth columns are the error rate and correct rate of each question. Table 2 shows the values of feature parameters of these 2-level scoring questions. The third, fourth, and fifth columns are respectively the degree of discrimination, difficulty, and guess coefficient of each question. Since greater discrimination degree value and smaller guess coefficient indicate that the design of the question is more reasonable, to get better evaluation results, the proportion of 2-level scoring questions with similar attribute as Question q25 should be increased.

Table 2. Values of feature parameters of 2-level scoring questions

Test question type	Test question number	Degree of discrimination	Difficulty	Guess coefficient
One-choice question	q ₁₁	0.652	-1.472	0.001
	q ₁₂	0.226	-4.483	0.021
	q ₁₃	0.953	-0.496	0.004
	q ₁₄	0.776	-0.622	0.003
	q ₁₅	0.435	-3.264	0.006
Cloze	q ₂₁	0.458	-2.717	0.003
	q ₂₂	0.596	-2.825	0.002
	q ₂₃	0.691	-1.845	0.001
	q ₂₄	0.832	-0.109	0.001
	q ₂₅	0.903	-0.211	0.001

Table 3. Score probabilities of multi-level scoring questions

Test question type	Test question number	Probability of scored 1 point	Probability of scored 2 points	Probability of scored 3 points	Probability of scored 4 points	Is there missing item?
English to Chinese translation	q ₃₁	0.0153	0.0994	0.5477	0.3376	0
	q ₃₂	0.0831	0.2645	0.6210	0.0314	0
Chinese to English translation	q ₄₁	0.0356	0.1751	0.5943	0.195	0
	q ₄₂	0.0576	0.2783	0.4846	0.1795	0

Table 3 shows the score probabilities of some multi-level scoring questions. The first and second columns are still the test question type and test question number. The third, fourth, fifth, and sixth columns are the probabilities of students scored 1 point, 2 points, 3 points, 4 points in these questions. Table 4 shows the values of feature parameters of multi-level scoring questions. The third column is the degree of discrimination of the question. The fourth, fifth, and sixth columns respectively give the difficulty for students to score 2 points, 3 points, and 4 points on the question. The multi-level scoring questions do not have guess coefficient. With the decrease of the probabilities of 1-2 points and the increase of the probabilities of 3-4 points on multi-level scoring questions, the evaluation value of students' English learning ability gradually increases.

Table 4. Values of feature parameters of multi-level scoring questions

Test question type	Test question number	Degree of discrimination	Difficulty 1	Difficulty 2	Difficulty 3
English to Chinese translation	q_{31}	1.441	-4.527	-2.534	1.583
	q_{32}	1.574	-2.417	-0.537	1.886
Chinese to English translation	q_{41}	2.237	-2.373	-0.985	1.831
	q_{42}	1.146	-3.287	-0.928	1.963

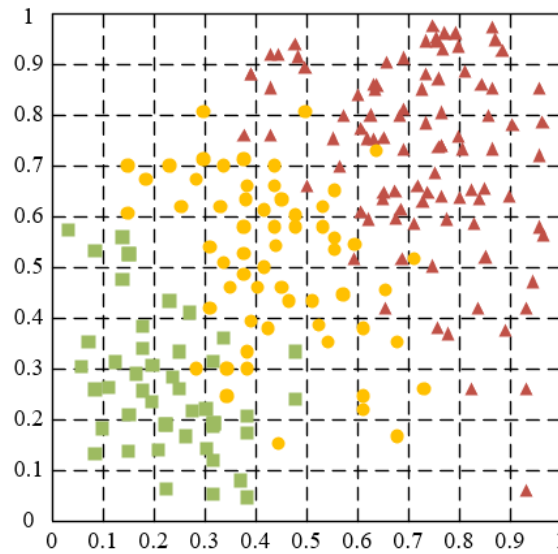


Fig. 3. The clustering results of the learning ability attributes of test questions of each learning project

Then, reasonable weights were assigned to describe the importance of the subjective and objective questions for the evaluation of student learning ability. Using the algorithm in this paper, the value range of student English learning ability was estimated to be [-5, +5]. Greater value indicates strong English learning ability, and vice versa. In this study, the WinGen data generation software was used to verify the effectiveness of the proposed learning ability evaluation method. Finally, the estimated learning ability value, degree of discrimination, difficulty, guessing coefficient and other parameters were consistent with the values generated by the software, which had verified that the learning ability evaluation values were scientific and reasonable.

For the learning ability attributes studied in this paper, clustering was carried out 8 times with the student English learning ability as the evaluation index. 3 clustering categories were set, corresponding to the three types of students with "weak ability", "average ability" and "strong ability". Figure 3 shows the clustering results of students based on the learning ability attributes of the test questions of each learning project.

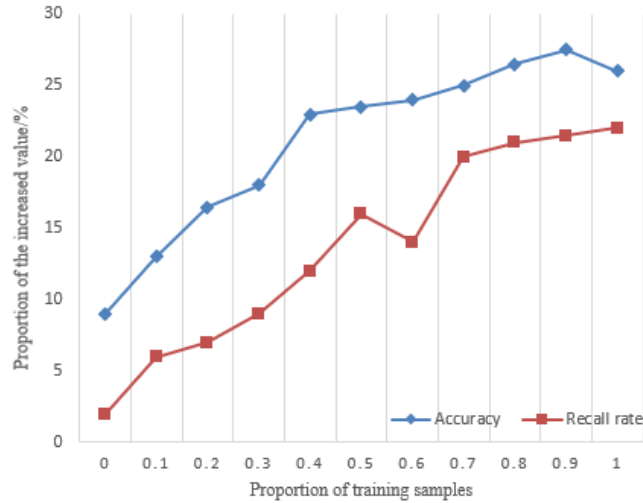


Fig. 4. Proportion of increased values of accuracy and recall rate

Compared with the traditional resource recommendation algorithm based on keyword matching, the resource recommendation algorithm based on student learning ability evaluation proposed in this paper is more reasonable in terms of recommendation ideas, and it can better master the learning features of students. According to Figure 4, as the proportion of training samples increases, the two evaluation index values of the accuracy and recall rate of the proposed algorithm gradually increase as well, which can further verify that the proposed algorithm has a better recommendation effect.

5 Conclusion

Based on student learning ability evaluation, this paper proposed a resource recommendation algorithm for online English learning systems. First, based on the design ideas of the proposed recommendation algorithm, this paper built a four-layer test system for the evaluation of English ability. Then, the paper elaborated the English learning ability evaluation method based on the maximum expectation algorithm, and realized the accurate estimation of parameters such as the learning ability value, degree of discrimination, difficulty, and guess coefficient. After that, this study performed experiments and adopted the WinGen data generation software to verify the effectiveness of the proposed learning ability evaluation method, and proved that the estimated learning ability values were scientific and reasonable. Also, the experimental results showed that the accuracy and recall rate of the proposed algorithm gradually increased with the increase in the proportion of training samples, which had verified that the proposed algorithm had a good recommendation effect.

6 Acknowledgement

This work was supported by Social Science Foundation of Hunan Province “Construction and application of college English online courses in Internet era” (No. 17WLH30) and a grant from teaching reformation projects of Education Department of Hunan Province “Construction and practice of college English online and offline blended courses of medical university under the background of comprehensive health” (No. HNJC-2020-1229).

7 References

- [1] Liu, H., Chen, R., Cao, S., Lv, H. (2020). Evaluation Of College English Teaching Quality Based On Grey Clustering Analysis, *International Journal of Emerging Technologies in Learning*, 16(2): 173-187. <https://doi.org/10.3991/ijet.v16i02.19727>
- [2] Li, N. (2020). A Fuzzy Evaluation Model of College English Teaching Quality Based on Analytic Hierarchy Process, *International Journal of Emerging Technologies in Learning*, 16(2): 17-30. <https://doi.org/10.3991/ijet.v16i02.19731>
- [3] Santana, B.S., Wives, L.K. (2020). Extraction and Use of Structured and Unstructured Features for the Recommendation of Urban Resources. In *International Conference on Computational Processing of the Portuguese Language*, 206-214. https://doi.org/10.1007/978-3-030-41505-1_20
- [4] Pudyastuti, K., Harmen, H., Djumantara, M., Fadhlurrahman, S. (2019). Thermal characteristics of the X geothermal resources, and a recommendation on utilization possibility. In *Journal of Physics: Conference Series*, 1402(2): 022057. <https://doi.org/10.1088/1742-6596/1402/2/022057>
- [5] Hajri, H., Bourda, Y., Popineau, F. (2018). Personalized recommendation of open educational resources in MOOCs. In *International Conference on Computer Supported Education*, 166-190. https://doi.org/10.1007/978-3-030-21151-6_9
- [6] Tana, G. (2019). Recommendation Algorithm for Potential Candidates in Human Resources Based on the Integration of Dynamic Decision Diagram. In *2019 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, 284-288. <https://doi.org/10.1109/ICSGEA.2019.00072>
- [7] Salehi, M. (2013). Application of implicit and explicit attribute based collaborative filtering and BIDE for learning resource recommendation. *Data & Knowledge Engineering*, 87: 130-145. <https://doi.org/10.1016/j.datak.2013.07.001>
- [8] Oh, Y., Kim, Y. (2019). A resource recommendation method based on dynamic cluster analysis of application characteristics. *Cluster Computing*, 22(1): 175-184. <https://doi.org/10.1007/s10586-018-2829-0>
- [9] Zhang, H., Huang, T., Lv, Z., Liu, S., Yang, H. (2019). MOOCRC: A highly accurate resource recommendation model for use in MOOC environments. *Mobile Networks and Applications*, 24(1): 34-46. <https://doi.org/10.1007/s11036-018-1131-y>
- [10] Kao, D., Lai, K.T., Chen, M.S. (2019). An Efficient and Resource-Aware Hashtag Recommendation Using Deep Neural Networks. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 150-162. https://doi.org/10.1007/978-3-030-16145-3_12

- [11] Jetinai, K. (2018). Rule-based reasoning for resource recommendation in personalized e-learning. In 2018 international conference on information and computer technologies (ICICT), 150-154. <https://doi.org/10.1109/INFOCT.2018.8356859>
- [12] Arias, M., Munoz-Gama, J., Sepúlveda, M., Miranda, J.C. (2018). Human resource allocation or recommendation based on multi-factor criteria in on-demand and batch scenarios. *European Journal of Industrial Engineering*, 12(3): 364-404. <https://doi.org/10.1504/EJIE.2018.092009>
- [13] Beldjoudi, S., Seridi, H., Zucker, C.F. (2017). Personalizing and improving resource recommendation by analyzing users' preferences in social tagging activities. *Computing and Informatics*, 36(1): 223-256.
- [14] Arias, M., Rojas, E., Lee, W.L.J., Munoz-Gama, J., Sepúlveda, M. (2016). ResRec: A Multi-criteria Tool for Resource Recommendation. In *BPM (Demos)*, 17-22.
- [15] Li, G.Q., Xia, J.L., Mei, S.Y. (2018). Learning resource recommendation method based on fuzzy logic. *Journal of Engineering Science and Technology Review*, 11(4): 146-153. <https://doi.org/10.25103/jestr.114.19>
- [16] Samia, B., Hassma, S. (2018). Improving resource recommendation in online collaborative learning. *Proceedings of the 29th French Knowledge Engineering Conference*, 267-269.
- [17] Sarra, B., Inès, S. (2018). Towards a personalized e-learning resource recommendation. *Proceedings of the 29th French Knowledge Engineering Conference*, 2: 265-266.
- [18] Mekni, M., Haynes, D. (2020). Smart Community Health: A Comprehensive Community Resource Recommendation Platform. In *HEALTHINF*, 614-624.
- [19] Wu, Y., Liu, Q., Chen, R., Li, C., Peng, Z. (2020). A Group Recommendation System of Network Document Resource Based on Knowledge Graph and LSTM in Edge Computing. *Security and Communication Networks*, 2020. <https://doi.org/10.1155/2020/8843803>
- [20] Biazus, M., dos Santos, C.H., Takeda, L.N., de Oliveira, J.P.M., Fantinato, M., Mendling, J., Thom, L.H. (2019). Software Resource Recommendation for Process Execution Based on the Organization's Profile. In *International Conference on Database and Expert Systems Applications*, 118-128. https://doi.org/10.1007/978-3-030-27618-8_9
- [21] Slimani, H., Hamal, O., El faddouli, N.E., Bennani, S., Amrous, N. (2020). The hybrid recommendation of digital educational resources in a distance learning environment: the case of MOOC. In *Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications*, 1-9. <https://doi.org/10.1145/3419604.3419621>
- [22] Agostino, F., Giuseppe, P. (2020). An agent-based platform for resources recommendation in internet of things. *ICEIS 2020 - Proceedings of the 22nd International Conference on Enterprise Information Systems*, 1: 775-779. <https://doi.org/10.5220/0009573207750779>
- [23] Yang, Z.Q., Yin, Z., Wang, H. (2014). Providing resource recommendation based on interactive behavior and resource content. *Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal of Computer-Aided Design and Computer Graphics*, 26(5): 747-754.
- [24] Arya, S., Subramani, T., Karunanidhi, D. (2020). Delineation of groundwater potential zones and recommendation of artificial recharge structures for augmentation of groundwater resources in Vattamalaikarai Basin, South India. *Environmental Earth Sciences*, 79(5): 1-13. <https://doi.org/10.1007/s12665-020-8832-9>
- [25] Rashidi, B., Fung, C., Vu, T. (2014). Recdroid: A resource access permission control portal and recommendation service for smartphone users. In *Proceedings of the ACM MobiCom workshop on Security and privacy in mobile environments*, 13-18. <https://doi.org/10.1145/2646584.2646586>
- [26] Gemmell, J., Mobasher, B., Burke, R. (2014). Resource Recommendation in Social Annotation Systems Based on User Partitioning. In *International Conference on Electronic*

Commerce and Web Technologies, 101-112. https://doi.org/10.1007/978-3-319-10491-1_11

- [27] Rashidi, B., Fung, C., Vu, T. (2015). Dude, ask the experts!: Android resource access permission recommendation with RecDroid. In 2015 IFIP/IEEE international symposium on integrated network management (IM), 296-304. <https://doi.org/10.1109/INM.2015.7140304>

8 Author

Zhenfang Zhou was born in February 1980 in Xiangxiang, Hunan province, China. She graduated from Hunan Normal University with a master's degree in foreign languages and literature in 2009. At present, she is working in the School of International Education, Hunan University of Medicine, mainly engaged in the research of college English education and teaching reform. Email: 330463327@qq.com

Article submitted 2021-02-13. Resubmitted 2021-03-21. Final acceptance 2021-03-22. Final version published as submitted by the authors.