ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2- 2026

Research on recommendation Model of English self-regulated Learning Based on Learning Behavior Data

Zhen Wang[1,*]

[1] School of Foreign Language Studies, Nanyang Institute of Technology, Nanyang, Henan, 473000, China

SUMMARY: *In recent years, learning analytics and intelligent recommendation technologies have continuously entered the foreign language education scene, making English autonomous learning gradually shift from experience support to data driven. Aiming at the problems of coarse granularity of resource push, insufficient recognition of learning status, and disassociation between recommendation results and process feedback in existing English learning platforms, this paper proposes an English self-regulation learning recommendation model based on learning behavior data. This paper extracts multi-dimensional features from login, browsing, practice, test, error correction and review behaviors, constructs learning state representation, resource matching and dynamic ranking mechanism, and further designs a comprehensive management and personalized recommendation system for English learning resources. Experimental results show that the proposed model achieves 0.412, 0.386, 0.447 and 0.521 on Precision@5, Recall@5, NDCG@5 and MRR indicators, respectively. The system still maintains good response performance under the condition of resource scale expansion and concurrent access. The results show that learning behavior data can provide effective support for English autonomous learning recommendation.*

KEYWORDS: *learning behavior data; Autonomous English learning; Self-regulated learning; Personalized recommendation*

1 Introduction

As English learning continues to become digital and platform oriented, learning activities have been extended from traditional classroom to course platform, mobile terminal and online resource environment. Especially in the scenario of college English autonomous learning, learners will continue to generate a large amount of behavioral data, such as login frequency, resource browsing path, practice stay time, test submission record, wrong question review times, and stage performance changes. These data can not only reflect the degree of learning engagement, but also reveal the differences of learners in task arrangement, strategy adjustment and feedback utilization. With the development of learning analytics, educational data mining and intelligent recommendation technology, how to extract stable features from continuously generated learning behavior data to support English autonomous learning has become an important issue in the cross research of foreign language education and computer technology.

The core of English autonomous learning is not only whether the number of resources is sufficient, but also whether learners can continuously adjust the learning content and learning

*hsgzwangzhen@163.com
https://doi.org/10.65102/is2026991

Published: 30 April, 2026

pace according to their own goals, ability level and process feedback. Previous studies have shown that self-regulated learning is closely related to learning engagement, learning effectiveness and continuous participation, and the feedback mechanism based on learning analysis can improve learners' strategy use and task completion quality to a certain extent. Therefore, behavioral data is not only a process record, but also an important basis for learning state recognition, ability profile construction, and personalized intervention generation. With the help of sequence modeling, feature representation learning and recommendation algorithms, the platform can shift from passive resource provision to active content matching, so as to generate more targeted learning paths for learners with different English levels and learning habits.

Although the current English learning platform and resource recommendation system have been widely used, there are still some shortcomings. Some systems focus on static label matching, and often push coarse-grained based on course category, vocabulary difficulty or historical click, which is difficult to capture learners' changes in the time dimension. Although some other systems introduce collaborative filtering, deep learning or reinforcement learning methods, they are more oriented to general course recommendation, and pay less attention to the continuous correlation between goal adjustment, behavior feedback and resource redistribution in English autonomous learning. In the real learning process, learners' state has obvious volatility, and different stages often show different concentration, strategy preference and content demand. If the recommendation model ignores the stage characteristics of the behavior sequence and the knowledge association of English learning tasks, the system output is prone to problems such as recommendation homogeneity, feedback lag and insufficient interpretability.

Based on this, this paper focuses on the problem of English self-regulated learning recommendation driven by learning behavior data, constructs a recommendation model that takes into account behavior feature representation, learning state recognition and personalized resource distribution, and further designs an integrated management and personalized recommendation system for English learning resources, in order to provide a technical solution with more adaptability, continuity and scalability for the English self-regulated learning platform.

2 Related Research

The existing research has formed a clear technical path around English autonomous learning, self-regulated learning support and educational recommendation system. One kind of research focuses on self-regulated learning itself, and mainly discusses the relationship between learners 'goal setting, strategy monitoring, process reflection and learning effectiveness. Yu[1] pointed out that self-regulation ability was an important factor affecting the effect of online second language learning. Yang et al. [2] believe that digital platforms and intelligent tools can improve the visualization degree of language learning process, but there is still a lack of a unified behavioral modeling framework. Mazandarani[3] proposed that self-regulation in language learning is not an abstract concept, but can be identified through observable behaviors such as platform interaction, task completion and strategy switching.

On this basis, some researches promote self-regulation support to a more specific computing implementation level with the help of intelligent classroom, learning analysis and feedback technology. Xu et al. [4] found that data-based teaching environment could strengthen the linkage relationship between self-regulation strategies, self-efficacy and learning engagement. Chen[5] proved that the feedback mechanism based on learning analysis could improve EFL learners' regulatory behavior and academic performance. Xia et al. [6, 7] pointed out that AI interactive tools can promote learners' active planning and continuous investment to a certain extent. Related research shows that English autonomous learning support is shifting from empirical judgment to data-driven, and continuous behavior records are gradually becoming an important basis for identifying learning states.

At the same time, the application of learning behavior data in the field of learning analysis continues to deepen. Ye and Pennisi[8] believe that trajectory data can restore the learner's adjustment path in detail. Lim et al. [9] pointed out that the personalized stent based on dynamic behavioral data is superior to the static unified intervention. Sun et al. [10] found that visiting, staying and task switching patterns in different periods were closely related to learning performance, cognitive load and participation status. Alhazbi et al. [11] and De Vreugd et al. [12] further emphasize that learning analysis should not stop at behavioral statistics, but also transform raw logs into interpretable, feedback-able, and intervenable information. Xu et al. [13] and Banihashem et al. [14] also show that the intervention supported by learning analysis is indeed helpful to improve learning effectiveness, but the premise is that the system can identify learners 'stage characteristics.

In parallel with it, the research of educational recommender system and personalized learning support is developed. Khor et al. [15] pointed out that the key of personalized learning lies in the joint modeling of learners' needs, interests, abilities and behavior patterns. Lampropoulos[16] believes that educational recommendation methods have gradually shifted from rule matching and collaborative filtering to composite patterns integrating deep learning, knowledge representation and behavior mining. Algarni and Sheldon[17], Narimani and Barbera [18] pointed out that although the existing recommender systems can realize resource screening and object matching, they are still insufficient in responding to the continuously changing learning needs. Ma et al. [19] proved that learners' historical behaviors, course characteristics and implicit preferences can be jointly represented by deep models, so as to improve the recommendation accuracy. The related studies are summarized in Table 1.

Table 1: Summary of existing related research content

Reference No.	Research Topic	Main Technologies or Methods	Main Conclusions	Existing Limitations
[1]–[3]	Research on English self-regulated learning	Self-regulated learning theory analysis and methodological reviews	These studies clarify the important influence of self-regulated learning on language learning outcomes	They focus more on theoretical and methodological discussion and lack concrete system implementation
[4]–[7]	Support for English self-regulated learning in intelligent environments	Structural equation modeling, learning feedback, and AI interactive tools	Data-driven environments can promote learning engagement and strategy adjustment	Continuous modeling of behavioral data remains insufficient
[8]–[14]	Learning analytics and behavioral trajectory research	Trajectory data analysis, temporal learning analytics, and learning dashboards	Behavioral data can be used to identify learning states and provide intervention support	Most studies emphasize analysis and feedback, while research on recommendation decision-making is relatively limited
[15]–[19]	Educational recommender systems and personalized learning	Collaborative filtering, deep learning, and learner profiling	These methods can improve resource matching accuracy and the level of personalization	They respond insufficiently to dynamic changes in learning needs
[20]–[24]	Research on English learning resource recommendation	Behavior detection, reinforcement learning, RNN, and text classification	English learning recommendation has begun to emphasize behavior awareness and intelligent delivery	System closed-loop capability remains weak, and integration with the self-regulated learning process is still insufficient

In the specific scenario of English learning, previous studies have begun to introduce behavior detection, text modeling and recommendation algorithms into the resource push task. Yuanfei[20] built an English teaching resource recommendation system based on learning behavior detection, which shows that behavior data can provide support for English resource matching. Zheng and Ding[21] introduced behavior perception and deep reinforcement learning mechanism to realize dynamic adjustment for English reading recommendation. Xian and Wu[22] combined the bag-of-words model with recurrent neural network for English vocabulary learning recommendation. Yan[23] optimized the English text reading recommendation model by fusing collaborative filtering and FastText classification method. Zhang[24], starting from the English online teaching platform, explained that platform logs, exercise records and process data have high value for identifying learning preferences and predicting learning performance. In general, the research on English learning recommendation has shifted from pure resource management to focus on both behavior perception and intelligent decision-making.

Although the existing research has made some progress, there are still several shortcomings: The research on self-regulated learning and recommendation system has not been truly connected, the stage change and strategy transfer in the behavior sequence are still not fully described, the knowledge association modeling between the resource difficulty, skill categories and learning objectives in English learning tasks is still rough, and the system-level resource organization and the algorithmic level dynamic recommendation lack a unified closed loop. Based on this, this paper takes learning behavior data as the core driving information, constructs a recommendation model for self-regulation process in the English self-regulated learning scenario, and co-designs behavior feature modeling, recommendation decision-making and resource management system to improve the platform's responsiveness to individual differences and process changes.

3 Design of English self-regulated learning recommendation system based on learning behavior data

3.1 Data analysis and feature representation of English self-regulated learning behavior

The value of English self-regulated learning behavior data lies not in simply recording "whether learners have studied", but in revealing how learners allocate time, choose resources, adjust strategies and respond to feedback during the learning process. For the English platform, the log left by learners is not a single click stream, but a multi-source heterogeneous sequence composed of access behavior, resource usage behavior, task completion behavior, test behavior and review behavior. For example, after entering the platform, learners browse vocabulary modules, switch listening units, complete reading exercises, view error analysis, repeat playback of audio, and modify learning plans. These operations together describe the way of goal advancement and adjustment ability in English learning. In order to make these discrete records enter the recommendation model, this paper uniformly represents the original behavior as a structured event with timestamp, object identification and result feedback. Let the record of action i of learner u be b_i, then there is:

$$b_i=(u,a_i,r_i,t_i,d_i,c_i) \quad (1)$$

where, a_i represents the type of behavior, r_i represents the object or resource number of the behavior, t_i represents the occurrence time, d_i represents the duration of stay, and c_i represents the result feedback corresponding to the behavior, such as whether the behavior is completed, the correct rate changes, or whether the review is triggered.

Due to the complex source of platform logs, there are usually problems such as repeated clicks, invalid stays, abnormal interrupts and time misplacement in the original data. If directly used for modeling, it is easy to cause misjudgment of learning status. Therefore, data cleaning, session segmentation, and time alignment need to be done before feature representation. In this paper, the 30-min no-operation interval is used as the session segmentation threshold to divide the continuous behavior of learners into multiple learning segments. For the set of actions of learner u in time window T, is defined as:

$$S_u^{(T)}=\{b_i \mid t_i \in T\} \tag{2}$$

This definition can reorganize daily fragmented learning into more analytical stage units, so that the model can retain fine-grained behavior information and not be disturbed by scattered noise. Figure 1 illustrates the flow of behavior data analysis and feature representation adopted in this section.

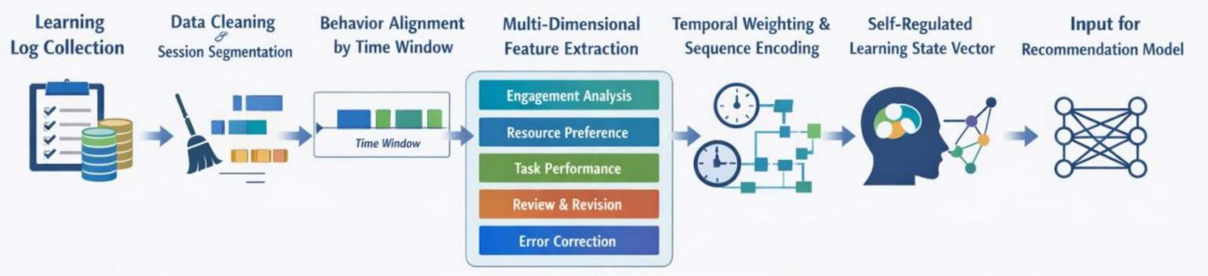

Figure 1: Process of data analysis and feature representation of English self-regulated learning behavior

After behavior segmentation, the system needs to convert different dimensions and different frequency indicators into comparable numerical features. This paper is represented from five dimensions: learning engagement, resource preference, task completion quality, review revision behavior, and rhythm stability. To avoid the disproportionate influence of high-frequency behaviors on the model, range normalization is used for continuous features, and the original value of learner u on the JTH feature is denoted as $f_{u,j}$, and the normalized result is denoted as:

$$x_{u,j}=\frac{f_{u,j}-f_j^{min}}{f_j^{max}-f_j^{min}+\varepsilon} \tag{3}$$

Here, f_j^{min} and f_j^{max} denote the minimum and maximum value of the JTH feature in all samples, respectively, and ε is a smoothing term that prevents the denominator from being zero. After this process, the characteristics such as login frequency, average stay time, resource switching times, and exercise completion rate can be compressed into a unified range, which is convenient for subsequent joint modeling.

A distinctive feature of self-regulated English learning is that short-term behavior often reflects current needs better than long-term behavior. Just because a learner browsed a lot of vocabulary resources two weeks ago does not necessarily mean that it still needs to be recommended. On the contrary, if they continue to listen to the wrong question in recent days and repeat the audio, it is more likely to indicate that listening comprehension is still the current weak point. Based on this consideration, this paper introduces time decay weights to assign differentiated contributions to events at different time positions in the behavior sequence. Let the current moment be tc, then the time weight of the action of article i is:

$$w_i = \exp\left[-\lambda\left(t_c - t_i\right)\right] \tag{4}$$

Here, λ is the time decay coefficient. This formula makes the recent behavior occupy a higher proportion in the state representation, so as to enhance the response speed of the recommendation system to the change of learning requirements.

In terms of behavior type encoding, this paper maps the operations such as browsing, bookmarking, repeated learning, exercise submission, wrong question review, and plan modification into the behavior embedding space, and combines the time weight to generate the temporal behavior representation of learners. Let $E(a_i)$ be the embedding vector of action type a_i, then the weighted action representation vector of learner u is defined as:

$$h_u = \frac{\sum_{i=1}^{n} w_i \, E(a_i)}{\sum_{i=1}^{n} w_i} \tag{5}$$

The vector preserves the overall outline of the learner's recent behavior structure, and can reflect whether the learner prefers input-based learning, training-based learning or revision learning. For example, if the weight of "reviewing wrong questions, redoing similar exercises, analyzing and browsing" continues to increase in a certain period of time, the system can identify that learners are in the stage of targeted compensation, rather than the stage of general resource browsing.

Behavior type and frequency alone are still not enough to support English learning recommendation, because different resources correspond to different language skill dimensions, and vocabulary, reading, listening, grammar and writing do not share the same difficulty structure. In order to enhance the teaching pertinence of feature representation, this paper further constructs skill mastery features based on resource labels and exercise results. Let $c_{u,k}$ denote the number of correct responses of learner u on skill dimension k, and $e_{u,k}$ denote the number of incorrect responses, then the mastery of this dimension is defined as:

$$m_{u,k} = \frac{c_{u,k}}{c_{u,k} + e_{u,k} + \varepsilon} \tag{6}$$

Through this formula, sub-abilities such as vocabulary recognition, reading comprehension, listening capture, and grammar judgment can be mapped into computable indicators. Compared with the single test score, this feature of mastery lays more emphasis on the cumulative performance of stages, which can more stably represent the relative strength of learners in different language skills.

Self-regulated learning is not equivalent to high-frequency learning. Although some learners have logged in many times, their task completion degree is low and the review chain is interrupted seriously, which does not form effective adjustment in fact. Therefore, this paper introduced the self-regulation index in addition to the behavioral statistics to comprehensively describe the four types of behaviors: plan execution, task persistence, feedback utilization and review return visit. Let p_u be the plan completion rate, q_u be the task persistence rate, f_u be the feedback response rate, v_u be the review return visit rate, then the self-regulation index of learner u can be expressed as:

$$R_u=\alpha p_u+\beta q_u+\gamma f_u+\delta v_u \tag{7}$$

where $\alpha+\beta+\gamma+\delta=1$, each parameter is set according to the performance of the validation set. The significance of this index is to converge the originally scattered regulation behaviors into a relatively stable state quantity, so that the system can distinguish two different types of learners: "high activity but low regulation" and "medium activity but high regulation".

In the actual modeling, the recommendation model does not need isolated indicators, but a comprehensive feature vector for unified input. Therefore, this paper concatenates normalized statistical features, temporal behavior embedding, skill mastery and self-regulation index to form the final representation:

$$z_u=[x_u \parallel h_u \parallel m_u \parallel R_u] \tag{8}$$

Here, x_u is the basic behavior statistical vector, h_u is the weighted behavior representation, m_u is the skill mastery vector, R_u is the self-regulation index, and the symbol "‖" represents the vector concatenation. This representation takes into account the quantitative features, sequence features, knowledge features and regulation features of learning behavior, which can provide more complete input information for subsequent recommendation models.

In order to facilitate system implementation, the core behavior fields and corresponding feature meanings are sorted out in Table 2. Each field in the table comes from the data interface that can be stably collected by the English autonomous learning platform, and does not rely on additional manual labeling, so it has good deployment feasibility.

Table 2: English self-regulated learning behavior data fields and feature meanings

Data Category	Main Fields	Feature Representation Content	Role in Recommendation
Platform access data	Number of logins, access time periods, session duration	Learning engagement intensity, distribution of study time	Used to assess learning activity level and recent study rhythm
Resource interaction data	Browsing, bookmarking, downloading, repeated playback	Resource interest preferences, depth of content usage	Used to identify preferred resource types and content needs
Practice behavior data	Number of submissions, accuracy rate, response time	Task completion quality, answering efficiency	Used to estimate the degree of difficulty matching
Error correction and review data	Wrong-question review, time spent on explanations, number of retries	Feedback utilization ability, revision tendency	Used to identify compensatory learning needs
Plan regulation data	Study plan setting, modification, completion status	Goal execution ability, self-monitoring level	Used to judge self-regulation status and recommendation intensity
Skill performance data	Vocabulary, reading, listening, grammar scores	Profile of mastery across different skills	Supports precise matching of category-specific resources

3.2 Recommendation Model design for English Self-regulated learning

After completing the data analysis and feature representation of learning behavior, the key tasks of the recommendation system have turned to the judgment of learners 'current learning status and the dynamic matching of English learning resources. Different from the recommendation in general e-commerce or information scenarios, English self-regulated learning recommendation does not simply pursue click-through rate maximization, but also considers the synergistic relationship between learning objectives, current ability, resource difficulty, skill shortcomings and periodic feedback. A click, a redo or a dropout of a learner in the platform may mean that his learning load, interest center or knowledge gap has changed. Based on this feature, this paper constructs a hierarchical recommendation model for English self-regulation learning, which organizes state representation, candidate resource generation, relevance matching, ranking adjustment and feedback update into continuous decision-making processes. The overall structure of the model is shown in Figure 2.

Figure 2: General structure of English self-regulated learning recommendation model

Figure 2 shows that the recommendation model is not a single-point predictor, but forms a closed loop around "learner status-resource attribute-interactive feedback". In order to depict the learner's comprehensive state at the current time, the behavior feature vector and the stage goal vector formed in Section 3.1 are jointly expressed as the learning state vector at time t:

$$s_u^{(t)}=[z_u^{(t)} \,||\, g_u^{(t)}] \tag{9}$$

Here, $z_u^{(t)}$ represents the behavior and ability representation of learner u at time t, and $g_u^{(t)}$ represents its current learning goal, such as vocabulary reinforcement, reading speed acceleration, listening compensation or comprehensive review. The function of this formula is to avoid the model only pushing similar resources according to historical behaviors, while ignoring learners 'recent task intentions, so as to make the recommendation results closer to the stage goal adjustment in English autonomous learning.

The second part of the recommendation model is candidate resource generation. English learning resources usually have the attributes of text topic, skill category, difficulty level, item structure and frequency at the same time. If the item-by-item matching is completed directly in the whole library, it is not only computationally expensive, but also easy to be interfered by low relevant resources. To this end, this paper first constructs a candidate set according to skill labels, resource types and recent needs. Let the original attribute vector of resource j be $[e_j \parallel d_j \parallel k_j]$, which represents the content semantic representation, difficulty representation and skill label representation respectively, then its unified encoding vector is:

$$q_j=W_r[e_j \parallel d_j \parallel k_j]+b_r \tag{10}$$

Here, W_r and b_r are trainable parameters. After this mapping, vocabulary cards, reading materials, listening segments, grammar exercises, and comprehensive quizzes are projected into the same vector space to facilitate uniform matching with learner states. The benefit of this processing is that instead of relying on a single label retrieval, the model can assimilate content similarity, skill attributes, and difficulty information simultaneously.

After the state modeling is completed, the system needs to determine which type of English ability the learner needs to compensate or improve the most at present. Since English self-regulated learning often presents a state of multi-skill intersection, the model should not simply use the highest frequency behavior as the judgment basis. In this paper, skill attention weights are introduced to dynamically focus the dimensions of vocabulary, reading, listening, grammar and writing. Let $m_{u,k}^{(t)}$ be the learner's current mastery on skill k and v_k be the trainable representation of this skill, then the weight is defined as:

$$\alpha_{u,k}^{(t)}=\frac{\exp\left(v_k^{\top} m_{u,k}^{(t)}\right)}{\sum_{l=1}^{K}\exp\left(v_l^{\top} m_{u,l}^{(t)}\right)} \tag{11}$$

Here, K represents the total number of skill dimensions. This formula enables the model to automatically adjust the attention center according to the learner's recent performance. If a learner's reading accuracy is stable and listening repetition is frequent, the listening related dimensions will obtain higher weights in the state matching, so as to promote the system to give priority to the targeted listening resources, rather than continue to repeat recommending the reading materials that the learner is familiar with.

In the candidate resource ranking stage, this paper uses the bilinear matching mechanism to calculate the base relevance score between the learner state and the resource. For a learner u and a candidate resource j, the matching score is denoted by:

$$\hat{y}_{u,j}^{(t)}=\sigma\left((s_u^{(t)})^{\top} M q_j+b\right) \tag{12}$$

where M is the state-resource interaction matrix, b is the bias term, and $\sigma(\cdot)$ is the Sigmoid function. The score reflects the basic probability of "whether the resource is worth recommending in the current state", but if the ranking is only based on this, the system is easy to continue to push high similar resources, which weakens the progressive nature of the training level. Therefore, learning gain judgments need to be added after the base matching.

The important difference between English learning recommendation and general interest recommendation is that whether a resource is "popular" does not equal whether it is "suitable for the current learning stage". Some resources have a high click-through rate, but the difficulty is too low to continue to improve learners' ability. Other resources, though less exposed, are just right for compensating for weak skills. To this end, this paper constructs a resource expected learning gain function, which brings capability gap, resource difficulty adaptation and history completion quality into the same judgment framework:

$$G_{u,j}^{(t)}=\eta_1\left(1-m_{u,k}^{(t)}\right)+\eta_2\,\mathrm{sim}(g_u^{(t)},k_j)-\eta_3\left|d_j-\bar{d}_u^{(t)}\right| \tag{13}$$

where η_1,η_2,η_3 are weight parameters, $\mathrm{sim}(g_u^{(t)},k_j)$ represents the similarity between the current learning goal and the resource skill label, and $\bar{d}_u^{(t)}$ represents the current appropriate difficulty level of the learner. This formula reflects three judgment directions: weak skill priority, consistent goal priority, too high difficulty or too low difficulty should be suppressed. This makes the list of recommendations less about what the user is likely to order and more about what the user is worth learning right now.

In the final output stage, we combine the base relevance score, expected learning gain, and novelty adjustment to form the ranking total score:

$$F_{u,j}^{(t)}=\beta_1\hat{y}_{u,j}^{(t)}+\beta_2G_{u,j}^{(t)}+\beta_3N_{u,j}^{(t)} \tag{14}$$

Here, $N_{u,j}^{(t)}$ represents the relative novelty of the resource to the learner, and β1+β2+β3=1. The reason for adding the novelty item is that in long-term learning, if the system always pushes too similar resources, learners are likely to have familiarity dependence, and the platform seems to be "accurate", but in fact, it will compress the learning scope. The moderate introduction of new resources and adjacent skills resources is conducive to expanding the scope of language input and improving the continuous learning experience. Figure 3 shows the specific process of recommendation decision and feedback update in this paper.

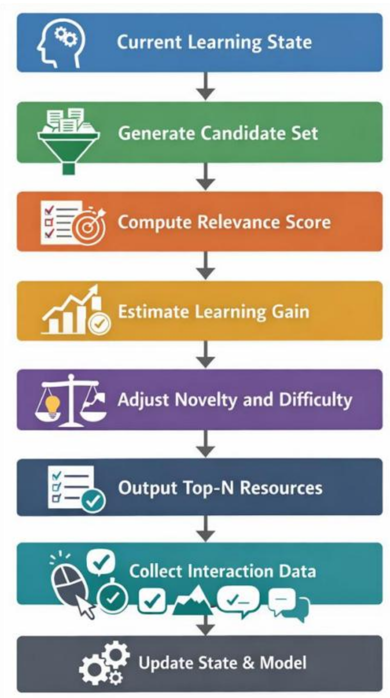

Figure 3: Dynamic recommendation and feedback update process of English self-regulated learning

After the recommendation result is output, the system does not regard a push as the end point, but continuously tracks learners 'click, completion, accuracy change and review behavior of the recommended resources, and updates the model parameters. Let $c_{u,j}^{(t)} \in \{0,1\}$ denote the learner's positive feedback labeling of resource j, then the online training loss can be written as:

$$L_t=-\sum_{(u,j)}\left[c_{u,j}^{(t)}\log\hat{y}_{u,j}^{(t)}+\left(1-c_{u,j}^{(t)}\right)\log\left(1-\hat{y}_{u,j}^{(t)}\right)\right] \quad (15)$$

On this basis, the model parameters are updated in gradient descent fashion:

$$\theta_{t+1}=\theta_t-\mu\nabla_\theta L_t \quad (16)$$

Here, θ represents the set of model parameters and μ is the learning rate. The update mechanism enables the system to revise the subsequent recommendation direction according to the immediate reaction of learners. For example, if the platform recommends several moderately difficult reading materials in a row, but the learner repeatedly interrupts and turns to vocabulary review, the system will lower the priority of such reading resources and increase the weight of vocabulary strengthening resources in the next update round.

3.3 Integrated management of English learning Resources and personalized recommendation system design

After completing the representation of learning behavior characteristics and the construction of the recommendation model, the focus of system design is no longer on the algorithm itself, but to further answer a more practical question: how to manage the scattered English learning resources in a unified way, and convert the output of the recommendation model into personalized resource services that learners can directly use in a stable and timely manner. For an English autonomous learning platform, resources are not a single text file, but a heterogeneous collection composed of vocabulary bag, reading materials, listening audio, grammar exercises, stage tests, parsing documents and learning task lists. These resources have obvious differences in format, granularity, skill attributes and frequency of use. Without a unified management structure, it is difficult to accurately push these resources even if the recommendation model can identify the learning needs. Therefore, in addition to the recommendation model, this paper further constructs a comprehensive management and personalized recommendation system for English learning resources, which integrates resource storage, metadata organization, authority control, online scheduling and feedback update into the same system framework. Its overall structure is shown in Figure 4.

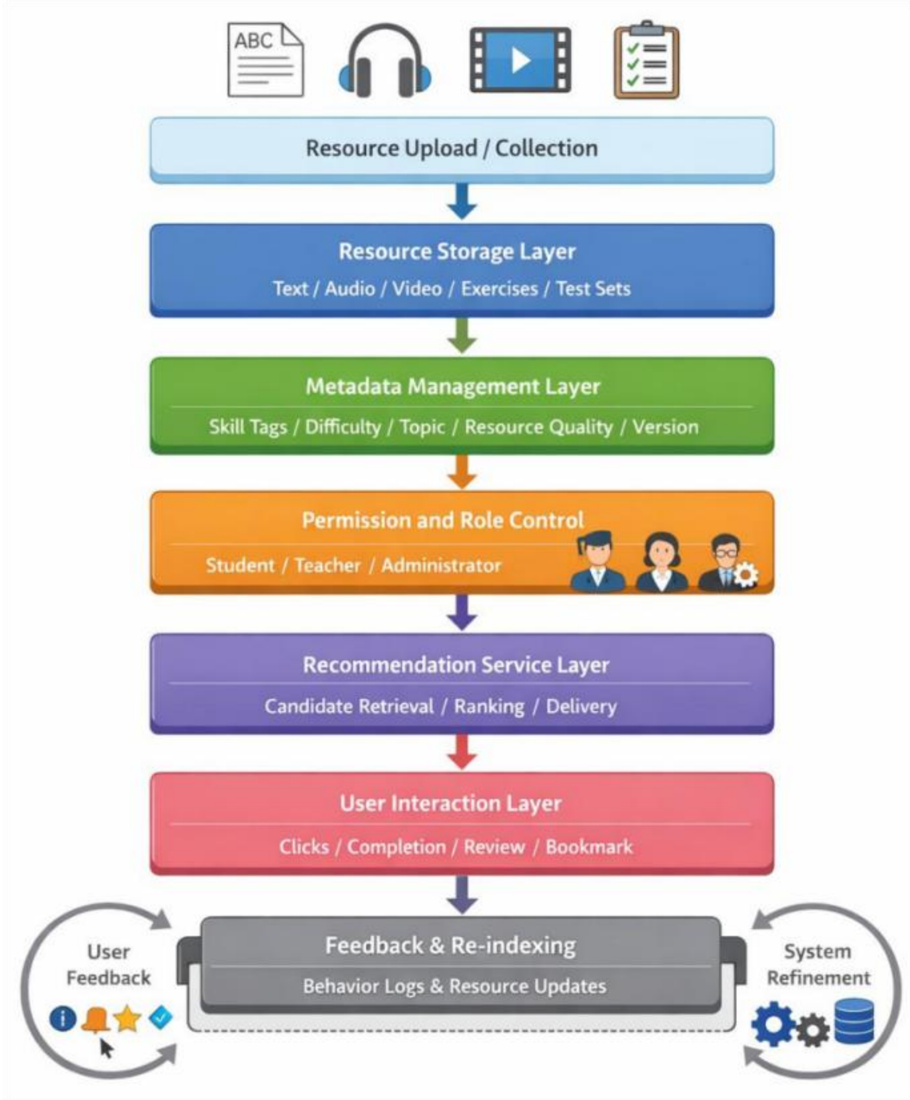

Figure 4: Overall structure of the integrated management and personalized recommendation system for English learning resources

As can be seen from Figure 4, the bottom layer of the system is not directly oriented to recommendation output, but first establishes a unified storage and description mechanism around resources. In this paper, each English learning resource record is represented as:

$$r_j=\left(id_j, type_j, skill_j, diff_j, topic_j, meta_j\right) \tag{17}$$

Among them, id_j is the resource number, $type_j$ represents the resource type, $skill_j$ represents the skill dimension, $diff_j$ represents the difficulty level, $topic_j$ represents the topic label, $meta_j$ represents the extended metadata including the duration, the number of questions, the version, the source and the applicable object. Through this structure, English resources with different formats are mapped into a unified data description space, which provides basic support for subsequent retrieval, filtering and recommendation scheduling.

In the stage of resource storage, the system needs to judge whether the resource has the basic conditions to enter the recommendation service. English learning resources cannot be uploaded and used. If there are problems such as missing labels, content distortion, inaccurate difficulty annotation or deviation from course objectives, the recommendation results will be directly affected. Therefore, this paper establishes quality assessment scores for resources. Let the completeness, content accuracy, pedagogical relevance and historical usage feedback of resource j be cj, aj, tj and uj, respectively, then its comprehensive quality score is defined as

In the stage of resource storage, the system needs to judge whether the resource has the basic conditions to enter the recommendation service. English learning resources cannot be uploaded and used. If there are problems such as missing labels, content distortion, inaccurate difficulty annotation or deviation from course objectives, the recommendation results will be directly affected. Therefore, this paper establishes quality assessment scores for resources. Let the completeness, content accuracy, pedagogical relevance and historical usage feedback of resource j be c_j、a_j、t_j and u_j, respectively, then its comprehensive quality score is defined as:

$$Q_j=\omega_1 c_j+\omega_2 a_j+\omega_3 t_j+\omega_4 u_j \tag{18}$$

Here, $\omega_1+\omega_2+\omega_3+\omega_4=1$. This score is used to control whether the resource enters the candidate library or not, and also for quality compensation in the subsequent recommendation ranking. The goal of this approach is to prevent the system from relying too heavily on behavior matching and pushing resources with average quality but high short-term popularity to the top.

In the resource management system, access control can not be ignored. English self-regulated learning platform usually contains three types of users, students, teachers and administrators. Different roles have different permissions in resource access, editing, auditing and recommendation. If there is no clear permission boundary, it will not only affect the security of resources, but also weaken the stability of system operation. Therefore, this paper adopts role-resource dual constraint mechanism. Let the role level of user u be $role_u$ and the access level of resource j be $level_j$, then its access decision function can be expressed as:

$$P(u,j)=\begin{cases}1, & role_u \geq level_j\\ 0, & role_u < level_j\end{cases} \tag{19}$$

When P(u,j), the system allows the user to access or call the corresponding resource. Otherwise, only the insufficient permission hint is returned. In this way, teachers can enter the resource audit and configuration interface, students mainly get the recommended results and learning record query permission, and administrators are responsible for system-level maintenance and resource life cycle management.

After completing the base management, the system needs to actually deploy the recommendation model in Section 3.2 as an online service. Specifically, the system does not directly calculate the recommendation score one by one in the full resource library, but first generates the candidate subset according to the skill tag, target topic, resource type and permission conditions, and then performs the fine ranking. Let the state of learner u at time t be denoted as $s_u^{(t)}$ and the encoding of candidate resource j be denoted as qj, then the candidate relevance score of resource retrieval phase is defined as:

$$C_{u,j}^{(t)}=\rho_1 \,\mathrm{sim}\left(s_u^{(t)},q_j\right)+\rho_2 Q_j+\rho_3 P(u,j) \tag{20}$$

where sim(·) represents the vector similarity function, ρ_1,ρ_2,ρ_3 are the weight parameters. This formula shows that whether a candidate resource should be ranked is not only determined by the matching degree between the user state and the resource semantics, but also the quality of the resource itself and the access legitimacy. In this way, invalid resources can be reduced into the ranking layer and the overall efficiency of the recommendation service can be improved.

The applicability of English learning resources also has obvious characteristics of timeliness. Although some resources are of high quality, if they have not been updated for a long time, the content expression, question structure or knowledge organization may not adapt to the current learning situation. Especially for vocabulary learning, reading training and real problem resources, the version update frequency will directly affect the recommendation value. Based on this, this paper introduces the resource timeliness factor. Let the current time be t_c and the last update time of resource j be t_j, then its timeliness score is:

$$F_j = \exp[-\kappa(t_c - t_j)] \tag{21}$$

Here, κ is the time decay coefficient. The score is not to simply suppress old resources, but to help the system to give priority to the content that is updated and more timely and adapted to the current teaching context in the case of similar quality.

In the output stage of personalized recommendation, the system needs to comprehensively consider resource correlation, resource quality, timeliness and system scheduling cost. Since the online service of the platform needs to strike a balance between the response speed and the recommendation effect, the final recommendation score is defined as in this paper:

$$R_{u,j}^{(t)} = \lambda_1 C_{u,j}^{(t)} + \lambda_2 F_j + \lambda_3 D_{u,j}^{(t)} \tag{22}$$

where, $D_{u,j}^{(t)}$ represents the difficulty adaptation term, which is used to measure the matching degree between the difficulty of resources and the current ability interval of learners, and $\lambda_1+\lambda_2+\lambda_3=1$. This ordering allows the system not only to recommend the "most relevant" resources, but also to avoid pushing content whose difficulty significantly deviates from the current state of learning. The online service process is shown in Figure 5.

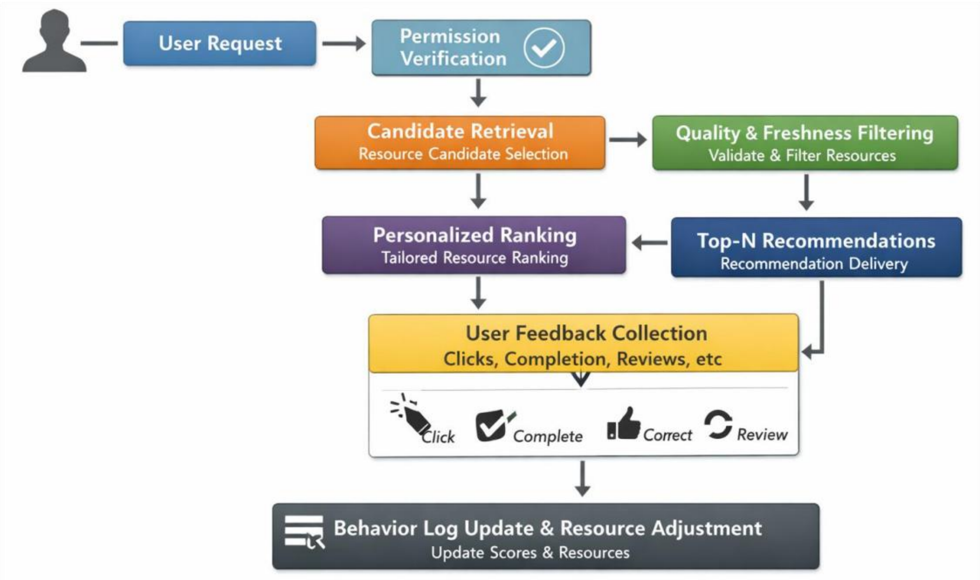

Figure 5: Personalized recommendation service and feedback reflux process

As can be seen from Figure 5, after the recommendation result is output, the system does not stop processing, but continues to track the resource usage and incorporate it into the subsequent resource management. For the English learning platform, resources "seen" does not mean resources "learned", the system is more concerned about whether the resource is completed after clicking, whether it improves the accuracy, and whether it triggers review visits. Therefore, in this paper, the feedback aggregation function is used to update the dynamic service weights of resources. Let the click rate, completion rate, correct rate increase and review return rate received by resource j in time window T be clk_j、fin_j、imp_j and rev_j, respectively, then the dynamic weight of resource j can be updated as:

$$W_j^{(T+1)}=\phi_1 clk_j+\phi_2 fin_j+\phi_3 imp_j+\phi_4 rev_j$$

Here, $\phi_1+\phi_2+\phi_3+\phi_4=1$. The significance of this formula is that the effectiveness feedback of resources in the real learning process is written back to the resource management system, so that the platform no longer only relies on the preset label of experts, but can continuously revise the resource value judgment according to the actual learning results.

In the engineering implementation level, the system is divided into four parts: resource storage layer, database management layer, recommendation service layer and front-end interaction layer. The resource storage layer is responsible for saving text, audio, video and exercise files. The database management layer is responsible for maintaining the relationship among user table, resource table, tag table, behavior log table and recommendation table. The recommendation service layer completes candidate retrieval, feature invocation and result ranking. The front-end interaction layer is responsible for presenting the recommendation results to learners and recycling the usage feedback. In order to ensure the stability of the platform, the system adopts asynchronous log writing and periodic index update mechanism, so that the recommendation call and background resource reconstruction are separated from each other to avoid service blocking in high concurrency scenarios.

4 Experimental results and analysis

4.1 Recommendation effect analysis of English self-regulated learning

In order to verify the effectiveness of the English self-regulated learning recommendation model constructed in this paper, this section carries out experimental analysis on three levels: recommendation accuracy, ranking quality and learning response performance. The experimental data comes from the real learning log of an English self-learning platform in a university for 12 consecutive weeks, which contains 428 learners, 18,764 resource records and 186,420 effective interactive behaviors. After removing duplicate clicks, invalid pauses and abnormal interrupts, the original data is divided into training set, validation set and test set according to the time order, and the ratio is 8:1:1. The experimental environment is deployed on Ubuntu 22.04, the server is configured with Intel Core i7-12700 processor, 32 GB RAM, and the development framework is Python 3.11, PyTorch 2.2 and PostgreSQL 14. To ensure the comparison is representative, four common methods, Popularity, ItemCF, GRU4Rec and DeepFM, are selected as the control models, and the evaluation indicators include Precision@5, Recall@5, NDCG@5 and MRR. At the same time, in order to test whether the recommendation results really contribute to English autonomous learning, this paper also counts two behavioral indicators: recommendation acceptance rate and task completion rate after recommendation.

Table 3 shows the comparison of the recommendation effect of different methods on the test set. It can be seen from Table 3 that the model in this paper has achieved the best results on four core indicators, and Precision@5 reaches 0.412, which is 0.146, 0.103, 0.067 and 0.031 higher than Popularity, ItemCF, GRU4Rec and DeepFM, respectively. Recall@5 reaches 0.386, indicating that the model can cover more real demand resources in the limited recommended bits. NDCG@5 is 0.447, indicating that highly relevant resources are more stably ranked at the top; The MRR reaches 0.521, which also reflects that the resources that learners will actually click on are higher in the recommended list on average. Compared with the method only relying on historical similarity, the ranking quality of the proposed model is improved more obviously, which indicates that the system can identify the phasic needs of English learning more accurately after the learning behavior sequence, self-regulation state and skill goal are jointly incorporated into the modeling.

Table 3: Comparison of the effects of different recommendation methods on the test set

Method	Precision@5	Recall@5	NDCG@5	MRR
Popularity	0.266	0.241	0.291	0.337
ItemCF	0.309	0.287	0.334	0.389
GRU4Rec	0.345	0.321	0.381	0.441
DeepFM	0.381	0.354	0.418	0.486
Proposed model	0.412	0.386	0.447	0.521

From the point of view of the training process, the model in this paper does not increase the complex structure in exchange for unstable local advantages, but shows better convergence characteristics in fewer iterations. Figure 6 shows the loss value and NDCG@5 change of the model under different training rounds. It can be seen that within the first 10 iterations, the loss value decreases rapidly from 0.624 to 0.431, and NDCG@5 increases from 0.301 to 0.401. After 20 rounds of iteration, the loss value further decreased to 0.356 and NDCG@5 reached 0.439. Around round 30, the curve flattens out and the final loss value stabilizes around 0.332 and NDCG@5 stabilizes around 0.447. Figure 6 shows that the proposed model can quickly complete effective learning, and there is no obvious oscillation on the validation set, indicating that there is a good synergy between the behavior feature representation and the state matching mechanism, rather than relying on accidental parameter fitting.

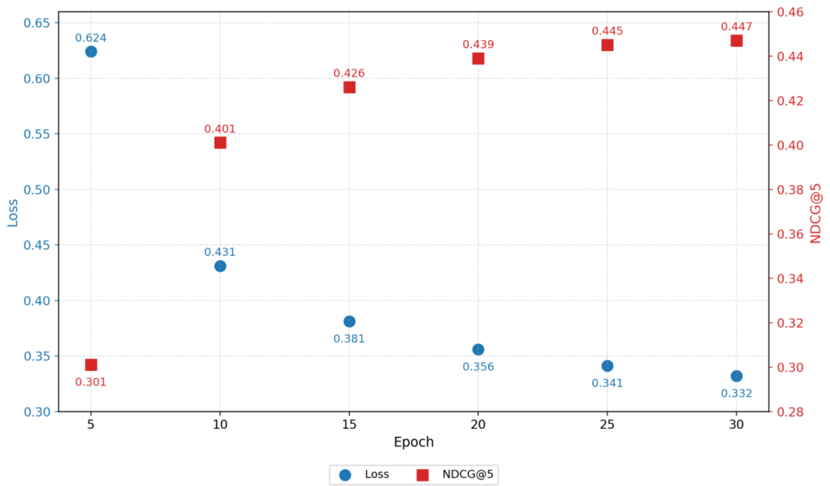

Figure 6: Model loss values and changes of NDCG@5 under different iteration rounds

Offline indicators alone are still not enough to illustrate the actual value of recommender systems in English self-regulated learning scenarios, so this paper further investigates the impact of recommendation results on learning behavior. Specifically, the test set learners were divided into three levels of low, medium and high according to the self-regulation index, and the behavioral responses generated by different groups after receiving the recommendation of the model in this paper were observed. Figure 7 shows the recommendation acceptance rate and task completion rate after recommendation for the three types of learners. The results showed that the recommendation acceptance rate was 61.8% and the task completion rate was 54.6% in the low self-regulation group. The moderate self-regulation group reached 68.9% and 62.7%, respectively; The high self-regulation group further increased to 74.3% and 69.8%. This result shows that the model in this paper has a positive support effect on learners with different levels of regulation, and the significance is especially obvious for the middle and low regulation groups. The reason is that these learners are more likely to hesitate or deviate from the goal in the resource selection stage, and the personalized recommendation driven by behavioral data reduces their decision-making cost to a certain extent, making it easier for them to enter the state of continuous learning.

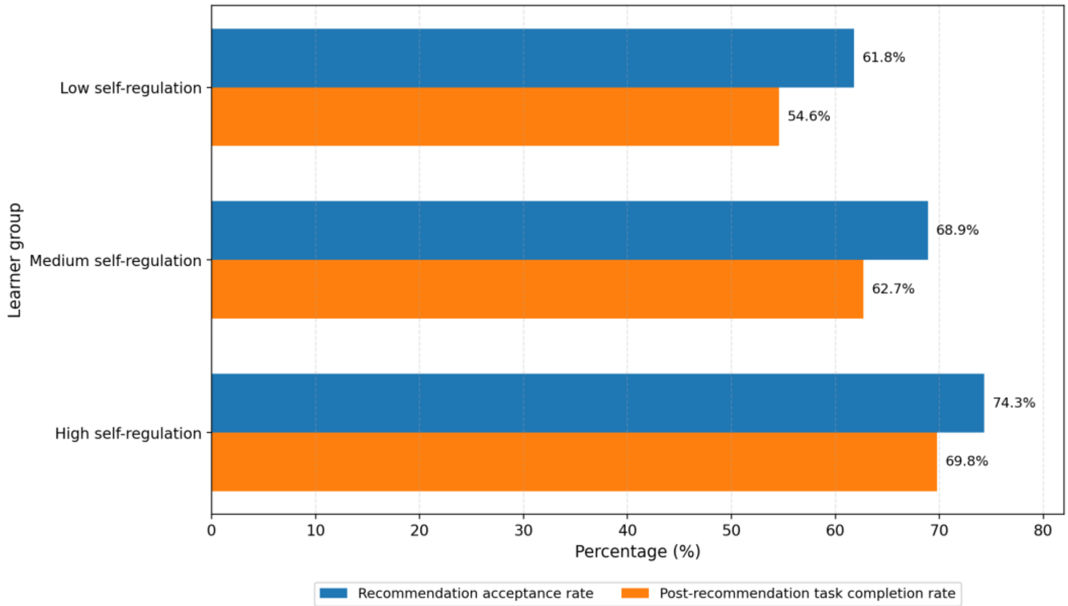

Figure 7: Recommendation response performance of learners with different levels of self-regulation

4.2 Experimental verification of comprehensive recommendation System for English Learning Resources

In order to test the usability and stability of the comprehensive recommendation system for English learning resources constructed in this paper under real operation conditions, this section no longer focuses on the recommendation ranking itself, but turns to the operation verification at the system level, focusing on the overall performance under the conditions of resource management, recommendation service invocation, feedback writeback and concurrent access. The experiment is still based on the above English autonomous learning platform, the test object is the comprehensive resource management and personalized recommendation system that has been deployed, and the control object is the traditional keyword retrieval resource management system running on the same resource library. When testing, the same hardware conditions and the same batch of resource data are retained to avoid additional impact on the results caused by environmental differences. The system deployment and test parameters are shown in Table 4.

Table 4: System deployment environment and test parameter Settings

Item	Configuration
Server environment	Ubuntu 22.04, Intel Core i7-12700, 32 GB RAM
Development and service framework	Python 3.11, FastAPI, PyTorch 2.2, Nginx
Database and cache	PostgreSQL 14, Redis 7.0
Resource library size	18,764 resources
Resource type composition	8,920 text resources, 4,386 audio resources, 3,947 practice resources, 1,511 test resources
Testing tool	Apache JMeter 5.6.3
Concurrent user settings	50, 100, 150, 200, 250
Observation metrics	Average response time, P95 response time, throughput, request success rate, recommendation refresh latency

It can be seen from Table 4 that this experiment is not completed under simplified samples, but on a resource scale close to the actual platform operating state. The system test is divided into two stages: one is to investigate the response changes of the system when the resource scale increases, and the other is to investigate the stability of the recommendation service under high concurrent access. Figure 8 shows the average response time variation curves of the two types of systems under different resource sizes. As the number of resources increases from 2,000 to 18,764, the average response time of the traditional retrieval system increases from 0.76 s to 2.73 s, while that of the proposed system increases from 0.71 s to 1.43 s. The overall growth rate is significantly slower. When the number of resources is less than 6,000, the gap between the two types of systems is not obvious. When the scale of resources exceeds 10,000, the advantages of the proposed system begin to enlarge continuously, which indicates that the organization of resource metadata, candidate caching and hierarchical retrieval mechanism can effectively reduce the additional overhead caused by the coupling of retrieval and recommendation in the scenario of medium and large scale English learning resources.

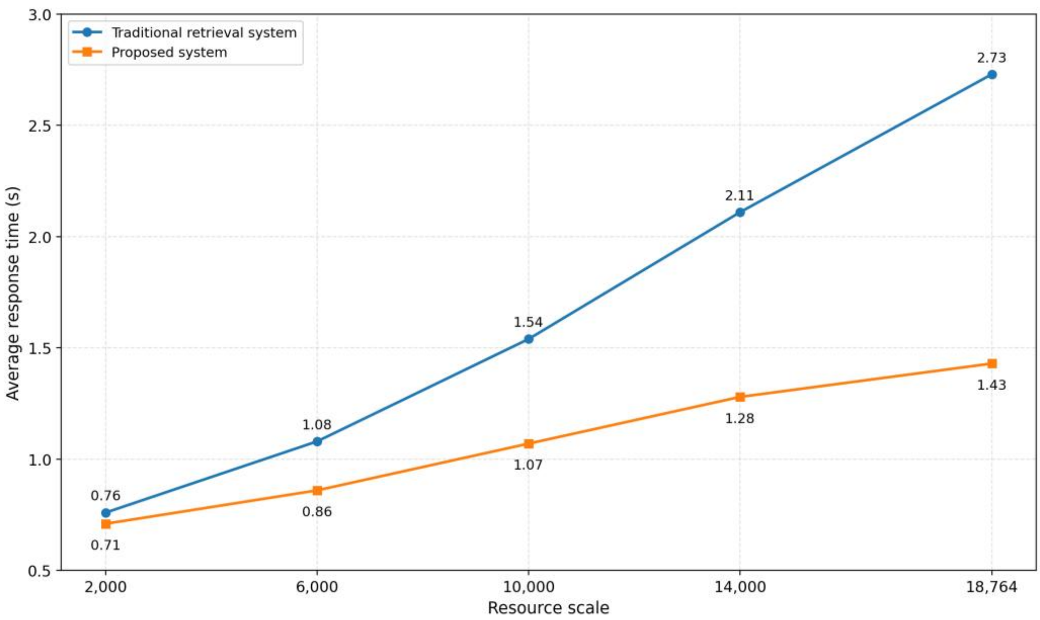

Figure 8: Average response time variation of the two types of systems under different resource scales

The trend reflected in Figure 8 shows that although the system in this paper takes on additional operations such as user permission verification, candidate recall, personalized sorting and feedback writeback, there is no obvious lag due to the increase of service chain. The reason is that the system separates resource organization, recommendation call and log writing into different processing layers, and uses cache to save high-frequency tag index and recent recommendation results, so that a large number of repeated queries do not have to repeatedly penetrate the database. In other words, the introduction of the recommendation function does not weaken the operating efficiency of the resource management system, but shows better stability when the resource scale is extended.

In the concurrent access scenario, this paper further selects 200 concurrent user conditions to compare the core performance indicators of the two types of systems, and the results are shown in Table 5. It can be seen that the average response time of the proposed system is 1.36 s, which is 1.05 s shorter than the traditional system. P95 response time is reduced from 3.62 s to 2.07 s, indicating that the delay fluctuation of tail requests is better controlled. The throughput is increased from 36.8req /s to 58.4req /s, and the request success rate is increased from 96.9% to 99.1%. More notably, the recommended refresh delay of the proposed system is only 6.8s, while the traditional system reaches 18.6s. This means that when the learner's new click, completion or error correction behavior is written back to the platform, the system in this paper can complete the status update and adjust the subsequent recommendation results in a shorter time, and the "dynamic" of the system is verified directly.

Table 5: Performance comparison of the two types of systems under the condition of 200 concurrent users

System	Average Response Time (s)	P95 Response Time (s)	Throughput ($req \cdot s^{-1}$)	Request Success Rate (%)	Recommendation Refresh Latency (s)
Traditional retrieval-based resource management system	2.41	3.62	36.8	96.9	18.6
Proposed system	1.36	2.07	58.4	99.1	6.8

The performance improvement shown in Table 5 is not just a superficial speedup from caching, but is directly related to the architectural design of the system. On the one hand, the resource management layer merges the metadata of skill tags, difficulty levels and topics in advance, so that the initial screening of candidate resources does not need to rely on full table scanning. On the other hand, the recommendation service layer uses the combination of phased recall and lightweight sorting to reduce the repeated calculation under high concurrency conditions. At the same time, the asynchronous queue mechanism was used to process the learning behavior log and the recommendation parameter update in time, so as to avoid the online response being blocked by the background update task. Because of this, the system can still maintain a high success rate when the resource scale expands and the access load rises, which is mutually confirmed with the smooth change of the response curve in Figure 8.

5 Discussion

Focusing on the problem of English self-regulated learning recommendation driven by learning behavior data, this paper constructs a technical solution consisting of behavior feature representation, recommendation model design and resource integrated management system. Its effectiveness is tested through recommendation effect experiments and system operation verification. The experimental results show that the proposed method is superior to the control model in indicators such as Precision@5, Recall@5, NDCG@5 and MRR, indicating that the system can more accurately identify learners 'phasic needs after jointly incorporating learning behavior sequence, skill mastery state and self-regulation characteristics into the modeling. Instead of staying at the general interest matching level. Compared with the recommendation ideas that focus on static portraits or single click history in related studies, this shows a stronger process adaptation ability. From the perspective of system implementation, the comprehensive recommendation system for English learning resources designed in this paper still maintains relatively stable response performance under the condition of resource scale expansion and concurrent access, which indicates that there is an effective collaborative relationship among resource management layer, recommendation service layer and feedback update mechanism. Compared with the platform with only resource retrieval function, the system not only shorted the path of resource acquisition, but also strengthened the closed-loop connection between recommendation results and real learning behavior, making personalized recommendation no longer just a front-end display function, but a sustainable updating learning support mechanism. However, there are still some limitations in this paper. First, the experimental samples are mainly from a single college English learning platform, and the learning task types and user structures are relatively concentrated, so the cross-platform transfer ability of the model still needs to be further tested. Second, the current recommendation goals mainly focus on click, completion and short-term learning response, and the tracking of long-term ability growth and policy transfer is still insufficient. The follow-up research can combine the long-term learning archive, knowledge graph modeling and reinforcement learning scheduling mechanism to further improve the continuous intervention ability and explanation depth of the English autonomous learning recommendation system.

6 Conclusions

This paper proposes an English self-regulated learning recommendation model based on learning behavior data to solve the problems of decentralized resource supply, dynamic change of learning needs, and the disconnection between recommendation results and learning process in the English self-regulated learning scenario. Starting from the platform log, the behaviors such as login, browsing, practice, error correction, review and plan adjustment were organized uniformly, and the feature representation method that could reflect the learning engagement, skill mastery and adjustment status was constructed. On this basis, a recommendation model for English learning tasks is further designed, which connects learning state recognition, resource candidate generation, relevance matching and ranking output. At the same time, the comprehensive management of English learning resources and the realization of personalized recommendation system are completed, so that the model results can form a service process that can be called, updated and fed back in the platform. The experimental results show that the joint modeling based on learning behavior sequence and self-regulation features can effectively improve the accuracy and ranking quality of English learning resource recommendation. Compared with the control methods, the proposed model achieves better performance on indicators such as Precision@5, Recall@5, NDCG@5 and MRR, indicating that the proposed method can more accurately capture the learning needs of learners at different stages. The system experiment also shows that the comprehensive recommendation system still maintains good response efficiency and operation stability under the condition of resource scale growth and concurrent access, indicating that the method proposed in this paper is not only feasible at the algorithm level, but also has the application value of platform landing. In general, this paper has completed a complete research chain from learning behavior data acquisition, feature modeling, recommendation decision and system implementation, which proves that learning behavior data can be used as an important basis for personalized support of English autonomous learning, and also shows that the combination of self-regulated learning mechanism and recommendation technology can provide more targeted and continuous resource service solutions for English learning platform.

References

[1] Yu B. Self-regulated learning: A key factor in the effectiveness of online learning for second language learners[J]. Frontiers in psychology, 2023, 13: 1051349.

[2] Yang Y, Wen Y, Song Y. A systematic review of technology-enhanced self-regulated language learning[J]. Educational Technology & Society, 2023, 26(1): 31-44.

[3] Mazandarani O. Self-regulated learning in ESL/EFL contexts: a methodological exploration[J]. Humanities and Social Sciences Communications, 2024, 11(1): 1-11.

[4] Xu J, Li J, Yang J. Self-regulated learning strategies, self-efficacy, and learning engagement of EFL students in smart classrooms: A structural equation modeling analysis[J]. System, 2024, 125: 103451.

[5] Chen J. Effects of learning analytics-based feedback on students' self-regulated learning and academic achievement in a blended EFL course[J]. System, 2024, 124: 103388.

[6] Xia Q, Chiu T K F, Chai C S, et al. The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot[J]. British Journal of Educational Technology, 2023, 54(4): 967-986.

[7] Xia Q, Chiu T K F, Chai C S. The moderating effects of gender and need satisfaction on self-regulated learning through Artificial Intelligence (AI)[J]. Education and Information Technologies, 2023, 28(7): 8691-8713.

[8] Ye D, Pennisi S. Using trace data to enhance Students' self-regulation: A learning analytics perspective[J]. The Internet and Higher Education, 2022, 54: 100855.

[9] Lim L, Bannert M, van der Graaf J, et al. Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning[J]. Computers in Human Behavior, 2023, 139: 107547.

[10] Sun J C Y, Liu Y, Lin X, et al. Temporal learning analytics to explore traces of self-regulated learning behaviors and their associations with learning performance, cognitive load, and student engagement in an asynchronous online course[J]. Frontiers in Psychology, 2023, 13: 1096337.

[11] Alhazbi S, Al-ali A, Tabassum A, et al. Using learning analytics to measure self-regulated learning: A systematic review of empirical studies in higher education[J]. Journal of Computer Assisted Learning, 2024, 40(4): 1658-1674.

[12] De Vreugd L, Van Leeuwen A, Jansen R, et al. Learning Analytics Dashboard Design and Evaluation to Support Student Self-Regulation of Study Behaviour[J]. Journal of Learning Analytics, 2024, 11(3): 249-262.

[13] Xu Z, Zhao Y, Zhang B, et al. A meta-analysis of the efficacy of self-regulated learning interventions on academic achievement in online and blended environments in K-12 and higher education[J]. Behaviour & Information Technology, 2023, 42(16): 2911-2931.

[14] Banihashem S K, Noroozi O, Van Ginkel S, et al. A systematic review of the role of learning analytics in enhancing feedback practices in higher education[J]. Educational Research Review, 2022, 37: 100489.

[15] Khor E T, K M. A systematic review of the role of learning analytics in supporting personalized learning[J]. Education sciences, 2023, 14(1): 51.

[16] Lampropoulos G. Recommender systems in education: A literature review and bibliometric analysis[J]. Advances in Mobile Learning Educational Research, 2023, 3(2): 829-850.

[17] Algarni S, Sheldon F. Systematic review of recommendation systems for course selection[J]. Machine Learning and Knowledge Extraction, 2023, 5(2): 560-596.

[18] Narimani A, Barberà E. Extracting course features and learner profiling for course recommendation systems: a comprehensive literature review[J]. The International Review of Research in Open and Distributed Learning, 2024, 25(1): 197-225.

[19] Ma Y, Ouyang R, Long X, et al. DORIS: Personalized course recommendation system based on deep learning[J]. PLoS One, 2023, 18(6): e0284687.

[20] Yuanfei Z. A personalized recommendation system for English Teaching resources based on learning behavior detection[J]. Mobile Information Systems, 2022, 2022(1): 4531867.

[21] Zheng T, Ding M. Behavior-aware english reading article recommendation system using online distilled deep q-learning[J]. Journal of Cases on Information Technology (JCIT), 2023, 25(1): 1-21.

[22] Xian H, Wu S. Design of an English vocabulary e-learning recommendation system based on word bag model and recurrent neural network algorithm[J]. Entertainment Computing, 2024, 50: 100692.

[23] Yan K. Optimizing an English text reading recommendation model by integrating collaborative filtering algorithm and FastText classification method[J]. Heliyon, 2024, 10(9).

[24] Zhang L. Data mining and analysis of learning behavior on English online teaching platforms[C]//Proceedings of the 2024 International Conference on Intelligent

Education and Computer Technology. 2024: 370-376.