International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-19

https://doi.org/10.70917/ijcisim-2025-0200
Article

Exploration and Application of Personalized Education Management Model Driven by AI Technology

Shasha Chen *

Sichuan University of Media and Communications, Chengdu, Sichuan, 610000, China; lisa_duoduo2025@163.com

Abstract: Students' learning behavior hides a large number of personalized laws that can't be identified by human beings, which need to be mined and utilized with the help of AI technology. And with the development of AI technology and the importance of national planning, the in-depth application of AI technology to the education management system to promote the personalization and intelligence of education management has become one of the current and future research priorities. This paper conforms to the development trend of the times and constructs a knowledge tracking model of the dual-head attention mechanism of the long and short-term memory network, which effectively mines and predicts the students' historical learning behaviors and current knowledge state data. After obtaining students' knowledge mastery through knowledge tracking, the CIFW personalized resource recommendation model is further used to achieve targeted resource pushing for teachers and students according to a certain cycle frequency and presentation method. After the knowledge tracking model is introduced into the 3 technology modules, the prediction correctness in the 2 datasets reaches 0.8254 and 0.8239. The HR and NDCG values of the resource recommendation model can reach up to 0.9419 and 0.6753, and the resource recommendation effect is good.

Keywords: AI technology; knowledge tracking; knowledge state prediction; CIFW resource recommendation; personalized education management

1. Introduction

In today's digital era, AI technology is penetrating into various fields at an astonishing speed, and the education field is no exception [1-2]. Education management, as an important part of ensuring education quality and improving education efficiency, how to achieve innovation and improvement with the help of AI technology has become an important topic worthy of in-depth discussion.AI technology brings unprecedented opportunities for education management, which can realize the deep mining and analysis of education data [3-4]. Education management involves a large amount of data, including students' academic performance, classroom performance, behavior, etc. In the past, the processing and analysis of these data often relied on manual labor, which is not only inefficient but also prone to bias [5-7]. In contrast, AI technology can quickly process and analyze huge amounts of data, providing accurate and comprehensive information for education administrators to help them better understand students' learning status and needs, so as to develop more scientific and reasonable education management strategies [8-11].

In addition, AI provides technical support for personalized education management. Each student has unique learning styles and needs, and the traditional “one-size-fits-all” education management model often fails to meet students' individual differences [12-14]. AI can build personalized learning models for each student based on their learning data and behavioral characteristics, and provide them with customized learning plans and tutoring programs [15-16]. More challenging learning tasks can be provided for students with faster learning progress, and more support and tutoring can be provided for students with learning difficulties. Xu [17] examined the application of AI optimization algorithms in higher education management and personalized teaching and emphasized that AI optimization algorithms effectively solved complex education management problems and provided personalized learning experiences. Mourtzis et al. [18] proposed a hybrid model under the concept of Instructional Factory framework, which, and successfully validated the feasibility of the model to effectively extend the instructional dimensions. Hu [19] aimed to assess the impact of AI-assisted personalized learning on students' learning outcomes based on a meta-analysis of the literature review, and the results pointed out that AI-assisted personalized learning has a positive impact on students' learning outcomes in terms of knowledge and competence. Hardaker and Glenn [20] aimed to identify the macro and micro level enablers of AI adoption in higher education institutions, emphasizing this “adoption” as a factor influencing the acceptance and use of AI as a personalized learning tool. Silva [21] analyzed the role of AI in creating personalized learning environments with the aim of evaluating the effectiveness of AI in improving learner engagement and outcomes through customized instructional approaches, and the results highlighted that personalized learning powered by AI significantly improves students' motivation and academic performance. Neuenschwander and Chiodo [22] illustrated the use of AI in education to help improve the quality of traditional curriculum, facilitate personalized tutoring, and help teachers to carry out educational administration more efficiently. Dembe [23] explored the key role of Educational Artificial Intelligence (EAI) in the advancement of personalized learning, emphasizing that by leveraging learning science and AI insights, EAI can personalize teaching, learning, and management. Igbokwe [24] highlighted the ability of AI in educational management to optimize the learning process, improve student performance, and streamline administrative tasks, but also challenges such as ethical issues, potential bias, and so on.

AI technology can also play an important role in the allocation of educational resources. By comprehensively evaluating and analyzing the school's educational resources, including faculty, teaching facilities, curriculum, etc., AI can help educational administrators to allocate resources more rationally and ensure that the resources are maximized [25-29]. Ding [30] proposed AI teaching resource recommendation and optimal resource allocation scheme in basic education, aiming to solve the problem of low accuracy of existing recommendation methods, and experiments proved that the proposed method can provide more accurate recommended content. Li et al [31] examined the application of machine learning technology in resource allocation and management in universities, and demonstrated the machine learning model's prediction accuracy and resource optimization effectiveness in terms of prediction accuracy and resource optimization. Wang et al [32] described how in the smart era, teaching modes are more flexible and teaching employs multiple intelligence systems, in which the acquisition of knowledge in the form of personalized education has become mainstream. Wu and Luo [33] examined the strategic integration of AI-powered chatbot technologies, especially optimal selection and resource allocation, to maximize educational effectiveness. In addition, AI technology plays an important role in improving the efficiency and accuracy of education management, automating administrative transactions such as student registration, course selection, and grade entry not only saves a lot of labor and time, but also reduces human errors [34-38].

The key to the in-depth reform of personalized education management lies in the use of AI technology to provide high-quality data and resource support from the perspective of individual student differentiation. This study focuses on the knowledge tracking model of the two-headed attention mechanism of the long and short-term memory network and the personalized resource recommendation model of CIFW based on the application of AI technology, with regard to the level of learning data processing and prediction performance of the students, as well as whether it is able to provide learning resources with greater differentiation. In the design of the knowledge tracking model of the two-headed attention mechanism of the long and short-term memory network, the actual visible learning behavior data of the students are projectively embedded with the skeleton network LSTM. Complemented by the long and short-term memory network and attention mechanism to mine the hidden changes in students' knowledge state under the surface data, to realize the overall prediction judgment from historical data-current knowledge state-future learning needs. In the CIFW personalized resource recommendation model, two technical modules, channel attention mechanism and bilinear feature interaction, are integrated to compute the matching degree between the recommended resources and the predicted students' needs layer by layer, so as to efficiently provide personalized learning resources.

2. Analysis of the way AI Technology is Used in Education

2.1. Problem Description and definition

Knowledge tracking aims to accurately predict and track learners' knowledge status with the help of AI technology tools to achieve personalized learning management and education services. By continuously optimizing the learning path and providing personalized learning support, it can help learners acquire knowledge more efficiently and improve their learning performance and learning experience. The knowledge tracking task consists of two main parts: 1) by sequential modeling of learners' learning history behaviors; and 2) by predicting learners' relevant performance in future learning tasks.

Specifically, knowledge tracking is mainly used to model the user's knowledge competence level by considering the user's historical interaction sequences through the knowledge tracking approach, where each user's interaction behavior sequence is represented as $H = \{(qi, ai) \mid i = 1, 2, 3, \ldots, m\}$, where $qi \in M$ denotes the knowledge points in the learning interaction behavior, M denotes the total number of knowledge points, $ai \in \{0,1\}$, when $ai = 0.0$, it means that there is no interaction behavior with the knowledge point when learning to answer the question; when $ai = 1.0$, it means that there is a corresponding interaction behavior with the knowledge point. Through supervised learning, the model is trained using students' historical learning sequences, so that it can predict students' current knowledge ability level and learning needs in the future. By analyzing students' answers and information about whether they answered correctly or incorrectly, the model is able to learn the state of students' knowledge and dynamically track changes in students' behavior. Ultimately, the model can provide educators with valuable reference information to help them better understand the learning status of students and provide targeted educational services according to students' needs.

2.2. Overall Architecture of the Knowledge Tracking Model

The knowledge tracking model with two-head attention mechanism for long and short-term memory network proposed in this paper is a model that uses LSTM network as a skeleton and defines an attention module for knowledge hidden state for modeling and prediction of students' knowledge state, and Fig. 1 shows the model architecture. The original embedded representation of the data is first performed using the word2vector method, then the attention module of the knowledge-hidden state is set up using LSTM and the attention mechanism, the embedded data is learned and trained, and finally the corresponding prediction is performed.

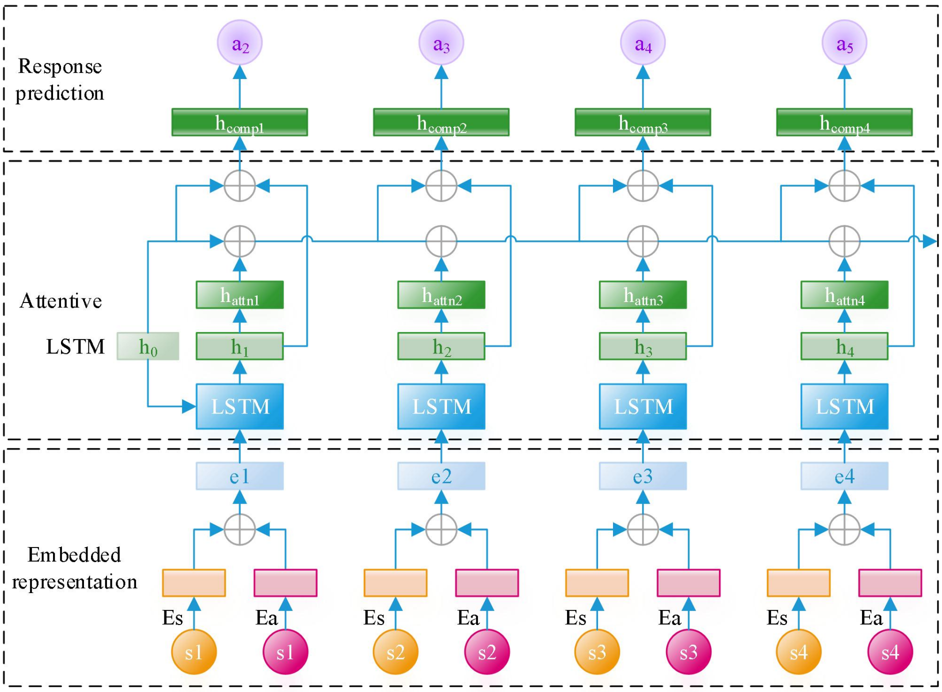


Figure 1. Knowledge tracking model architecture.

2.2.1. Data Embedding Representation

The original knowledge points and the user's interaction behavior indexes on the knowledge points are projected into the embedding space using the LSTM model. Assume that the knowledge point and response embedding matrices are represented as $E_s \in R^{|S| \times ds}$ and $E_a \in R^{2 \times da}$, where d_s and d_a

denote the corresponding embedding dimensions, and R denotes the embedding space (also known as the representation space). The E_s denotes that each row represents a particular knowledge point, while the two rows in E_a represent wrong and right answers, respectively. Given the interaction of a certain student user s_j with a knowledge point a_j, by finding E_s and E_a, their corresponding embedding representations $e_j^s \in R^{1\times ds}$ and $e_j^a \in R^{1\times da}$. In order to obtain a complete representation of the student interaction and to emphasize the difference between correct and incorrect responses, the following definitions are made:

$$e_j = \begin{cases} e_j^s \oplus e_j^a & if\ a_j = 1.0 \\ e_j^a \oplus e_j^s & if\ a_j = 0.0 \end{cases} \tag{1}$$

where $\oplus$ denotes the connectivity operation, which is related to the order of operations of the arithmetic formula, with different order before and after resulting in different outputs, and $e_j \in R^{1\times(ds+da)}$ denotes the complete interaction behavior of the student.

2.2.2. Long and short-term memory networks and attentional mechanisms

The proposed long and short-term memory network model and attention mechanism serve as the skeleton of knowledge tracking, which consists of knowledge hidden state modeling and knowledge hidden state attention mechanism modules.

1) Knowledge-hidden state modeling: from the complete sequence of students' interaction behaviors $\{e_1, e_2, e_3, \ldots, e_t\}$ defined above, an attempt is made to capture the knowledge-hidden state t at each time $h_t \in R^{d_h \times 1}$, d_h denotes the dimension of h_t. Based on previous approaches to knowledge tracking, a conventional RNN temporal modeling approach is used to obtain the corresponding hidden state at each step. In this study, the Long Short-Term Memory Network Model (LSTM) will be used to update the hidden state h_t cyclically by obtaining the current e_t and the previous h_{t-1}, described as follows:

$$h_t = \mathrm{LSTM}(e_t, h_{t-1}) \tag{2}$$

When $t = 1$, h_{t-1} is equivalent to h_0, which is referred to as the initial state of the hidden state.

2) Knowledge Hidden State Attention Mechanism Module: using the updated h_{t-1} it is possible to directly predict the probability of learning user's correct knowledge of the next knowledge point. However, since the learning user's interaction at each time point has a different degree of influence on the h_{t-1} update. Based on this factor consideration, the influence of weight coefficients of each interaction on the hidden sequence $\{h_1, h_2, h_3, \ldots, h_t\}$ is obtained in real time by using the attention mechanism, which can be described as:

$$u_i = \tanh(W_w h_i + b_w) \tag{3}$$

$$\alpha_t = \frac{\exp(u_i^T u_w)}{\sum_t \exp(u_i^T u_w)} \tag{4}$$

where $\exp()$ is an exponential function with a natural constant (e) base, denoting the exponential part of $\mathrm{softmax}()$ used to compute the exponential part. Specifically, for a node h_i, it is first fed into a multilayer perceptual machine (MLP) that contains a learned weight matrix $W_w \in R^{d_w \times d_h}$ and a bias $b_w \in R^{d_w \times 1}$, a hidden representation vector u_i is output. Then, the similarity between this hidden representation vector u_i and the hidden representation vector u_w of the reference node is computed by the click operation, and after $softmax()$, the importance coefficient α_t can be generated. Since the previous knowledge-hiding state tensors $\{h_1, h_2, h_3, \ldots, h_{t-2}\}$ will have a different impact on the current h_{t-1} prediction, define their total impact as follows:

$$h_{atmj} = \alpha_j h_j \tag{5}$$

$$\hat{h}_{t-2} = \sum_{j=1}^{t-2} h_{atmj} \tag{6}$$

In addition, the importance of h_{t-1} must be emphasized here because it implicitly represents the current knowledge mastery, i.e., the level of knowledge competence, of the learning user and is closely related to the prediction of the next state. Therefore, h_{t-2} and h_{t-1} are used as the final prediction representation after a concatenation operation in the research process, described as follows:

$$h_{comt_{-1}} = \hat{h}_{t-2} \oplus h_{t-1} \tag{7}$$

2.2.3. Response prediction

This part is used to predict the learning user's response to the next knowledge point linkage through the composite representation $h_{com_{t-1}}$. First, $h_{comt_{-1}}$ is processed by a multilayer perceptual machine (MLP), and then $sigmoid()$ produces the predicted probability $\hat{a}_t \in [0.0, 1.0]$. Next, a binary cross-entropy loss function is used to compute the loss between the predicted probability $\hat{a}_t$ and the true response a_t, defined as follows:

$$Loss(\hat{a}_t, a_t) = -\left(a_t \log \hat{a}_t + (1 - a_t)\log(1 - \hat{a}_t)\right) \tag{8}$$

Overall, the Long and Short-Term Memory Network Attention Mechanism Knowledge Tracking Model is able to effectively model students' learning history and make predictions about future learning performance based on it by using the Long and Short-Term Memory Network as the overall skeleton. Meanwhile, through the introduction of the attention mechanism, the model is able to better understand the important patterns and key information in the sequence of students' behaviors, which improves the ability to accurately model and predict students' knowledge status.

2.3. Resource recommendation modeling

With the development of personalized smart education, the exponential growth of user information and interaction information makes traditional recommendation algorithms unable to make recommendations efficiently when faced with large amounts of sparse data. More and more companies adopt deep learning-based recommendation algorithms to replace traditional recommendation algorithms. And there are various types of problems in the commonly used deep learning algorithms, such as Wide&Deep does not consider explicit higher-order feature interactions and DCN does not use vector-level feature interactions. This chapter proposes the CIFW personalized resource recommendation model, which is based on the XDeepFM algorithm, retains the compressed interaction network for explicit feature interactions, uses the channel attention mechanism with the bilinear feature interaction module to enhance the input features, and performs implicit feature interactions through the multilayer perceptual machine (MLP). The main techniques in the CIFW model are described below.

2.3.1. Channel Attention Mechanisms

The main idea of the channel attention mechanism is to correct the original channel by establishing the relationship between the channels as a way to improve the performance of the neural network. The channel attention mechanism has three main parts of work, which are Squeeze, Excitation, and Reweight.Figure 2 shows the overall architecture of the channel attention mechanism.

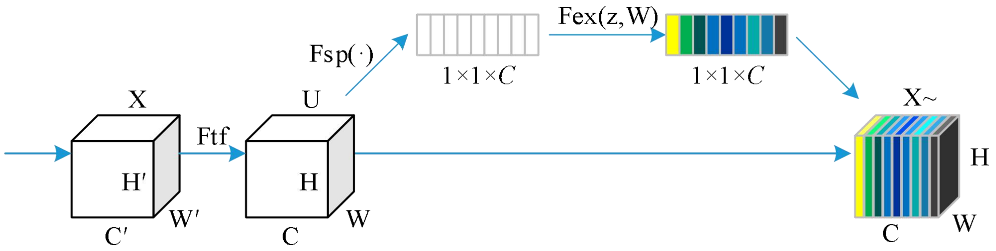


Figure 2. Channel Attention Mechanism.

The channel attention mechanism starts with U. It first does a global rating pooling operation, which is the process of Squeeze, on U. The computation of Squeeze is shown in Equation (9).

$$Z_C = F_{sq}\left(u_c\right) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \tag{9}$$

The data is compressed to a data size of $1 \times 1 \times \mathrm{C}$ after the Squeeze process, and then passes through a two-stage fully-connected network, with the activation functions chosen to be Relu and sigmoid, respectively, and the whole operation is known as the Exciation process.The Exciation process is illustrated in Equation (10).

$$s = F_{ex}(z,W) = \sigma(\mathrm{g}(z,W)) = \sigma\left(W_2\delta\left(W_l z\right)\right) \tag{10}$$

After the Exciation process, a C-dimensional vector is output and the value of the vector is the weight of the channel. The input features are multiplied with the weights and the important features are enhanced and the useless features converge to 0.0.The process of assigning weights to the inputs is known as Reweight and the Reweight process is shown in equation (11).

$$\widetilde{\mathrm{X}} = F_{scale}\left(u_c, s_c\right) = s_c \cdot u_c \tag{11}$$

2.3.2. Bilinear feature interaction

Figure 3 shows the calculation process of bilinear feature interaction, bilinear feature interaction introduces the parameter cross matrix to realize the interaction between features, in the calculation process feature i is firstly inner product with parameter matrix, and the calculation result is then hadamard product with feature j. And there are three models according to the number of parameter cross matrices:

1) Field All Type:

All cross features share one parameter matrix.

2) Field Each Type:

Each Field exclusively shares one parameter matrix.

3) Field Interaction Type

Each Filed combination enjoys one parameter matrix.

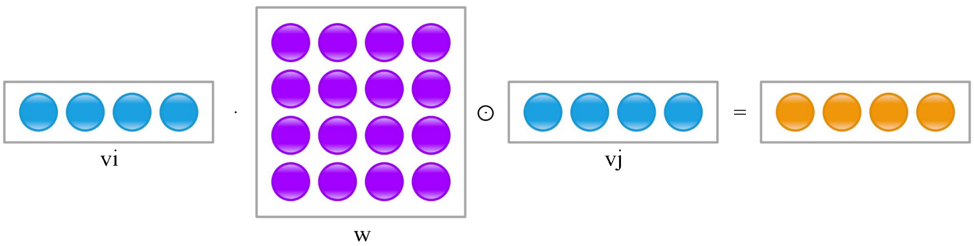


Figure 3. Process of Bilinear Feature Interaction Computation.

The bilinear eigeninteraction formula is shown in equation (12):

$$p_{ij} = v_i W \circ v_j \tag{12}$$

The bilinear eigeninteraction formula is shown in equation (12):

In this paper, the CIFW personalized resource recommendation model is based on the XDeepFM model, retains the original explicit higher-order interaction part, and adds the channel attention mechanism to the original DNN module to enhance the weight of the valid information in the input features, so as to improve the network performance. And the weight sharing technique of bilinear feature interaction module is used to alleviate the problem of feature sparsity.

1) Input Layer

The input layer serves as the first layer of the network structure, and the inputs are divided into two types, the sparse input x_{sparse} and the dense input x_{dense}, the sparse inputs are the discrete features, $x_{sparse} = \left[x_1, x_2, \ldots, x_n\right]$, n for the number of discrete features. The sparse inputs need to be processed by unique thermal encoding before they can be input into the model, while the dense inputs are numerical features that can be directly input into the model.

2) Embedding layer

After the features are encoded in the input layer, they will form high-dimensional sparse data, which cannot be directly applied to the neural network. The embedding layer can map the high-dimensional sparse features to the low-dimensional dense feature space in order to better utilize the neural network for feature processing. The recommendation algorithm model presented in this paper maps all the discrete features to the embedding layer and splices all the mapped features as the output of the embedding layer. The embedding mapping is shown in equation (13):

$$e_i = v_i x_i \tag{13}$$

where e_i is the embedding vector of the i th feature and v_i is the embedding matrix of the i th feature. The output of the embedding layer is formed by splicing the embedding vectors of the input features, $E = \left[e_1, e_2, \ldots, e_n\right]$, where n is the number of features.

3) Explicit Interaction Layer

The input of the explicit interaction layer $x_{cin} = \left[E, x_{dense}\right]$ is the output of the embedding layer spliced with the dense input, and the output of the explicit interaction layer is p^+.

4) Feature Enhancement Layer

The input of the feature enhancement layer is the output of the embedding layer. After the data enters the feature enhancement layer, it is divided into two parts according to the direction of data flow. The first part will directly input the input data into the bilinear feature interaction module and get the output cross-features. The second part will first dynamically learn the importance of features through the channel attention mechanism to get the enhanced embedding vector, and then input the enhanced embedding vector to the bilinear feature interaction module to get the output cross-feature. These two parts of the output cross-features are midlinked together and fed into the multilayer perceptron network, and the output of the multilayer perceptron network is shown in equation (14):

$$h_f = MLP\left(Concat\left(E_{bilinear1}, E_{bilinear2}\right)\right) \tag{14}$$

5) Prediction Layer

The prediction layer generates predicted values based on the output of the interaction layer, and the recommendation algorithm model used in this paper uses the sigmoid function as the activation function of the prediction layer. The predicted value is shown in equation (15):

$$y_{pred} = \sigma\left(W_{linear} x_{linear} + \left[p^+, h_f\right] W_o + b_o\right) \tag{15}$$

2.4. Personalized Learning Resources Push Process Design

In order to improve the usability of the recommendation model in frontline teaching practice, so that students can effectively access personalized learning resources, this study designed a personalized learning resources push process based on the analysis of assessment data, and established a push mechanism and specification for both the specific push cycle frequency and push presentation.

Figure 4 shows the push process of the push model in frontline teaching. First, the teacher breaks down the test into corresponding knowledge points according to the content of the test, calls the corresponding test questions in the test question resource library, and prepares the subject quiz in accordance with the requirements for quiz compilation proposed in this study; second, the subject quiz is distributed to the students to complete, and the intelligent assessment data collection system is used to obtain the student assessment data and analyze the student assessment data to get the learner's knowledge structure and level of comprehension; and then, the personalized learning resources push model based on assessment data analysis is used to push personalized learning resources for learners; finally, students' assessment data are collected iteratively, analyzed, processed and pushed in a new round. In the push implementation process, researchers and teachers need to design and control the cycle frequency and presentation of pushing personalized resources.

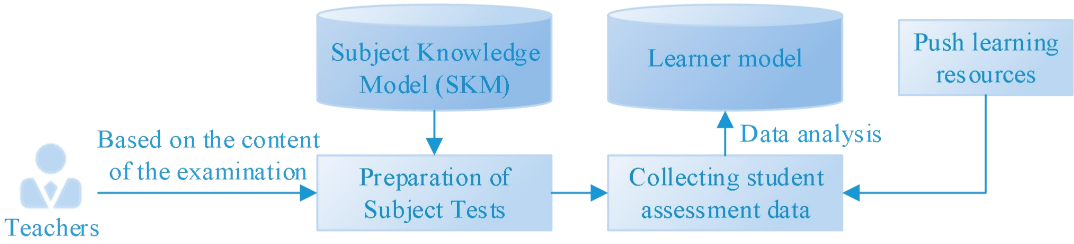


Figure 4. Personalized Learning Resource Push Based on Assessment Data Analysis.

2.4.1. Push resource cycle frequency

The frequency of personalized learning resources push cycle is determined by the instructor according to the teaching needs, and this study suggests that the instructor should provide prepared subject quizzes after each new lesson, input the quiz data into the push system, and then the push system will automatically provide personalized learning resources for the learners; as for other nature of the courses, such as review classes and exercises, etc., the instructor can decide whether or not to provide personalized learning resources for the students according to the teaching needs. Personalized learning resources.

Although personalized learning resources delivery with the support of computer technology has become possible, in the field of basic education, the penetration rate of computers and mobile devices is still relatively low, and most of the schools and parents do not support students to use electronic devices to carry out learning, which creates a certain obstacle to our personalized learning resources delivery. Therefore, in the actual teaching process, the cycle frequency of personalized learning resources is determined by the cycle frequency of student assessment data collection. According to the teaching needs, teachers prepare the subject quizzes in the system before class, print them into paper copies, and distribute the quizzes to students after class. After collecting students' assessment data, the data is input into the push system, and the personalized learning resources output from the system are provided to students within 8h.

2.4.2. Pushing the presentation of resources

To ensure that learners have access to personalized learning resources, this study provides learners with both electronic and paper presentations. If students have electronic devices, they can access the electronic version of the personalized learning resources directly on their devices and get quick feedback on their consolidation exercises. If the students are in boarding schools and do not have access to electronic learning resources, the teacher prints out a paper version of each learner's personalized learning resources, distributes it to the students, and collects the assessment data from the consolidation exercise section after they have completed it to iteratively update the learner model. And then they would distribute each student's answer parses for them to revise and refine on their own.

3. Model Performance Validation and Personalized Education Management Model Exploration

3.1. Knowledge tracking model performance test and effect analysis

3.1.1. Experimental Analysis of Knowledge Tracking Model Ablation

In order to gain a deeper understanding of the differences in the performance of the model proposed in this paper between different sections and to prove the effectiveness of the key components of the model, this paper conducts several ablation experiments to dig out the contribution of different model modules to the effectiveness of the model. Using the historical test data of first-year students from a university in the past 2 years as the research data, we construct the question-answering datasets Dataset 1 and Dataset 2, and complete the ablation experiments in this section on these 2 datasets. Table 1 shows the results of multiple ablation experiments. “1” indicates that the module is involved in the computation of the model, and “0” indicates that the module is not involved in the computation of the model. When all three modules are not involved in the computation, the model has an AUC value of 0.8073 and a correctness rate of 0.8004 in the Dataset 1 dataset, and an AUC value of 0.8016 and a correctness rate of 0.8031 in the Dataset 2 dataset.When 1 of the modules is added, the model's AUC and ACC values in the 2 datasets are improved. When 2 modules are added, the values of the model are further improved by adding 1 model. With the addition of 3 modules, the model achieves an AUC value of 0.8261 and a correct rate of 0.8254 on the Dataset 1 dataset, and an AUC value of 0.8212 and a correct rate of 0.8239 on the Dataset 2 dataset.In comparison, the model of this paper with the addition of 3 modules has the highest AUC value and correct rate ACC on the 2 datasets. Through the ablation experiments, it can be proved that the knowledge tracking model of the two-headed attention mechanism of the long and short-term memory network designed in this paper has obvious performance improvement on the basis of the initial model, and has the potential for in-depth application.

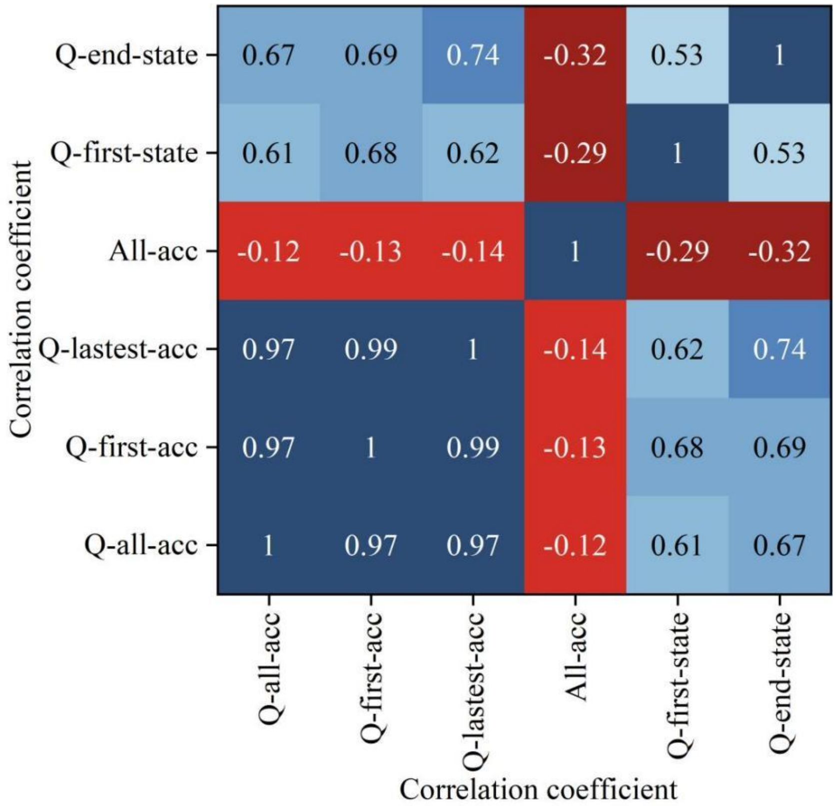
Table 1. Results of multiple melting experiments.

Model			Dataset 1		Dataset 2	
Data embedding representation	Long Short-Term Memory Network	Attention mechanism	AUC	ACC	AUC	ACC
0	0	0	0.8073	0.8004	0.8016	0.8031
0	0	1	0.8154	0.8036	0.8137	0.8112
1	1	0	0.8192	0.8197	0.8191	0.8199
0	1	0	0.8165	0.8163	0.8162	0.8168
1	0	0	0.8148	0.8146	0.8142	0.8144
1	1	1	0.8261	0.8254	0.8212	0.8239

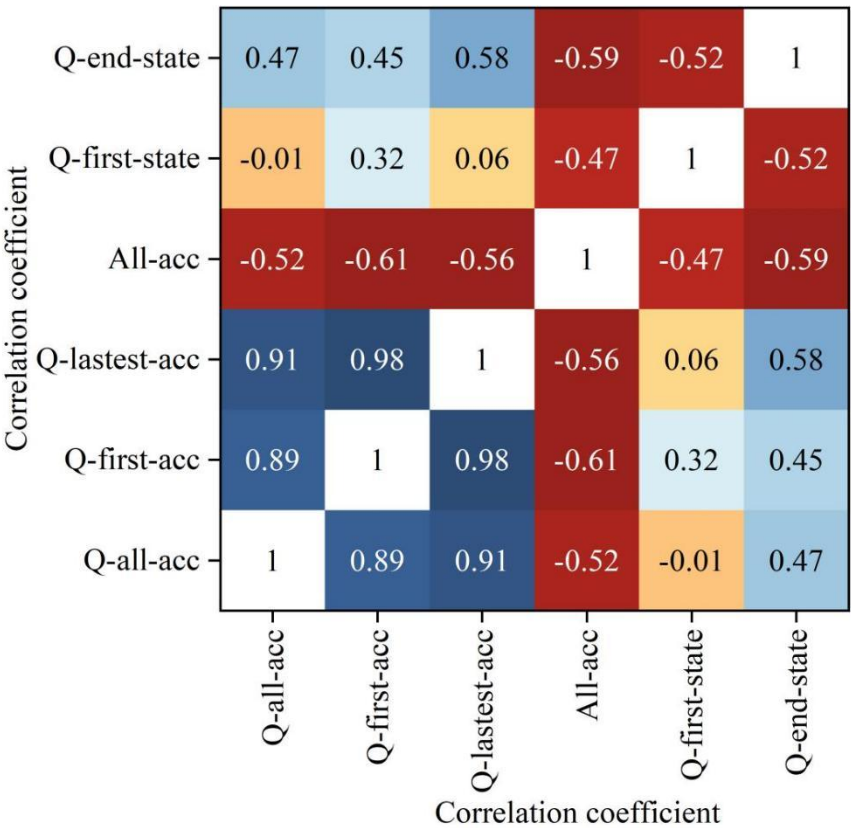
3.1.2. Exploration of Knowledge Tracking Model Fitting Effectiveness

After examining the effect of model performance improvement, in order to understand the effect of the model in generating historical learning ability and predicting the current knowledge state (i.e., the students' forgetting level of a certain knowledge point), this section utilizes the Dataset 1 and Dataset 2 datasets as the comparison datasets, and this section compares the relevant features of historical learning ability generation and current knowledge state prediction, and the heat map is made for the two datasets by calculating the correlation of each learning feature with the other variables. variable correlation between them to do the heat map of the two datasets, and the same method is used to compare the relevant features of current knowledge state prediction in the two datasets.

Figure 5 shows the performance of the model in this paper to generate history learning capability on 2 datasets. On Dataset 1, the correlation between Q-lastest-acc and Q-first-acc is 0.99, indicating that the correlation between the two is very high on this dataset; in Dataset 2, the correlation between the two correlation coefficients is 0.98. It can be seen that the correlation between the students' initial performance and the recent performance on both datasets is very high, which indicates that the students' historical learning performance is stable. In the correlation of Q-first-state and Q-first-acc variables, the correlation coefficient in Dataset 1 dataset is 0.68, and the heat map results show that there is a large correlation between the students' first answers and the first few performances in this dataset, while the correlation coefficient of the two variables is 0.32 in Dataset 2 dataset, and the correlation coefficient is higher compared with that of Dataset 1. Dataset 1 has a smaller correlation coefficient compared to Dataset 1, indicating that different dataset situations also affect the degree of correlation between students' performance on questions. In Q-first-state and Q-all-acc, i.e., first-time question performance and overall correctness of Q, the correlation between the two is stronger in the Dataset 1 dataset, with a correlation coefficient of 0.61, whereas in the Dataset 2 dataset, the correlation coefficient is only -0.01, i.e., first-time question performance does not affect correctness. The model in this paper is good at processing students' historical question answering data to find the correlation between answering sequences and to analyze students' question answering.



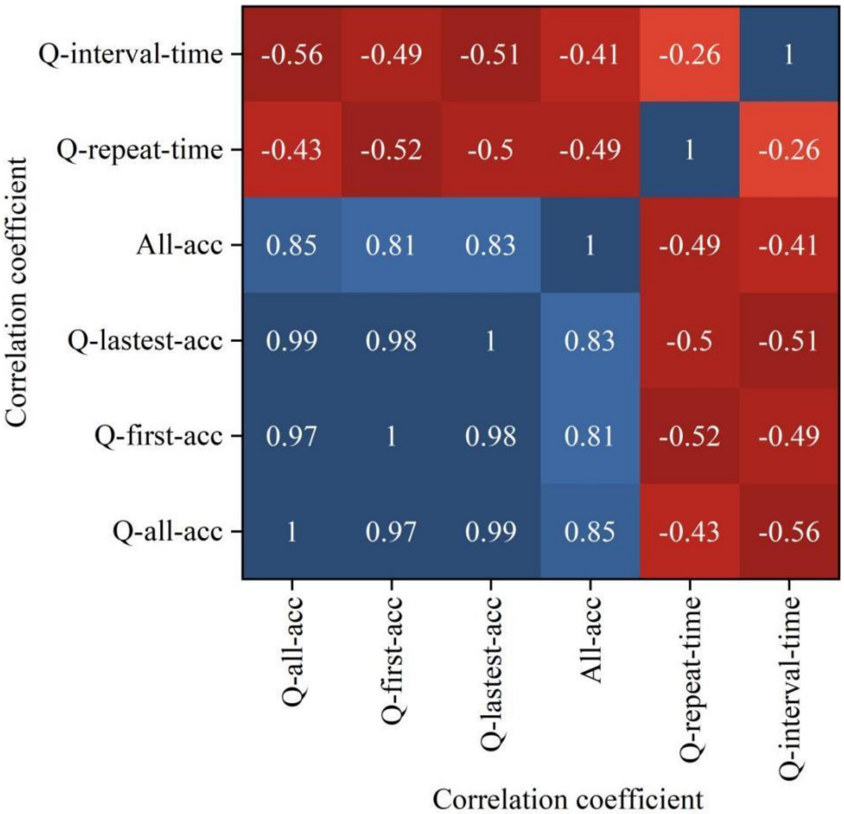
(a) Dataset 1



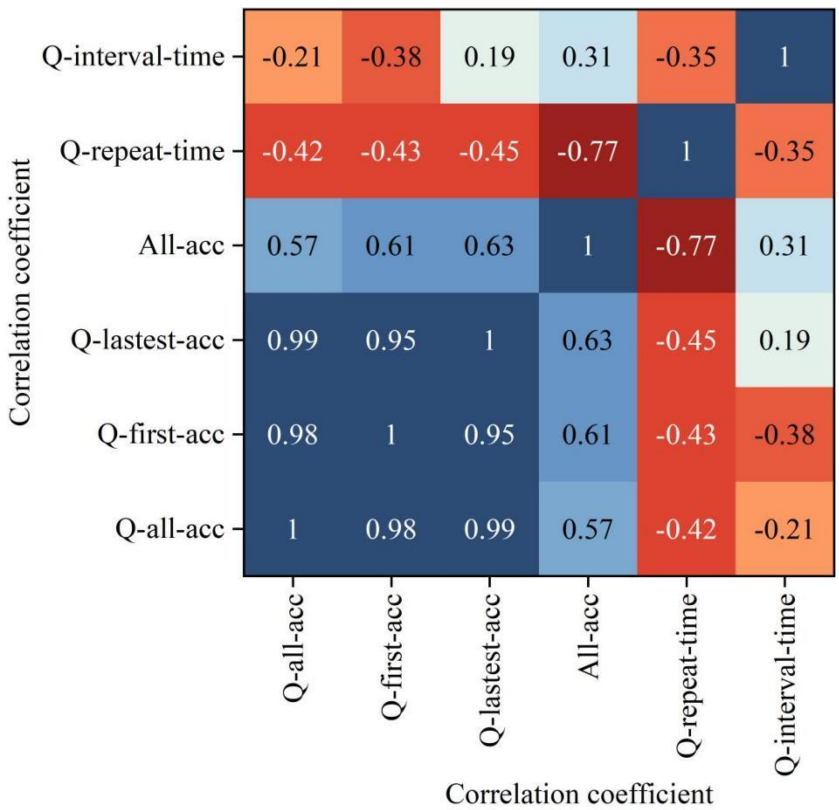
(b) Dataset 2

Figure 5. Correlation coefficient of historical learning ability.

Figure 6 shows the performance of this paper's model to predict the current knowledge state on the 2 datasets. On Dataset 1, the positive correlation of Q-lastest-acc, Q-first-acc and Q-all-acc are 0.99 and 0.97 respectively, with correlation coefficients close to 1. Q-repeat-acc, Q-interval-acc and Q-all-acc are negatively correlated with correlation coefficients of -0.43 and - 0.56.This means that the value of all-acc tends to decrease as the interval increases. In Dataset 2, Q-lastest-acc, Q-first-acc and Q-all-acc are still positively correlated, with correlation coefficients of 0.99 and 0.98 respectively, which are also close to 1. Q-repeat-acc, Q-interval-acc and Q-all-acc are also negatively correlated but the degree of negative correlation is weaker than that of Dataset 1. Dataset 1 is weaker, with correlation coefficients of -0.42 and -0.21, respectively.Overall, the heat maps of the 2 datasets of the model reveal the correlation between the answer sequences, and it can be observed that the correlation evolves with the depth of the students' answers. It also shows that the model can well predict students' current knowledge status, which provides a basis for subsequent learning resources recommendation.



(a) Dataset 1



(b) Dataset 2

Figure 6. Current knowledge state prediction correlation coefficient.

Students' level of forgetting of knowledge points is an important factor in measuring the current state of knowledge of students. Eight topics were taken from Dataset 1 and the real-time forgetting level of students was calculated using the model. Table 2 shows the change in the state of students' forgetting level. The first row shows the 8 topics practiced by the students and the first column represents the change in the students' time spent on the topics. The data in the table are the students' forgetting level of a certain knowledge point, the higher the value the greater the level of forgetting, and the bolded data indicate that the students answered the question correctly, and vice versa, they answered the question incorrectly. Through the model's learning of students' historical question-answering interaction data and the prediction of current knowledge status, we can see from the table the students' forgotten knowledge points in these eight questions. For example, the student's forgetting level for practice question 2 is 6.39 at the very beginning, and after answering the questions incorrectly at moments 1-5, the forgetting level increases to 6.45 at moment 5, and after the student answers the questions correctly consecutively at moments 6-7, the forgetting level decreases to 6.04, and after answering the questions incorrectly consecutively at moments 8-9, the forgetting level increases to 6.45, and after answering the questions correctly consecutively at moments 10-12, the forgetting level finally decreases to 6.02. The model predicts that students' knowledge of these eight questions will be reduced by the model to 6.02. The model is more accurate in predicting the students' current knowledge status, and can provide technical support for resource recommendation.

Table 2. The changes in students' levels of memory retention.

T	2	9	18	25	40	49	56	60
1	6.39	6.48	**6.05**	6.32	**6.05**	6.45	6.34	**6.07**
2	6.40	6.49	6.32	**6.04**	6.12	6.47	6.42	**6.06**
3	6.41	6.53	**6.01**	6.45	**6.04**	6.48	6.47	6.47
4	6.42	6.60	6.45	6.47	6.45	6.49	6.54	6.45
5	6.45	6.65	6.72	6.49	6.43	6.53	6.58	6.42

6	**6.05**	6.68	6.78	6.53	6.41	**6.05**	6.59	**6.05**
7	**6.04**	6.71	6.79	**6.03**	6.40	6.44	6.63	6.43
8	6.36	**6.09**	**6.00**	6.48	**6.03**	6.46	**6.04**	6.42
9	6.45	**6.07**	6.38	6.37	**6.02**	6.51	**6.02**	6.40
10	**6.04**	**6.06**	6.46	6.32	6.35	6.53	6.47	**6.04**
11	**6.03**	**6.05**	6.57	6.24	6.32	6.56	**6.01**	6.37
12	**6.02**	**6.04**	6.59	**6.02**	6.30	**6.04**	**6.00**	**6.01**

3.2. Resource recommendation model performance test and effect analysis

3.2.1. Resource recommendation model ablation experiment analysis

In order to verify the effectiveness of the key techniques of the resource recommendation model proposed in this paper, ablation experiments are conducted in this section. The 2 datasets Dataset 1 and Dataset 2 used in 3.1 are subdivided into 4 datasets, Dataset A, Dataset B, Dataset C, and Dataset D, according to the class as the division basis. The ablation experiments of the resource recommendation model are completed on these 4 datasets. Table 3 lists the exercise recommendation results of the resource recommendation model and its 2 variants on the 4 datasets. Each variant removes a key technology module and evaluates the performance. The variants of the resource recommendation model with the channel attention mechanism or bilinear feature interaction module ablated in the 4 datasets show a decreasing pattern in the 2 metrics, HR and NDCG. The CIFW resource recommendation model in this paper has a high HR value of 0.9419 and a high NDCG value of 0.6753 in the dataset Dataset D. In the other 3 datasets, the HR and NDCG values are also consistently greater than 0.9 and 0.6, respectively.The 2 modules, the channel attention mechanism and the bilinear feature interactions, have a better effect on the resource recommendation effect of the CIFW model.

Table 3. The results of the ablation experiment.

Dataset	Indicators	Model		
		CIFW	CIFW-No channel attention mechanism	CIFW-No bilinear feature interaction
Dataset A	HR	0.9198	0.7244	0.6051
	NDCG	0.6657	0.3655	0.2533
Dataset B	HR	0.9206	0.7314	0.6274
	NDCG	0.6558	0.4776	0.3634
Dataset C	HR	0.9406	0.5691	0.4503
	NDCG	0.6242	0.3047	0.1841
Dataset D	HR	0.9419	0.6298	0.4212
	NDCG	0.6753	0.2487	0.1348

3.2.2. Experimental analysis of hyperparameter determination for resource recommendation models

Since the CIFW resource recommendation model is based on the XDeepFM algorithm, and also the bilinear feature interaction module needs to realize the interaction between the features by introducing the parameter crossover matrix, and the model introduces the attention mechanism to enhance the weight of valid information in the input features. Therefore, the optimal number of parameter cross matrix nodes, number of attention heads, number of model hierarchical structures, and algorithmic parameter p(T) need to be determined so that the model's behavior and performance are in the optimal state.

Figure 7 shows the effect of the number of parameter cross matrix nodes on the model performance.

Fig. 8 shows the effect of number of attention heads on model performance. Fig. 9 is the effect of number of model hierarchical structures on model performance. Figure 10 is the effect of p(T) parameter on model performance.

When the number of nodes in the parameter crossover matrix is 6, the HR@15 values of the 4 datasets are 0.85, 0.89, 0.62, and 0.74, and the NDCG@15 values are 0.55, 0.62, 0.58, and 0.46. At this time, the model performance is the best. When the number of attention heads is 8, the HR@15 values of the 4 datasets are 0.78, 0.65, 0.72, and 0.78, and the NDCG@15 values are 0.53, 0.52, 0.45, and 0.46. At this time, the model performance is the best. When the number of model hierarchical structures is 5, the HR@15 values of the 4 datasets are 0.79, 0.82, 0.88, and 0.78, and the NDCG@15 values are 0.65, 0.54, 0.57, and 0.42. At this time, the model performance is the best. When the algorithm parameter p(T) is 0.3, the accuracy rates of the four datasets are 0.84, 0.74, 0.85, and 0.70 respectively. At this point, the model performance is the best. Therefore, the number of nodes in the parameter crossover matrix, the number of attention heads, the number of model hierarchical structures, and the algorithm parameter p(T) are respectively set to 6, 8, 5, and 0.3 to achieve the best level of resource recommendation effect and provide teachers with higher-quality recommended exercises.

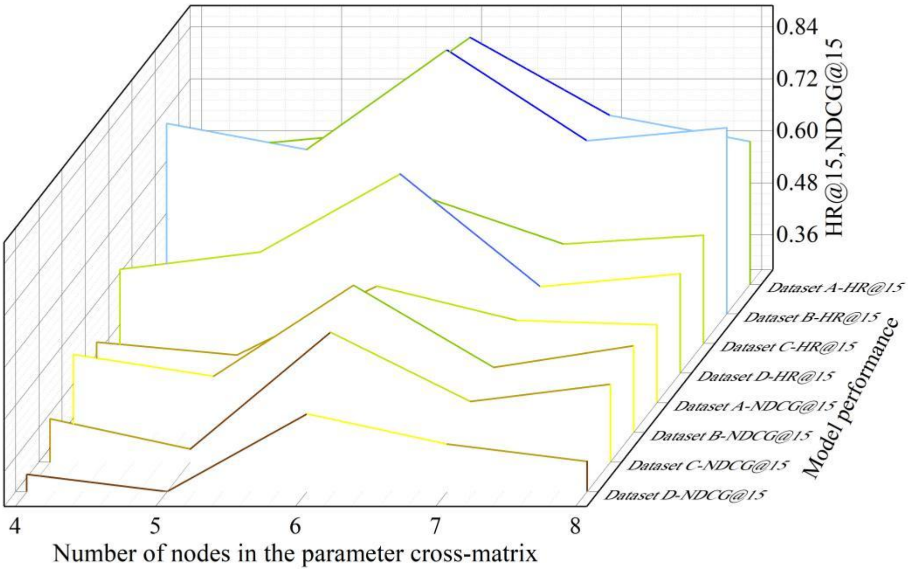


Figure 7. Influence of number of nodes in parameter cross-matrix on model.

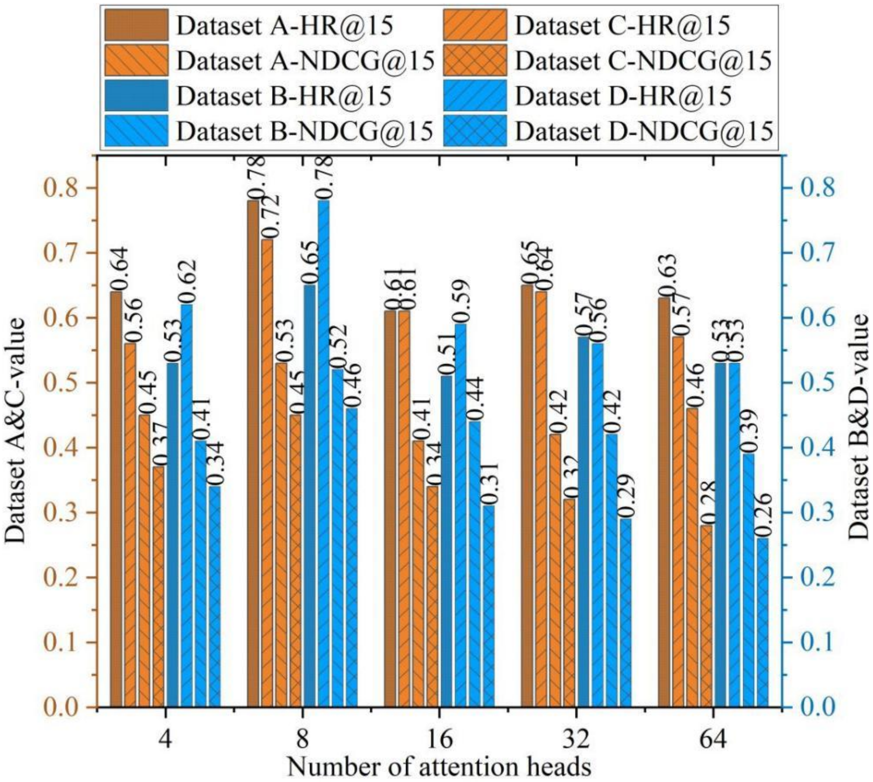


Figure 8. Impact of number of attention heads on the performance of the model.

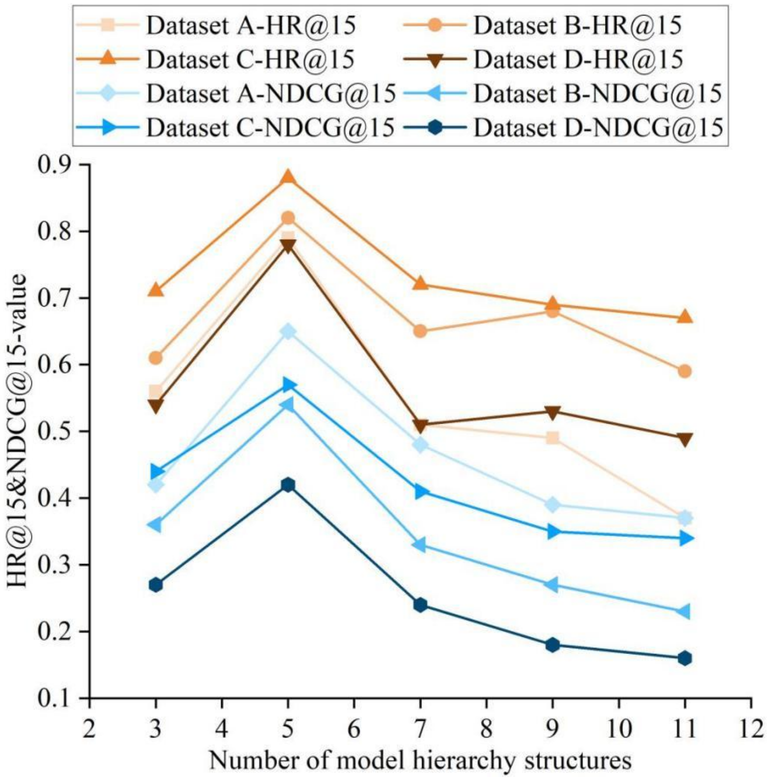


Figure 9. Influence of number of model hierarchy structures on model.

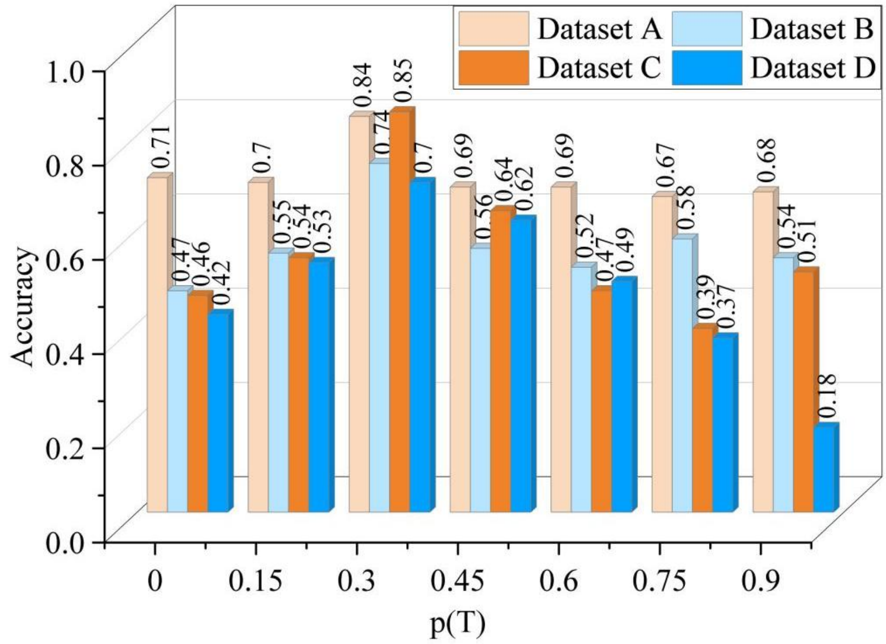


Figure 10. The influence of the p(T) parameter on the model performance.

4. Directions for the development of personalized education management in colleges and universities with AI technology support

4.1. Enhancing technology maturity and adaptability

First of all, it is necessary to commit to upgrading the maturity of the technology. The designed long and short-term memory network two-headed attention mechanism knowledge tracking model and CIFW personalized resource recommendation model, even though they have a certain application effect, still need to be upgraded continuously in order to adapt to the diversification of educational scenarios. At the same time, attention should be paid to the technological development of model-related modules as much as possible, such as the optimization and upgrading of the long and short-term memory network module, etc., so as to improve the local modules and the overall model in a timely manner, and then develop a more mature technology. On this basis, the establishment of a database specialized in students' learning characteristics should be initiated, and the model should be used for continuous learning in order to improve the generalization performance of the model.

Second, efforts should be made to improve the technology adaptability. The designers of the Long and Short-term Memory Network Dual Attention Mechanism Knowledge Tracking Model and CIFW Personalized Resource Recommendation Model need to work closely with the teachers of various majors and courses in colleges and universities, and carry out the research and development of the customized models based on the teaching objectives and requirements of different majors and courses. At the same time, a common technical framework for personalized education should be constructed and interface standards should be set up. In view of the need to promote the deep integration of AI technology with the teaching scene in colleges and universities, the relevant departments or industry associations should formulate a unified common technical framework and interface standards in advance to enhance the support value of AI technology for education management.

4.2. Enhance the ability of teachers and students to adapt to their roles

On the one hand, the adaptive capacity of teachers should be promoted. In terms of the adaptability challenges faced by teachers, universities should build a support system for the improvement of teachers' adaptability ability. In terms of technical integration ability, special training of AI education tools is carried out, such as the training and explanation of the specific application logic and application method of the knowledge tracking model of the two-headed attention mechanism of the long and short-term memory network, so as to enhance the teachers' operational ability of AI education tools; in terms of teaching reconstruction ability, a workshop on designing the “problem a competence and a knowledge” map is constructed to enhance the teachers' teaching reconstruction ability with the support of AI technology. In terms of teaching reconstruction ability, the “Question, Ability, Knowledge” mapping design workshop is constructed to improve teachers' teaching reconstruction ability with AI technology support; in terms of role change support, the teacher learning community, i.e., the interdisciplinary teaching and research community, is established to improve teachers' role change ability; and in terms of the mechanism for continuous development, the “Training, Practice, Reflection” spiral enhancement model is implemented to improve teachers' capacity for continuous development in applying AI technology.

On the other hand, we focus on cultivating students' role adaptation ability. When students encounter various problems in the process of trying to adapt to the new learning mode driven by AI technology, college administrators or teachers need to take corresponding measures to cultivate students' role-adaptive ability. Carry out students' independent inquiry training, set up graded project tasks, such as from the principle of AI knowledge prediction to the selection of AI recommended resources, in order to enhance students' independent inquiry ability in the face of the change of AI technology; enhance students' human-computer collaboration literacy, set up a course on the ethical and critical use of AI tools, and systematically teach students how to clarify their learning goals, and how to formulate a reasonable study plan according to their own actual situation; collaborative innovation ability: to build a new learning mode driven by AI technology, and to create a new learning mode for students to adapt to. Collaboration and innovation ability: create a “teacher-student-AI-student” triad mechanism, i.e., teacher guidance, student exploration, AI support, make full use of the advantages of AI technology, and provide students with targeted personalized learning guidance; build a learning monitoring system, and develop a personalized learning path recommending and dynamic evaluation system, in order to further cultivate students' human-computer collaboration literacy driven by AI technology. Further cultivate students' personalized learning ability driven by AI technology.

5. Conclusion

In this paper, the knowledge tracking model and the resource recommendation model were chosen to be designed for the personalized development of students in higher education in the innovative exploration of education management.After the completion of the design of the 2 models, both of them were examined for their performance through module ablation experiments. The AUC values of the knowledge tracking model in the 2 datasets were 0.8261 and 0.8212, and the correct rates were 0.8254 and 0.8239, both above 0.8. The HR and NDCG values of the resource recommendation model in the 4 datasets are over 0.9 and 0.6.The introduction of each module has enabled the model's performance to be effectively improved.

Meanwhile, the knowledge tracking model passed the experimental requirements in 3 aspects, namely, generating historical learning ability, current knowledge state prediction ability, and real-time tracking of forgetting level, with better fitting effect. After the experiments of hyperparameter determination, the number of parameter cross matrix nodes, the number of attention heads, the number of model hierarchical structures, and the algorithmic parameter p(T) are finally set to 6, 8, 5, and 0.3, which maximizes the personalized recommendation performance of the resource recommendation model. Based on the good application effect of AI technology, it is recommended to promote the deep application of AI technology in education management from the 2 dimensions of technology maturity and adaptability, and the improvement of teacher and student role-adaptability ability.

References

1. Al-Zahrani, A. M. (2024). Unveiling the shadows: Beyond the hype of AI in education. Heliyon, 10(9).
2. Ayeni, O. O., Al Hamad, N. M., Chisom, O. N., Osawaru, B., & Adewusi, O. E. (2024). AI in education: A review of personalized learning and educational technology. GSC Advanced Research and Reviews, 18(2), 261-271.
3. He, A., Yuan, W., Lee, L. S., & Tian, T. (2025). AI-driven predictive models for optimizing mathematics education technology: Enhancing decision-making through educational data mining and meta-analysis. Smart Learning Environments, 12(1), 1-42.
4. López-Meneses, E., Mellado-Moreno, P. C., Gallardo Herrerías, C., & Pelícano-Piris, N. (2025). Educational Data Mining and Predictive Modeling in the Age of Artificial Intelligence: An In-Depth Analysis of Research Dynamics. Computers, 14(2), 68.
5. Anas, I. (2025). The impact of artificial intelligence on educational leadership: Challenges and opportunities in school management. EduYorasaki: Journal of Islamic Education Management, 2(01).
6. Selwyn, N. (2016). ‘There’s so much data’: Exploring the realities of data-based school governance. European Educational Research Journal, 15(1), 54-68.
7. Zhang, J. (2024). Application of Artificial Intelligence Tools in Educational Assessment: Enhancing Teaching and Learning Management in Beijing Junior High Schools. Eurasian Journal of Educational Research (EJER), (112).

8. Sorensen, L. C. (2019). “Big Data” in educational administration: An application for predicting school dropout risk. Educational Administration Quarterly, 55(3), 404-446.
9. Jamaica, N. J., & Tagbo, S. (2025). Artificial Intelligence in Educational Management for Enhanced Administrative Effectiveness in Rivers State Universities. International Journal of Educational Management, Rivers State University., 1(2), 213-236.
10. Sain, Z. H., Sain, S. H., & Serban, R. (2024). Implementing artificial intelligence in educational management systems: A comprehensive study of opportunities and challenges. Asian Journal of Managerial Science, 13(1), 23-31.
11. Yilmaz, A., Koç, A., Tekin, M. Z., Altun, R., & Aydin, M. (2025). Artificial intelligence in educational management: Current research. International Journal of Eurasian Education and Culture, 10(29), 157-179.
12. Lan, Q. (2021, February). Construction of personalized education model for college students driven by big data and artificial intelligence. In Journal of Physics: Conference Series (Vol. 1744, No. 3, p. 032022). IOP Publishing.
13. Frank, E. (2024). The Influence of artificial intelligence on education: Enhancing personalized learning experiences. EasyChair Preprint, 14675.
14. Yekollu, R. K., Bhimraj Ghuge, T., Sunil Biradar, S., Haldikar, S. V., & Farook Mohideen Abdul Kader, O. (2024, February). AI-driven personalized learning paths: Enhancing education through adaptive systems. In International Conference on Smart data intelligence (pp. 507-517). Singapore: Springer Nature Singapore.
15. Barrera Castro, G. P., Chiappe, A., Ramírez-Montoya, M. S., & Alcántar Nieblas, C. (2025). Key barriers to personalized learning in times of artificial intelligence: A literature review. Applied Sciences, 15(6), 3103.
16. Ellikkal, A., & Rajamohan, S. (2025). AI-enabled personalized learning: empowering management students for improving engagement and academic performance. Vilakshan-XIMB Journal of Management, 22(1), 28-44.
17. Xu, X. (2025). AI optimization algorithms enhance higher education management and personalized teaching through empirical analysis. Scientific Reports, 15(1), 10157.
18. Mourtzis, D., Panopoulos, N., & Angelopoulos, J. (2023). A hybrid teaching factory model towards personalized education 4.0. International Journal of Computer Integrated Manufacturing, 36(12), 1739-1759.
19. Hu, S. (2024). The effect of artificial intelligence-assisted personalized learning on student learning outcomes: A meta-analysis based on 31 empirical research papers. Science Insights Education Frontiers, 24(1), 3873-3894.
20. Hardaker, G., & Glenn, L. E. (2025). Artificial intelligence for personalized learning: a systematic literature review. The International Journal of Information and Learning Technology, 42(1), 1-14.
21. Silva, G., Godwin, G., & Jayanagara, O. (2024). The impact of AI on personalized learning and educational analytics. International Transactions on Education Technology (ITEE), 3(1), 36-46.
22. Neuenschwander, B. E., & Chiodo, A. (2024). Benefits of Artificial Intelligence in Education: Personalized Learning and Enhanced Classroom Management. Ann Chiodo.
23. Dembe, A. (2024). Advancing personalized learning through educational artificial intelligence: Challenges, opportunities, and future directions. Res Invent J Eng Phys Sci, 3(1), 89-101.
24. Igbokwe, I. C. (2023). Application of artificial intelligence (AI) in educational management. International Journal of Scientific and Research Publications, 13(3), 300-307.
25. Uchenna, M. A. (2025). FINANCIAL MANAGEMENT AND RESOURCE ALLOCATION IN EDUCATIONAL LEADERSHIP IN THE ERA OF AI: LEGAL IMPLICATIONS. UNIZIK JOURNAL OF EDUCATIONAL RESEARCH, SCIENCE AND VOCATIONAL STUDIES, 2(1).
26. Ingavélez-Guerra, P., Robles-Bykbaev, V. E., Perez-Muñoz, A., Hilera-González, J., & Oton-Tortosa, S. (2022). Automatic adaptation of open educational resources: an approach from a multilevel methodology based on students’ preferences, educational special needs, artificial intelligence and accessibility metadata. IEEE Access, 10, 9703-9716.
27. Song, Y., & Wang, N. (2025). Application of Big Data Analytics in Education Management: Enhancing Teaching Quality and Resource Allocation Efficiency. International Journal of High Speed Electronics and Systems, 2540637.
28. Liu, D. (2025). Research on Balanced Allocation of English Educational Resources Based on Multi-objective Optimisation and NSGA-II. Journal of Information & Knowledge Management, 24(02), 2550015.
29. Farhad, A., & Pyun, J. Y. (2023). AI-ERA: Artificial intelligence-empowered resource allocation for LoRa-enabled IoT applications. IEEE Transactions on Industrial Informatics, 19(12), 11640-11652.
30. Ding, L. (2025). Intelligent Teaching Resource Recommendation and Optimal Allocation in Basic Education. In Design Studies and Intelligence Engineering (pp. 524-533). IOS Press.
31. Li, Y., Liu, J., & Qiu, J. (2024, April). Using Machine Learning to Optimize Resource Allocation and Management Strategies in Colleges and Universities. In 2024 13th International Conference of Information and Communication Technology (ICTech) (pp. 253-257). IEEE.
32. Wang, B., Liu, H., An, P., Li, Q., Li, K., Chen, L., ... & Gu, S. (2019). Artificial intelligence and education. In Reconstructing our orders: Artificial intelligence and human society (pp. 129-161). Singapore: Springer Singapore.
33. Wu, S., & Luo, M. (2025). Selection and Resource Allocation Strategies for Chatbot Technologies in Higher Education: An Optimization Model Approach. IEEE access.
34. Suryanarayana, K. S., Kandi, V. P., Pavani, G., Rao, A. S., Rout, S., & Krishna, T. S. R. (2024). Artificial intelligence enhanced digital learning for the sustainability of education management system. The Journal of High Technology Management Research, 35(2), 100495.

35. Ufomba, E. R., Judith, E. O., & Daniel, O. C. (2024). The role of artificial intelligence in enhancing administrative and learning efficiency in higher education: Insights from Abia State University, Uturu. International Academy Journal of Administration, Education and Society, 6(3), 189-202.
36. Han, X., Xiao, S., Sheng, J., & Zhang, G. (2025). Enhancing efficiency and decision-making in higher education through intelligent commercial integration: Leveraging artificial intelligence. Journal of the Knowledge Economy, 16(1), 1546-1582.
37. Anshori, I., Yansyah, D., Nasiroh, N., Urfi, U., & Ismail, H. A. B. H. (2025). Integration of Artificial Intelligence in Education Management Systems to Improve the Quality of Islamic Education. Thawalib: Jurnal Kependidikan Islam, 6(2), 209-224.
38. Ahmad, S. F., Alam, M. M., Rahmat, M. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and administrative role of artificial intelligence in education. Sustainability, 14(3), 1101.