International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-19

https://doi.org/10.70917/ijcisim-2026-0053
Article

The Application of Big Data Analytics Methods in Higher Education Management to the Design of Individualized Student Pathways

Shucheng Li *

Law School, Tiangong University, Tianjin 300387, China; lishucheng2001@126.com

Abstract: The arrival of big data technology brings the possibility of realizing personalized education. This study applies big data analysis methods to students' personalized cultivation path, proposes a neurocognitive diagnosis method that integrates forgetting and knowledge point importance, and on this basis, constructs a personalized resource recommendation model based on cognitive diagnosis, which integrates convolutional neural network technology into the joint probability matrix decomposition model, and digs deeply into the information of personalized learning resources. Experiments show that the FAINCD cognitive diagnostic model in this paper shows better student performance prediction effect on different data sets, and its DOA value is above 0.85, which is significantly higher than that of the comparison method.The F-value of the CUPMF resource recommendation model is improved by 5.89%-11.89%, and the MAE value is decreased by 5.96%-14.84%, and the students' learning performance using this paper is improved by 15.81%, which is higher than that of the other methods. 15.81%, which is 2.56% to 5.29% higher than other methods. Therefore, the proposed model has a better effect of resource recommendation and learning facilitation, which can provide a guarantee for the personalized training of students.

Keywords: big data analysis; cognitive diagnosis; resource recommendation model; students' personalized training

1. Introduction

Higher education is the main position of talent cultivation and output, but with the increasing difficulty of higher education management, the problem of education and teaching is also increasingly significant [1-2]. For example, the traditional operation methods applied in student management are not only inefficient and difficult to cope with large-scale data processing needs, but also prone to errors, resulting in poor data accuracy [3-5]. In addition, due to the lack of unified data management standards and norms, it is difficult to share data between school management departments, forming many “information islands”, which greatly reduces the value of data utilization [6-7]. The above problems seriously affect the efficiency and quality of student management work, constrain the innovation and development of student management work, and make it difficult to meet the growing personalized needs of students [8-10].

In order to better carry out education management, colleges and universities must introduce and apply big data technology. By collecting, organizing and analyzing a large amount of data, higher education institutions can more accurately understand the needs of students [11]. At the same time, big data analysis can also provide a scientific basis for higher education information management and help the development of educational decisions [12-13]. Based on this, colleges and universities can use big data to strengthen the effect of personalized education, use the network platform to collect student learning information to effectively adjust the curriculum teaching plan, properly open some courses of interest to students, and take targeted education counseling methods, so as to enhance the comprehensive quality of students, and promote the overall development of students [14-17]. Therefore, colleges and universities should actively combine big data with education management, continuously improve the original management methods, create a new path of higher education management, improve the quality and efficiency of higher education management, and cultivate more excellent talents for the society.

In this paper, we design students' personalized training path based on big data technology, and take students' cognitive diagnosis and resource recommendation as the entry point for specific analysis. Aiming at the problem that the cognitive diagnosis model ignores some important factors in the students' learning process, we introduce the forgetting information to improve the prediction accuracy of cognitive diagnosis, and then use the Attention mechanism to obtain the importance of the test questions on the examination of the knowledge points, update the factors of the test questions in the cognitive diagnosis, and put forward the neurocognitive diagnosis model that integrates the forgetting and the importance of the knowledge points. Combined with the knowledge level of students obtained by cognitive diagnosis, the convolutional neural network technology is integrated into the joint probability matrix decomposition model, and the resource recommendation list is generated according to the cognitive ability level of students and their personalized needs, so as to realize the personalized resource recommendation model based on cognitive diagnosis. On this basis, the comparison experiments of different models and the reasonableness test of diagnostic results are carried out based on real data sets, and the knowledge mastery level of two students is selected for case study to test the prediction effect of the proposed cognitive diagnostic model. Then the comparison experiment of the resource recommendation model is conducted, where learners are divided into four groups, and different resource recommendation methods are used for learning, and the achievement score improvement rates of the four groups of students are compared to analyze the recommendation accuracy of the resource recommendation model and its effect on the overall learning of the students.

2. Individualized Training Pathways Based on Big Data

The arrival of big data technology has brought achievable possibilities for realizing personalized higher education management. The path of personalized training based on big data is shown in Fig. 1, and the research on the collection and analysis technology of campus teaching data, online data, IoT data, image data and video data, etc., is carried out according to the different sources of educational data. Specifically, various types of data such as classroom teaching, various information management systems, online learning platforms, mobile APPs, one-card passes, image data, and surveillance videos are first collected through multiple ways, and the collection technologies that are appropriate to them are used to carry out regularized data collection to support education big data. Then, based on the collected data, the collected data are cleaned and made into data in a unified format that can be directly utilized. Finally, these large-scale data are analyzed by designing analytical methods to analyze students' learning information, teachers' teaching strategies, and mine the laws of students' personalized training from the big data warehouse, so as to provide appropriate ways and methods for realizing students' personalized training.

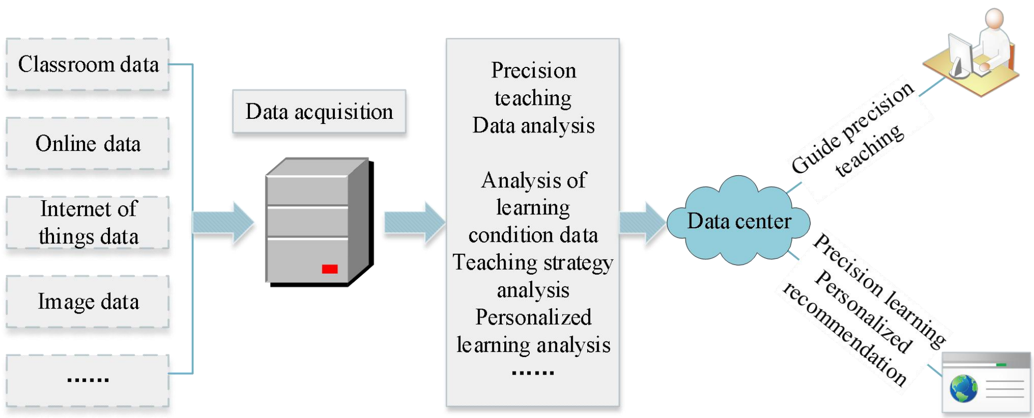

Figure 1. Personalized training path based on large numbers.

2.1. Data Acquisition

At present, relatively perfect information systems have been established in colleges and universities, including teaching management information systems, one-card systems, course selection systems and other types of information systems. At the same time, it is also possible to develop corresponding mobile applications to collect data on students' interaction trajectories and the use of students' cell phones. It is also possible to collect image and surveillance video data from all kinds of classroom teaching. The data in these systems constitute a source of big data for accurate teaching of students.

Depending on the source of data, different technologies such as Internet of Things (IoT) technology, video analytics, image recognition technology, and online platform collection technology can be used to collect a variety of different data. There are generally two methods that can be used in the process of big data processing, i.e., offline processing methods and online processing methods.

2.2. Data Extraction

The collected data are processed to extract suitable and useful data under the premise of quality assurance, and then further analysis is carried out on the basis of this data. In the project, the data collected in the Hadoop Distributed File System (HDFS) are of various sources and types, belonging to irregular data, and the format and data fields of the data itself, do not meet the needs of the analysis, and there exists more useless or redundant data. At this stage, this irregular data format does not fully meet the basic requirements for data processing, so it is necessary to pre-process the data, so that it is transformed into a more regular data needed in the later links, and then selectively select the useful information in the data according to the needs of the study of personalized learning of the students, screening and processing of the data, and the useless information is given to the elimination.

2.3. Data Analysis

Based on the analysis of big data, the online learning analysis model suitable for personalized learning can be continuously constructed from four dimensions: data and environment, stakeholders, methods and goals. The analysis can be carried out according to the requirements of the course in terms of the process structure of student learning, the visualization of the learning process and the empirical evidence of the learning effect. Then, through the results of data analysis, a reasonable learning path and learning resources of appropriate difficulty are appropriately recommended, while timely feedback on students' learning effects is realized and personalized services are provided in order to promote teaching and learning.

The key to personalized learning lies in the analysis of data, and only by providing effective analysis can we support students' personalized learning and promote their learning. When analyzing, it is necessary to combine the problems to be solved, analyze the impact of data on the results, and find out the reasons affecting students' learning based on the correlation between data and data, or provide certain data support or suggestions for students' learning and teachers' teaching.

3. Personalized Resource Recommendation Model Based on Cognitive Diagnosis

According to the theoretical analysis of personalized cultivation path based on big data, this paper analyzes the personalized resource recommendation based on cognitive diagnosis. The study of personalized learning resource pushing mechanism based on cognitive diagnosis is to study the key issues of learning resource pushing mechanism under the guidance of students' cognitive diagnosis, so as to improve students' learning efficiency and learning effect.

3.1. Neurocognitive Diagnosis of Students

At this stage there is a large amount of work on personalized educational resources recommendation based on students' individual knowledge, and cognitive diagnosis plays an irreplaceable role in extracting students' cognitive level. In this paper, we propose the Fusion of Forgetting and Knowledge Point Importance Neurocognitive Diagnosis (FAINCD).

3.1.1. Knowledge Point Importance Acquisition Module

In this paper, the attention mechanism is used to obtain the focus of the test questions on the knowledge points. The calculation method is shown in equations (1) and (2):

$$s(E_i, K) = v^T \cdot \tanh(W \cdot E_i + U \cdot K) \tag{1}$$

where: E_i is the stacking matrix of the embedding vectors e_i of the test question i, $K = \{k_1, k_2, ..., k_j\}$ is the knowledge embedding matrix. The association $s(E_i, K)$ between test question i and each knowledge point is computed by the attention model, v, W, U is the weight, · is the dot product, and tanh is the activation function.

Get the knowledge point attention weights of test question i as:

$$w_i = softmax(s(E_i, K)) \tag{2}$$

where: w_i is the vector of examination weights of test questions i related to each knowledge point.

After obtaining the knowledge point examination weights, the new vector of test question representations with knowledge point examination importance can be expressed as equation (3):

$$q_i^{'} = q_i \cdot w_i \tag{3}$$

where: $q_i \in \{0,1\}$, obtained directly from the Q matrix, is the i th row of the Q matrix. The $q_i^{'}$ denotes the test factor that incorporates the importance of the knowledge point test.

3.1.2. Oblivion Information Acquisition Module

In order to better access the cognitive state of students, student forgetting information was introduced, and student forgetting information was obtained by combining the frequency of students' responses to the knowledge point and the difficulty of the knowledge point.

The frequency of students' responses to a knowledge point is determined by dividing the number of times students responded to the knowledge point by the total number of knowledge points students responded to. See formula (4) for details:

$$p_k^u = \frac{U_{uk}^K}{\sum_{j=0}^{J} U_{uj}^K} \tag{4}$$

where: p_k^u represents the frequency of students u responding to knowledge point k, U_{uk}^K represents the number of times students u responded to knowledge point k, and J is the total number of knowledge points.

This yields a vector of the frequency of students' answers to each knowledge point $p^u = \{p_1^u, p_2^u, ..., p_J^u\}$. Then the students' knowledge point response frequency vector p^u is combined with the knowledge point difficulty vector h^{diff} to obtain the degree of students' forgetting of each knowledge point. The knowledge point forgetting information is combined with the Q-matrix information to obtain the degree of students' forgetting of each test question. The ultimate goal of the model is to predict the positive response rate of the test questions, so it is necessary to obtain the degree of forgetting of the test questions by the students. The degree of forgetting for test question i can be obtained from the frequency of students' responses to the knowledge points examined in that test question $p_{\varepsilon i}^u$ and the difficulty of the knowledge points examined in that test question h_{ei}^{diff}. As in equations (5) and (6):

$$p_{ei}^u = p^u \times q_i \tag{5}$$

$$h_{ei}^{diff} = h^{diff} \times q_i \tag{6}$$

where: p_{ei}^u and h_{ei}^{dif} are obtained by interacting p^u and h^{diff} with the i th row of the Q matrix.

Since the two vectors are not suitable for direct arithmetic fusion to form a forgetting vector, the features are extracted using a convolutional neural network. Here the two vectors are spliced into $[p_{ei}^u, h_{ei}^{diff}]$ using column-wise splicing and fed into a convolutional neural network for feature extraction with the formula shown in (7):

$$c_{mk} = f(w_m \cdot [p_{ei}^u, h_{ei}^{diff}]_k + b) \tag{7}$$

where: w_m is the convolution kernel, b is the bias, and f is the activation function, here the sigmoid function is used. c_{mk} denotes the convolutional kernel w_m for the convolution of knowledge points k to extract features.

The feature vector of all knowledge points extracted by convolution using the convolution kernel w_m is $c_m = [c_{m1}, c_{m2}, ..., c_{mJ}]$. Feature extraction is performed using multiple convolutions, $w = \{w_1, w_2, ..., w_M\}$ The feature matrix obtained is $C = [c_1, c_2, ..., c_M]$. The features extracted from the convolution kernels of different channels are linearly combined and then normalized to obtain the forgetting factor vector f . As shown in equation (8):

$$f = softmax(Cw) \tag{8}$$

where: $w \in \mathbb{R}^{M\times 1}$, is the weight vector.

3.1.3. Information Fusion and Score Prediction

The diagnostic factors such as test question factor containing the importance of knowledge points, student factor, knowledge point difficulty factor and test question differentiation were fused in the form of equation (9):

$$x_{nj} = q_j^{'} \circ (a_n - h^{diff}) \times h_j^{disc} \tag{9}$$

where: x_{nj} represents the fusion of diagnostic factors of students n interaction to test questions j , and h_j^{duc} represents the differentiation of test questions j .

In order to obtain deeper interaction characteristics and improve the diagnostic accuracy rate, considering the temporal influence of students' interactions with test questions, the Transformer network structure is used as the basis for generating the predicted answer results with a reasonable design in the input part and the introduction of forgetting information in the output part. The network model contains four parts: input part, encoder part, decoder part and predicted score part.

Input section: the input of the encoder section is the fusion matrix of diagnostic factors X for the interaction between students and test questions, and the input of the decoder is the embedding matrix R of the students' response results to the test questions and the output O of the encoder section. In this case, for the embedding of the vectors in the matrix R , the method of initializing the parameters of the normal distribution with different expectations is set according to the results of the students' answers, i.e., $\mu > 0$ when the students answer the test questions correctly, and vice versa $\mu < 0$. Input and output formulas for encoder and decoder:

$$O = Encoder(X) \tag{10}$$

$$\hat{r} = Decoder(O, R) \tag{11}$$

The encoder and decoder are combinations of multi-head attention networks, which are the core part, followed by feed-forward neural networks. In this paper, inputs corresponding to future information are masked for all multi-head attention networks to prevent invalid participation. This ensures that the computation of $\hat{r}_k$ depends only on the previous practice interactions $X_1, X_2, \ldots, X_k$ and the answering results $R_1, R_2, ..., R_{k-1}$.

Positional encoding: when using the attention mechanism to deal with the sequence problem, the positional information of the sequence needs to rely on positional encoding to obtain. Here the sine-cosine function position encoding is used. As shown in Eqs. (12) and (13):

$$PE_{pos,2i} = \sin(pos / 10000^{2i/d}) \tag{12}$$

$$PE_{pos,2i+1} = \cos(pos / 10000^{2i/d}) \tag{13}$$

where: pos represents the position of the length of the sequence, i represents the corresponding feature dimension of the sequence, e.g., features at even positions will be computed and features at odd positions will be computed cos. This sequence of values forms the position encoding vector. Where $i \in \{0,1,2,...,d_{model}/2\}$.

Multi-Headed Attention Networks: Multi-Headed Attention Networks contain Q_{in}, K_{in} and V_{in} , which denote the query, key and value, respectively. A multi-head attentional network is an attentional network that applies different mapping matrices multiple times to the same input sequence. The attention layer maps each Q_{in}, K_{in} and V_{in} to a potential space using the matrices w_i^Q, w_i^K and w_i^v . This is shown in Eqs. (14) to (16):

$$Q_i = Q_{in} W_i^Q \tag{14}$$

$$K_i = K_{in} W_i^K \tag{15}$$

$$V_i = V_{in} W_i^V \tag{16}$$

In addition the attention network modeled in this paper requires a masking mechanism to prevent leakage of subsequent information during prediction. The masking mechanism replaces the upper triangular part of the $Q_i K_i^T$ matrix with -0, which has the effect of zeroing out the attentional weights of the subsequent positions after the softmax operation. HEAD, is the value V_i multiplied by the masked attentional weights. As shown in equation (17):

$$head_i = Softmax(Mask(\frac{Q_i K_i^T}{\sqrt{d}}))V_i \tag{17}$$

where: d is the dimension of the vector in Q_i and K_i .

Multiple heads of attention are concatenated and multiplied by w^o to aggregate the outputs of different heads of attention. The final output of the multi-head attention network is obtained as shown in Equation (18):

$$MultiHead(Q_{in}, K_{in}, V_{in}) = Concat(head_1, ..., head_h)W^o \tag{18}$$

Feed-forward network: a feed-forward network is applied to the multi-head attention output to increase the nonlinear complexity of the mapping of the model. As shown in Eqs. (19) and (20):

$$F = (F_1, ..., F_k) = FFN(M) \tag{19}$$

$$F_i = ReLU(M_i W_1^{FF} + b_1^{FF})W_2^{FF} + b_2^{FF} \tag{20}$$

Encoder: The encoder contains multiple network layers, where each network layer consists of multiple attention networks and feedforward neural networks as shown in Eqs. (21) and (22):

$$M = \$SkipConct(Multihead(LayerNorm(Q_{in}, K_{in}, V_{in}))) \tag{21}$$

$$O = SkipConct(FFN(LayerNorm(M))) \tag{22}$$

where: SkipConct and LayerNorm are applied to each sub-layer.

Decoder: The decoder is similar in structure to the encoder and also consists of multiple network layers, each of which consists of multiple attention networks and feed-forward neural networks shown in equations (23), (24) and (25):

$$M_1 = SkipConct(Multihead(LayerNorm(Q_{in}, K_{in}, V_{in}))) \tag{23}$$

$$M_2 = SkipConct(Multihead(LayerNorm(M_1, O, O))) \tag{24}$$

$$L = SkipConct(FFN(LayerNorm(M_2))) \tag{25}$$

where: O is the final output of the encoder.SkipConct and LayerNorm are used for each sub-layer.

The model uses a two-layer multi-head self-attention module, and in order to enhance the representation of the network, residual links are applied to the attention module and the feed-forward neural network part, i.e., SkipConct. And a layer normalization Layer Norm is used between each layer to stabilize the gradient of the neural network and to speed up the training convergence. Before the output of the last layer, the predicted score probability is output after a linear transformation layer and a sigmoid layer after introducing the student's forgetting factor for the test questions.

The cross-entropy between the model output and the real label is used as the loss function. The vector of students' cognitive level after the completion of training α is the target to be obtained for cognitive diagnosis. This vector represents the cognitive level of the student.

3.2. Personalized Resource Recommendation Based on CUPMF

3.2.1. Overall Framework

The personalized resource recommendation method based on convolutional joint probability matrix factorization (CUPMF) model proposed in this paper contains three main parts: the first part is the student cognitive diagnostic modeling part based on the FAINCD model, and the cognitive level of the students is obtained as the a priori condition for the joint probability matrix factorization part of the recommendation algorithm of the CUPMF model. The second part is the convolutional neural network module of the CUPMF model, and the last part is the joint probability matrix decomposition part.

3.2.2. Convolutional Neural Networks

In this paper, we choose to use CNN to mine the teaching resource data, the CNN framework is mainly responsible for mining the potential features of the test question teaching resources, generating the test question implicit feature vectors, and constructing the implicit feature matrix representation of the test question with CNN weight parameters, which is used in the joint probability matrix decomposition model for the training of the solution.

The convolutional network framework of the CUPMF model consists of the following four layers:

1) Word embedding layer

The word embedding layer converts the original test question information into a dense numerical matrix as the input to the next convolutional layer. Specifically, the test question information mainly contains three parts: test question stem, test question answer, and test question paraphrase, which are processed through the word splitting technique, and each word is converted into a word vector by means of random initialization value or through the pre-trained word embedding model. Finally, the test question is represented as a dense numeric matrix $T_j \in R^{p \times l}$ by connecting the word vectors in the test question information, as shown in equation (26):

$$T_j = \begin{bmatrix} \cdots & | & | & | & \cdots \\ \cdots & w_{i-1} & w_i & w_{i+1} & \cdots \\ \cdots & | & | & | & \cdots \end{bmatrix} \tag{26}$$

where p is the dimension of the vector and l denotes the number of word vectors.

2) Convolutional Layer

Convolutional layer is mainly used to extract the feature information of the test questions, the nature of the test question feature information is different from the contextual information of the picture, audio or video, which requires certain modifications to the convolutional network to be suitable. A test question context feature $c_i^j \in R$ is extracted by the j th shared weight $W_c^j \in R^{p \times w}$, whose window size w indicates the number of surrounding words, i.e., it satisfies equation (27):

$$c_i^j = fun(W_c^j \otimes D_{(:,i:(i+w-1))} + b_c^j) \tag{27}$$

where " $\otimes$ " denotes the convolution operation, $b_c^j \in R$ is the deviation corresponding to W_c^j , and fun is a nonlinear excitation function. Considering that the gradient vanishes during the execution of the gradient descent operation, which makes the optimization convergence process of the model slow and even the procedure appears uncontrollable, in addition, it may also lead to the model stopping when the training reaches a local minimum and cannot continue to be optimized, this paper uses the ReLU to avoid the problem of vanishing gradient. Then, the context feature vector $c^j \in R^{l-w+1}$ with weight W_c^j is constructed by equation (27) as shown in equation (28):

$$c^j = [c_1^j, c_2^j, ..., c_i^j, ..., c_{l-w+1}^j] \tag{28}$$

3) Pooling layer

This model extracts representative features from each test question feature vector through the pooling layer and reduces the representation of the test question document to n_c fixed-length feature vector by constructing a fixed-length feature vector by the subsumption operation as shown in Equation (29):

$$d_f = [\max(c^1), \max(c^2), ..., \max(c^j), ..., \max(c^{n_c})] \tag{29}$$

4) Output Layer

In the output layer is mainly responsible for making a nonlinear mapping of the output of the previous layers. Therefore, it is necessary to map d_f on the k -dimensional space of the joint probability matrix decomposition model to accomplish the recommendation task, i.e., generating the trial latent matrices by using the regular nonlinear mapping as shown in Equation (30):

$$D_j = \tanh(W_{f_2}\{\tanh(W_{f_1} d_f + b_{f_1})\} + b_{f_2}) \tag{30}$$

where $W_{f_1} \in R^{f \times n_c}$ and $W_{f2} \in R^{k \times f}$ are the mapping matrices, and b_{f1} and b_{f2} are the mapping matrices of W_{f1}, W_{f1} and $D_j \in R^k$ are the deviation vectors. Finally, through the convolution and nonlinear transformation processing of the above layers of hidden layers, the convolution part of the CUPMF model is approximately intertwined into a nonlinear function, which takes the test word vectors as inputs, and the outputs are the implicit feature vectors corresponding to each test question as shown in Eq. (31):

$$D_j = Cnn(W, T_j) \tag{31}$$

where W denotes all weight and bias variables, T_j denotes the word vector of trial j after word diving through, and D_j denotes the implicit feature vector of trial j .

3.2.3. Joint Probability Matrix Decomposition

In this paper, we combine the convolutional neural network with the joint probability matrix factorization model and propose the convolutional joint probability matrix factorization (CUPMF) model. The interpretation of each matrix is as follows:

Matrix A : student-knowledge mastery matrix calculated by TDNIA model, $A \in \mathbb{R}^{m \times l}$.

Matrix R : matrix of student-test scores obtained from preprocessing of student-response data collected by the platform, $R \in \mathbb{R}^{m \times n}$.

Matrix Q : test-question-knowledge-point association matrix labeled by domain experts, $Q \in \mathbb{R}^{n \times l}$.

Matrix K : student-knowledge-point matrix and test-question-knowledge-point matrix by matrix decomposition to obtain the implicit feature matrix of knowledge points, $K \in \mathbb{R}^{l \times v}$.

Matrix U : student-knowledge-point matrix and student-test-question matrix are decomposed by matrix to obtain student-user implicit eigenmatrix, $U \in \mathbb{R}^{m \times v}$.

Matrix D : implicit feature matrix of test questions generated by convolutional neural network, $D = Cnn(W, TS), D \in \mathbb{R}^{n \times v}$.

The prior probability of the initialization matrix U obeys a Gaussian distribution with mean 0 variance σ_U^2 as in equation (32), and the prior probability of the matrix K obeys a Gaussian distribution with mean 0 variance σ_k^2 as in equation (33). The initialization of matrix D is different from the traditional joint probability matrix decomposition, which is mainly determined by three variables: 1) the weight W between neurons in the convolutional network, and the probability distribution of weight W is as in Eq. (34). 2) the word vector T_j representing the test question j generated by the word embedding technique. 3) the Gaussian noise $\varepsilon_j \sim N(0,\sigma_D^2)$ variable. Therefore, the implicit feature vector D_j generated by the CNN network for the test question j is represented by Eq. (35), and the probability distribution of the matrix D is obtained by the above equation as Eq. (36):

$$p(U \mid \sigma_U^2) = \prod_{i=1}^{m} G(U_i \mid 0, \sigma_U^2 I) \tag{32}$$

$$p(K \mid \sigma_K^2) = \prod_{i=1}^{l} G(K_i \mid 0, \sigma_K^2 I) \tag{33}$$

$$p(W \mid \sigma_W^2) = \prod_{k} G(w_k \mid 0, \sigma_W^2) \tag{34}$$

$$D_j = Cnn(W, T_j) + \varepsilon_j \tag{35}$$

$$p(D \mid W, T, \sigma_D^2) = \prod_{j=1}^{n} G(D_j \mid Cnn(W, T_j), \sigma_D^2 I) \tag{36}$$

From the implicit vector U_i of student i and the implicit vector $D_j = Cnn(W, T_j)$ of test j, it can be obtained that the probability r_{ij} of student i scoring on test j is distributed as a Gaussian distribution with mean $h(U_i^T Cnn(W, T_j))$ and variance σ_R^2 and independent, respectively, and the mathematical expression of its conditional probability distribution is shown in Eq. (37):

$$p(R \mid U, D, \sigma_R^2) = \prod_{i=1}^{m} \prod_{j=1}^{n} [G(r_{ij} \mid h(U_i^T Cnn(W, T_j)), \sigma_R^2 I)]^{I_{ij}^R} \tag{37}$$

where I_{ij}^R is the indicator function, $I_{ij}^R = 1$ if student i has done test j, otherwise $I_{ij}^R = 0$. $h(x)$ is a sigmoid function, i.e., it is the value of $U_i^T Cnn(W, T_j)$ that is mapped to the range (0, 1).

Similarly, from the implicit eigenvector U_i of student i and the implicit eigenvector K_j of knowledge point j, it can be obtained that: student i's mastery of knowledge point j, $\propto_{ij}$, satisfies a Gaussian distribution with mean $h(U_i^T K_j)$ and variance σ_A^2 and is independent, and the mathematical expression of its conditional probability distribution is shown in Eq. (38):

$$p(A \mid U, K, \sigma_A^2) = \prod_{i=1}^{m} \prod_{j=1}^{l} [G(\alpha_{ij} \mid h(U_i^T K_j), \sigma_A^2 I)]^{I_{ij}^A} \tag{38}$$

where I_{ij}^A is the indicator function, if student i masters knowledge point j, then $I_{ij}^A = 1$, otherwise $I_{ij}^A = 0$.

Similarly, from the implicit feature vector D_i of the test question i and the implicit feature vector

K_j of the knowledge point j, it can be obtained that: the association case q_{ij} of the test question i with the knowledge point j satisfies the mean value of $h(Cnn(W,T_j)^T K_j)$. The Gaussian distribution with variance σ_Q^2 and independent has the conditional probability distribution shown in equation (39):

$$p(Q \mid D,K,\sigma_Q^2)=\prod_{i=1}^{m}\prod_{j=1}^{l}[G(\alpha_{ij} \mid h(Cnn(W,T_j)^T K_j),\sigma_Q^2 I)]^{l_{ij}^Q} \tag{39}$$

where l_{ij}^q is the indicator function, if the test question i examines the knowledge point j, then $l_{ij}^Q=1$, otherwise $I_{ij}^Q=0$.

Combined with the above equation of prior probability distribution, the posterior probability distribution of matrix U,D,W,K can be obtained from Bayesian criterion as shown in equation (40):

$$\begin{aligned}
&p(U,D,W,K \mid R,A,Q,T,\sigma_U^2,\sigma_W^2,\sigma_D^2,\sigma_K^2,\sigma_R^2,\sigma_A^2,\sigma_Q^2)\\
&=p(Q \mid D,K,\sigma_Q^2)\cdot p(K \mid \sigma_K^2)\cdot p(D \mid W,T,\sigma_D^2)\cdot p(W \mid \sigma_W^2)\\
&\cdot p(R \mid U,D,\sigma_R^2)\cdot p(U \mid \sigma_U^2)\cdot p(D \mid W,T,\sigma_D^2)\cdot p(W \mid \sigma_W^2)\\
&\cdot p(A \mid U,K,\sigma_A^2)\cdot p(U \mid \sigma_U^2)\cdot p(K \mid \sigma_K^2)
\end{aligned} \tag{40}$$

Bringing the above equation into Equation (40) taking logarithms on both sides gives Equation (41):

$$\begin{aligned}
&\ln p(U,D,W,K,R,A,Q,T,\sigma_U^2,\sigma_W^2,\sigma_D^2,\sigma_K^2,\sigma_R^2,\sigma_A^2,\sigma_Q^2\rangle\\
&=-\frac{1}{2\sigma_Q^2}\sum_{i=1}^{n}\sum_{j=1}^{l} l_{ij}^Q (q_{ij}-h(Cnn(W,T_i)^T K_j))^2\\
&-\frac{1}{2\sigma_R^2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^R (r_{ij}-h(U_i^T Cnn(W,T_j)))^2\\
&-\frac{1}{2\sigma_A^2}\sum_{i=1}^{m}\sum_{j=1}^{l} l_{ij}^A (\alpha_{ij}-h(U_i^T K_j))^2-\frac{1}{2\sigma_U^2}\sum_{i=1}^{m} U_i^T U_i\\
&-\frac{1}{2\sigma_D^2}\sum_{j=1}^{n}(D_j-Cnn(W,T_j))^2-\frac{1}{2\sigma_W^2}\sum_{i=1}^{|W_i|} W_i^T W_i\\
&-\frac{1}{2\sigma_K^2}\sum_{i=1}^{l} K_i^T K_i-\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^R \ln\sigma_R-\sum_{i=1}^{m}\sum_{j=1}^{l} I_{ij}^A \ln\sigma_A\\
&-\sum_{i=1}^{n}\sum_{j=1}^{l} I_{ij}^Q \ln\sigma_Q-p\sum_{i=1}^{m}\ln\sigma_U-p\sum_{i=1}^{l}\ln\sigma_K\\
&-p\sum_{i=1}^{n}\ln\sigma_D-p\sum_{i=1}^{|W_i|}\ln\sigma_W+C
\end{aligned} \tag{41}$$

where p denotes the dimension of the implicit feature vector and C is a constant. The maximization formula (41) can be regarded as an unconstrained optimization problem, which is equivalent to the minimization formula (42):

$$
\begin{aligned}
&E(U,D,W,K,R,A,Q) \\
&=\frac{\varphi_Q}{2}\sum_{i=1}^{n}\sum_{j=1}^{l} I_{ij}^{Q}(q_{ij}-h(Cnn(W,T_i)^T K_j))^2 \\
&+\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^{R}(r_{ij}-h(U_i^T Cnn(W,T_j)))^2 \\
&+\frac{\varphi_A}{2}\sum_{i=1}^{m}\sum_{j=1}^{l} I_{ij}^{A}(\alpha_{ij}-h(U_i^T K_j))^2+\frac{\varphi_U}{2}\sum_{i=1}^{m}U_i^T U_i \\
&+\frac{\varphi_D}{2}\sum_{j=1}^{n}(D_j-Cnn(W,T_j))^2+\frac{\varphi_w}{2}\sum_{i=1}^{|W_i|}W_i{}^T W_i \\
&+\frac{\varphi_K}{2}\sum_{i=1}^{l}K_i^T K_i
\end{aligned} \tag{42}
$$

The optimal solution of Eq. (42) can be solved by stochastic gradient descent method, which in turn computes the user implicit feature matrix, the test question implicit feature matrix, and the knowledge point implicit feature matrix. Then, the student-user's score for the test question is predicted by multiplying the user implicit feature matrix and the test question implicit feature matrix. Finally, based on the score combined with the user's knowledge point mastery, teaching resources with appropriate difficulty coefficients can be recommended for the student users.

4. Experimental Validation and Analysis

4.1. Analysis of Students' Cognitive Judgments

In order to validate the diagnostic performance of the Fusion of Forgetting and Knowledge Point Importance Neurocognitive Diagnostic (FAINCD) model, this paper employs a variety of methods for validation. First, the performance of the FAINCD model in terms of its ability to predict student performance is verified by comparing the FAINCD model with other models. Next, model interpretability experiments are used to assess whether the model's diagnostic results are reasonable. Finally, the cognitive diagnostic results were analyzed using a student case study.

4.1.1. Data Sets

In the experiments, this paper uses five real public datasets for the experiments, including three small-scale datasets: FrcSub, Math1, and Math2, and two large-scale datasets: ASSISTment09, and ASSISTment17.

The FrcSub dataset is about elementary school fraction addition and subtraction tests, which provides records of each student's scores on these tests and the knowledge points examined in each test question.Math1 and Math2 are datasets of a high school final math test, which also consists of two parts: students' scores records and knowledge points examined in the test questions. Due to the small data volume of the above three open datasets, and in order to more accurately verify the diagnostic effect of the model, this paper uses the ASSISTment09 dataset and ASSISTment17 dataset, which are open datasets collected by the ASSISTments online tutoring system, providing the records of the students' answers as well as the concepts of the knowledge points. 80% of them are used for training set and 20% for testing set.

4.1.2. Prediction of Student Performance

In order to verify the effectiveness of the proposed FAINCD model on students' cognitive diagnosis, in this paper, FAINCD is compared with cognitive diagnostic methods and collaborative filtering methods in a comparative experiment based on different datasets.

On the small-scale datasets FrcSub, Math1 and Math2, the methods involved in the comparison are cognitive diagnostic model DINA, FuzzyCDF, NeuralCDM, and collaborative filtering-based method PMF model, and the comparison results of each model index under the three datasets are shown in Fig. 2, Fig. 3 and Fig. 4. From the experimental results, it can be seen that on the three small-scale datasets, the Accuracy of the FAINCD model is 0.843, 0.685, and 0.716, the AUC is 0.897, 0.746, and 0.788, respectively, and the RMSE is 0.359, 0.447, and 0.437, respectively, which are achieved more than that of the DINA model, the FuzzyCDF model, the N euralCDM model and PMF model, indicating that compared with the traditional cognitive diagnostic methods, the neurocognitive diagnostic method integrating forgetting and knowledge point importance can better fit the complex interaction process between students and test questions, thus improving the diagnostic effect.

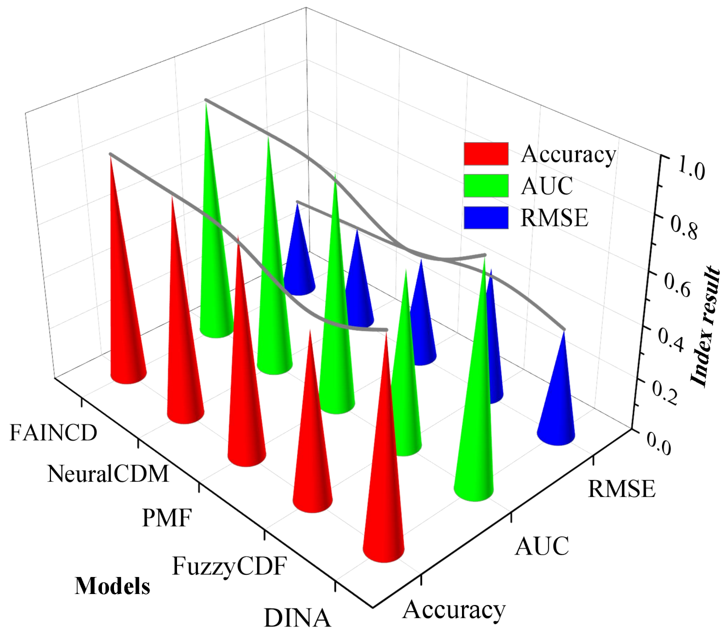

Figure 2. The comparison of each model in the FrcSub data set.

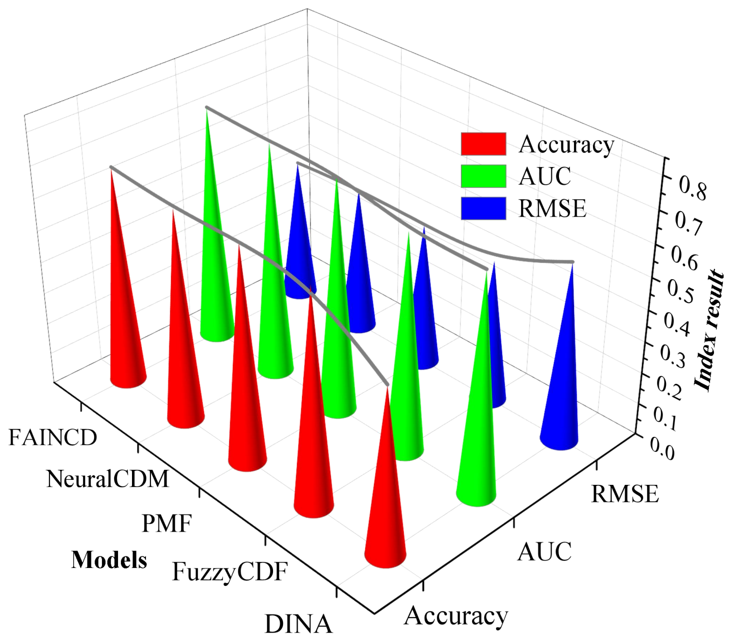

Figure 3. The comparison of each model in the Math1 data set.

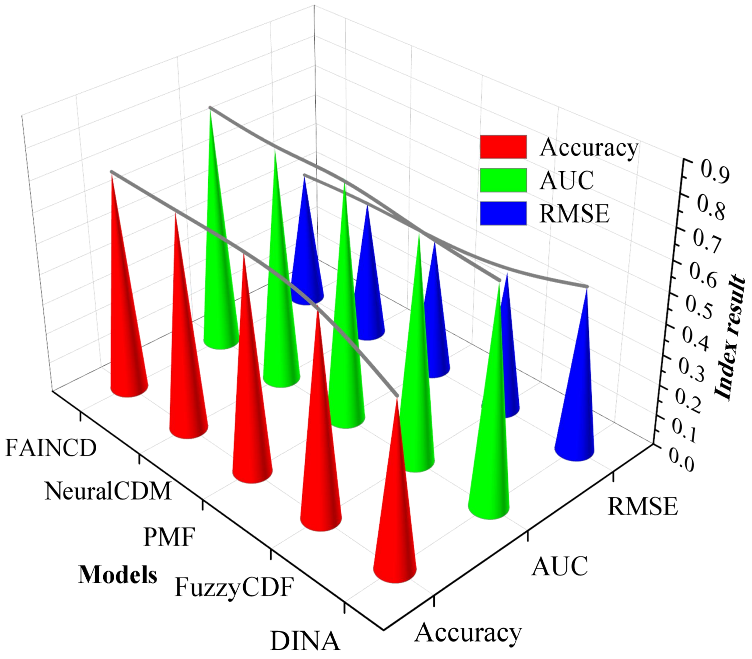

Figure 4. The comparison of each model in the Math2 data set.

Comparative analysis of student achievement prediction on large-scale datasets ASSISTment09 and ASSISTment17, FAINCD model was compared only with NeuralCDM model and PMF. This is because in DINA and FuzzyCDF models, the complexity of the model is exponentially related to the number of knowledge concepts, and neither of them is applicable to large-scale datasets, and the comparison of the model metrics under the ASSISTments dataset is shown in Fig. 5.The FAINCD model, compared to the NeuralCDM model and the PMF model, shows an increase in Accuracy and AUC metrics. Compared with the NeuralCDM model and PMF model, the FAINCD model improved in Accuracy and AUC metrics, which were 0.728 and 0.762 on ASSISTment09 and 0.692 and 0.756 on ASSISTment17, respectively, and the RMSE metrics decreased, which were 0.445 and 0.448, respectively, suggesting that the FAINCD method can improve the accuracy of cognitive diagnosis effectively.

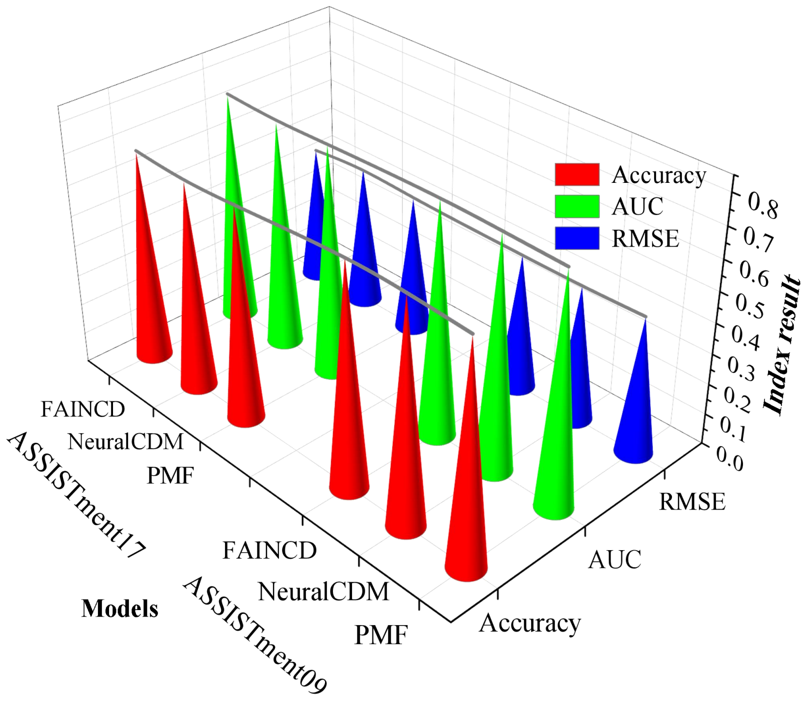

Figure 5. The comparison of each model in the ASSISTments data set.

4.1.3. Cognitive Diagnostic Rationality

The student performance prediction experiment indirectly verified the validity of the cognitive diagnostic model, but there is still a need to further evaluate the reasonableness (interpretability) of the cognitive diagnostic results, which can help students' self-assessment as well as future recommendation tasks. In order to assess the interpretability of the cognitive diagnostic model, the degree of agreement (DOA) was used as an evaluation index.The higher the value of the DOA evaluation index, the more the cognitive state diagnosed by the model conforms to the assumption of monotonicity, i.e., the better the interpretability of the model.

In this experiment, the cognitive diagnostic method DINA, FuzzyCDF, NeuralCDM model, and PMF model in the recommendation domain are still used as comparative models for model reasonableness experiments on FrcSub, Math1, and Math2 datasets, respectively. In addition, since the PMF model is unable to generate student knowledge point mastery vectors for DOA metrics computation, in order to simulate the knowledge point mastery vectors, the experiments were conducted by setting the dimensions of the potential feature vectors of the students and the test questions to be the same size as the number of knowledge point concepts for the DOA metrics computation.

The results of the DOA experiments of different models are shown in Fig. 6. The DOA values of the FAINCD model as well as the FuzzyCDF model on the FrcSub dataset reached more than 0.9, with 0.937 and 0.923, respectively, and on the Math1 and Math2 datasets, the DOA values of the FAINCD model were higher than those of the traditional cognitive diagnostic models, DINA and PMF, and of the FuzzyCDF, and slightly higher than the NeuralCDM model. Specifically, since the PMF model only considers the student score matrix, which is calculated using the simulated knowledge point mastery vector, the PMF model has the lowest DOA value in all three datasets, and the simulated knowledge point mastery obtained is of low quality. In the three experimental datasets, the DOA values of the FAINCD model are 0.937, 0.852, 0.850, respectively, and the DOA values of the NeuralCDM model are 0.872, 0.849, 0.841, respectively. Although the difference between the DOA values of the two models is not large, it can also explain to a certain extent that the cognitive diagnostic results of students obtained by the FAINCD model are reasonable. Although the difference in DOA values between the two models is large, it can also indicate to some extent that the cognitive diagnosis results of students obtained by FAINCD model are reasonable.

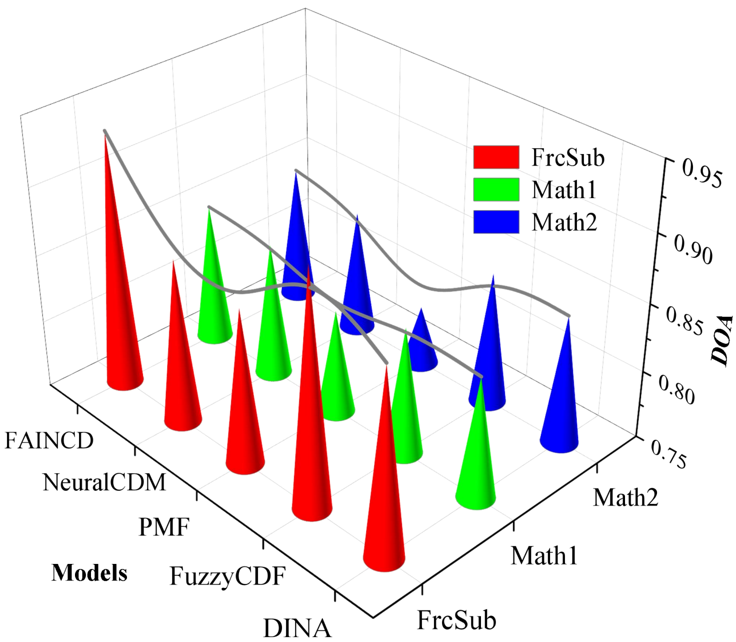

Figure 6. DOA experimental results of different models.

4.1.4. Cognitive Diagnostics Case Study.

In order to further illustrate the interpretability and effectiveness of the cognitive diagnostic method proposed in this paper, Figure 7 shows a comparison of the mastery level of two students selected from the FrcSub dataset for the 10 knowledge point attributes, and the higher the mastery level of the students on the knowledge points indicates that the students have a better grasp of the knowledge points.

Student A's mastery of knowledge points S1 and S4 is relatively strong, with a mastery level of 0.75 or above, but the mastery level of knowledge points S2 and S7 is less than 0.45, which indicates that he or she is not proficient in this knowledge point and needs to strengthen the training. In addition, by comparing Student A and Student B, it can be concluded that the overall mastery level of Student A is better than that of Student B, and the average value of the mastery water of the two knowledge points is 0.67 and 0.57, respectively. From the case study, it can be seen that the neurocognitive diagnostic model proposed by this paper, which integrates forgetting and the importance of the knowledge points, can help students to more intuitively determine the degree of their own mastery of the knowledge points, and thus make more targeted learning plans and strategies. Targeted development of learning plans and strategies to achieve personalized talent training for students.

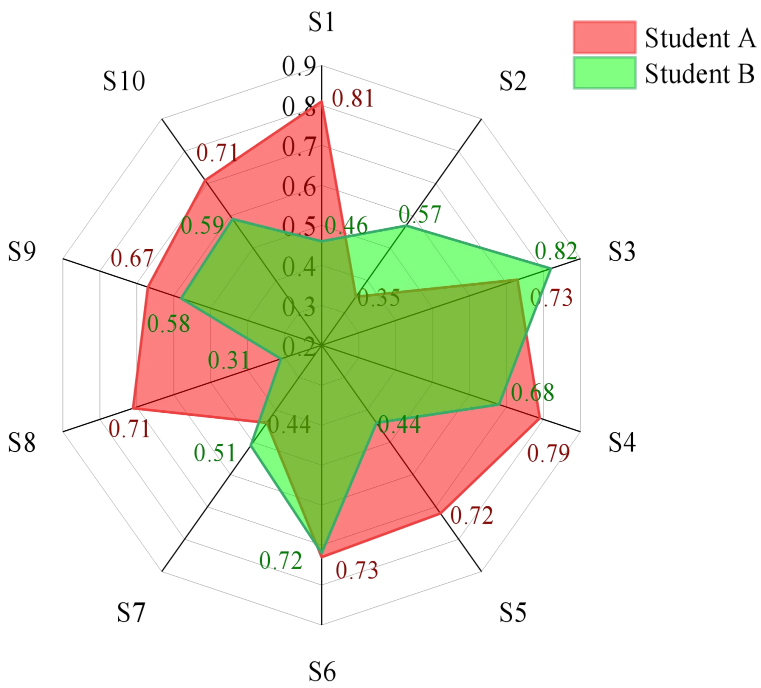

Figure 7. Students' knowledge level.

4.2. Analysis of the Effectiveness of Resource Recommendations

This experiment contains two parts to validate the effectiveness of the personalized learning resource recommendation model and the actual improvement effect on learners' academic performance, respectively.

4.2.1. Experimental Data Sets

In order to verify the effectiveness of the personalized learning recommendation model, this experiment is divided into two parts, the first part of the experiment is to verify the effectiveness of the model by comparing it with other traditional personalized learning resource recommendation models. The second part of the experiment is to select learners whose scores are between 60% and 70% (inclusive) as edge learners, so as to verify the improvement of edge learners' learning performance.

In this paper, the data information dataset1, dataset2, dataset3 of a course in a university are collected, dataset1 is the score results of the final test scores of previous learners in the university in the course, dataset2, dataset3 are the score results of the two scores of the organization's test on the learners in the university, which as well as the previous learners' exams on the course Test dataset records.

Test data dataset2 and dataset3:

In order to verify the actual enhancement effect of the model on the learners, the organizational test was conducted on 88 students from the relevant majors in the university. Firstly, the first test is conducted without using the model scheme and the learners' score dataset dataset2 is collected.Then, the learners are recommended to study using different personalized resource recommendation model schemes and the learners are tested twice and the learners' score dataset dataset3 is collected.

4.2.2. Evaluation Criteria

In personalized recommender systems, accuracy and recall are often used as the two most commonly used evaluation metrics, and the F1 value is used as a combined evaluation metric for accuracy and recall, with MAE (Mean Absolute Error) for the closeness of the learner's predicted knowledge point scores to the true scores. In addition to the above evaluation metrics, there is also an indicator rate IncreaseRate used to validate the learner's score improvement, which is obtained by dividing the learner's improved score by the learner's score on the first test calculation.

4.2.3. Validity of the Model

In order to validate the effectiveness of the CUPMF model in this paper, the model is analyzed in a comparative experiment with the traditional collaborative filtering (CF) method, the improved collaborative filtering algorithm incorporating learners' learning styles (CF-SPM), and the SVD-based collaborative filtering algorithm (CF-SVD), respectively, on the same test data set In this experiment, 90% of the learners are used as the training set and 10% of the learners are used as the validation set.

Through the experimental analysis, the data values of the experimental results of different recommendation methods are shown in Fig. 8. The effectiveness of CUPMF model in this paper is better than the other three recommendation models. Among them, in the key index F-value, the CUPMF model of this paper is 11.89%, 7.88% and 5.89% higher than the three recommendation models CF, CF-SPM and CF-SVD, respectively.

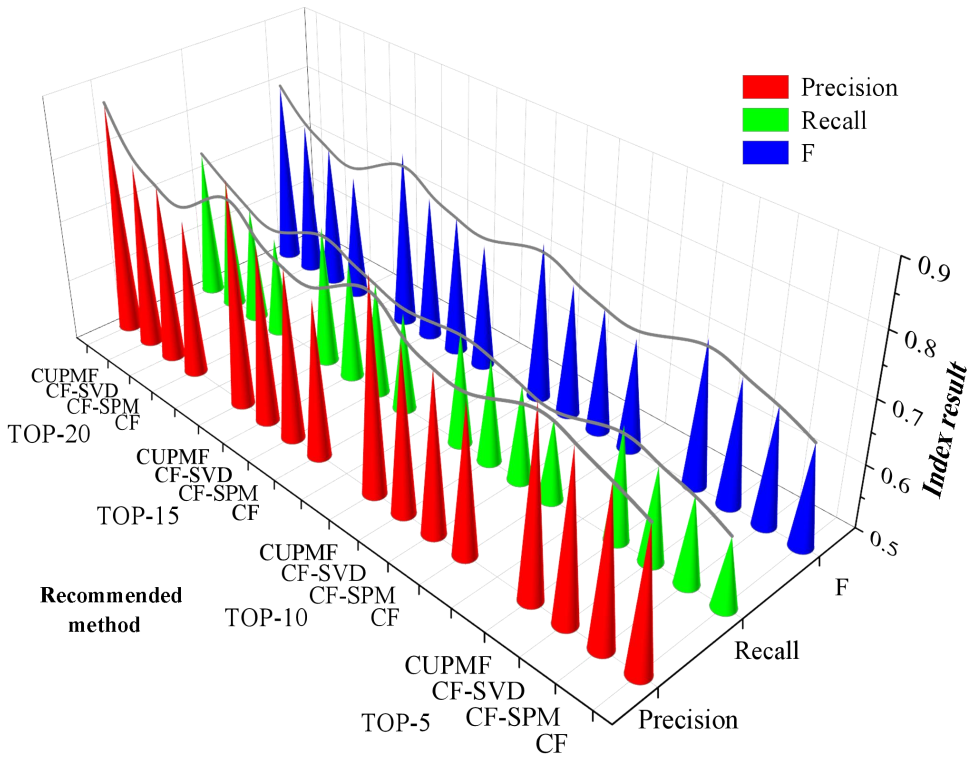

Figure 8. Data values of experimental results of different recommendations.

Among them, in addition to the above indicators, the MAE value is also an indicator to verify the accuracy of the recommendation model, the MAE value is a measure of the average of the absolute error between the learner's predicted knowledge point scores and the learner's actual scores, if the MAE value is smaller, it indicates that the learner's prediction of the knowledge point scores in the recommendation model is closer to the actual value. The MAE value after normalization of knowledge point scores is lower, and the lower the MAE value, the more accurate the prediction of knowledge point scores. The MAE values of each recommended method are shown in Figure 9 below, and it can be clearly seen that the model proposed in this paper has an obvious advantage over the other three models. The MAE value of the recommended method CUPMF is 14.84%, 13.91%, and 5.96% lower than that of the recommended methods CF, CF-SPM, and CF-SVD, respectively, thus it can be seen that the closer the predicted scores of the knowledge points of the exercises are to the actual scores, the more accurately the scores of the predicted knowledge points are.

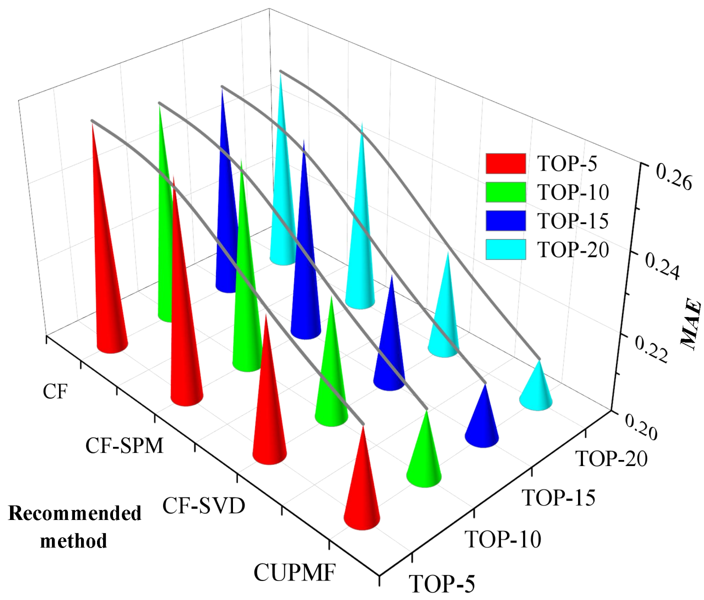

Figure 9. The MAE values of the recommended methods.

4.2.4. Actual Lifting Effect

In order to verify the actual enhancement effect of the model on the learners' academic performance, 88 learners in a university were organized to conduct experiments, and the learners were divided into four groups (A, B, C, D), with 22 students in each group, and different groups of learners, respectively, were taken to adopt different recommendation models for recommendation verification.

The first test experiment was conducted on the four groups of learners, using four personalized learning resource recommendation schemes, and after a period of time, the second experimental test was conducted on the four groups of learners, and the score comparison results in the two tests are shown in Figure 10 below. The scores of the second experimental test are improved by 6.37, 7.11, 8.03, and 9.59 compared to the scores of the first experimental test, respectively, $Increase\ Rate$ indicators are 10.52%, 11.74%, 13.25%, and 15.81%, respectively, it can be concluded that the enhancement rate of the recommendation effect of the CUPMF personalized resource recommendation model is better than the other three recommendation models.

Through the comparative experimental analysis, it is found that under the learning of different recommendation schemes, the average scores of learners in each group have been improved to a certain extent, especially under the guidance of the CUPMF model, the learners are more clear about their own learning deficiencies, and they will be more targeted to learn, so as to achieve the purpose of accurate and efficient learning.

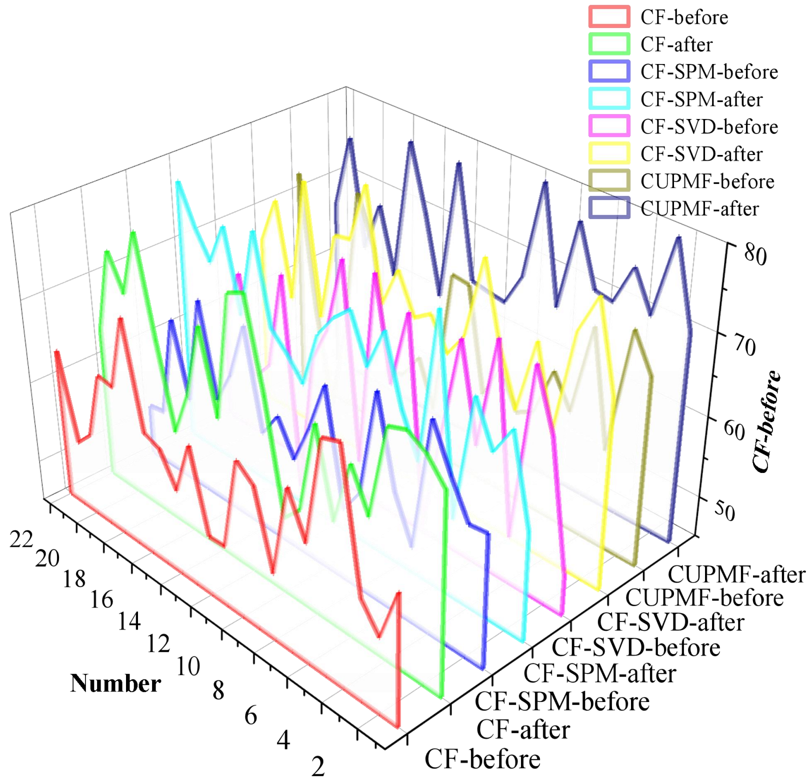

Figure 10. The comparison of scores in two tests.

5. Conclusion

The development of education informatization promotes the development of personalized education, this study uses big data analysis methods to explore the path of students' personalized cultivation, and proposes a personalized resource recommendation model based on cognitive diagnosis, and verifies the effects of the proposed students' cognitive diagnosis model and personalized resource recommendation model through experiments. The main research results are as follows:

(1) By comparing the student achievement prediction of different models on five datasets, the Accuracy, AUC and RMSE results of the FAINCD cognitive diagnosis model in this paper are better than those of the comparison models, which are 0.685 to 0.843, 0.746 to 0.897, and 0.359 to 0.448, respectively. The DOA value of the FAINCD model is greater than 0.85 on all three datasets, which verifies the rationality and superiority of the model's cognitive diagnosis. In addition, through the case study, the mastery level of two students on different knowledge points is obtained, which shows the diagnostic effect of the proposed FAINCD model on students' knowledge level.

(2) The F-value of personalized resource recommendation of CUPMF model is improved by 5.89% to 11.89% and the MAE-value is reduced by 5.96% to 14.84% compared with other recommendation models, which reflects the accuracy of the CUPMF model in recommending the learning resources, and it helps to personalize the training of the students. At the same time, the improvement rate of students' performance using the personalized resource recommendation method in this paper reaches 15.81%, which is 2.56% to 5.29% higher than other methods. That is, with the help of the personalized resource recommendation model based on cognitive diagnosis, the overall learning effect of students has been significantly improved.

The personalized resource recommendation model based on cognitive diagnosis improves the effect of learning resource recommendation to a certain extent, which can be effectively applied to the personalized training of students and enhance the management efficiency of higher education. In addition, oriented to student learning, big data technology will be able to provide teachers with the most realistic information about each student's characteristics in order to tailor teaching to students' needs and further promote students' personalized training.

References

[1] Khamdamov, U., Abdullayev, A., Elov, J., & Sultanov, D. (2020). Conceptual model of the education management information system for higher education institutions. International Journal of Advanced Trends in Computer Science and Engineering, 9(5).

[2] Psomas, E., & Antony, J. (2017). Total quality management elements and results in higher education institutions: The Greek case. Quality Assurance in Education, 25(2), 206-223.

[3] Yang, H., & Zhang, W. (2022). Data mining in college student education management information system. International Journal of Embedded Systems, 15(3), 279-287.

[4] Fernandes, E., Holanda, M., Victorino, M., Borges, V., Carvalho, R., & Van Erven, G. (2019). Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. Journal of business research, 94, 335-343.

[5] Rodrigues, M. W., Isotani, S., & Zarate, L. E. (2018). Educational Data Mining: A review of evaluation process in the e-learning. Telematics and Informatics, 35(6), 1701-1717.

[6] Silva, N. J., Costa-Lobo, C., Pereira, C. S., & Durão, N. S. (2017). Evaluation of teaching performance in higher education: Students perspectives and teaching management indicators. In INTED2017 Proceedings (pp. 4565-4571). IATED.

[7] Alamleh, H. (2020, December). Private and Secure Students' Data Sharing in Educational Systems. In 2020 Sixth International Conference on e-Learning (econf) (pp. 158-161). IEEE.

[8] Hills, K., Andersen, K., & Davidson, S. (2018). Personalized learning and teaching approaches to meet diverse needs: a prototype tertiary education program. Reimagining Christian education: Cultivating transformative approaches, 233-257.

[9] Akyuz, Y. (2020). Personalized learning in education. American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS), 69(1), 175-194.

[10] 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.

[11] Martins, J., Branco, F., Gonçalves, R., Au-Yong-Oliveira, M., Oliveira, T., Naranjo-Zolotov, M., & Cruz-Jesus, F. (2019). Assessing the success behind the use of education management information systems in higher education. Telematics and Informatics, 38, 182-193.

[12] Tahvildarzadeh, M., Moghaddasi, H., & Hosseini, M. (2017). A framework for quality management of university educational information: a review study. Research and Development in Medical Education, 6(1), 3-11.

[13] Pinho, C., Franco, M., & Mendes, L. (2018). Web portals as tools to support information management in higher education institutions: A systematic literature review. International Journal of Information Management, 41, 80-92.

[14] Dishon, G. (2017). New data, old tensions: Big data, personalized learning, and the challenges of progressive education. Theory and Research in Education, 15(3), 272-289.

[15]Abad-Segura, E., González-Zamar, M. D., Infante-Moro, J. C., & Ruipérez García, G. (2020). Sustainable management of digital transformation in higher education: Global research trends. Sustainability, 12(5), 2107.

[16] Fetaji, B., & Zeqiri, I. (2024). Application of Big Data Analytics for University Management: Improving Decision-Making and Supporting Personalized Learning. Editorial Board, 52.

[17] Dipace, A., Loperfido, F. F., & Scarinci, A. (2018). From Big Data to Learning Analytics for a personalized learning experience. REM, 3-9.