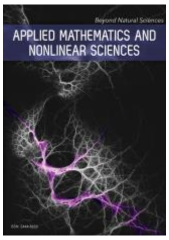
sciendo Applied Mathematics and Nonlinear Sciences, 9(1) (2024) 1-17



Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

Research on educational applications based on diagnostic learning analytics in the context of big data analytics

Naimin Zhang[1], Linlin Zhang[1,†]

1. Hebei Oriental University, Langfang, Hebei, 065000, China.

Submission Info

Communicated by Z. Sabir
Received January 15, 2024
Accepted January 25, 2024
Available online February 26, 2024

Abstract

In the context of the significant data era, this paper explores the educational applications based on diagnostic learning analytics technology to improve personalized learning and teaching effects in the educational process. The study adopts a multidimensional feature fusion approach to construct a cognitive diagnostic model to predict learners' knowledge status and future learning performance. Through actual data testing, the model can effectively predict the students' knowledge mastery state and analyze the students' learning process in depth. The experimental results show that the diagnostic model exhibits high efficiency and accuracy in predicting students' knowledge mastery status, with an accuracy rate of 92.97%, significantly better than traditional teaching methods. In addition, the study explores the encoding method of learners' multidimensional features and constructs a dynamic diagnostic model of test factors and student factors based on graph attention network. The study provides a new learning analysis and diagnostic method in the education field, which helps improve the effect of personalized learning.

Keywords: Diagnostic learning; Multidimensional features; Graph attention network.
AMS 2010 codes: 94A08

†Corresponding author.
Email address: jkllmyt@163.com

ISSN 2444-8656
https://doi.org/10.2478/amns-2024-0624

1 Introduction

In the rapid development of information technology, big data analytics is increasingly used in education. Educational applications based on diagnostic learning analytics (DLA), especially in personalized learning and student performance assessment, show great potential [1-3].

The application of big data analytics in education focuses on the tracking of students' learning behaviors, the assessment of learning outcomes, and the optimization of teaching methods [4]. By collecting and analyzing student interaction data on learning platforms, educators can better understand students' learning habits, mastery progress, and difficulties. DLA is a technology that uses data analytics to diagnose problems in students' learning processes and provide personalized feedback [5-7]. It is able to provide customized learning resources and instructional strategies to help students overcome learning barriers based on their specific performance [8-9]. The core of DLA technology lies in its ability to identify students' learning needs in a timely manner and provide targeted support.

In educational practice, DLA can help teachers achieve more accurate teaching. By analyzing students' homework and test data, DLA can identify students' mastery of specific knowledge points, thus helping teachers adjust their teaching plans and target tutoring [10-11]. In addition, DLA can also help educational institutions optimize the curriculum design to better meet students' learning needs. Although DLA shows great potential in the field of education, it also faces some challenges in practical application. Data privacy and security issues, data quality and accuracy of analysis, and how to effectively integrate DLA results into teaching practices [12-13]. In the future, as technology continues to advance and educators gain a deeper understanding of big data and its applications, the application of DLA technology in education will become more widespread and deeper [14].

Educational applications based on diagnostic learning analytics technology are gradually becoming an important direction for educational reform and development. By accurately analyzing and responding to students' learning needs, DLA is expected to achieve a more personalized and efficient teaching mode [15-17]. With the continuous development and improvement of technology, future education will focus more on data-driven decision-making and the realization of personalized learning.

First, the research team collected multidimensional data from learners, including knowledge characteristics, interaction characteristics, behavioral characteristics, and temporal characteristics. Then, graph attention network and memory network techniques are used to construct a dynamic model of students and test factors so as to comprehensively analyze and predict students' learning status. Finally, the validity of the model is verified through actual cases, and data analysis is carried out so as to explore the application effect of learning analytics technology in the field of education.

2 Multidimensional feature fusion for learning analytics

2.1 Learning cognitive diagnostic model characterization

Assuming that a large sequence of observable multidimensional extrinsic features is generated during the learner's interaction with the program (notation $EF=\{EF_1,EF_2,\ldots,EF_n\}$) and because the learner's performance conforms to the cognitive laws of cognitive processing theory (notation CL), a mapping from extrinsic features to cognitive laws needs to be constructed (notation $CL_\theta(EF)$). The goal of the cognitive diagnostic model is twofold:

1) To predict the learner's probability of answering correctly on future test questions based on modeling the cognitive laws of the extrinsic features.

2) To hypothesize the learner's potential state on each knowledge skill.

Thus, the objective function of cognitive diagnosis (i.e., the item response function) can be expressed as follows:

$$P\left(r_{ij}=1 \mid \alpha_i\right)=\Phi_{\alpha_i}\left(CL_\theta(EF)\right) \tag{1}$$

Where Φ_{a_t} represents the item response function of the learner's knowledge state based on cognitive laws, $CL_\theta(\cdot)$ represents different cognitive law mapping functions, θ represents speed, learning, forgetting, or some other internal cognitive laws and synthesizing ability, etc. r_{ij} represents the learner's response to the item, which is 1 for a correct answer and 0 for the opposite. α_i is the learner's potential knowledge mastery state. The larger α_i is, the higher the probability of its learners answering correctly. The parameters are trained by comparing the predicted responses with the real responses, and ultimately, the level of the learner's knowledge status is inferred based on the performance prediction of the future responses.

For the $CL_\theta(\cdot)$ function, different cognitive features modeling has different mapping functions with different features. For example, speed modeling requires the use of knowledge features and interaction features represented by reaction times, which are represented by the formula:

$$CL_{\theta=s}=\Psi(K, I) \tag{2}$$

Where K denotes the associated features of items and knowledge, and I denotes the interaction features of learners and items, including learners' reaction time, which is used to recommend learners' reaction speed and thus affect the prediction of their reaction results. And the memory modeling for learning and forgetting requires the use of sequence features such as behavioral features and time intervals, which are expressed by the formula:

$$CL_{\theta=m}=\Psi\left(r_{ij}, B, P\right) \tag{3}$$

Where m represents the cognitive laws of memory such as learning and forgetting. Here I focuses on the characteristics of learner-item response-outcome interactions, B represents the behavioral characteristics that arise when learners interact with items, and P represents the characteristics of time intervals in the time-series data of learner-item interactions.

2.2 Multidimensional characteristics of learners and their encoding

The multidimensional characterization of student learning analysis constructed in this paper is shown in Figure 1. From the learner self-regulation framework, learners are generally coded in terms of cognitive measures, behavioral measures, and affective measures.

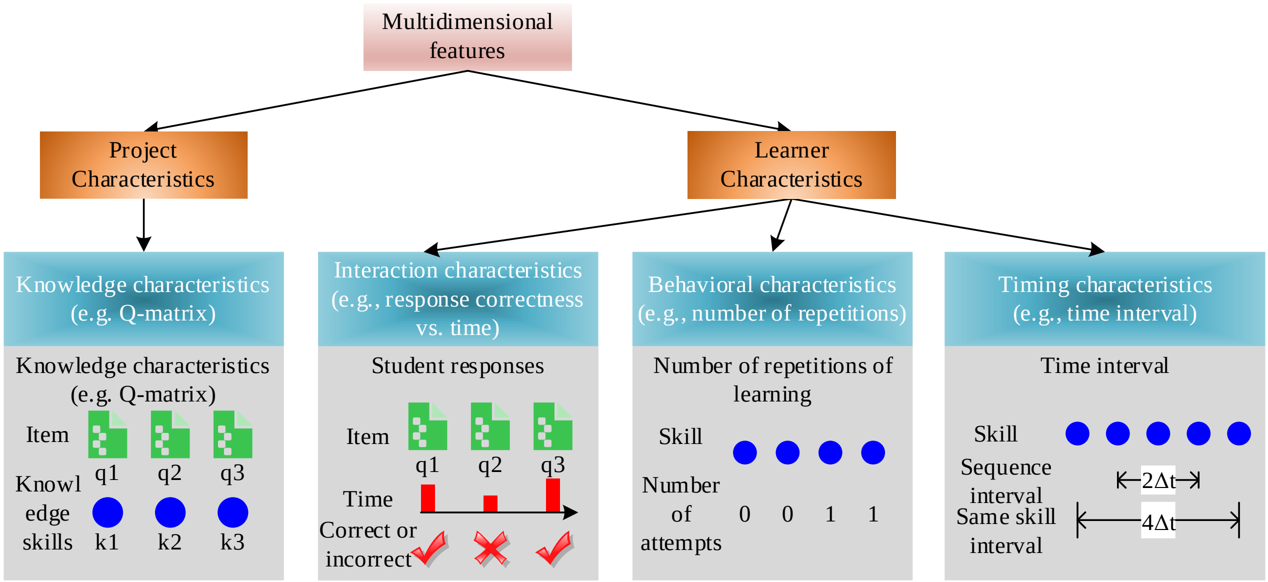


Figure 1. Students diagnose multidimensional characteristics

From the multidimensional feature analysis, it can be seen that the multidimensional features include knowledge features (items, knowledge points associated with the items), interaction features (corrections and errors obtained by the learner when interacting with the items, and the completion time), behavioral features (the number of repetitions of the same skill in the long term assessment), and temporal features (the sequence time interval and the repetition time interval). Thus, a learner's multidimensional extrinsic profile EF is constructed from the four dimensions of knowledge profile K, interaction profile I, behavioral profile B, and temporal profile P, denoted as a quaternion $EF = (K, I, B, P)$.

2.2.1 Encoding of knowledge features

Knowledge features represent the association between items and knowledge points; by using methods such as Q matrix or directly using knowledge points instead of items, the set of knowledge points is represented by K, and $K = \{k_1, k_2, k_3, \ldots, k_n\}$ means that there are n knowledge points, and the following formula is the correlation matrix between the test questions and the knowledge points in the figure, i.e., the Q matrix:

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{4}$$

Where 1 means that the item is associated with a knowledge point and 0 means that the item is not associated with a knowledge point, thus the Q matrix can be directly input to the cognitive diagnostic model represented by DINA. The Q matrix allows an item to be associated with more than one knowledge point, i.e., there may be more than one 1 on an item in a row. If an item is associated with only one knowledge point, there is only one 1 in the row, and all others are 0s.

2.2.2 Interactive feature coding

Interaction features are represented using I . In general, during the assessment process, the interaction feature consists of the learner's positive and incorrect response results R and the response time T , i.e., $I \in \{R,T\}$. Positive and incorrect responses $R=\{r_{ij}\}$, which denote the response of the i th learner on the j item. $r_{ij} \in \{0,1\}$, i.e., the learner i answered item j correctly (i.e., $r_{ij}=1$) or incorrectly (i.e., $r_{ij}=0$). Reaction time $T=\{t_{ij}\}$, where t_{ij} denotes the time spent by the learner i in doing the answer to item j .

Depending on the computational methods of different diagnostic models, the representation of positive and incorrect responses is slightly different in probabilistic statistics-based and deep learning-based approaches.

In probabilistic statistics-based cognitive diagnostic models, a matrix is often used to represent the positive and incorrect responses of learners, i.e., the response vector of learner $S1$ in the above figure can be represented as:

$$R_{s1}=\begin{bmatrix}1 & 0 & 1 & 1 & 1 & 0 & 1 & 0\end{bmatrix} \tag{5}$$

In knowledge tracking models based on deep learning methods, the inputs to the deep learning model are more uniform because the most common representation of learner responses in deep knowledge tracking is shown below:

$$(e_t,r_t)=\begin{cases}[e_t,0], & if r_t=1\\ [0,e_t], & if r_t=0\end{cases} \tag{6}$$

Where r_t represents t moments when the learner answered correctly or incorrectly on e_t the item, by which the item and the corresponding response are spliced into a vector for easy input into the deep learning method.

2.2.3 Encoding of timing features

The temporal sequence feature is used to be represented by P , including equidistant time intervals and unequal time intervals. In the cognitive diagnostic domain, on the other hand, the learners do not participate in the assessment items at each time equally spaced, so the sequence time interval is a very important assessment timing feature, which is commonly represented by Δs_t .

The former Δs_t is the time interval between the current and the last assessment (two adjacent assessments). Learners are more likely to perform better on shorter sequence time intervals due to the potential correlation between test questions from different assessments. For example, a learner with the same knowledge $K1$ has a shorter sequence time interval on test $e7$ than test $e5$, then $\Delta S_{t=2,k=K1,e=e5}=\Delta t_{21}=t_2-t_1=21 days, \Delta S_{t=3,k=K1,e=e7}=\Delta t_{32}=t_3-t_2=5 days$.

Thus, the likelihood of answering correctly on $e7$ is greater than the likelihood of answering correctly on $e5$.

The latter is the minimum time interval between current and previous assessments of the same knowledge concept, as mentioned in previous studies. Δr_t For example, the learner's repetition interval for knowledge $K1$ is shorter than $K2$ as described by the following equation:

$$\begin{aligned}\Delta r_{t=3,k=K1,e=e7} &= \Delta t_{32} = t_3 - t_2 = 5days \\ \Delta r_{t=3,k=K2,e=e8} &= \Delta t_{31} = t_3 - t_1 = 26days\end{aligned} \tag{7}$$

Therefore, learner $S1$ is more likely to get it right on $e7$ (corresponding to knowledge $K1$) than on $e8$ (corresponding to knowledge $K2$).

2.3 Cognitive modeling

The long-cycle process time-series data of learners' participation in the assessment implies the iso-memory law of learners' learning and forgetting, which is represented by m. The learner's knowledge state will be forgotten as the time interval grows, and their knowledge mastery will gradually decline, which is the law of forgetting and is represented by f. When the learner repeats the recollection and re-recognition of knowledge by doing problems or studying, his memory of knowledge gradually deepens, so his knowledge mastery state will rise. This is the law of learning, which is represented by l. When the learner repeats a knowledge point, the learner's memory of this knowledge point will deepen, and the degree of forgetting will become lower. Therefore, the law of memory includes the laws of forgetting and learning, i.e., $m \in \{f, l, \ldots\}$. Cognitive diagnostic modeling based on learning and forgetting aims to predict the probability of a learner's correct answer through the representation of the learner's cognitive laws of forgetting and learning, which can be expressed by the formula:

$$P\left(r_{ij} = 1 | \alpha_i\right) = \Phi\left(CL_{\theta=m}\left(r_{ij}, f, l\right)\right) \tag{8}$$

And again, since the law of learning and forgetting relies on temporal characteristics represented by time intervals and behavioral characteristics represented by the number of repetitions of learning, the law of memory can be represented as:

$$CL_{\theta=m} = \Psi\left(r_{ij}, B, P\right) \tag{9}$$

3 Learning cognitive diagnosis incorporating multidimensional features

3.1 Cognitive Diagnostic Modeling

By coding the multidimensional features of students' learning cognition, this paper proposes an attention-based dynamic cognitive diagnostic model (ACD) in order to realize a dynamic cognitive diagnosis and improve the accuracy of cognitive diagnosis as much as possible while retaining better interpretability. The analytical model of learning cognitive diagnosis is shown in Figure 2. Simulating the interactive process when students answer the test questions, the whole cognitive diagnosis process is divided into the "student-test question factor embedded layer based on memory and figure attention" and the "student-test question factor interactive layer based on forgetting self-attention "Two components.

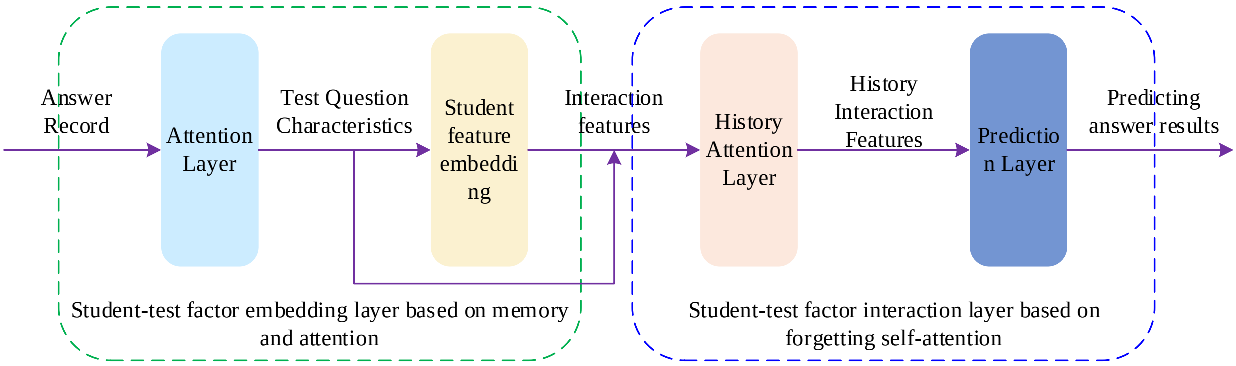


Figure 2. Learning cognitive diagnostic analysis model

3.2 Trial factors based on graph attention

The trial factors in ACD include trial difficulty h^{diff}, trial differentiation h^{disc}, and trial correlation w, which are common in traditional cognitive diagnostics. In common knowledge tracking models such as DKT, DKVMN, etc., the trial factors are usually only trial knowledge points, which is a more comprehensive and accurate diagnostic result in ACD compared with these models. In NeuralCDM, its test question knowledge point correlation Q_e comes from the expert labeling Q matrix, which contains only 0 and 1 and can only show whether the knowledge point is included in the test question or not but cannot distinguish the primary and secondary relationship between these knowledge points. At the same time, its test question difficulty factor is only related to the existence of corresponding knowledge points in the test questions, and no distinction can be made between test questions with the same knowledge points.

In this model, the test-question knowledge point correlations come not only from the Q matrix but from a graph attention network on a graph of correlations from the test-question knowledge points. In the K-layer structure of the graph attention network, each layer of the attention mechanism will learn to derive the correlation α_i^k of the test questions e_i with the knowledge points and derive the final weights w_i of the test knowledge points. Instead of consisting of 0s and 1s, as in the Q matrix, w_i in the model consists of a series of real numbers with a sum of 1s, where each dimension w_{ij} represents the weight of the corresponding knowledge point c_j in the test question e_i.

While obtaining the weights of the relevance of the knowledge points of test e_i, by obtaining the eigenvectors h'_{ei} of test e_i, from which the eigenvectors can be obtained the test difficulty factor h_i^{diff} and the test differentiation factor h_i^{disc}. where h_i^{diff} denotes the overall difficulty of test e_i, which is calculated by the formula:

$$h_i^{diff} = Sigmoid\left(h_{ei}' \times W_{diff}\right) \tag{10}$$

Where $W_{diff} \in \mathbb{R}^{d \times d}$ is the matrix of trainable parameters. The test question differentiation factor h_i^{disc} is a real number between $(0,1)$. It has the same meaning as in IRT and MIRT and indicates the ability of the test question to differentiate between students with a high level of mastery of the knowledge point and those with a low level of mastery of the knowledge point. Its calculation formula is shown in (11):

$$h_i^{disc} = Sigmoid\left(h'_{ei} \times W_{disc}\right) \tag{11}$$

Where $W_{disc} \in \mathbb{R}^{d\times 1}$ is the trainable parameter matrix.

3.3 Student Factors Based on Memory Networks

3.3.1 Memory reading process

As the most important component of cognitive diagnosis, the form of student factors directly determines whether the cognitive diagnostic model can accurately diagnose the students' knowledge mastery. In traditional cognitive diagnostic models, student factors can be in the form of unidimensional, multidimensional, discrete, and continuous quantities, each with its own advantages and disadvantages. Among the knowledge tracking models, the knowledge tracking models based on recurrent neural network RNN such as DKT, etc., the student's knowledge mastery is expressed in the form of potential vectors in RNN, which indicates the student's overall mastery of all the knowledge points through the high-dimensional vectors and lacks the embodiment of the specific knowledge points, and such as knowledge tracking based on self-attention such as SAKT, etc., whose student factors are hidden in the history of students' question answering records, and also cannot visualize students' mastery of knowledge points.

The ACD represents the students' mastery of all knowledge points in the form of a memory network. Memory matrix M^V is a $d \times |C|$-dimensional real matrix, where $|C|$ is the modulus of the set of all knowledge points, and each column of d-dimensional vectors represents the students' mastery of the corresponding knowledge points.

When the test question e_i is input, the relevant weights w_i of the knowledge points contained in the test question e_i can be obtained after the processing of the graph attention network. w_i is a $1\times|C|$-dimensional vector, where each dimensional w_{ij} value indicates the weight of the knowledge point c_j in the test question e_i, and thus the degree of mastery of the students on the knowledge points h_i^s contained in the current test question can be obtained by using Equation (12):

$$h_i^s = w_i \times M_{i-1}^V \tag{12}$$

Where M_{i-1}^V is the state that the student is in at the end of the $t-1$ moment, and h_i^s is a $1\times d$-dimensional vector that is the weighted sum of the relevant knowledge points in M_{i-1}^V.

After obtaining all the student factors and test factors, the student-test factor feature vectors can be constructed according to the form in IRT and NeuralCDM, and here the d-dimensional vector y_i is used to represent the interaction equation between the student and the test e_i, as shown in Eq. (13):

$$y_i = \left(h_i^s - h_i^{diff}\right) \times h_i^{disc} \tag{13}$$

3.3.2 Memory Renewal Processes

After obtaining the student factor h_i^s, the system needs to update the student's knowledge mastery M^V from moment $i-1$ to the latest moment i to ensure that the student's knowledge mastery remains up-to-date at moment $i+1$.

First, the tuple (e_i, r_i) of students' answers in the current test question is obtained based on their answer history, where r_i denotes the students' answers. Embed the tuple (e_i, r_i) as a $2|C|$-dimensional one-hot vector v_i and multiply it with the learnable parameter matrix W_g to embed it as a knowledge growth vector g_i, as shown in (14) in Eq:

$$g_i = W_g \times v_i \tag{14}$$

Where W_g is the learnable parameter matrix of $d \times 2|C|$ and g_i is a d-dimensional column vector. When updating the memory matrix M_{i-1}^V, similar to the forgetting gate in LSTM, the memory matrix M^V needs to delete a part of memories before updating and then add new memories. The formula for the forgetting vector z_i is shown in (15):

$$z_i = Sigmoid\left(W_z^T \times g_i + b_z\right) \tag{15}$$

Where W_z is the $d \times d$ learnable parameter matrix and z_i is the d-dimensional column vector. After obtaining the forgetting vector z_i, combined with the weights of the knowledge points obtained in the graph attention network w_i, the form of the memory matrix after forgetting is calculated $\tilde{M}_{i-1}^V$, and the formula is shown in (16):

$$\tilde{M}_{i-1}^V(j) = M_{i-1}^V(j) \circ \left[1 - w_{ij} z_i\right] \tag{16}$$

Where $M_{i-1}^V(j)$ denotes the j rd knowledge point in the M_{i-1}^V-matrix, is the product of elements. After going through the forgetting matrix, the knowledge points in the M_{i-1}^V matrix that are related to the ones in z_i will decrease, while the irrelevant knowledge points will remain unchanged. After going through forgetting, the learning vector a_i is obtained from the knowledge growth vector g_i. The formula is shown in (17):

$$a_i = \mathrm{Tanh}\left(W_a^T \times g_i + b_a\right) \tag{17}$$

Where W_a is the matrix of parameters that can be learned by $d \times d$. Afterwards, the newly learned content is added to the knowledge points corresponding to $\tilde{M}_{i-1}^V$ in combination with the knowledge point weights w_i.

4 Learning Diagnostic Analysis Example Exploration

The previous section proposed a cognitive diagnostic model based on graph attention on the basis of multidimensional feature coding for learning analysis. In this chapter, the model should be applied to a study of teaching and learning in the eighth grade of a middle school, and The Hook Theorem is used as the target content of cognitive diagnosis. Firstly, the cognitive attributes of The Hook Theorem are established as follows:

R1: The Collinear Theorem and basic applications.

R2: Inverse Theorem of the Pythagorean Theorem and its applications.

R3: Shortest path on a three-dimensional figure.

R4: The collinear theorem and equations.

R5: Combined application of the collinear theorem and the inverse theorem.

R6: Folding of graphs.

R7: Expansion of graphs.

R8: Equation thinking.

Based on the above cognitive attributes, test papers were developed to develop a cognitive diagnosis of learning. A random class was selected from one of the 8th-grade classes of the school, and the students of the class were randomly coded.

4.1 Cognitive attribute scores and parameterization

The test papers were recovered to obtain the students' scores, as shown in Figure 3. The process of test preparation was carried out under the guidance of the math teacher to ensure the scientific validity of the test. Excluding the number of absentees, a sample of 32 students in the class was selected for this study to take the actual test. A 0-1 scoring system was used for the evaluation, in which 0 points were awarded for incorrect answers or no answers, and 1 point was awarded for correct answers. The advantage of using 0-1 scoring is that the average score is the students' overall accuracy. Thus, the average accuracy of the eight cognitive attributes for which the collinearity theorem is easily obtained is [0.531, 0.406, 0.594, 0.563, 0.531, 0.500, 0.469, 0.344]. It can be initially judged that the property R8 is the least likely to be grasped by the students and is a difficult one to teach. This is because R8 represents the equation idea in the student's previous learning that has not really been exposed to not only the mastery of knowledge points but, more importantly, the formation of mathematical thinking. This is a learning difficulty for most students.

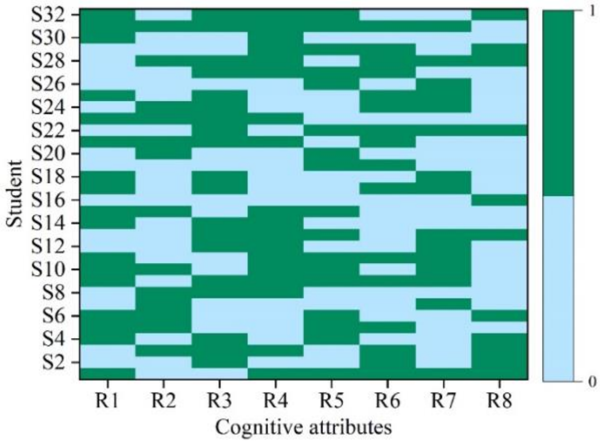


Figure 3. Student scores

In order to ensure that the developed test paper on The Hook Theorem meets the requirements of cognitive diagnostic analysis, this paper also uses a two-parameter logistic model to calculate the relevant parameters of the sample. The parameters of the test paper and answer scores of The Hook Theorem are shown in Figure 4. Where the parameter α indicates the differentiation and β indicates the difficulty. According to the requirements of cognitive diagnostic analysis, it is stipulated that the values of differentiation and difficulty of the items should be in (-3,3), and from the figure, it is known that the calculation of the parameters of the eight items meets the requirements, so the set assessment is reasonable.

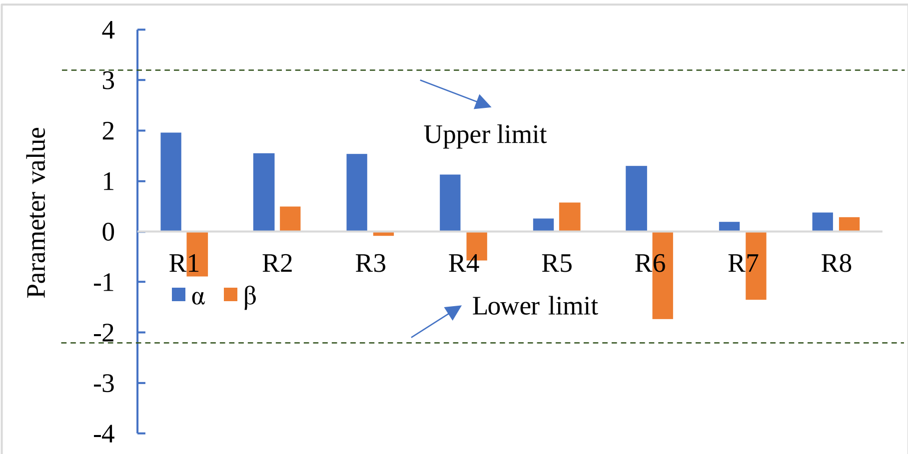


Figure 4. The question of the Pythagorean paper and the parameters of the answer

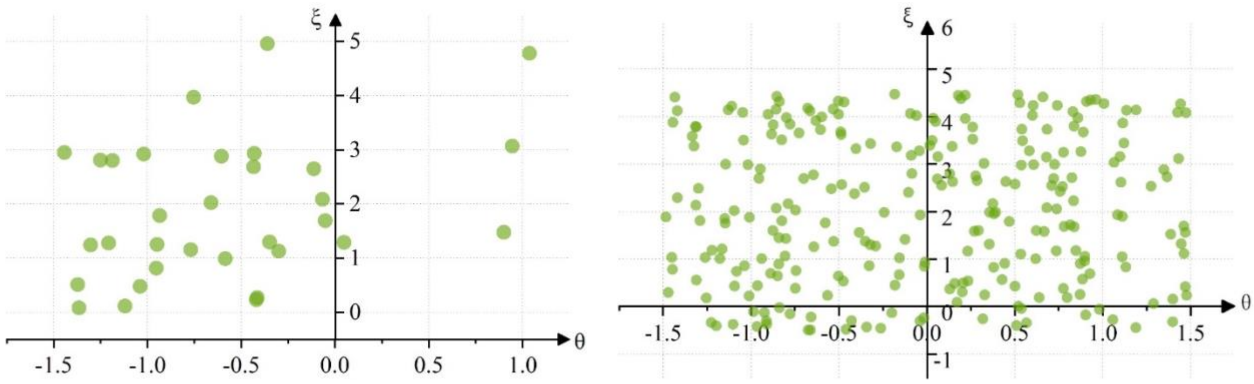
4.2 Ideal Learning Cognitive Rule Point Coordinates

According to the scores of cognitive attributes and parameter calculations, combined with the test factors and student factors given by the graph attention and memory networks, respectively, the cognitive rule point coordinates of the ideal model are solved as shown in Table 1, where θ denotes the value of ability and ξ denotes the degree of bias. According to the rule, point coordinates to map the actual cognitive rule points and subject responses for the whole school sample of eighth-grade students (the effective sample size is 256) to obtain the actual cognitive rule point distribution of The Hook Theorem and the distribution of the subjects are shown in Fig. 5(a) and (b), respectively. Compared with the ideal model, the distribution of rule points is more dispersed, indicating that the responses of items with different attribute mastery models vary widely. Based on the measurement of the distance from the ideal points, 238 of the 256 subjects selected were successfully categorized into the 15 ideal response patterns, and only 25 subjects could not be categorized, with a categorization rate of 92.97%, which is more than 90%. This indicates that the cognitive diagnostic evaluation of the content of the eighth-grade Hook's Theorem using an attention-based dynamic cognitive model showed validity on the mathematics test.

Among them, several ideal mastery patterns with the highest percentage were typical item response patterns such as (01011111), (01010111), (00011101), (00011111), and (00011111). Subjects belonging to the ideal attribute mastery pattern (01011111) only mastered the application of the collinear theorem and its inverse and the collinear theorem and its inverse, and mastery of the four attributes of shortest paths on three-dimensional graphs, the collinear theorem and equations, the folding and unfolding of graphs, and the idea of equations were lacking. Subjects in the Ideal Attribute Mastery Pattern (01010111) mastered the collinear theorem and its inverse, application of the collinear theorem and its inverse, and folding and unfolding of graphs but lacked mastery on shortest paths on three-dimensional graphs, the collinear theorem and equations. Subjects in the attribute mastery mode (00011101) mastered the collinear theorem and its inverse theorem, shortest paths on three-dimensional graphs, applications of the collinear theorem and its inverse theorem, and folding and unfolding of graphs but did not master the attributes of the collinear theorem with equations and equation thinking. Subjects in the desirable attribute mastery pattern (00011111) mastered the collinear theorem and its inverse, the collinear theorem with equations and equational ideas, the application of the collinear theorem and its inverse, the two attributes of shortest paths on three-dimensional graphs and the folding and unfolding of graphs were not mastered, the determination of parallelograms, and the integrated application of parallelograms. Students belonging to the 27th and 31st attribute mastery patterns accounted for 7.21% and 4.25% of the total number of students, respectively. These students belonged to the group of students who had a better mastery of the collinearity theorem and its inverse and its application and who were also able to meet the objectives of the Mathematics Curriculum Standards for the students and the requirements of the questions on the midterm examination for the students.

Table 1. The ideal pattern of cognitive rule point coordinates

Ideal	R1	R2	R3	R4	R5	R6	R7	R8	θ	ξ
S1	0	1	0	1	1	1	0	1	-0.349	1.298
S2	0	0	0	0	0	0	0	1	-0.585	0.991
S3	1	0	0	0	0	0	1	1	-1.12	0.115
S4	1	1	0	1	0	0	1	1	-1.443	2.95
S5	0	1	0	0	0	1	1	1	-1.207	1.279
S6	1	1	1	1	0	1	1	0	-0.935	1.786
S7	0	0	1	1	1	0	0	0	-0.3	1.128
S8	0	0	0	0	0	0	1	1	-0.113	2.649
S9	0	1	0	0	0	0	1	0	-1.303	1.244
S10	0	0	0	1	1	1	0	1	-0.434	2.686
S11	0	1	0	1	1	1	1	1	0.945	3.068
S12	0	1	1	1	1	0	1	0	-1.365	0.083
S13	1	0	1	0	0	0	1	1	-0.361	4.959
S14	0	1	0	1	0	1	1	1	-0.661	2.025
S15	1	0	0	1	1	1	0	1	-0.431	2.932
S16	0	1	1	1	1	0	0	1	-0.415	0.272
S17	0	0	0	0	1	1	0	1	0.9	1.477
S18	0	0	0	1	1	1	0	1	-1.187	2.803
S19	0	1	1	0	0	0	0	1	-0.949	1.252
S20	0	1	1	0	1	1	0	0	-0.42	0.231
S21	0	0	0	1	1	1	1	1	-0.754	3.971
S22	0	0	0	1	1	1	0	0	-0.769	1.156
S23	1	1	1	0	1	1	1	0	-1.039	0.479
S24	0	0	1	0	1	0	1	1	-0.052	1.693
S25	0	1	0	0	1	0	0	1	0.048	1.291
S26	0	1	0	1	1	1	1	0	-1.251	2.811
S27	1	1	1	0	1	0	1	0	-1.372	0.512
S28	1	1	0	0	1	0	0	1	-0.606	2.881
S29	0	0	0	1	0	0	0	0	-1.02	2.921
S30	1	0	1	1	0	0	0	0	-0.953	0.817
S31	1	1	1	0	1	1	1	0	1.037	4.783
S32	0	0	0	0	1	1	0	0	-0.067	2.086



(a) Actual cognitive rule points (b) The point of the subect

Figure 5. Cognitive rule point distribution

4.3 Learning Characterization

In order to study more deeply the interaction characteristics of students' learning in different courses, the 8th-grade courses of the school were still used as an example. Statistically, the top 2 courses with the most knowledge units, C1 and C2, were selected. Table 2 shows the calculated values of the correlation of the learning interaction activities of C1 and C2 courses; the lower-left distribution value is the activity correlation of C1, and the upper-right distribution value is the activity correlation of C2. The Homepage and Forumng correlations of C1 and C2 are both extremely strong, with correlations of 0.852 and 0.813, respectively, which is a clear positive correlation. Positive correlation. Subpage and Resource have extremely strong positive correlations for the C1 course and strong positive correlations for the C2 course. The distribution of courses in the same cycle of the eighth grade in terms of participation in learning interactions is approximate, with scatter matrix and correlation coefficient values indicating positive correlations of varying strength.

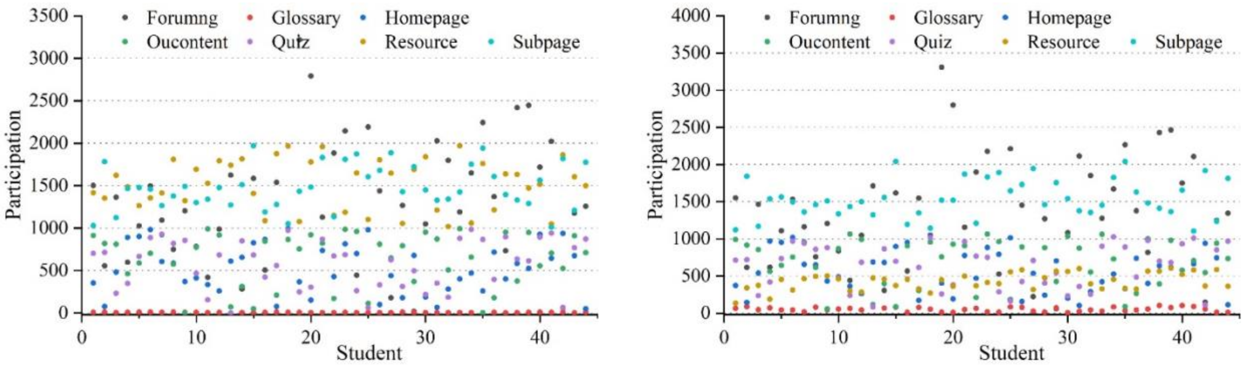
Table 2. The study of the interaction activity of C1 and C2 courses

C1¥C2	Forumng	Glossary	Homepage	Oucontent	Quiz	Resource	Subpage
Forumng	1	0.265	0.813	0.579	0.486	0.627	0.392
Glossary	0.722	1	0.368	0.575	0.4	0.519	0.523
Homepage	0.852	0.535	1	0.478	0.355	0.641	0.517
Oucontent	0.337	0.378	0.344	1	0.42	0.349	0.315
Quiz	0.283	0.593	0.280	0.685	1	0.684	0.733
Resource	0.562	0.746	0.730	0.732	0.615	1	0.36
Subpage	0.646	0.632	0.321	0.426	0.734	0.409	1

From the results of the correlation analysis of learning interaction activities, it can be seen that the same course has a certain approximation between different weekly periods, and there is also an approximate correlation between courses of the same category in the same cycle for learning interaction activities. Based on the correlation value, the analytical prediction of learning interaction activities is realized.

4.4 Effectiveness of the application of learning analytics

The random field collaborative inference strategy for learning interactions in Section 4.3 is applied to the actual online teaching and learning process. Course C3, which is close to courses C1 and C2, was chosen as the learning content, and two semesters were used as the statistical cycle of learning behaviors. Since the teaching goal of the course is to improve the learners' ability to combine theory and practice, the course design of the experimental process is incorporated into the delivery process, and the learning process takes place by both offline and offline means. In order to strengthen the communication of the course learning process, the interactive and collaborative part of the learning process is set as an online operation, the online interactive mode of the course is created, and the course web disk is designed to realize the interaction between learners and the web disk. The auxiliary materials related to the course experiments, etc., are realized to be managed online with the learning interactive activities of Content, Dualpane, and Dataplus, respectively. The assessment of the course consists of three parts: the final quiz, the daily project, and the course paper. The distribution of participation in learning interaction activities in two semesters is shown in Fig. 6, (a) and (b) for the first and second semesters, respectively. It is clear that the scatter distribution is more active, and the participation of Forumng and Quiz is fully mobilized. Among them, Forumng's interactive participation degree from the first semester to the second semester increased by 86.5 on average. The motivation of interactive learning activities' random field collaboration prediction strategy makes the participation in different semesters of the same learning content change more obvious. Strengthening the evaluation measures of the learning process and the mechanism of interaction and collaboration has a better The enhancement of evaluation measures and interactive and collaborative mechanisms in the learning process has a better effect on improving learners' interest in participation and sense of independent learning.



(a) First semester (b) Second semester

Figure 6. Learning interactive activity participation distribution

5 Conclusion

This study successfully implemented an educational application study based on diagnostic learning analytics technology, demonstrating the efficiency and accuracy of this method in the field of education. The cognitive diagnostic model constructed was able to accurately predict students' knowledge mastery status and future performance, with an accuracy rate as high as 92.97% in the experiment, which far exceeded that of traditional teaching methods. This result shows that the learning analysis method integrating multidimensional features can provide educators with more comprehensive and in-depth information about students' learning status.

The dynamic diagnostic model constructed by using graph attention network and memory network technology can effectively deal with the complex interaction between students and test questions, making learning diagnosis more refined and personalized. Especially when dealing with students' long-term learning data, the model can dynamically reflect the changes in students' knowledge mastery and provide educators with more effective teaching strategies.

The study demonstrates the practical effects of diagnostic learning analytics in education in application examples. Through the diagnostic analysis of the learning of the knowledge points of the eighth-grade Theorem of the Hook and Strand Theorem, the study reveals the students' mastery level and learning difficulties in each knowledge point, which provides teachers with targeted teaching references. Meanwhile, by analyzing learners' multidimensional interactions, the study reveals the correlation between different learning activities, which has a positive effect on improving learners' participation and awareness of independent learning.

References

[1] Lin, Y. T. (2019). Impacts of a flipped classroom with a smart learning diagnosis system on students' learning performance, perception, and problem solving ability in a software engineering course. Computers in Human Behavior, 95(JUN.), 187-196.

[2] Safadi, R. S. (2021). Learning from self-diagnosis activities when contrasting students' own solutions with worked examples: the case of 10th graders studying geometric optics. Research in science education, 51(2).

[3] Potter, & Carol, A. (2017). "i received a leaflet and that is all": father experiences of a diagnosis of autism. British Journal of Learning Disabilities, 45(2), 95-105.

[4] Reinstein, I., Hill, J., Cook, D. A., Lineberry, M., & Pusic, M. V. (2021). Multi-level longitudinal learning curve regression models integrated with item difficulty metrics for deliberate practice of visual diagnosis: groundwork for adaptive learning. Advances in Health Sciences Education, 1-32.

[5] Albert, M., & Uhlig, M. (2022). Education for sustainable development at chemnitz university of technology. International journal of sustainability in higher education.

[6] Stork, M. G., & Heller, M. F. (2017). Is technology good for education?. Journal of Educational Research.

[7] Jeffrey, B., Niall, S., Donal, C., & Lena, G. (2018). Visualization, inductive reasoning, and memory span as components of fluid intelligence: implications for technology education. International Journal of Educational Research, 90, 64–77.

[8] Kure, A. E., Brevik, L. M., & Blikstad-Balas, M. (2023). Digital skills critical for education: video analysis of students' technology use in norwegian secondary english classrooms. Journal of computer assisted learning.

[9] Thompson, K., & Lodge, J. M. (2020). 2020 vision: what happens next in education technology research in australia. Australasian Journal of Educational Technology(4).

[10] Alcala, D. H., Garijo, A. H., Villora, S., & Extremera, A. (2020). "cooperative learning does not work for me": analysis of its implementation in future physical education teachers. Frontiers in Psychology.

[11] Lin, Y. L., & Wang, W. T. (2023). Analysis of the social interaction of perceived problem-based learning performance in internship courses. Journal of computer assisted learning.

[12] Schneider, S., Beege, M., Nebel, S., & Rey, Günter Daniel. (2017). A meta-analysis of how signaling affects learning with media. Educational Research Review, 23, 1-24.

[13] Chang, M. M., & Hung, H. T. (2019). Effects of technology-enhanced language learning on second language acquisition: a meta-analysis. Educational Technology & Society, 22.

[14] Disessa, A. A. (2017). Conceptual change in a microcosm: comparative learning analysis of a learning event. Human Development, 60(4), 201-201.

[15] Williamson, B. (2020). Education technology seizes a pandemic opening. Current history (New York, N.Y.: 1941), 120(822), 15-20.

[16] Sun, X. (2022). Application of data mining technology in college mental health education. Frontiers in psychology, 13, 974576.

[17] Lee, J. Y., & Park, S. (2017). Analysis of critical success factors of online international learning exchange of korean school pupils with english-speaking counterparts. British Journal of Educational Technology.

About the Author

Naimin Zhang, Graduated from Northeast Electric Power University in 2014. Work at Hebei Oriental University.Engaged in the research of internet of things and automation.

Linlin Zhang, Graduated from Chang'an University in 2016. Work at Hebei Oriental University. Engaged in the research of logistics and transportation.