# A Study on Distance Personalized English Teaching Based on Deep Directed Graph Knowledge Tracking Model

Lianmei Deng
Department of Economics and Trade, Chongqing College of Finance and Economics, Chongqing, 402160, China

***Abstract*—Despite the continuous development of online education models, the effectiveness of online distance education has never been able to meet people's expectations due to individual difference of learners. How to personalize teaching in a targeted manner and stimulate learners' independent learning ability has become a key issue. In this study, the multidimensional features of the learning process are mined with the help of the BORUTA feature selection model, and the DKVMN-BORUTA model incorporating multidimensional features is established. This optimized deep knowledge tracking method is combined with graph structure rules. Then, an intelligent knowledge recommendation algorithm based on reinforcement learning is used to construct a fusion approach-based model for distanced personalized teaching and learning of English. The results show that the research proposed fused deep-directed graph knowledge tracking with graph structure rules for remote personalized English teaching model has the lowest AUC value of 0.893 and the highest AUC value of 0.921 on each dataset. The prediction accuracy of the research model is 94.3% and the F1 score is 0.92, which is the highest among the studied models, indicating that the proposed model has a strong performance. The fusion model proposed in the study has a higher accuracy rate of knowledge personalization recommendation than the traditional deep knowledge tracking model, and it can help learners save revision time effectively and improve their overall English performance.**

***Keywords*—Distanced personalized English teaching model; knowledge tracking; deep learning; graph structure rules; DKVMN***

## I. Introduction

With the booming development of internet information technology, online educational resources are shared and disseminated nationwide [1]. Online distanced English teaching is particularly outstanding in terms of resource sharing and students are very fond of being taught by foreign teachers via online interactive communication. However, as the online teaching method is quite different from the traditional face-to-face teaching method, it put forwards higher requirements for students' independent learning ability and self-motivation. Therefore, a more scientific and effective way is required to track students' status in the classroom, attract students' attention, achieve knowledge tracking, and improve the quality of distanced English learning [2-3]. In previous studies, some scholars have proposed to apply deep learning to the prediction of knowledge and to build Deep Knowledge Tracing (DKT) models by combining Long Short-Term Networks (LSTM) with history department sequences [4-5]. However, the DKT model is not very accurate in analyzing students' learning data and predicting knowledge, and its performance is weakly stable. At present, the improvement of distance personalized teaching quality mainly depends on the knowledge tracking of learners. To accurately analyze the knowledge state of learners, many technologies have been applied in the field of knowledge tracking. Among them, the most advanced and widely used ones mainly include two types, namely, the knowledge tracking method based on probability graph and the knowledge tracking method based on deep learning. However, the knowledge tracking method based on probability graph lacks the consideration of the relationship between learner forgetting factors and knowledge, so its accuracy is low. However, the knowledge tracking method based on deep learning is difficult to obtain the specific information of learners' knowledge state. For this reason, the study optimizes the DKT model based on data mining and improves the DKT model in terms of multidimensional features of the learning process and feature fusion. The optimized model is combined with graph structure rules, and then the distanced personalized English teaching model based on deep directed graph knowledge tracking is proposed. This paper first reviews the application of deep learning model in teaching, and then summarizes the existing defects of current methods through the study of previous methods. Then, the research methods are elaborated, and the deep learning knowledge tracking model is proposed, which is applied to English distance personalized education. Secondly, in the part of result analysis, the research algorithm is compared with the current more advanced knowledge tracking algorithm, and the effectiveness and accuracy of the research algorithm are verified. Finally, in the conclusion part, the advantages and disadvantages of the research algorithm are summarized, and the future research direction is proposed. It aims to precisely locate the learning status of learners and their mastery of knowledge points so as to achieve good results of distanced personalized English teaching.

With the rapid development of deep learning and neural network algorithms, the role of computer technology in the teaching and learning process has been explored, and scholars have begun to focus on related topics, and substantial progress has been achieved. GERVETT et al. applied deep learning DKT models to knowledge tracking and analyzed the conditions under which their deep models can make the most accurate predictions [6]. SAPOUNTZIA et al. analyzed the Bayesian Knowledge Tracing (BKT) model, the DKT model, and the Dynamic Key-Value Memory Network (DKVMN) model from a technical and educational perspective and compared the modeling techniques, evaluation, and performance in terms of optimization of the three models, and the results showed superior performance of the DKVMN model [7]. Hassan et al. focused on student dropout risk prediction by using deep long and short-term memory models. Multidimensional data can help teachers predict students' learning status, and comparative testing results on real datasets show superior application performance of the model over logistic regression and artificial neural networks [8]. Mubarak et al. analyzed the role of video clickstream data in prediction of online learners' learning performance, and deep neural network algorithms incorporating implicit features were found to have excellent performance in real-world applications, which outperformed super vector machines and logistic regression methods [9]. Kaser et al. proposed a dynamic Bayesian knowledge tracing (DBKT) model that uses DBNs to combine multiple knowledge tracing (KC) in one model, and the results showed that the model had validity and strong contractual performance [10].

DEONOVICB et al. proposed a method to improve the prediction accuracy of knowledge tracking by combining probability maps and cognitive diagnostic techniques in order to improve the performance of knowledge tracking, and the experimental results showed that the method effectively improved the performance and prediction accuracy of knowledge tracking [11]. Yang et al. optimized the deep knowledge tracking technique in tracking students' knowledge states. They enriched the original deep knowledge tracking model by adding heterogeneous feature implicatures to determine the probability of students' correct answers in the exercises. The study evaluated the optimized model using two different education-related datasets, and the evaluation results indicated the superior performance of the model in the relevant domain [12]. Wang et al. designed the DKTS method using the feature relationship between topics and the linkage of knowledge points, and the experimental results showed that the method is effective in improving the service performance of the knowledge tracking method [13]. Vie et al. concluded that tracking the evolution of students' knowledge can help teachers' instructional optimization, and factor decomposer can be used as a regression and classification model. The findings show that the model can handle multi-feature high-dimensional student learning data in a large number of real datasets, and it has significant superiority over existing models [14]. This is the first work to integrate competency-based tracking into MOOC course recommendations. Extensive experiments on real-world datasets demonstrate that capacity tracking-enhanced course recommendations improve the effectiveness and interpretability of MOOCs. Tian et al. made an attempt to integrate capacity tracking into course recommendations for MOOCs. Experiments on a large amount of real-world data demonstrate that deep knowledge tracking can improve the validity and interpretability of MOOCs [15].

In summary, deep learning and neural network algorithms have been widely used in education and a large number of research findings have enriched current educational approaches, where deep knowledge tracking has also been widely used to track student knowledge evolution. However, there has been no reports investigating directed graphs as an input data source. Therefore, the research will focus on the construction of graph attention neural networks on this basis to obtain a student knowledge tracking model with better performance.

## II. Deep Directed Graph Knowledge Tracking Model based on Personalized English Distance Teaching

### *A. Design of a Deep Learning Knowledge Tracking Model based on Multidimensional Feature Fusion*

Knowledge tracking automatically tracks changes in the state of knowledge of learners based on their historical learning trajectories and predicts learners' mastery of knowledge in the future process [16]. The study describes a learner's answer interaction as $R_i = \{s_i, q_i, r_i, b_i\}$ , where $b_i$ is a sequence of learning behaviors. Considering the large number of feature dimensions and the need for knowledge tracking methods to mine the historical learning data collected by the learning platform, this research proposes a knowledge tracking method incorporating multidimensional features of the learning process, as shown in Fig. 1.

As shown in Fig. 1, the knowledge tracking framework consists of data composition, knowledge tracking optimization, and knowledge tracking model. Among them, the answer history and other data constitute the data framework, and the knowledge tracking optimization includes data mining and improvement methods for the deep knowledge tracking model to optimize the model and improve the prediction accuracy and interpretability of the model. Since deep knowledge tracking and other learning models have the problem of weak interpretability, some studies have pointed out that the interpretability can be divided into Ante-hoc and Post-hoc perspectives. Based on Post-hoc as the perspective for global interpretation, the multidimensional features that affect the learning results are explored from many features of the learning process, and then redundant features are removed to complete the input composition of the model, i.e., the BORUTA feature selection method [17]. Moreover, the knowledge tracking model mainly utilizes the Dynamic Key-Value Memory Networks (DKVMN) model to ensure the performance and interpretability of the method [18]. Since DKVMN has a strong expansion performance, the core of the research is the optimization of the deep knowledge tracking model incorporating multidimensional features, as shown in Fig. 2.

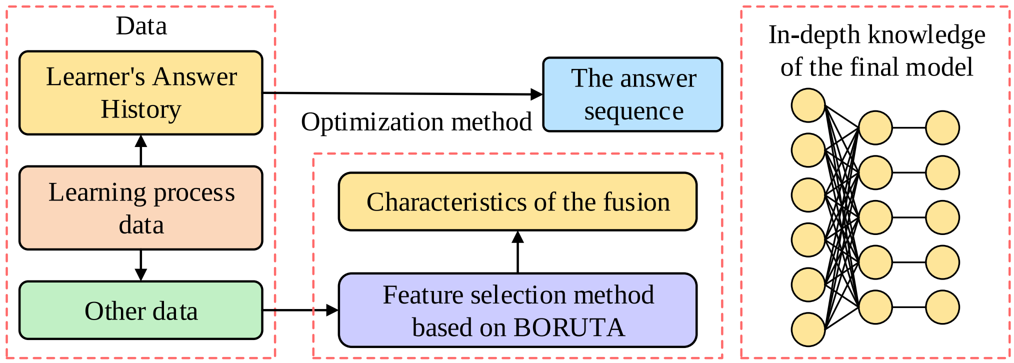



Fig. 1. Knowledge tracking framework based on learning process characteristics.

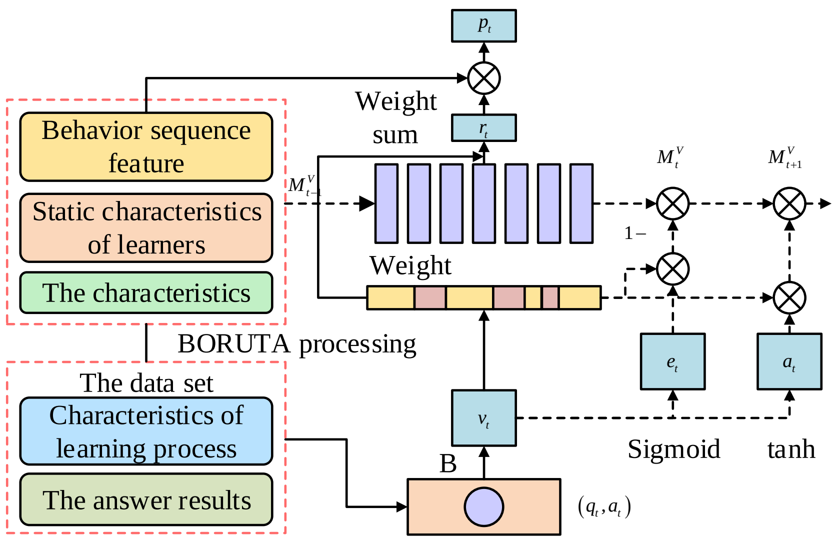



Fig. 2. Depth knowledge tracking model with multi-dimensional features.

As shown in Fig. 2, the processed data is used as the input of the model, and three networks, namely learning behavior sequence features, learner static features, and topic features, are built at the output location of this model, and then the optimization method is incorporated into the DKVMN framework, i.e., the deep knowledge tracking optimization method DKVMN-BORUTA is obtained .The study obtains the Embedding matrix of the topic by constructing its vector representation as in equation (1) for the knowledge concept weights[19].

$$w_t(i) = soft\max\left(M_i^k \cdot k_t\right) \tag{1}$$

Where, the problem $q$ encountered by the learner at the moment $t$ is first multiplied by the embedding matrix $A$ that has been trained to obtain the embedding vector $k_t$. Then $k_t$ is computed by $soft\max$ to obtain the attention weight vector $w_t$. The ease of expansion of DKVMN creates the conditions for the incorporation of multidimensional features in the learning process. Specifically, the DKVMN model receives learning records that will have some impact on the learner's feature vector $f_t$. Since $f_t$ is obtained by fusing the read vector $r_t$ with the embedding vector $k_t$, it contains both information about the state of the learner's knowledge of the topic $q_t$ and the embedding information of $q_t$. By processing the Fourier transform through the neural network, it is possible to infer the learner's ability on $q_t$. And the difficulty level of $q_t$ can be obtained by passing $k_t$ to the neural network. Equation (2) is the formula for the read vector of. $r_t$

$$r_t = \sum_{i=1}^{N} w_t(i) M_t^v \tag{2}$$

Where, $M_t^v$ represents the read value memory. The BORUTA features are processed through the embedding matrix to obtain the vector representation and then the embedding matrix is spliced with the topic content feature vector $v_t$. Eq. (3) is the formula for calculating the probability of a learner correctly answering the question $q_t$ $p_t$.

$$p_t = \varnothing\left(\tanh\left(W_O^T\left[r_t, m_t, l_t, v_t\right]\right) + b_{it}\right) \quad (3)$$

Where, $l_t$ denotes the learner static features, $m_t$ denotes the topic feature limit. $b_{it}$ represents the conditioning factor, and $\varnothing$ is the Sigmoid activation function. Where $b_{it} = Sigmoid\left|b_{it}\right|$, $b_t$ is calculated by first mining the effective learning behavior through the learning behavior sequence feature vector. If the learner first watched the video and then discussed in the discussion forum and watched the learning video to answer the question, the learning behavior sequence of this learner is $b_t = (2,1)$. The overall vector is a fusion of the BORUTA feature set with the current topic content features and the learner's knowledge state. The knowledge state matrix of the deep knowledge tracking optimization model is updated by inputting answer records $(q_t, a_t)$ and $w_t$ to collaboratively update $M_t^v$, and by erasing weights $E$ and adding weights $D$, respectively, so as to obtain the memory erasure vector $e$ and the memory addition vector $a$. The memory erasure vector is calculated by equation (4).

$$e_t = \text{Sin}\, gmoid\left(W_e v + t\right) \quad (4)$$

Where, $W_e$ represents the erasure weights and $t$ represents the discretized value of the learner's time to deal with the problem. The memory addition vector is calculated by Eq. (5).

$$a_t = \tanh\left(W_a v_t + t\right) \quad (5)$$

Where, $W_a$ stands for adding weights. The process of dynamic update of the learner's knowledge state is expressed by Eq. (6).

$$M_{t+1}^v = M_t^v\left[1 - w(i)e_t\right]\left[1 + w(i)a_t\right] \quad (6)$$

Where $x_t = (q_t, a_t)$ represents the learner's answering behavior after the $t$ moment and the value of the dynamic matrix $M$, which is transformed from $M_t^v$ to $M_{t+1}^v$. The goal of model optimization is to minimize the difference between $p_t$ and $a_t$ with the minimum spread loss function. The model optimization is conducted using the momentum gradient descent method, as shown in equation (7).

$$L = -\sum_t\left[a_t \log p_t + (1 + a_t)\log(1 - p_t)\right] \quad (7)$$

In this study, the BORUTA feature selection model is used to explore the multi-dimensional features in the online learning process. Based on DKVMN model, the multi-dimensional feature network is constructed. Secondly, the deep knowledge tracking model with multi-dimensional learning features is designed. Finally, a deep knowledge tracking optimization model is constructed.

### B. A Model Design for Distanced Personalized English Learning incorporating Deep Knowledge Tracking and Graph Structure Rules

In the design process of this research model, the deep knowledge tracking model was first introduced, allowing the construction of a directed graph of knowledge concepts for the probability values of the outcomes, and then directed graph was fed into the deep learning neural network as a data source, which in turn formed the attention neural network [20]. The probability value of the association relationship between concepts is actually the weight value of the network nodes. The structure of the deep directed knowledge tracking model is shown in Fig. 3.

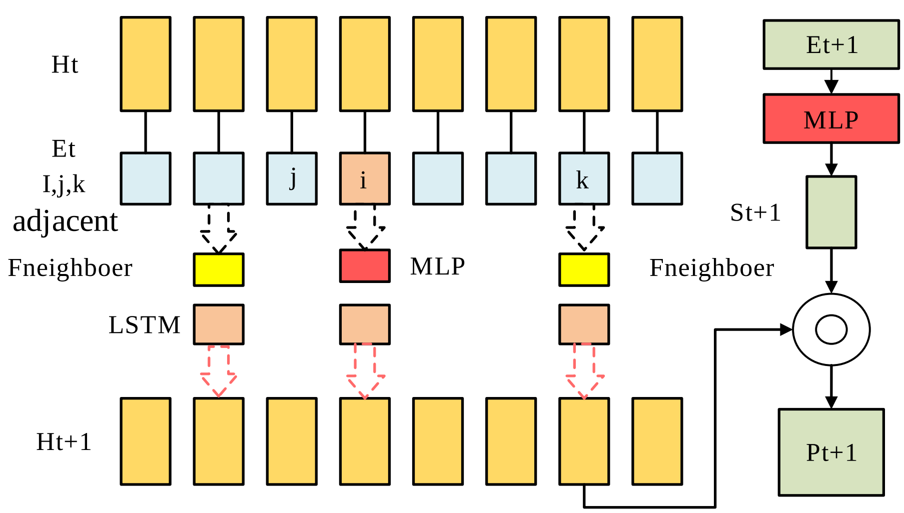



Fig. 3. Overall structure of knowledge tracking model integrating depth directed graph.

As shown in Fig. 3, the model first undergoes a Manifold Embedding popular embedding, setting up the memory linear matrix $H$, which holds mainly the basic state of the knowledge concept, and the corresponding $N$ vector values of relatively independent concepts related to the problem. For the vectors, $N\ d$ are used to represent their dimensions. The memory values in the matrix $H$ at the time of $t$ are $H_t$, and $E$ is used to identify a matrix of processing result values after popular embedding. In the model processing, the extremely strong answer to the question at the $t$ moment is processed as a one-bit valid code of length $2N\ x_t$, and the code is processed using the embedding matrix of $2N*e$ to downscale the dimensionality to a vector of length $e$. For example, equation (8) is the vector matrix of. $N*e$

$$E_k^t = \begin{Bmatrix} xtEx(k=i) \\ Ec(k)(k \neq i) \end{Bmatrix} \tag{8}$$

Where $x_t$ represents the heat code and $e_x$ represents the vector value. From Fig. 3, it can be seen that the model input value at the moment of $t$ is $(q_t, a_t)$, and after that knowledge concept is defined as $i$, the directed graph connecting the adjacent nodes of $i$, $j,k$ and the vector corresponding to $i,j,k$ in the matrix of synchronous negative memory concepts $h$ is established, as shown in Eq. (9).

$$H_t = [H_t, E_t] \tag{9}$$

Where the vector $h_t^i$ is spliced with the vector $e_t^i$, processed by a fully connected neural network containing hidden and input layers, and then fed into the LSTM network to obtain the newly generated memory matrix, where the $i$-th vector expression is shown in Eq. (10)[21].

$$H_k^{t+1} = \begin{cases} RNN\left(f_{MLP}\left(h_k^t\right)\right) \\ RNN\left(f_{neighbor}\left(h_i^t, h_k^t\right)\right) \end{cases} \tag{10}$$

Where $RNN$ represents the fully connected recurrent neural network, $h_k^t$ represents the vector values at the moment, and $f_{MLP}$ is the LSTM network. The process focuses on the nearest neighbor function of vector $j$ and vector $k$, and this attention algorithm is formulated in Eq. (11).

$$f_{neighbor}\left(h_i^t, h_k^t\right) = \frac{1}{K}\sum_{k\in K} a_{ij}^k f_k\left(h_i^t, h_k^t\right) \tag{11}$$

Where $a_{i,j}^k$ represents the threshold value for simplifying the graph structure. The study uses a deep knowledge tracing model to obtain vector graphs, which are acyclic structural graphs with positive-valued full connectivity between the nodes of the graph, and a threshold parameter will be introduced to simplify the graph structure. When the topic concept at the moment $t$ is $K$, only the vector $i$ adjacent to it needs to be considered, and the graph model is defined using the connection relations between directed graphs. Finally, the model needs to encode the knowledge concepts. The knowledge concept sailing in at $t+1$ is set to $e_{t+1}$, and forward neural network processing is performed to obtain $S_{t+1}$. This encoding long queue is $d$, which is calculated by equation (12).

$$S_{t+1} = \sigma\left(W_1 \cdot \sigma\left(W_O \cdot e_{t+1} + b_o\right) + b_1\right) \tag{12}$$

Where the forgetting gate in the structure of the feedforward neural network is represented with the previous memory gate input parameter [22]. At the moment of $t+1$, the memory matrix vector $H$ is multiplied by the vector points of the knowledge encoding $S$ to obtain the true mastery probability value for moment $P_{t+1}$. The study developed a methodological model of reinforcement learning by mimicking the teacher's learning behavior towards the students and using a deep knowledge tracking model process sample training. In the process of using the knowledge tracking model, the recommended questions are first selected and then the input of the questions is completed. The predicted answers are obtained and the concept of knowledge is encoded, and this encoding is used as the input learning state vector value for training the reinforcement learning method. The model structure is shown in Fig. 4.

As shown in Fig. 4, when the algorithm is executed, a random state vector of fixed length is first initialized and the basic level of knowledge point concept mastered by the student is identified, where each element of the vector is labeled as the degree of learning to the knowledge value. On this basis, a convolutional neural network is built. Then, the process of learning action selection is executed, where three outputs exist in this process. When the return value is 1, then the learning is sufficient and no subsequent selection of knowledge-related topics is required. When the return value is 0, the old and new states, actions and their corresponding reward values need to be combined and saved to the experience pool according to the fixed collocation principle as the basis for subsequent learning and optimization. Continuous action selection for the immediate moment is performed, and the deep knowledge tracking model is trained for deep reinforcement learning after several iterations of the loop. The reinforcement learning method designed for this research uses the Nature DQN (Deep Q Network, DQN) network structure [23]. Two neural networks with the same number of layers and nodes were set up to address the correlation between the data samples and the network before training. One of the $Q$ networks is used as the training network with the input value of the state value of the environment mention training and the output corresponding to it is the recommended action. The second $Q$ network is used as the target network to reduce the overfitting of the neural network, so that this network serves as the final target value for reinforcement learning, thus enabling an optimal update of the corresponding network parameters. The study evaluates the model in terms of average accuracy, average completeness and accuracy, where the average completeness $(AP)$ is calculated by equation (13).

$$AP = \sum_{u=1}^{r} P(u) \tag{13}$$

Where $r$ denotes the number of categories and $u$ denotes the category labels. The average finding rate $(AR)$ is expressed by Eq. (14).

$$AR = \sum_{u=1}^{r} R(u) \tag{14}$$

The accuracy of $(ACC)$ is calculated by Eq. (15) [24].

$$ACC = \frac{(TP + TN)}{(TP + FN + FP + TN)} \tag{15}$$

Where $TP, FP$ denote the number of positive classes and the number of negative classes for positive class prediction, respectively. $FN, TN$ denote the number of positive classes and the number of negative classes for negative class prediction, respectively. In addition, the study determines the performance of the model by comparing the Area Under Curve (AUC) of each dataset completion curve.

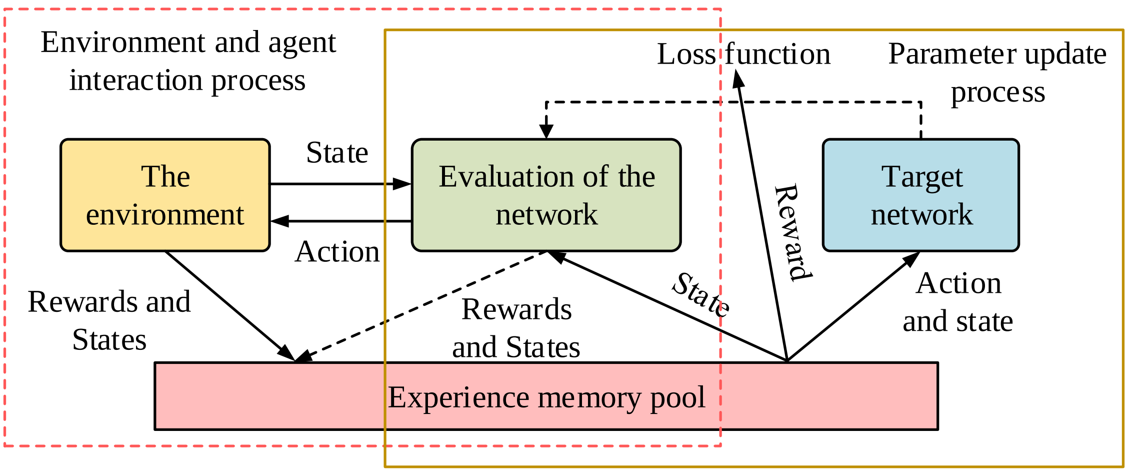



Fig. 4. Structure diagram of training reinforcement learning method model.

## III. Analysis of Experimental Results of a Deep Directed Graph Knowledge-Tracking-Based Model for Distanced Personalized English Teaching

### A. *Analysis of the Effectiveness of a Deep Learning Knowledge Tracking Model based on Multidimensional Feature Fusion*

The validity, interpretability analysis of the deep knowledge tracking model incorporating multidimensional features was conducted. In this experiment, the Adam optimizer was used to train the model, and all datasets were sliced 7:3 to test set and training set with 10 training times. Based on tensor flow and keras, the methods in Table I were completed in NVIDIA 1080 Ti GPU environment. At present, knowledge tracing methods widely used include BKT (Bayesian knowledge tracing), DKT (Deep knowledge tracing), and DKVMN (Dynamic key tracing) value memory networks) these three knowledge tracking methods. BKT can construct the knowledge state of learners as a set of binary variables, but its prediction accuracy depends on the experience of the teacher, and the degree of automation is low. DKT is the current mainstream knowledge tracking method without a lot of teaching experience and manual labeling, but it is difficult to obtain more specific knowledge state of learners. DKVMN can solve the problems existing in BKT and DKT with high prediction accuracy, but it ignores the factors and characteristics existing in the learning process of learners. AUC (Area Under Curve) ranges from 0.5 to 1.0. The closer the AUC is to 1.0, the higher the authenticity of the detection method. Table I shows the AUC values of multiple algorithms on 6 datasets.

As can be seen from Table I, the DKVMN-BORUTA model proposed in the study has an AUC value of around 0.83, which is higher than that of other three models, indicating the superior performance of the studied model. Since the model incorporates multidimensional features in the learning process, it can to some extent solve the problem of traditional knowledge tracking models for modeling simple mathematical logic in learning and achieve better simulation of the real state of the learner as well as better assessment of the learner's knowledge state.

TABLE I. AUC VALUES OF FOUR METHODS ON DATASETS

| Data set | Number of students | Knowledge tag | AUC | | | |
|---|---|---|---|---|---|---|
| | | | BKT | DKT | DKTVMN | DKVMN-BORUTA |
| **ASSIST2014** | 4208 | 112 | 0.632 | 0.806±0.02 | 0.815±0.01 | 0.832±0.01 |
| **ASSIST2015** | 19850 | 110 | 0.646 | 0.726±0.01 | 0.748±0.01 | 0.804±0.02 |
| **Statics2011** | 334 | 158 | 0.71 | 0.803±0.02 | 0.799±0.02 | 0.832±0.01 |
| **Synthetic** | 2000 | 6 | 0.761 | 0.804±0.01 | 0.806±0.01 | 0.841±0.02 |
| **Khan Academy** | 17825 | 108 | 0.689 | 0.802±0.02 | 0.814±0.02 | 0.822±0.01 |
| **KDD Cup** | 3996 | 180 | 0.743 | 0.783±0.01 | 0.811±0.01 | 0.831±0.02 |

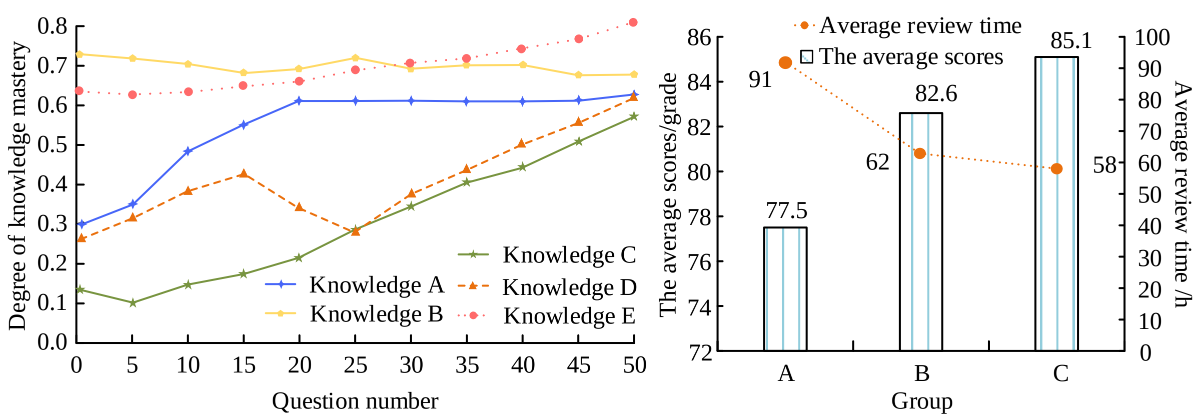
The study collected data from the first half semester of a university's personalized English course instruction in 2021 and organized that data into a dataset, which was used to train the DKVMN-BORUTA model. The questions were set with good differentiation and moderate difficulty, and if the mastery rate of the knowledge point under this question set was less than 40%, it indicates that the learners did not master the knowledge point. To verify the reliability of the proposed method, a random sample of learners was selected for the topic mastery statistics, and the changes in the knowledge state of the learners are shown in Fig. 5(a). Also, to verify whether the proposed method can achieve the goal of personalized instruction, the randomly selected learners were divided into three groups. No knowledge state information was provided for group A, knowledge state information but no precise instructional support service was provided for group B, and knowledge state information with precise instructional support service was provided for group C. Then, the learning performance and learning review time of the learners in the three groups were recorded, as shown in Fig. 5(b).

As can be seen from Fig. 5(a), the overall trend of learners' mastery of the knowledge points was on the rise, indicating that this learner's learning status became gradually become after being taught by the proposed method, indicating that the proposed knowledge tracking method has better readability. As shown in Fig. 5(b), the average test scores of Group C and Group B were higher than that of Group A, and the average review time of both groups was less than that of Group A. This indicates that the deep learning knowledge tracking model based on multidimensional feature fusion proposed by the study is effective and can find learners' weak points for learning knowledge more precisely. The average score of group C is 2.5 points higher than group B, and the average revision time is 4h less than group B. This indicates that the deep knowledge tracking model has better interpretability and can fully combine the multidimensional features of learners' learning process to reach accuracy teaching and improve the English learning efficiency.

## B. *Analysis of the Effectiveness of a Distanced Personalized English Teaching Model F incorporating Deep Directed Graph Knowledge Tracing and Graph Structure Rules*

After verifying the effectiveness of the deep learning knowledge tracking model based on multidimensional feature fusion, an experimental analysis of this deep learning knowledge tracking model combined with graph structure rules for distanced personalized English teaching and learning was conducted. The same dataset as before was used in this experimental phase and its basic situation is shown in Table II.

As can be seen from Table II, the lowest AUC value of the proposed model is 0.893 and the highest is 0.921, which shows the effectiveness and superior performance of the proposed model. Then, the performance of the proposed model was compared with other three knowledge tracking models in terms of prediction effectiveness by measuring the AUC of each dataset and its variation with the number of training sessions, and the results of the comparison are shown in Fig. 6(a).

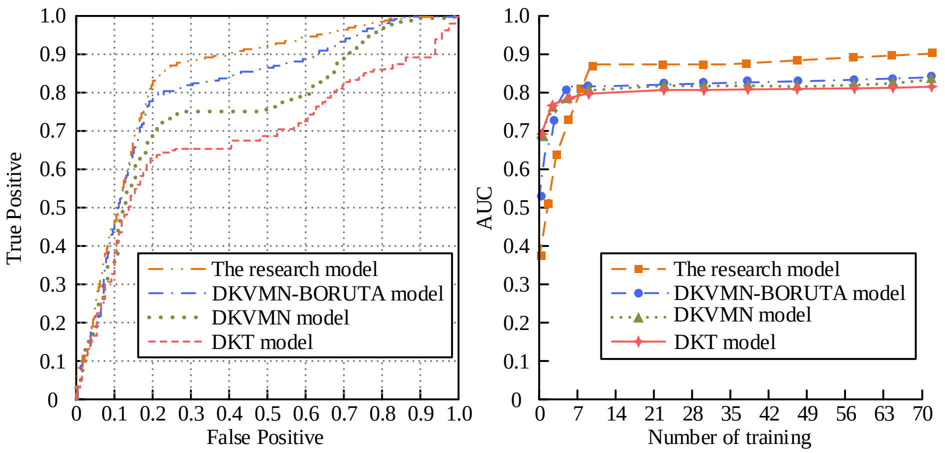



(a) The changing process of the learner's knowledge state (b) Empirical study control experimental results

Fig. 5. The changes of learners' knowledge state and the results of the empirical study.

TABLE II. AUC VALUES OF SIX PUBLIC DATA SETS AND FOUR MODELS

| Data set | Number of students | Knowledge tag | Number of interactions | Number of valid result records | The AUC of the study model |
|---|---|---|---|---|---|
| **ASSIST2014** | 4296 | 120 | 325682 | 314126 | 0.914±0.1 |
| **ASSIST2015** | 18246 | 116 | 598521 | 565247 | 0.893±0.2 |
| **Statics2011** | 400 | 188 | 3548 | 3264 | 0.912±0.1 |
| **Synthetic** | 2100 | 10 | 14621 | 13346 | 0.911±0.1 |
| **Khan Academy** | 18364 | 127 | 236489 | 216721 | 0.921±0.1 |
| **KDD Cup** | 4010 | 186 | 11000 | 10834 | 0.898±0.2 |




(a) Comparison of AUC curves of four deep knowledge tracking models

(b) The AUC curve of four deep knowledge tracking models varies with the number of training rounds

Fig. 6. AUC curves of four deep knowledge learning models.

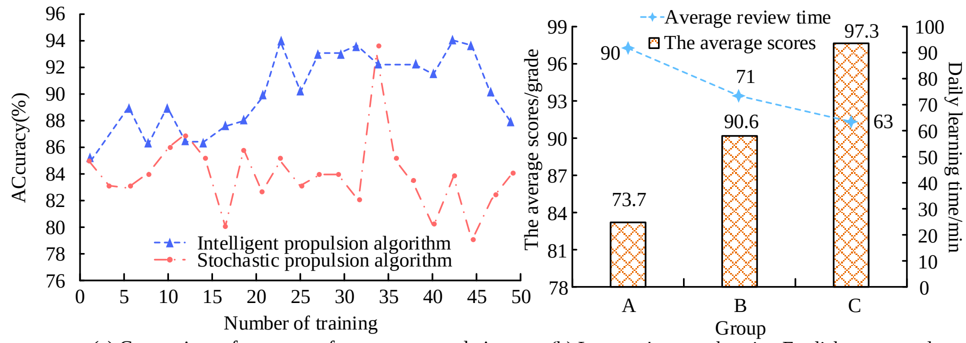
From Fig. 6(a), it can be seen that the model proposed in the study had an average AUC value of 0.92 in various datasets, while the AUC value of the DKT model was only 0.73. Through comparing the boundary baseline on the KDD dataset, the model proposed in this study improved the gain by a factor of about four. As can be seen from Fig. 6(b), it can be found that the proposed model tended to be stable after the 9th training session, and the AUC value of the proposed model was consistently higher than that of the other three models after the 9th training session, indicating the superiority of the proposed model. The study collected the learners' mastery degree of each knowledge point in the personalized English instruction course, and the collected data were processed to obtain the empirical findings, as shown in Table III.

As shown in Table III, the prediction accuracy of the proposed method was 94.3%, which was 4.9%-26.0% higher than that of the other models. The mean square error of the proposed model was 0.1623, which was the smallest among the tested models. In addition, the F1 score of the proposed model was 0.92, which was the highest among the tested models. The research data indicated that the proposed method had strong performance. Finally, the study tested the variation of the reward value and the number of training sessions, as shown in Fig. 7(a). The comparison of the overall English scores of the learners in each group and the time spent on revision is shown in Fig. 7(b).

From Fig. 7(a), it can be seen that the overall accuracy of the intelligent personalized recommendation algorithm based on reinforcement learning proposed in this study tended to increase with the increase of training times, and the difference between it and the random recommendation algorithm was gradually increasing, with the difference in the range of 6%-9%. This result indicates that the proposed method had more accurate recommendation capability, which can meet the demand of distanced personalized English teaching. It can be seen from Fig. 7(b) that the comprehensive score of the study group using the study method was 97.3, and the average daily review time of the students in this group was 63min. Research shows the proposed method can effectively improve learners' overall English performance and save learning time.

TABLE III. EMPIRICAL RESEARCH RESULTS

| Model type | AP (%) | ACC (%) | F1-score | MSE |
|---|---|---|---|---|
| **BKT** | 87.3 | 68.3 | 0.73 | 0.2491 |
| **DKT** | 86.2 | 77.9 | 0.79 | 0.1892 |
| **DKVMN** | 88.4 | 82.6 | 0.82 | 0.1834 |
| **DKVMN-BORUTA** | 88.6 | 89.4 | 0.89 | 0.1729 |
| **The research model** | 90.7 | 94.3 | 0.92 | 0.1623 |




(a) Comparison of accuracy of two recommendation algorithms
(b) Learners' comprehensive English scores and learning time

Fig. 7. Comparison of recommendation algorithm accuracy with learners' English scores.

## IV. Conclusion

The rapid development of Internet information technology has led to the continuous innovation and progress of English online education models. This study proposes a distanced personalized English teaching model incorporating deep directed graph knowledge tracking. Firstly, with the help of BORUTA feature selection model, the multidimensional features of the learning process are mined and a deep knowledge-tracking model of the multidimensional features of the learning process is fused. This optimized deep knowledge tracking method is then combined with graph structure rules to construct a fused approach-based model for distanced personalized English teaching and learning. The results show that the proposed model has the lowest AUC value of 0.893 and the highest of 0.921 on each dataset, indicating its effectiveness and superior performance. The prediction accuracy of the proposed model is 94.3% and its mean square error is 0.1623, which is the smallest among the models. The F1 score of the proposed model is 0.92, which is the highest among the models indicating that the model has strong performance. Moreover, the overall accuracy of the proposed method shows an increasing trend and is higher than that of the random recommendation algorithm. The optimized deep knowledge tracking method based on the multi-dimensional features of the learning process has practical teaching application value. The fusion model has higher recommendation accuracy and better recommendation effect than the traditional method. And it can meet the needs of personalized remote English teaching and targeted recommendation teaching, so as to improve the quality of English teaching. This study provides a practical theoretical basis and reference direction for online teaching methods in the future. However, there is still room for improvement in the interpretability of the model, and future study will focus on the improvement and enhancement of the performance of the model.

## References

[1] M. D. Labitad, L. S. Lomibao, The Experiences, Challenges and Perception on Online Lesson Study (O-LS) as Teacher Professional Development Program. American Journal of Educational Research, vol. 9(10), pp. 639-646, 2021.

[2] A. Cully, Y. Demiris, "Online Knowledge Level Tracking with Data-Driven Student Models and Collaborative Filtering". institute of Electrical and Electronics Engineers (IEEE), vol. 32(10), pp. 2000-2013, 2020.

[3] Wang M M. Educational Probation Under the Tutorial System and the Construction of English Student Teachers' Knowledge. teaching English in America and China: English version, vol. 16(2), pp. 75-83, 2019.

[4] T. Y. Zhao, M. Zeng, J. H. Feng, "An Exercise Collection Auto-Assembling Framework with Knowledge Tracing and Reinforcement Learning", Journal of Computer Science and Technology, 2022, vol. 37(5), pp. 1105-1117, 2022.

[5] T. Dani, R. Priyatikanto, A. Winarko, P. Gerhana, "Solar Wind Speed Time-Series Forecasting Based on Long Short-Term Memory (LSTM) Neural Network Model". vol. 275, pp. 193-198, 2022.

[6] Gervett, Koedingerk, Schneider J. When is deep learning the best approach to knowledge tracing. journal of Educational Data Mining, vol. 12(3), pp. 31- 54, 2020.

[7] Sapountzia, Bhulais, Corneliszi. "Dynamic knowledge tracing models for large-scalead aptive learning environments" International Journalon Advances in Intel ligentSystems, vol. 12, pp. 93-110, 2019.

[8] [8] S. U. Hassan, H. Waheed, N. R. Aljohani, M. Ali, S. Ventura, F. Herrera, "Virtual learning environment to predict withdrawal by leveraging deep learning", International Journal of Intelligent Systems. vol. 34(8), pp. 1935-1952, 2019.

[9] A. A. Mubarak, H. Cao, S. A. Ahmed, "Predictive learning analytics using deep learning model in MOOCs' courses videos", Education and Information Technologies. vol. 26(1), pp. 371-392, 2021.

[10] Käsert, Klinglers, Schwingag. "Dynamic Bayesian networks for student modeling". IEEE Transactions on Learning Technologies, vol. 10(4), pp. 450-462, 2019.

[11] Deonovicb, Yudelsonm, Bolsinovam. "Learning meets assessment. behaviormetrika," vol. 45(2), pp. 457-474, 2018.

[12] H. Yang, L. P. Cheung, "Implicit heterogeneous features embedding in deep knowledge tracing." cognitive Computation. vol. 10(1), pp. 3-14, 2018.

[13] Z. Wang, X. Feng, J. Tang, "Deep knowledge tracing with side in formation", Proceedings of the International Conference on Artificial Intelligence in Education. cham: springer, vol. 2, pp. 303-308, 2019.

[14] J. J. Vie, H. Kashima, "Knowledge tracing machines: factororization machines for knowledge tracing. inProceedings of the AAAI Conference on Artificial Intelligence", vol. 33(1), pp. 750-757, 2019.

[15] X. Tian, F. Liu, "Capacity tracing-enhanced course recommendation in MOOCs", IEEE Transactions on Learning Technologies. vol. 14(3), pp. 313-321, 2021.

[16] Y. Zou, X. Yan, W. Li, "Knowledge Tracking Model Based on Learning Process", Computers and Communications, vol. 10, pp. 7-17, 2020.

[17] Z. Zhang, X. Liu, Z. Li, H. Hu, "Outburst prediction and influencing factors analysis based on Boruta-Apriori and BO-SVM algorithms", Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, vol. 41(2), pp. 3201-3218, 2021.

[18] A. Es, B. Rw, B. Aks, "Securing color image transmission using compression-encryption model with dynamic key generator and efficient symmetric key distribution", Digital Communications and Networks, vol. 6(4), pp. 486-503, 2020.

[19] F. M. Faulstich, R. Kim, Z. H. Cui, "Pure State v-Representability of Density Matrix Embedding Theory", Journal of chemical theory and computation: JCTC, vol. 18(2), pp. 851-864, 2022.

[20] J. Zhang, Q. Xu, "Attention-Aware Heterogeneous Graph Neural Network", Big Data Mining and Analytics, vol. 4(4), pp. 233-241, 2021.

[21] X. Shi, Y. Li, Y. Yang, B. Sun, F. Qi, "Multi-models and dual-sampling periods quality prediction with time-dimensional K-means and state transition-LSTM network", Information Sciences: An International Journal, vol. 580, pp. 913-933, 2021.

[22] M. Zheng, J. Luo, Z. Dang, "Feedforward Neural Network Based Time-Varying State-Transition-Matrix of Tschauner-Hempel Equations", vol. 69(2), pp. 1000-1011, 2021.

[23] Y. S. Lee, A. Masood, W. J. Noh, S. G. "Cho DQN based user association control in hierarchical mobile edge computing systems for mobile IoT services". Future generation computer systems, vol. 137, pp. 53-69, 2022.

[24] I. Skrzypczak, G. Oleniacz, A. Leśniak, K. Zima, M. Mrówczyńska, K. Kazak Jan "Scan-to-BIM method in construction: Assessment of the 3D buildings model accuracy in terms inventory measurements", Building Research & Information, 2022, vol. 50(8), pp. 859-880, 2022.