International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 1412-1422

https://doi.org/10.70517/ijhsa463109

Prediction of Students' Knowledge Acquisition in Civic Education Based on Deep Knowledge Tracking and Knowledge Mapping

Haozhe Bao[1,*], Shunli Hong[1] and Xuejun Wen[1]
[1] Zhejiang Institute of Communications, Hangzhou, Zhejiang, 311112, China
Corresponding authors: (e-mail: 18368821912@163.com).

Abstract Aiming at the problem of poor accuracy in tracking the knowledge status of students' civic education, this study proposes a deep knowledge tracking model incorporating domain features with the support of educational knowledge graph, the DKT-KG model. The model filters important assessment behavioral features through a decision tree and incorporates the knowledge dependencies characterized by the educational knowledge graph to solve the problem of poor prediction accuracy of the original deep knowledge tracking model due to the lack of domain features. The experimental results on the ASSISTments09, Junyi, and KDD datasets show that the DKT-KG model can more accurately track knowledge points of the mastery level, and the AUC index and F1 scores are higher than those of other comparison models. Empirical analysis was conducted using historical data of question answering in four Civics courses taken by learners on the online education platform. The average knowledge mastery probability and its corresponding prediction accuracy derived when learners do different amounts of questions show that the deep knowledge tracking model constructed in this paper can accurately predict students' knowledge mastery.

Index Terms Knowledge Graph, Deep Knowledge Tracking, Decision Tree, Civics Education

I. Introduction

Civic education is an important educational activity in the Chinese education system, aiming to strengthen students' ideological and political awareness, moral and ethical awareness, awareness of the rule of law, and sense of social responsibility, to cultivate qualified socialist builders and successors, and it is an important way to cultivate students' correct worldview, outlook on life and values [1]-[4]. This concept is of great significance in Chinese education, especially in primary and secondary schools, colleges and universities and other educational institutions where it is widely implemented [5], [6]. However, in the teaching of Civics, it is difficult for teachers to understand the students' mastery of Civics knowledge through teaching to the extent that they cannot better improve the quality of Civics education [7].

Prediction refers to the measurement of learning outcomes in accordance with certain methods and laws on the basis of mastering students' existing information in order to understand the direction of students' learning development in advance [8], [9]. In recent years, with the rapid development of artificial intelligence technology, teachers can use deep knowledge tracking and knowledge mapping to predict the knowledge of students' mastery of Civic and Political Education, in order to adopt corresponding personalized teaching strategies according to the prediction results [10]-[12]. Deep knowledge tracking technology is an intelligent technology born with the growing demand for personalized learning, which provides personalized learning paths for students by analyzing their dynamic knowledge changes and helps teachers optimize their teaching strategies, which is becoming more and more important in students' Civics education [13]-[16]. And knowledge mapping is a data structure based on semantic network, which forms a complete knowledge system by associating and integrating the relationship between multiple knowledge points [17], [18]. In Civics education, knowledge mapping technology can help students better understand knowledge and improve learning efficiency, and it can also help teachers better manage and organize teaching content and improve teaching quality [19], [20].

This paper explores knowledge graph and deep knowledge tracking model, and then proposes a deep knowledge tracking model that incorporates domain features such as assessment behavioral features and knowledge dependency relationships, the DKT-KG model, to address the lack of integration of domain features in the deep knowledge tracking model. The core idea of the model is to incorporate assessment behavioral features and knowledge dependencies into the deep knowledge tracking model by means of decision trees and “ripple diffusion”.

In order to verify the validity of the proposed model, the ASSISTments09, Junyi and KDD datasets were used as examples for experimental validation. The knowledge mastery prediction of the model was also empirically analyzed by the historical data of real Civic Education answers in School A.

II. Method

II. A.Knowledge mapping

Knowledge graphs are structured semantic webs for describing knowledge entities and their interrelationships in knowledge engineering in the form of graphs. Here, the triad “entity-relationship-entity” or “entity-attribute-attribute-value” is introduced for knowledge representation, where each entity can be identified by a globally unique ID, the attribute-attribute-value describes the intrinsic characteristics of the knowledge entity, and the knowledge entities form a mesh knowledge structure through relationships [21]. In the learning process is mostly the course as the smallest unit of the knowledge system, so this paper will be the Civic Education course entity as the object of study to build a knowledge map. Let the ternary form used for knowledge expression be $T_F = \{E, R, E_T\}$, where E is the set of all the Civic and Political Education course entities, R is the set of relationships between the entities, and E_T is the set of intrinsic attributes of the course entities. The connection between entities can form a complex network structure. Where E_i denotes a specific course entity. R_{ij} denotes the relationship between the course entity E_i and E_j corresponding to the set R. E_{Ti} denotes the set of intrinsic attributes of the course entity E_i.

The systematic framework of the knowledge model is shown in Figure 1. The dotted line part is the initial building process of the knowledge graph of ideology and political education, which is externally used to realize the function of dynamic updating of the knowledge model. Firstly, the knowledge data of ideological and political education are subjected to entity extraction, relationship extraction and attribute extraction, which can also be collectively referred to as information extraction, to form highly structured data. Then the extracted data are subjected to ontologized knowledge expression to complete the description of the concept hierarchy. Finally, the reviewed and approved specific concepts are added to the ontology library to form the initial knowledge map. In order to guarantee the systematicity of the knowledge model, new knowledge concepts need to be iterated continuously on the basis of the existing knowledge map. New knowledge concepts need to be structured before being introduced into the model, and then fused with the knowledge concepts of the existing knowledge map to remove knowledge concepts that overlap or have a low confidence level, so as to avoid data redundancy. In addition, it is also necessary to pay attention to the fitness of new knowledge concepts with the original ontology model, and the whole passes the quality assessment link as the latest knowledge model.

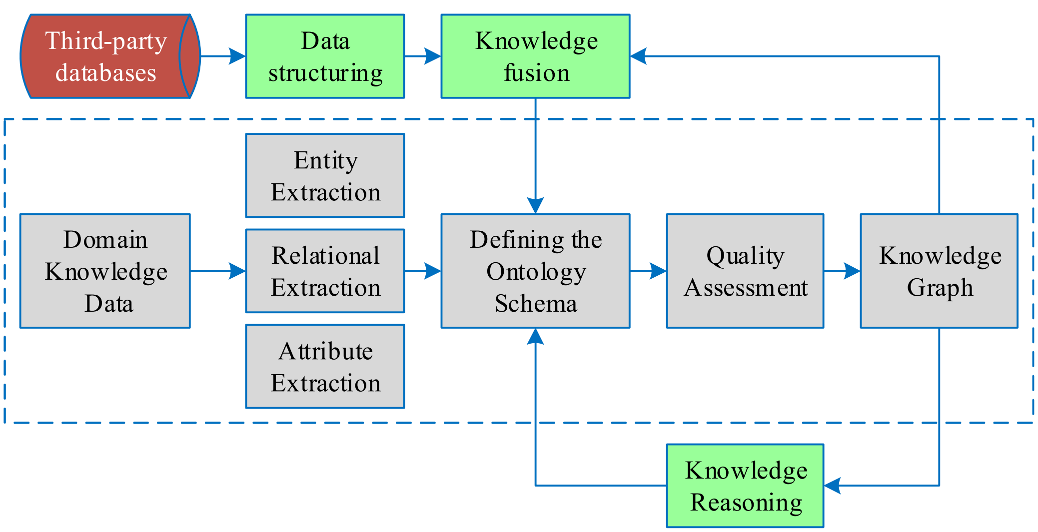

Figure 1: The system framework of knowledge model

II. B.Deep knowledge tracking

Deep knowledge tracking models can use both directed graphs to connect individual knowledge nodes and clustering to associate knowledge points. However, both approaches have obvious advantages and disadvantages, directed graph is not suitable for visual expression of the relationship between the problems, and clustering is unable to show the relationship between the problems under the same knowledge point on a local level [22]. Therefore, this paper adopts the combination of directed graph and clustering, using clustering globally to divide all the problems into multiple relatively independent branch problems, and using directed graph to associate all the knowledge points locally in order to show the relationship between the knowledge points.

The classical DKT model is based on Long Short-Term Memory (LSTM), compared with RNN, LSTM has one more cell state state, which carries the information of all the previous states, and whenever a new moment is reached, there will be a corresponding operation to determine whether to discard some old information and add some new information. During the update process, the state will be updated slowly, which is quite different from the rapid update of the hidden layer state h. This network structure is well suited for long sequential inputs such as online education data, because students tend to go through a long learning process, and then the network structure also needs to model the memorization and forgetting of knowledge over time.

The input of the model is encoded as a sequence of $\{0,1\}^{2N}$, where N is the number of problem concepts, and there is only one 1 in a sequence obtained by this encoding. For example, if a sequence "00100000" represents "the third question is answered correctly", the third element of the one-hot encoding sequence is 1, and the rest are 0, where N is 4. If the sequence becomes "00000010", it means that the third question is answered incorrectly, then the $7(N+6)$ element of the sequence is 1, and the rest of the position is 0. The sequences encoded in this way have the same length and contain only one "1".

The output Y_t of DKT is a sequence of length N. During the training process, the unique heat code X_t of the input at moment t is passed through the LSTM network and saved, and then the network reads the knowledge concept scalar q^{t+1} at moment $t+1$, and the value at the corresponding position of y^{t+1} in the output sequence is the prediction value of the DKT for the next moment. In LSTM, latent variables retain their values until they are explicitly removed by the action of a “forgetting gate”. Thus, they more naturally retain much of the time-step information. In addition, the hidden units are updated using cross-multiplication interactions, so they can perform more complex transformations for the same number of potential units.The update equations of the LSTM are shown in Eqs. (1) to (7), which represent the input gates, the output gates, the long memories, the short memories, the oblivion gates, the oblivion gates memories, and the outputs of the model at the moment t, respectively. In Eq. (1), h_t and C_t represent the long memory and short memory, W and b represent the implicit parameters in the input gate, output gate, and forgetting gate, and in each basic unit of LSTM, h_t directly determines the output y_t, and C_t influences h_t. Thus the output at moment t is always jointly influenced by the current input and the long and short memory vectors:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_t\right) \tag{1}$$

$$o_t = \sigma\left(W_o [h_{t-1}, x_t] + b_c\right) \tag{2}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{3}$$

$$f_t = \sigma\left(w_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{4}$$

$$\tilde{C} = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_c\right) \tag{5}$$

$$h_t = o_t \times \tanh\left(C_t\right) \tag{6}$$

$$y^t = \sigma\left(W'h^t\right) \tag{7}$$

II. C.Deep Knowledge Tracking Models Supported by Knowledge Graphs

This chapter proposes a deep knowledge tracking model that incorporates domain features, the DKT-KG model. First, the decision tree algorithm [23] is applied to mine the characteristics of learners' assessment behaviors in assessment big data. Second, the Civic and Political Education Knowledge Graph is processed into serialized sequences by combining the idea of “ripple diffusion” and topological ordering. Third, the logical relationship between the knowledge characterized in the Civic and Political Education Knowledge Atlas and the behavioral features extracted from the assessment behavioral data, as well as the student-question matrix are encoded together with one-hot, and then input into the long and short-term memory network for in-depth knowledge tracking after dimensionality reduction of the feature vectors to output the mastery state of each student in each Civic and Political Knowledge point at the end.

II. C. 1) Decision Tree Based Behavioral Feature Selection

The decision tree-based behavioral feature selection process mainly includes the following three basic steps:

(1) Gini coefficient calculation stage

In this paper, the Gini index in information theory is used as a quantitative strategy to screen them.The formula of Gini coefficient is as follows:

$$Gini(p)=\sum_{K=1}^{K}p_k(1-p_k)=1-\sum_{K=1}^{K}p_k^2 \tag{8}$$

For the assessment data in this study, assume that the dataset of answering behaviors is denoted as D, and since binary categorization (right and wrong answers) is used for the answering results, the Gini index of this dataset D is: $2p(1-p)$. Further, suppose the set of behavioral features in the dataset D is $\{A_1,A_2,A_3...,A_4\}$ then the Gini index of the dataset D under the condition of behavioral characteristics A is:

$$Gini\_Gain(D,A)=\sum_{v=1}^{V}\frac{|\mathrm{D}^{v}|}{|D|}=Gini(\mathrm{D}^{v}) \tag{9}$$

where V denotes the number of value spaces of the feature, and D^v denotes the subset into which the dataset D is divided by the behavioral feature A. The smaller the value of $Gini\_Gain(D,A)$, the higher the purity of the dataset D, and the lower the information entropy of the node split, so the behavioral feature with the smallest Gini index should be selected as the basis of node split.

(2) Decision tree generation stage

Taking the answer result as the target variable and various behavioral characteristics as the predictors, the decision tree is built recursively until the depth of the tree reaches the set depth H by adopting the idea of "top-down" and selecting the minimized Gini index for splitting from the root node layer by layer downward.

(3) Decision tree pruning stage

In order to improve the generalization ability of the decision tree and reduce the risk of overfitting, the decision tree is finally pruned by minimizing the loss of the decision tree as a whole, and the loss function of the subtree is calculated during the pruning process as:

$$C_a(T)=C(T)+a(T) \tag{10}$$

where $C(T)$ the Gini index of the subtree T and $|T|$ is the number of leaf nodes of the subtree $C(T)$.

II. C. 2) Knowledge Mapping into Knowledge Tracking

In this paper, based on the idea of structural feature embedding method of knowledge graph, a principle similar to "ripple diffusion" is used to integrate the educational knowledge graph into the deep knowledge tracking model, and the specific process is shown in Figure 2.

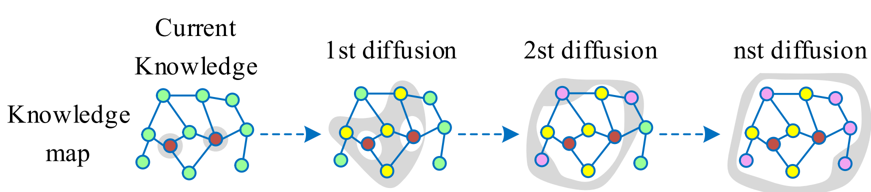

Figure 2: The knowledge map structure is serialized based on the "ripple diffusion"

The process takes the vectorized processed educational knowledge graph as input, diffuses for the current knowledge point, and outputs the sequence of knowledge points associated with the current knowledge point after topological sorting.

The set composed of neighboring knowledge points at the k th diffusion of the current knowledge point u is defined as the set of associated knowledge points, denoted as follows:

$$\varepsilon_u^k=\left\{t\,|\,(h,r,t)\in G \text{ and } h\in\varepsilon_u^{k-1}\right\},k=1,2,\cdots,H \tag{11}$$

The set consisting of all the knowledge points covered by the current knowledge point u after k diffusions is defined as a diffusion set, which is a correlation ternary with the set of $k-1$ times related knowledge points as the head vector, denoted as follows:

$$S_u^k = \left\{(h,r,t) \mid (h,r,t) \in G \ and \ h \in \varepsilon_u^{k-1}\right\}, k = 1,2,\cdots,H \tag{12}$$

For each diffusion set, the features are classified using the softmax function and embedded in the set of related knowledge points ε_u^k and topologically sorted to obtain a sequence of knowledge points.

II. C. 3) Self-encoder based feature dimensionality reduction

Self-encoder based feature dimensionality reduction is shown in Fig. 3. Constructing a self-encoder requires two parts: an encoder and a decoder. The encoder compresses the input into potential spatial representations, which can be represented by the function $f(x)$, and the decoder reconstructs the potential spatial representations into outputs, which can be represented by the function $g(x)$. Both the encoding function $f(x)$ and the decoding function $g(x)$ are neural network models.

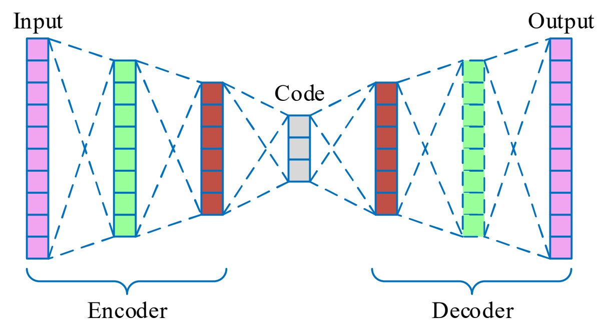

Figure 3: Based on the characteristics of self-code

Before feeding the data obtained from the educational knowledge map into the self-encoder, it is first necessary to perform feature crossover, on the basis of which one-hot coding, coded feature cascade is performed, and this process is expressed as follows:

$$v_t = O\left(C(m_1,a)\right) \sim O\left(C(m_2,a)\right) \sim \cdots \sim O\left(C(m_n,a)\right) \tag{13}$$

where the O function denotes one-hot coding of the vector, the C function denotes crossover operation on the features, and the ~ operation denotes cascading of the feature coding.

On the basis of the above, the feature vectors are downscaled using stack self-encoder, and the hidden layer and output layer functions are respectively:

$$v_t' = \tanh\left(W_{ed} \cdot v_t + b_{ed}\right), \ y_t = \tanh\left(W_{ed}^T \cdot v_t' + b_{dd}\right) \tag{14}$$

II. C. 4) Knowledge tracking based on LSTM networks

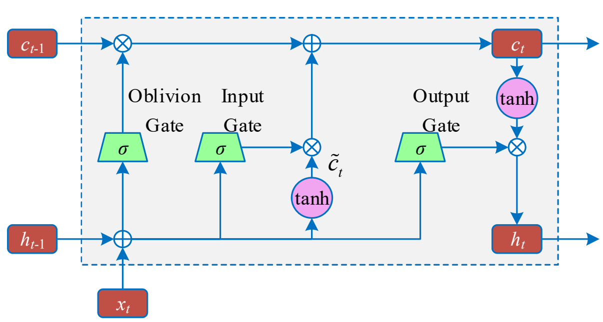

Figure 4: Long term memory neural network structure

The main task of the knowledge tracking process is to train a deep knowledge tracking model based on recurrent neural networks and incorporating domain features by using the topological ordering and problem-knowledge matrix obtained above as inputs to the network after one-hot coding. The LSTM network used in this study mainly consists of an input gate i, a forgetting gate f, an output gate o, and a cell unit c [24]. Its network structure is shown in Fig. 4.

The specific functions of each gate unit are as follows:

(1) Input Gate

The input gate control consists of two parts: one part is the sigmoid function that determines what information will be updated. The other part is the output of a candidate value vector by the tanh function, and the formulas for these two parts are as follows:

$$i_t = \sigma\left(W_i * \left[h_{t-1}, x_t\right] + b_i\right) \tag{15}$$

$$\bar{C}_t = \tanh\left(W_c * \left[h_{t-1}, x_t\right] + b_c\right) \tag{16}$$

(2) Oblivion Gate

The inputs to the forgetting gate are h_{t-1} and x_t, and a sigmoid function is generally used to output a value between 0 and 1 to determine the information to be discarded, with 1 indicating complete retention and 0 indicating complete discard. The formula for the forgetting gate is as follows:

$$f_t = \sigma\left(W_f * \left[h_{t-1}, x_t\right] + b_f\right) \tag{17}$$

(3) Cell state renewal

The formula for this process is:

$$C_t = f_t * C_{t-1} + i_t * \bar{C}_t \tag{18}$$

(4) Output Gate

The output gate uses the sigmoid function to determine the information to be output. And then, the output gate is processed using the tanh function to determine which part of the information is finally output. Its calculation formula is:

$$o_t = \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{19}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{20}$$

II. C. 5) Choice of loss function

Assuming that the predicted value of the model is $f(x)$ and the true value is Y, the loss function is usually expressed using $L\left(Y, f(x)\right)$. The loss function of the final model is:

$$\tilde{L} = L + \lambda_r r + \lambda_{w_1} w_1 + \lambda_{w_2} w_2^2 \tag{21}$$

$$w_1 = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T_i - 1} \left\| y_{t+1}^i - y_t^i \right\|_1}{M \sum_{i=1}^{n} \left(T_i - 1\right)} \tag{22}$$

$$w_2^2 = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T_i - 1} \left\| y_{t+1}^i - y_t^i \right\|_2^2}{M \sum_{i=1}^{n} \left(T_i - 1\right)} \tag{23}$$

$$r = \frac{1}{\sum_{i=1}^{n} \left(T_i - 1\right)} \left(\sum_{i=1}^{n} \sum_{t=1}^{T_i - 1} l\left(y_t^i \cdot \delta\left(q_t^i\right), a_t^i\right) \right) \tag{24}$$

where the role of r is to reduce the error in reconstructing the input sequence, and the goal of w_1, w_2 is to make the prediction result smoother.

III. Results and discussion

In this chapter, after first conducting a performance comparison experiment on the DKT-KG model constructed in the previous section to verify the validity of the model, an empirical analysis was conducted using the historical data of students' Civic Education responses in two natural semesters, spring and fall of 2024 at School A. The data were analyzed by the DKT-KG model.

III. A. Comparative Experiments on Knowledge Tracking Models

III. A. 1) Experimental data set

In this paper, the real datasets ASSISTments09, Junyi and KDD, which are common in the field of knowledge tracking, are used for the study.

The ASSISTments09 dataset collects learner response data from the online education platform ASSISTments from the academic year 2009 to 2010. This dataset contains important information such as student ID, knowledge concept ID, correct and incorrect information of exercise responses, etc. It includes a total of 325,964 historical records of learner responses to exercises by 4,200 learners on 125 knowledge concepts, which are screened for experimental validation of Civics and Political Science knowledge points.

For the KDD dataset, the field information used in this study is the same as that of ASSISTments09, which contain knowledge point information.

The Junyi dataset is the real data of learner-practice interactions collected by the Junyi Education website in Taiwan, China, which records important information such as learner ID, practice ID, knowledge concept name, and correct and incorrect information of practice responses, and includes a total of 1056 learners on 811 Civic Education knowledge concepts of 404,903 historical records of learners' responses to practice questions.

III. A. 2) Analysis of experimental results

In this study, two model performance evaluation metrics are used to measure the model effectiveness, which are the area under the receiver operating characteristic curve (ROC) enclosed with the axes (AUC) and the F1 score.

The experiments in this section focus on comparing and analyzing the performance of the knowledge tracking model supported by knowledge graph (DKT-KG) proposed in this study with the existing models in the domain. The excellent models in the knowledge tracking domain are selected for the experiment: the deep knowledge tracking model (DKT), the graph embedding-based knowledge tracking model (GKT), and the dual model based on time effect and spatial effect (SKT). The performance of the above models is compared experimentally with DKT-KG. The classical models are all experimented on the same real dataset using the model code publicly available at GitHub, with the hyperparameters of the models kept the same.

The experimental results of each model on the Assistments09, Junyi and KDD datasets are shown in Figures 5 to 7, respectively. Compared with the original DKT model, the AUC values of the DKT-KG model on the Assistments09, Junyi and KDD datasets are improved by 4.2%, 5.5%, and 4.7%, respectively.

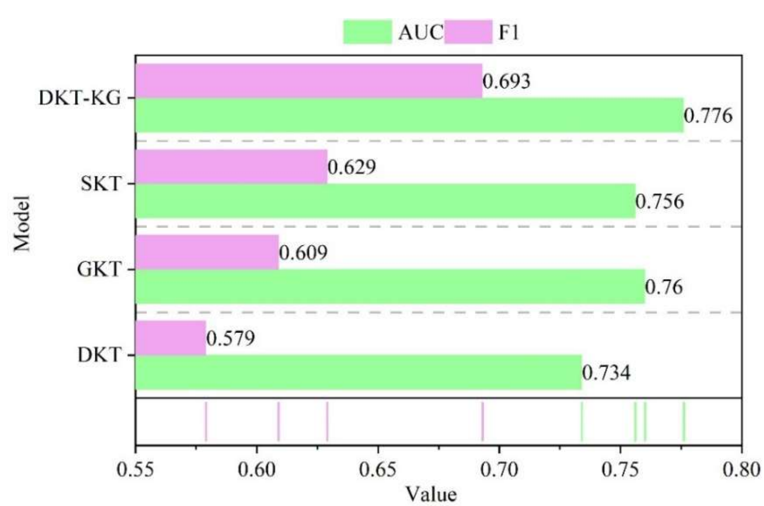

Figure 5: The results of the Assistments09 data set

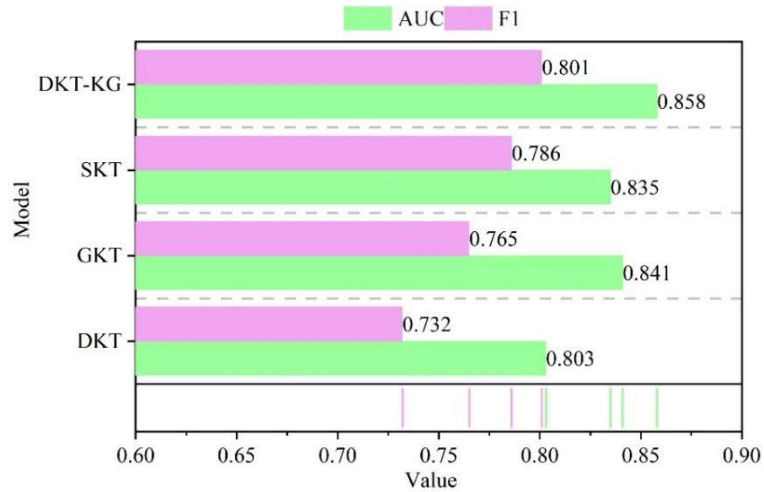

Figure 6: The results of the Junyi data set

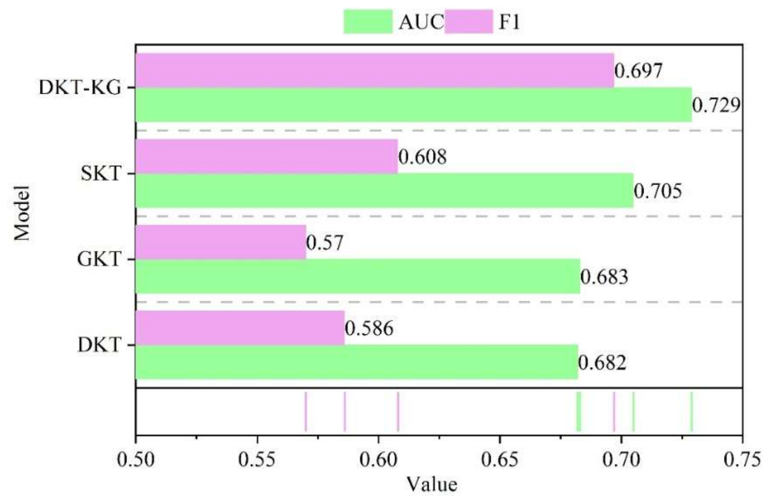

Figure 7: The results of the KDD data set

From Fig. 5 to Fig. 7, the following conclusions can be obtained:

(1) In terms of model prediction accuracy, among all models, the present model demonstrates better prediction ability on ASSISTments09 dataset, Junyi dataset and KDD dataset. Compared to the DKT model, the prediction ability of the present model is improved, this is the incorporation of relational information between the knowledge points, so that the model adds reasonable constraints in performing the prediction task and the model performance is improved. Compared with GKT and SKT, the present model considers more different kinds of relationships between knowledge points and refines the effect of knowledge structure on learners' responses, resulting in more accurate prediction performance.

(2) From the analysis of model structure, GKT, SKT and the present model consider knowledge domain features, i.e., information about relationships between knowledge points, relative to DKT, and the experimental results prove that all three models demonstrate better prediction performance than the traditional knowledge tracking model on all three datasets.GKT, from the graph neural network, proposes a new method of incorporating single relationships between knowledge points, but the division of knowledge point relationships is too monotonous and does not pay attention to the effects brought by other knowledge structures.SKT considers both temporal and spatial effects on predictability through knowledge propagation theory, but the model only considers knowledge points in the neighborhood of the current knowledge point, which leads to a limited propagation of relationships. This model proposes a new method of fusing knowledge structure graphs, feeding the representation of knowledge points in the whole graph at the input layer, so that the model can track the mastery level of knowledge points more accurately and obtain the learners' future learning under the condition that the learners' historical learning records are known.

III. B. Predictive Analysis of Students' Knowledge Acquisition of Civic and Political Education

III. B. 1) Data sources

The data for this study are derived from four ideological and political courses taken by learners on the online education platform, including "Introduction to the Basic Principles of Marxism", "Theory and Practice of Socialism with Chinese Characteristics in the New Era", "Outline of Modern Chinese History", and "Ideological and Moral Cultivation and Legal Foundation". Historical data of learners' answers in the two natural semesters of Spring 2022 (data set is D) and Fall 2022 (data set is D'). The statistical attributes of each dataset are the number of learner users and the number of answering interactions, and the relevant information of the four course datasets is shown in Table 1, with a total of 10,352 participants.

Table 1: Four courses data set related information

Data	D		D'	
	Number of learners	Answer the number of times	Number of learners	Answer the number of times
Introduction to the basic principle of Marxism	2195	17763	3560	21782
The theory and practice of socialism with Chinese characteristics in the new era	2648	13581	2132	19712
China modern history	1628	23295	1669	32609
Moral cultivation and legal basis	3881	37028	2223	17205

III. B. 2) Pre-processing

Firstly, the original datasets D and D' are cleaned, if the percentage of missing values of a data is greater than or equal to 50%, the data is removed, otherwise the data is retained and filled with missing values. Secondly, the outliers in the two datasets are removed, for the same learner ID the number of times the same practice question is answered more than 10 times is considered as outlier data. Again, the two datasets are normalized. Finally, multiple duplicate data for the same learner are merged and identified according to the learner ID to get a total of 3500 learner users who are studying four courses at the same time and have a valid answer count of 12000.

III. B. 3) Analysis of average knowledge acquisition and corresponding prediction accuracy

In this section, the deep knowledge tracking model supported by knowledge graph is applied to analyze the average knowledge mastery probabilities and their corresponding prediction accuracies (AUCs) derived from learners in datasets D and D' when they do different amounts of questions. The specific results are shown in Figures 8 and 9. From these two figures, it is found that the learners' knowledge mastery and model prediction accuracy predicted by the deep knowledge tracking model float and change as the amount of questions learners do changes. Especially when the answer data is long enough, the deep knowledge tracking model can better distinguish between these two performances. The curves of learner knowledge mastery and model prediction accuracy with the amount of questions done are consistent with learner learning patterns and model prediction timeliness. Therefore, it is considered that the learner cognitive diagnosis results given by the deep knowledge tracking model on these two datasets are reasonable.

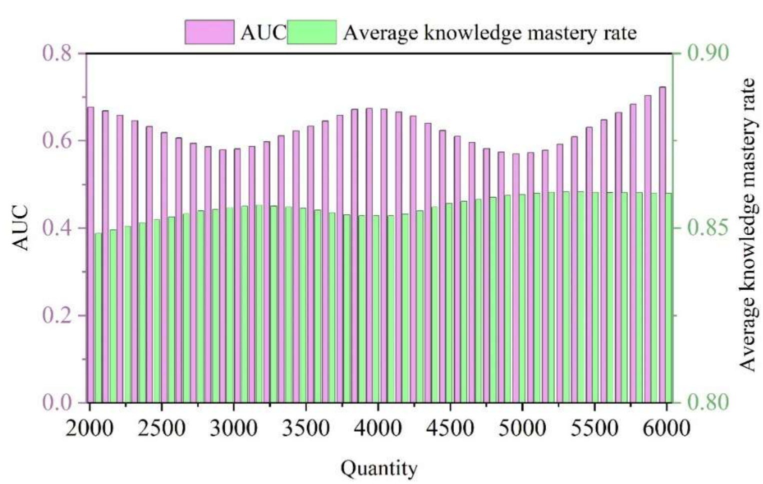

Figure 8: Average knowledge control and prediction accuracy on data set D

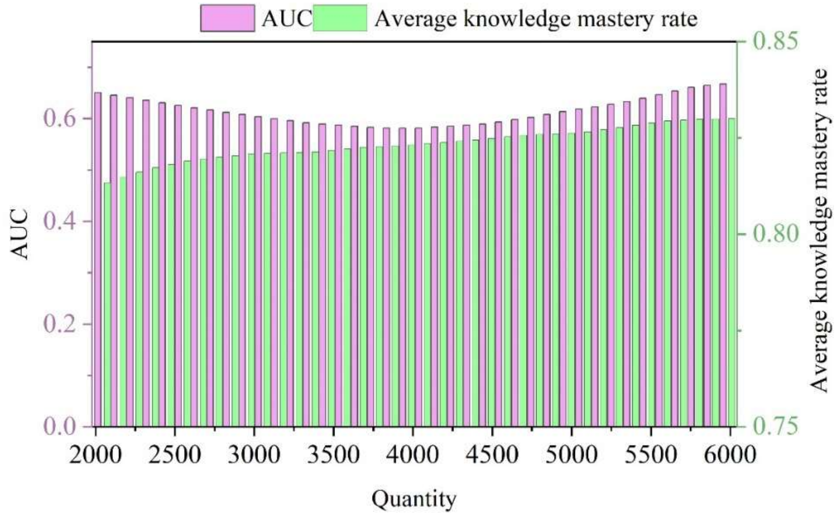

Figure 9: Average knowledge control and prediction accuracy on data set D′

IV. Conclusion

In this paper, we propose a knowledge graph-supported deep knowledge tracking model (DKT-KG model), which extracts assessment behavioral features through a decision tree, and integrates inter-knowledge correlations into the deep knowledge tracking model through the idea of “ripple diffusion”, so as to realize the integration of domain features. Experimental results on ASSISTments09, Junyi and KDD datasets show that the model in this paper has high effectiveness and accuracy in acquiring learners' future learning. Compared with the original DKT model, the predictive effectiveness of the DKT-KG model is improved by 4.2%, 5.5%, and 4.7% on the three datasets, respectively. An empirical analysis was conducted using the effective number of questions answered by 3500 users in School A. The curves of learners' knowledge mastery and model prediction accuracy with the change of the number of questions done are in line with the learners' learning pattern and the model prediction timeliness, thus proving the effectiveness of the model proposed in this paper.

References

[1] Yaocan, Z., & Qinghua, S. (2025). Historical Experience of the Research on the Principles of Ideological and Political Education. Teaching and Research, 59(2), 49.

[2] Chengyuan, W. (2025). Research on the Integration Path of College Students’ Mental Health Education and Ideological and Political Education. International Educational Research, 8(2), p44-p44.

[3] Xisheng, W. (2025). Innovation of Ideas in the Research on the Principles of Ideological and Political Education. Teaching and Research, 59(2), 58.

[4] Lei, Z. (2025). Building Ideological and Political Educational Practices and Civic Engagement of Students in Chongqing, China. Journal of Education, Humanities, and Social Research, 2(1), 38-49.

[5] Shiqi, D., Li, K., & Yuqing, J. (2025). Philosophical Foundations and Cultural Implications of Ideological and Political Education in Higher Education: A Marxist Perspective. Cultura: International Journal of Philosophy of Culture and Axiology, 22(1).

[6] Fan, Y. (2025). Research on the Innovation Path of Network Ideological and Political Education in Colleges and Universities. Curriculum Learning and Exploration, 2(4).

[7] Zhou, M., & Li, S. (2025). Research on the Ideological and Political Education in the Principles of Management of College Students. Frontiers in Humanities and Social Sciences, 5(1), 123-126.

[8] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student’performance prediction using machine learning techniques. Education Sciences, 11(9), 552.

[9] Xu, Z., Yuan, H., & Liu, Q. (2020). Student performance prediction based on blended learning. IEEE Transactions on Education, 64(1), 66-73.

[10] Bujang, S. D. A., Selamat, A., Ibrahim, R., Krejcar, O., Herrera-Viedma, E., Fujita, H., & Ghani, N. A. M. (2021). Multiclass prediction model for student grade prediction using machine learning. Ieee Access, 9, 95608-95621.

[11] Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. International Journal of Educational Technology in Higher Education, 20(1), 4.

[12] Hussain, M., Zhu, W., Zhang, W., Abidi, S. M. R., & Ali, S. (2019). Using machine learning to predict student difficulties from learning session data. Artificial Intelligence Review, 52, 381-407.

[13] Gervet, T., Koedinger, K., Schneider, J., & Mitchell, T. (2020). When is deep learning the best approach to knowledge tracing?. Journal of Educational Data Mining, 12(3), 31-54.

[14] Liu, Z., Guo, T., Liang, Q., Hou, M., Zhan, B., Tang, J., ... & Weng, J. (2025). Deep Learning Based Knowledge Tracing: A Review, A Tool and Empirical Studies. IEEE Transactions on Knowledge and Data Engineering.

[15] Song, X., Li, J., Cai, T., Yang, S., Yang, T., & Liu, C. (2022). A survey on deep learning based knowledge tracing. Knowledge-Based Systems, 258, 110036.

[16] Dai, M., Hung, J. L., Du, X., Tang, H., & Li, H. (2021). Knowledge tracing: A review of available technologies. Journal of Educational Technology Development and Exchange (JETDE), 14(2), 1-20.
[17] MD, S., & Krishnamoorthy, S. (2025). Knowledge Graphs for Representing Knowledge Progression of Students across Heterogeneous Learning Systems. International Journal of Artificial Intelligence in Education, 1-29.
[18] Wu, W., Yang, S., Tian, F., Wei, X., & Pan, Y. (2025). Knowledge graph enhanced recommendation for semantic structure construction and feedback in online learning. Physical Communication, 69, 102590.
[19] Li, H., & Yang, Y. (2025). Literature Review on the Role of Generative Artificial Intelligence Technologies in Education--Knowledge Graph-based Visual Analytics. International Journal of Social Science and Education Research, 8(1), 118-127.
[20] Aytekin, M. C., & Saygın, Y. (2024). ACE: AI-Assisted Construction of Educational Knowledge Graphs with Prerequisite Relations. Journal of Educational Data Mining, 16(2), 85-114.
[21] Jie Hu,Hongqun Yang,Fei Teng,Shengdong Du & Tianrui Li. (2024). A knowledge graph completion model based on triple level interaction and contrastive learning. Pattern Recognition,156,110783-110783.
[22] Deliang Wang1,Zhi Zhang1,Jiachen Song2,Yu Lu1,2 & *. (2021). Traditional Knowledge Tracing Models for Clustered Students. The Educational Review, USA,4(12).
[23] Xiaoyin Yang. (2023). Design of online learning behaviour feature mining method based on decision tree. International Journal of Continuing Engineering Education and Life-Long Learning,33(2-3),269-281.
[24] Yujia Huo,Derek F. Wong,Lionel M. Ni,Lidia S. Chao & Jing Zhang. (2020). Knowledge modeling via contextualized representations for LSTM-based personalized exercise recommendation. Information Sciences,523,266-278.