FDKT: Towards an Interpretable Deep Knowledge Tracing via Fuzzy Reasoning

FEI LIU, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
CHENYANG BU, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
HAOTIAN ZHANG, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
LE WU, Hefei University of Technology, Hefei, China
KUI YU, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
XUEGANG HU, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China

In educational data mining, knowledge tracing (KT) aims to model learning performance based on student knowledge mastery. Deep-learning-based KT models perform remarkably better than traditional KT and have attracted considerable attention. However, most of them lack interpretability, making it challenging to explain why the model performed well in the prediction. In this paper, we propose an interpretable deep KT model, referred to as fuzzy deep knowledge tracing (FDKT) via fuzzy reasoning. Specifically, we formalize continuous scores into several fuzzy scores using the fuzzification module. Then, we input the fuzzy scores into the fuzzy reasoning module (FRM). FRM is designed to deduce the current cognitive ability, based on which the future performance was predicted. FDKT greatly enhanced the intrinsic interpretability of deep-learning-based KT through the interpretation of the deduction of student cognition. Furthermore, it broadened the application of KT to continuous scores. Improved performance with regard to both the advantages of FDKT was demonstrated through comparisons with the state-of-the-art models.

CCS Concepts: • **Applied computing** → **Education**; • **Information systems** → *Information systems applications;*

Additional Key Words and Phrases: Educational data mining, knowledge tracing, model interpretability, fuzzy reasoning, deep learning

This work was supported by the National Science and Technology Major Project (under grant 2021ZD0111802), the National Natural Science Foundation of China (under grants 61806065, 62376087, 62076085, 62376085, 62106262, 72188101, 62120106008), and the Fundamental Research Funds for the Central Universities (under grant JZ2022HGTB0239).

Authors' addresses: F. Liu, C. Bu (Corresponding author), H. Zhang, K. Yu, and X. Hu, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, Anhui, China; e-mails: feiliu@mail.hfut.edu.cn, chenyangbu@hfut.edu.cn, haotianzhang521@gmail.com, yukui@hfut.edu.cn, jsjxhuxg@hfut.edu.cn; L. Wu, Hefei University of Technology, Hefei, Anhui, China; e-mail: lewu@hfut.edu.cn.

ACM 1046-8188/2024/05-ART139
https://doi.org/10.1145/3656167

ACM Reference Format:
Fei Liu, Chenyang Bu, Haotian Zhang, Le Wu, Kui Yu, and Xuegang Hu. 2024. FDKT: Towards an Interpretable Deep Knowledge Tracing via Fuzzy Reasoning. *ACM Trans. Inf. Syst.* 42, 5, Article 139 (May 2024), 26 pages.
https://doi.org/10.1145/3656167

1 INTRODUCTION

Online education systems such as MOOC, ASSISTments, and Khan Academy are being increasingly used, producing large amounts of student learning data [1–7]. **Knowledge tracing (KT)** [8–10] focuses on predicting future performance based on estimating the over-time knowledge mastery of students from the learning logs, as shown in Figure 1. KT is one of the important tasks of educational data mining [11, 12] and can be applied to various scenarios, such as facilitating better personalized learning resource recommendations [13, 14].

Model interpretability recently has attracted increasing attention in the field of educational data mining, including the KT task. **Interpretability** is defined as *the ability to provide explanations in understandable terms to a human* [15, 16]. Being able to explain the reasons why the model was able to achieve good prediction performance in an interpretable KT model is as crucial as achieving desirable performance [17]. To obtain this understanding, interpretability can be improved from both intrinsic and post hoc aspects, as shown in Figure 2. **Intrinsic** interpretability explains *how the model works*, the interpretability comes from the model-specific constraints based on the domain knowledge [18]. The way to construct an intrinsically interpretable model, for example, is by using interpretable models such as linear regression, decision tree, and decision rules. **Post hoc** interpretability provides answers to the question *what else can the model tell us*. It refers to improving interpretability using model-agnostic methods [18], such as visualization of the features and effects. KT is more concerned with the cognitive state of the student, and the accuracy of its assessment cannot be directly measured. Instead, the accuracy of performance predictions is measured. Therefore, interpretability is significant in KT to explain the process of obtaining the predicted results and the relationship between the predicted results and the cognitive state.

Since it is difficult to measure students' knowledge mastery, most existing KT models use end-to-end learning to measure the accuracy of prediction performance [19]. Therefore, interpretability is typically not the major focus of most existing models, especially for those deep-learning-based KT models from the intrinsic aspect [19, 20]. This can be analyzed from the following three aspects. (1) The first deep-learning-based KT model, the **deep knowledge tracing (DKT)** [21], applied recurrent neural networks to KT. Estimating student cognition is difficult for DKT, since there is no interpretable parameter to inspect [22]. It has achieved excellent prediction accuracy owing to its large vectors of 'neurons' which are hard to interpret. (2) With an increasing amount of attention being paid to interpretable machine learning approaches, some studies have attempted to improve the interpretability of DKT using post hoc methods such as layer-wise relevance propagation [20] and visual methods [23]. They have attempted to answer the question of what else the DKT can tell but have not been able to explain the intrinsic process. (3) Some deep-learning-based KT models have denoted hidden layers as student cognition to enhance the interpretability to some extent [24–27]. However, the lack of intrinsic interpretability can also be attributed to how the model is constructed, its parameters, non-linear activation functions, and so on. In other words, such models cannot be explained as the interpretable ones, such as decision trees or rules.

Fuzzy theory is a powerful tool to represent human knowledge and mimic human reasoning capabilities, which is demonstrated as successful applications in education data mining [28, 29]. **Fuzzy Cognitive Diagnosis Framework (FuzzyCDF)** [28] is a typical cognitive diagnosis framework that leveraged fuzzy theory to model students' abilities to continuous score scenarios. The

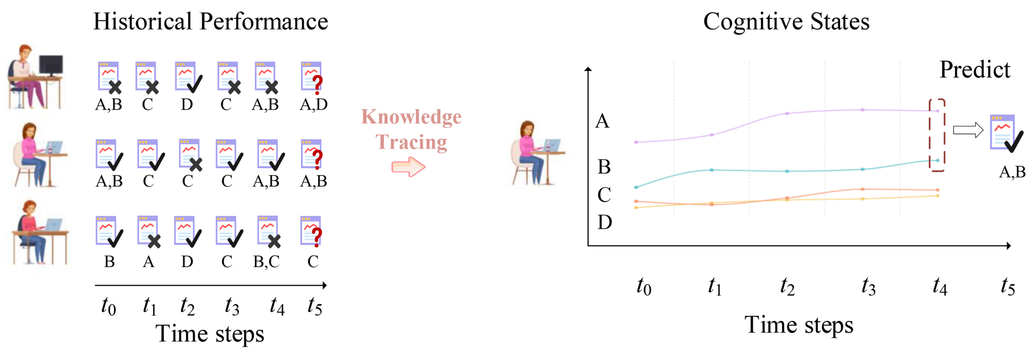

Fig. 1. Schematic of knowledge tracing. A, B, C, and D represent four knowledge components, examined by various exercises. Knowledge tracing estimates the over-time cognitive states of students and predicts future performance based on their cognitive states.

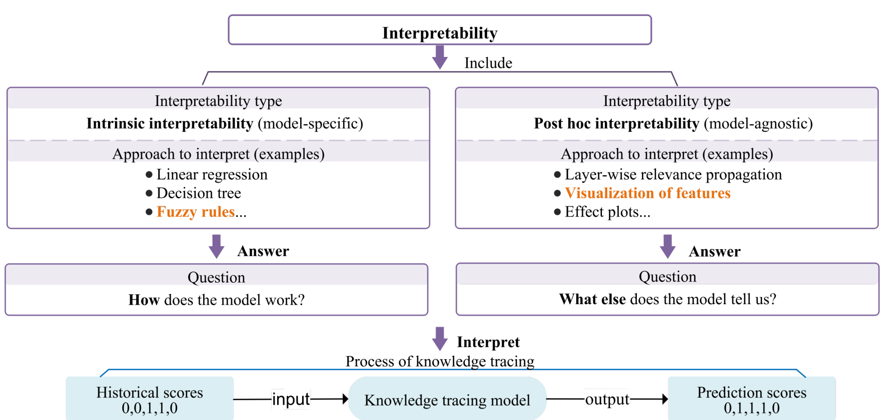

Fig. 2. Schematic of model interpretability, using knowledge tracing as an example. Models with interpretability better reason the obtained prediction than models lacking interpretability, in terms of both intrinsic and post-hoc aspects. The former usually explains the working of the model; the latter usually interprets further workability of the model.

temporal characteristics of the learning logs were not considered in cognitive diagnosis (in other words, KT can be regarded as a dynamic cognitive diagnosis task). In our previous work, FBKT [29] reported effective performance fuzzifing the continuous scores into the type-1 and type-2 fuzzy sets in Bayesian KT. However, as is the case with the traditional Bayesian KT [30], they must classify students' learning logs by the knowledge components related to the exercises [31]. For example, $(A_1, A_2, B_1, A_3, C_1, B_2)$ is the original exercising sequence of a student, where A_1 denotes the first exercise related to the knowledge component A. FBKT cannot directly deal with them, and instead, it preprocessed the sequence into three portions: (A_1, A_2, A_3), (B_1, B_2), and (C_1). As a result, it changed the temporal information in the original sequence. Furthermore, fuzzy reasoning offers a better framework for interpretability considerations owing to its rules [32]. Based on fuzzy reasoning, the above applications have not utilized fuzzy rules to reason such that they owned inadequate intrinsic interpretability.

Contributions. The major contributions of this study are as follows, shown in Figure 3. In this paper, we propose **fuzzy deep knowledge tracing (FDKT)**, which introduces **fuzzy neural**

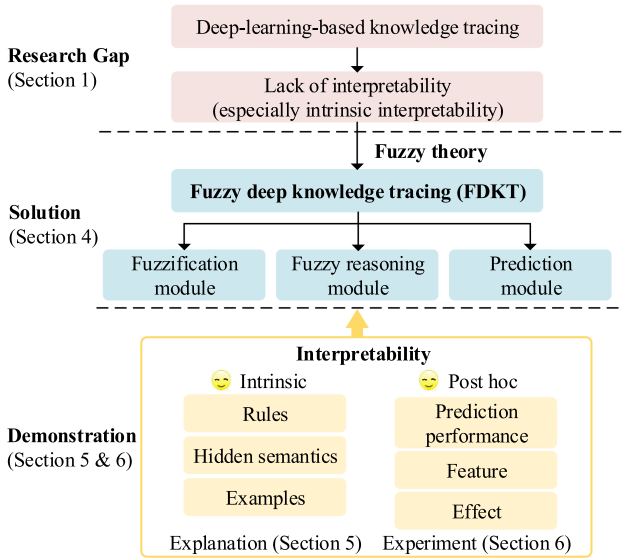

Fig. 3. Schematic of the limitations and contributions.

networks (FNNs) [33] to enhance the interpretability of existing deep-learning-based KT models. Specifically, FDKT contains three main modules, i.e., the fuzzification, fuzzy reasoning, and prediction modules. First, the continuous scores on the historical exercises are fuzzified into several fuzzy scores, rather than hard encoding similar to a black box. Subsequently, the current fuzzy cognition is deduced according to the fuzzy reasoning module, which is the core of the proposed model improving intrinsic interpretability. Finally, the performance is predicted. It is remarkable that the proposed model has demonstrated interpretability in terms of both the intrinsic and post hoc aspects.

— To improve the interpretability (especially in the intrinsic aspect) of the traditional deep-learning-based KT models, we explored the utility of fuzzy reasoning in the field of KT. The proposed model combines the advantages of both fuzzy theory and neural networks, i.e., the ability to combine language-based knowledge (e.g., expert experience) and the ease of training the model parameters (e.g., backpropagation).
— To deal with the uncertainty in the KT task, i.e., uncertainty regarding the levels of continuous scores of students and their cognitive states, we extend the application of the most deep-learning-based KT models in continuous scenarios.
— The above-mentioned two benefits are demonstrated as follows. (a) Its intrinsic interpretability is explained through the rules and hidden semantics (Section 5), and post hoc interpretability is experimentally visualized (Section 6.3). (b) Better prediction performance, in the continuous-score application, is achieved when compared to 14 state-of-the-art models using four real-world datasets (Section 6.2).

This paper is organized as follows. Related work is reviewed in the next section. In Section 3, background material is presented, including the deep knowledge tracing and fuzzy neural networks. In Section 4, the framework of FDKT is detailed. In Section 5, the intrinsic interpretability of FDKT is presented. In Section 6, the experiments are discussed. Finally, Section 7 concludes the paper.

Table 1. Comparison between FDKT and Some Representative Models[1]

Models	BKT [30]	DKT [21]	FuzzyCDF [28]	FBKT [29]
Fuzzy sets	✗	✗	✓	✓
Fuzzy reasoning	✗	✗	✗	✗
Dynamic data	✓	✓	✗	✓
Continuous scores	✗	–	✓	✓
Mixing KCs	✗	✓	✓	✗
Interpretability	–	–	Visualization	Example
Models	DKVMN [24]	DeepIRT [25]	SAKT [34]	KQN [27]
Fuzzy sets	✗	✗	✗	✗
Fuzzy reasoning	✗	✗	✗	✗
Dynamic data	✓	✓	✓	✓
Continuous scores	–	–	–	–
Mixing KCs	✓	✓	✓	✓
Interpretability	Attention & Visualization	Combination[2] & Example	Attention & Visualization	Visualization
Models	AKT [35]	CKT [36]	**FDKT**	
Fuzzy sets	✗	✗	✓	
Fuzzy reasoning	✗	✗	✓	
Dynamic data	✓	✓	✓	
Continuous scores	–	–	✓	
Mixing KCs	✓	✓	✓	
Interpretability	Attention & Visualization	Visualization	Rules & Hidden semantics & Example & Visualization	

[1]– refers that the item has not been demonstrated in the paper.
[2]*Combination* refers to a combination of the KT model and the traditional model in education.

2 RELATED WORK

The related work is introduced including the KT models and the interpretability in educational data mining. Several representative models mentioned in this section are compared with the proposed FDKT in Table 1.

2.1 Knowledge Tracing

With regard to the two mainstream types, that is, Bayesian and deep-learning-based KT models, the former rely on intrinsic interpretable first-order Markov models [30]. However, their prediction performance is not satisfactory as they are less representative in terms of the complexity of the human brain and human knowledge. The latter have shown remarkable improvement in terms of prediction accuracy using deep learning methods, with the strong characterization capabilities. Therefore, they have attracted a significant amounts of attention [21, 24–27, 37].

In the type of deep-learning-based KT models, DKT uses **recurrent neural networks (RNNs)** to model student learning and achieves an excellent AUC in prediction performance [21]. Using a memory-augmented neural network, DKVMN exploits the relationships between concepts [24]. [38] proposed three distributed memory networks to model student performance, i.e., DMN, ADMN, IADMN. To enhance the predictive consistency in DKT, [39] introduced regularization terms to propose DKT+. DKT_DSC assigns students to a distinct classification to improve the accuracy of DKT [40]. Deep-IRT is a synthesis of the **item response theory (IRT)** model and DKVMN. Thus, it retains the prediction performance of the DKVMN and interpretability of IRT [25]. The self-attention-based approach SAKT captures a complex representation of human learning [34]. KQN introduces the probabilistic skill similarity of the knowledge components [27]. AKT uses a novel monotonic attention mechanism and the Rasch model to regularize the concept and question embeddings [35]. CKT models individualization in KT [36]. The federated DKT collectively trains high-quality DKT models for multiple silos using the federated learning method [37]. Based on the dual-attentional mechanism, MF-DAKT [41] enriches question representations and utilizes multiple factors to model the knowledge tracing process. CL4KT [42] uses four data enhancement methods and hard negatives to reveal the learning history of similar and dissimilar semantics. With the evolution of graph neural networks [43–46], researchers have begun delving into the graph structural relationships within KT tasks. GIKT [47] employs the graph convolutional network to effectively integrate the problem-skill correlation.

2.2 Interpretability in Educational Data Mining

In recent years, model interpretability has attracted more attention by researchers in the field of educational data mining, including student models. A model is expected to be easy to understand with satisfactory prediction performance [20]. In this subsection, the existing studies towards interpretable student models are introduced (KT is considered as a type of student model).

For the traditional student models like DINA [48], IRT [49], LFA [50], PFA [50], and BKT [30], they provide better understanding based on interpretable probabilistic statistics or Markov models, and the like. To improve the prediction performance of the traditional models, deep-learning-based models spring up. However, it is an open problem of model interpretability because there are large vectors of artificial 'neurons' [21, 31].

To alleviate this problem, we classify the subsequent work into the following categories. (1) Some introduced the educational theory into the models. For example, NeuralCD [51] placed a monotonicity assumption taking from an educational property on the framework to enhance its interpretability, where the monotonicity assumption is described as follows: the probability of correct response to the exercise is monotonically increasing at any dimension of the student's knowledge proficiency. DIRT [52] and Deep-IRT [25] combined deep learning with IRT to make the model more explainable. (2) Attention-based methods also offer some interpretability to student models. For example, References [24, 35, 53, 54] utilized attention mechanism to make the models more interpretable. (3) Many studies also take advantage of visualization towards interpretability [24, 27, 28, 34–36, 55–57]. They vividly demonstrated part of the results of the models via visualization.

The above studies have made a certain effort towards interpretable KT models, owing to the methods like visualization and attention. However, the KT models with intrinsic interpretability still need to be further explored.

3 BACKGROUND

The backgrounds in DKT, fuzzy theory, and fuzzy neural networks (FNN) are introduced.

3.1 Deep Knowledge Tracing

KT models the students' performance on exercises in a time-varying prediction task, where each exercise is related to a knowledge component.

We use the DKT model as an example to explain the KT process. As shown in Figure 4, the student answers an exercise at each time step. $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_T$ denotes the input vector at each time step, where $\boldsymbol{x}_t$ contains the following two aspects of information: (1) the knowledge components

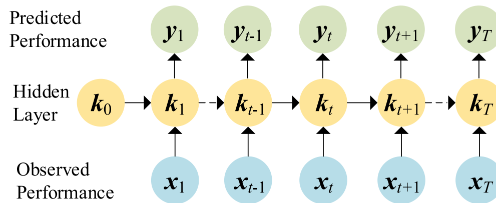

Fig. 4. Framework of the DKT models [21]. In the model, there are input, hidden, and output layers, where input and output layers corresponding to the observed performance and predicted performance, respectively. $\boldsymbol{x}_t$, $\boldsymbol{k}_t$, and $\boldsymbol{y}_t$ are the representation vectors of observed performance, hidden variable, and predicted performance at time step t, respectively.

of the exercise that the student answers at time step t; and (2) the score of the exercise that the student achieves at time step t. In particular, the scores of exercises in the traditional DKT model were only taken in $\{0, 1\}$. $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_T$ denotes the output vector at each time step, where $\boldsymbol{y}_t$ represents the predicted probability vector that the student would respond with correct answers to the exercises, related to each knowledge component at time step t. $\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_T$ denotes the hidden vector in the network that temporarily stores information. The objective of the DKT model is to minimize the negative log-likelihood of the observed sequence of the students' scores.

3.2 Fuzzy Theory

Fuzzy logic [58] is an expansion of binary logic. It was developed to address ambiguities that exist in the real world, such as hot and cold, fast and slow, and large and small. In classical two-valued logic, all objects are assumed explicit [59]. For example, in a classification task, an object may or may not belong to this class. Fuzzy logic solves many problems in reality that cannot be clearly described.

Fuzzy Sets and Membership Functions. Fuzzy sets [60] are a fundamental concept in fuzzy logic theory. Fuzzy sets allow for the representation of uncertainty and vagueness by assigning degrees of membership to elements. In a fuzzy set, each element of the universe of discourse can have a membership value ranging from 0 to 1, indicating the degree to which the element belongs to the set. The membership function defines this mapping of elements to membership degrees. Various types of membership functions can be used, such as triangular, trapezoidal, Gaussian, or sigmoidal functions, depending on the nature of the problem and the desired representation. A formal description of the fuzzy sets and their operations is as follows: Suppose there exist fuzzy sets $\tilde{M}_i$ and $\tilde{S}_j$. The membership functions $mf_i^{(k)}$ and $mf_j^{(k)}$ denote the degrees to which element k belongs to $\tilde{M}_i$ and $\tilde{S}_j$, respectively.

T-norm Fuzzy Logics. The main objective of t-norm fuzzy logics [61] is to extend classical two-valued logic by introducing intermediary truth values between 1 (representing truth) and 0 (representing falsity). These intermediary truth values serve to quantify the degrees of truth associated with propositions. The degrees of truth in t-norm fuzzy logics are considered to be real numbers within the range of the unit interval $[0, 1]$. Prominent examples include the minimum t-norm, product t-norm, and Lukasiewicz t-norm, among others. For example, the fuzzy intersected set of $\tilde{M}_i$ and $\tilde{S}_j$ is denoted by $\tilde{M}_i \sqcap \tilde{S}_j$. When using minimum t-norm logics, the membership function $mf_{i,j}^{(k)}$ is defined as $mf_{i,j}^{(k)} = \min\{mf_i^{(k)}, mf_j^{(k)}\}$. When using product t-norm logics, $mf_{i,j}^{(k)} = mf_i^{(k)} \cdot mf_j^{(k)}$. Due to the widespread application of the minimum t-norm in fuzzy logic, this calculation method will be used in the subsequent sections.

Table 2. Notations

Notation	Description
I	Number of the fuzzy cognition sets
J	Number of the fuzzy score sets
K	Number of the knowledge components
T	Total time steps
$x_t^{(k)}$	Observed score on the knowledge components k at time step t, $k \in \{1, 2, \dots, K\}$, $t \in \{1, 2, \dots, T\}$
$y_t^{(k)}$	Target score of the knowledge components k at time step t, $k \in \{1, 2, \dots, K\}$, $t \in \{1, 2, \dots, T\}$
$\hat{y}_t^{(k)}$	Predicted score of the knowledge components k at time step t, $k \in \{1, 2, \dots, K\}$, $t \in \{1, 2, \dots, T\}$
$m_t^{(k)}$	Student's cognition of the knowledge components k at time step t, $k \in \{1, 2, \dots, K\}$, $t \in \{1, 2, \dots, T\}$
$\tilde{M}_i$	The i-th fuzzy cognition set, $i \in \{1, 2, \dots, I\}$
$\tilde{S}_j$	The j-th fuzzy score set, $j \in \{1, 2, \dots, J\}$
$mc_{t,i}^{(k)}$	Membership value (probability) of $m_t^{(k)} \in \tilde{M}_i$
$ms_{t,j}^{(k)}$	Membership value (probability) of $x_t^{(k)} \in \tilde{S}_j$
$RN_{i,j}$	Fuzzy rule node when $m_{t-1}^{(k)} \in \tilde{M}_i$ and $x_t^{(k)} \in \tilde{S}_j$
$r_{i,j}$	Output of the fuzzy rule node $RN_{i,j}$
μ_j	Mean value in the Gaussian membership function of $\tilde{S}_j$
σ_j	Standard deviation value in the Gaussian membership function of $\tilde{S}_j$
$\boldsymbol{w}_1$	Weight vector in the prediction module
$\boldsymbol{w}_2$	Weight vector in the fuzzy reasoning module

Fuzzy Rules. A fuzzy system is essentially a rule-based expert system consisting of a set of linguistic rules and one of the most commonly used fuzzy rules in the form of IF-THEN [62]. A formal description of the fuzzy rules is as follows: R: IF x_1 is $\tilde{M}_1$, and ..., x_i is $\tilde{M}_i$, THEN y_1 is $\tilde{S}_1$, and ..., and y_j is $\tilde{S}_j$, where $\tilde{M}_1, \dots, \tilde{M}_i$ and $\tilde{S}_1, \dots, \tilde{S}_j$ are fuzzy sets.

3.3 Fuzzy Neural Networks

FNN is gradually turning into a research hotpot, because it combines the powerful calculation and representation capabilities of the neural networks with the heuristic expert knowledge of the fuzzy system. For example, IF-THEN [62] (introduced in Section 3.2) expresses the output preferences under the specified conditions, which is a kind of knowledge.

The traditional FNN is limited to static problems due to its feedforward network structure [33]. To address this shortcoming, Lee and Teng [63] proposed the **recurrent FNN (RFNN)** by capturing the dynamic response of the system through its internal feedback loop, which is more suitable for describing dynamic systems as compared to the FNN.

In the RFNN, there are four layers: input, membership, rule, and output layers. The input nodes are fuzzified into the membership layers that contain the memory terms storing the past information of the network. The membership nodes enter the rule layer through the application of fuzzy intersection operation (detailed in Section 3.2). Finally, the output nodes are obtained through a linear combination of each rule node. The RFNN can be shown to be a universal uniform approximator for continuous functions over compact sets if it satisfies a certain condition [63].

4 FRAMEWORK OF FDKT

FDKT is proposed to enhance the interpretability of the deep-learning-based KT models, owing to the reasoning of the fuzzy rule-based module. In this section, we first formulate the task and then present the model of the FDKT containing the fuzzification, fuzzy reasoning, and prediction modules. Subsequently, the rules of fuzzy reasoning and the layered operation of FDKT are detailed. Finally, the time complexity is analyzed.

The notation used in this paper is listed in Table 2.

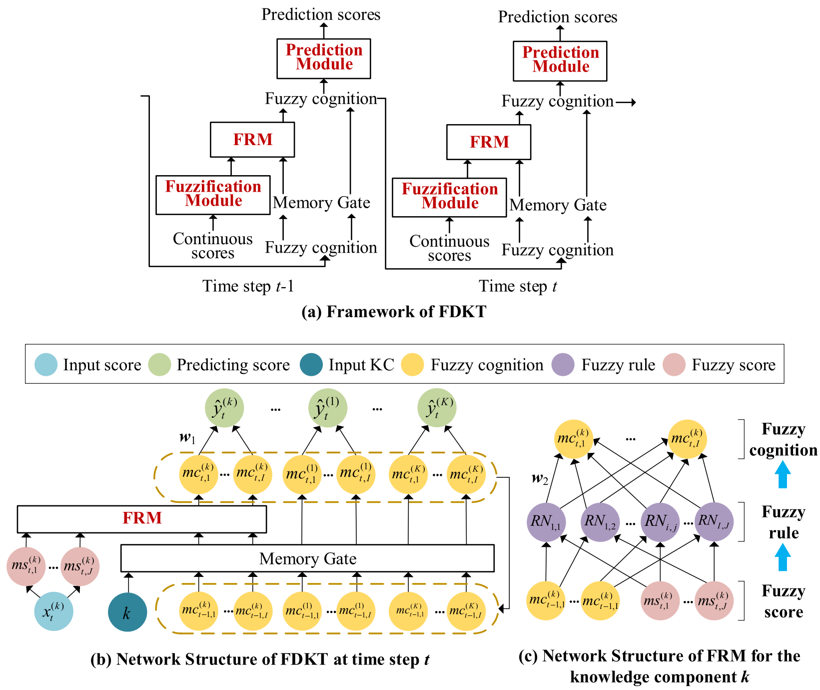

Fig. 5. Schematic of the FDKT. (a) is the framework of FDKT, which includes three main modules: fuzzification, fuzzy reasoning, and prediction. (b) and (c) depict the network structures of FDKT and FRM at time step t, respectively.

4.1 Formulation

FDKT aims to estimate student cognition of the knowledge components and predict their future performance on exercises based on previous performance. Notably, the input is the performance which is continuous; however the input has been bisected in most existing studies. We denote I, J, K as the numbers of fuzzy cognition sets, fuzzy score sets, and knowledge components, respectively. Further, T is the total time step. The input consists of the input continuous score x_t and knowledge component $k, k \in \{1, 2, \ldots, K\}$. FDKT estimates the current cognition $(m_t^{(1)}, \ldots, m_t^{(K)})$ and predicts the next-time-step performance $y_t^{(k')}, k' \in (1, 2, \ldots, K)$ on k' based on $m_t^{(k')}$. For the current cognition of k, $m_t^{(k)} \in \tilde{M}_1, \ldots, m_t^{(k)} \in \tilde{M}_I$ with probabilities of $mc_{t,1}^{(k)}, \ldots, mc_{t,I}^{(k)}$, respectively, where $\{\tilde{M}_1, \tilde{M}_2, \ldots, \tilde{M}_I\}$ denotes I fuzzy cognition sets. For clarity, the notation used is listed in Table 2, in order of appearance in this paper.

Optimization. The objective of FDKT is to minimize the loss $\mathcal{L}_f = l(\boldsymbol{y}, \hat{\boldsymbol{y}})$ between the ground truth and prediction scores, optimized through gradient descent on batches. $l(\cdot)$ denotes the mean absolute error.

4.2 Model

The framework of FDKT is shown in Figure 5(a). In the framework, the fuzzy score sets and fuzzy cognition sets are defined as follows. **Fuzzy score sets** are the fuzzy sets defined for the continuous scores obtained by students answering the exercises. Different fuzzy score sets represent different score levels and continuous scores belong to fuzzy score sets with a certain probability. **Fuzzy cognition sets** are the fuzzy sets defined for the students' cognitive states of knowledge

ALGORITHM 1: Fuzzy deep knowledge tracing

Input: $I, J, K, T, x_t^{(k)}$
Output: $\boldsymbol{y}$
Initialize $(\boldsymbol{mc}_0^{(1)}, \boldsymbol{mc}_0^{(2)}, \ldots, \boldsymbol{mc}_0^{(K)})$ on K knowledge components at initial time step.
Let $t = 1$.
while $t <= T$ **do**
  Fuzzify $x_t^{(k)}$ into J fuzzy scores $\tilde{S}$ with membership values $\boldsymbol{ms}_t^{(k)}$ through the **fuzzification module**.
  Obtain $\boldsymbol{mc}_t^{(k)}$ of k from **FRM** (conducted by Algorithm 2).
  Obtain $\boldsymbol{mc}_t^{(1)}, \ldots, \boldsymbol{mc}_t^{(k-1)}, \boldsymbol{mc}_t^{(k+1)}, \ldots, \boldsymbol{mc}_t^{(K)}$ $(\boldsymbol{mc}_t^{(k')} = \boldsymbol{mc}_{t-1}^{(k')}, k' \neq k)$.
  Predict the performance $\hat{\boldsymbol{y}}_t = (\hat{y}_t^{(1)}, \hat{y}_t^{(2)}, \ldots, \hat{y}_t^{(K)})$.
end while
return $\hat{\boldsymbol{y}} = (\hat{\boldsymbol{y}}_1, \hat{\boldsymbol{y}}_2, \ldots, \hat{\boldsymbol{y}}_T)$.

components. Different fuzzy cognition sets represent different levels of cognition. As shown in Figure 5(a), FDKT contains three main modules: the fuzzification, fuzzy reasoning, and prediction modules.

Specifically, the network structure of FDKT at time step t is shown in Figure 5(b). The **fuzzification module** addresses the input of the continuous score $x_t^{(k)}$ into the fuzzy scores (denoted as $\tilde{S} = \{\tilde{S}_1, \tilde{S}_2, \ldots, \tilde{S}_J\}$). The **fuzzy reasoning module (FRM)** determines the current fuzzy cognition $m_t^{(k)}$ based on the fuzzy scores from the fuzzification module and the historical cognition $m_{t-1}^{(k)}$ of k, where $m_{t-1}^{(k)}$ is obtained from $(m_{t-1}^{(1)}, \ldots, m_{t-1}^{(K)})$ through the memory gate. FRM promotes the interpretability of FDKT because it can estimate the student cognition on knowledge components through the use of fuzzy rules, thereby explaining the result of the prediction performance. Finally, the **prediction module** obtains the future performance $\boldsymbol{y}_t$ based on $(m_t^{(1)}, \ldots, m_t^{(K)})$.

The pseudo-code of FDKT is detailed in Algorithm 1. The remainder of this section details the three modules.

4.2.1 Fuzzification Module. The fuzzification module fuzzifies the continuous scores into several fuzzy scores. The continuous score has a certain probability belonging to each fuzzy score, where the probability is referred to as the membership. The Gaussian fuzzy logic system is applied to describe the membership function of the fuzzy scores, as detailed in Equation (1). It is worth noting that, after calculating the membership degrees of an individual to different fuzzy sets, we performed probability normalization on these membership degree values to ensure their sum is equal to 1. This normalization was done to transform all membership degree distributions into a standardized form, allowing for a more intuitive representation of the relative sizes and proportions of probabilities.

$$ms_{t,j}^{(k)} = \exp\left\{-\frac{(x_t^{(k)} - \mu_j)^2}{\sigma_j^2}\right\}, \tag{1}$$

where $x_t^{(k)}$ denotes the continuous score at time step t related to the knowledge component k. $ms_{t,j}^{(k)}$ denotes the membership of $x_t^{(k)}$ belongs to the fuzzy score $\tilde{S}_j$. μ_j and σ_j denote the mean and std of $\tilde{S}_j$, respectively.

4.2.2 Fuzzy Reasoning Module. The network structure of FRM is designed as follows, and the process of fuzzy reasoning is detailed in Section 4.3. FRM determines the current cognition at each

ALGORITHM 2: Fuzzy reasoning module

Input: $\boldsymbol{mc}_{t-1}^{(k)} = (mc_{t-1,1}^{(k)}, mc_{t-1,2}^{(k)}, \ldots, mc_{t-1,I}^{(k)})$ and $\boldsymbol{ms}_{t}^{(k)} = (ms_{t,1}^{(k)}, ms_{t,2}^{(k)}, \ldots, ms_{t,J}^{(k)})$
Parameter: $\boldsymbol{w}_2$
Output: $\boldsymbol{mc}_t^{(k)}$
Calculate $\boldsymbol{r} = (r_{1,1}, r_{1,2}, \ldots, r_{I,J})$ according to Equation (2).
Let $u = 1$.
while $u <= I$ **do**
    Calculate $mc_{t,u}^{(k)}$ that $m_t^k \in \tilde{M}_u$ according to Equation (3).
end while
return $\boldsymbol{mc}_t^{(k)} = (mc_{t,1}^{(k)}, mc_{t,2}^{(k)}, \ldots, mc_{t,I}^{(k)})$.

time step, based on both last cognition (factor A) and current performance of the exercise (factor B). The former is obtained from the fuzzy cognition at time step $(t-1)$ and the latter is the output of the fuzzification module at time step t. As shown in Figure 5(c), different combinations of factors A and B lead to different fuzzy cognitions. Therefore, there are $I * J$ fuzzy rules corresponding to $I * J$ combinations of factors A and B. The pseudo-code of FRM is detailed in Algorithm 2, where $r_{i,j} \in \boldsymbol{r}$ is the output of the fuzzy rule node $RN_{i,j}$, given by Equation (2).

$$r_{i,j} = ms_{t,j}^{(k)} * mc_{t-1,i}^{(k)}. \tag{2}$$

Subsequently, the probability $mc_{t,u}^{(k)}$ that the current cognition $m_t^{(k)} \in \tilde{M}_u$ is given by Equation (3).

$$mc_{t,u}^{(k)} = f_{\boldsymbol{w}_{2,u}}(\boldsymbol{r}) = \sum_{i=1}^{I}\sum_{j=1}^{J} w_{u,i,j} * r_{i,j}. \tag{3}$$

$\boldsymbol{w}_2 = (\boldsymbol{w}_{2,1}, \boldsymbol{w}_{2,2}, \ldots, \boldsymbol{w}_{2,I})$ denotes the adjustable weight, where $\boldsymbol{w}_{2,u} = (w_{u,1,1}, w_{u,1,2}, \ldots, w_{u,I,J})$, $u \in \{1, 2, \ldots, I\}$.

4.2.3 *Prediction Module.* The prediction module predicts the future performance of the students on exercises based on their current cognition of the knowledge components. Specifically, there are K outputs in the prediction process, as shown in Equation (4), corresponding to the prediction performance on the exercise related to K knowledge components. The performance on k is calculated using a linear function as expressed in Equation (5).

$$\hat{\boldsymbol{y}}_t = (\hat{y}_t^{(1)}, \ldots, \hat{y}_t^{(K)}), \tag{4}$$

where $\hat{y}_t^{(k)}, k \in \{1, 2, \ldots, K\}$ satisfies Equation (5).

$$\hat{y}_t^{(k)} = f_{\boldsymbol{w}_{1,k}}(\boldsymbol{mc}_t^{(k)}) = \boldsymbol{w}_{1,k} \cdot \boldsymbol{mc}_t^{(k)}. \tag{5}$$

$\boldsymbol{w}_1 = (\boldsymbol{w}_{1,1}, \boldsymbol{w}_{1,2}, \ldots, \boldsymbol{w}_{1,K})$ denotes the adjustable parameter, where $\boldsymbol{w}_{1,k} = (w_{1,k,1}, w_{1,k,2}, \ldots, w_{1,k,I})$, $k \in \{1, 2, \ldots, K\}$. $\boldsymbol{mc}_t^{(k)} = (mc_{t,1}^{(k)}, mc_{t,2}^{(k)}, \ldots, mc_{t,I}^{(k)})$.

4.3 Fuzzy Reasoning

The process of fuzzy reasoning is detailed, to deduce the current fuzzy cognition from the last fuzzy cognition and the current performance. This is the core of the interpretability of FDKT.

4.3.1 Reasoning in Memory Gate. According to the network structure of FDKT at time step t in Figure 5(b), the current fuzzy cognition is obtained from the FRM or directly from the last fuzzy cognition through the memory gate. In other word, $\boldsymbol{mc}_t^{(k)}$ satisfies the decision rules as $\boldsymbol{R}1 = \{R1^{(1)}, R1^{(2)}, \ldots, R1^{(K)}\}$, where $R1^{(k)}(k \in \{1, 2, \ldots, K\})$ is given by Equation (6).

$$\begin{aligned} R1^{(k)}\text{: } & \textbf{if } k = k', \ \textbf{then } \boldsymbol{mc}_t^{(k)} \text{ satisfies } FRM(\boldsymbol{mc}_{t-1}^{(k)}), \\ & \textbf{else } \boldsymbol{mc}_t^{(k)} = \boldsymbol{mc}_{t-1}^{(k)}, \\ & \textbf{when } k' \text{ is the knowledge component at time step } t. \end{aligned} \tag{6}$$

The antecedent is the knowledge component k whether related to the exercise at time step t, and the consequent is the probability of the current cognition $\boldsymbol{mc}_t^{(k)}$. $FRM(\boldsymbol{mc}_{t-1}^{(k)})$ is denoted as the current cognition obtained from the FRM.

4.3.2 Reasoning in FRM. According to the network structure of the FRM in Figure 5(c), the current fuzzy cognition is obtained from the current performance on the exercise and the last fuzzy cognition through $I * J$ fuzzy rules, where I and J denote the numbers of the fuzzy cognition and fuzzy score sets, respectively. For each fuzzy rule node $RN_{i,j}$, its effects on $m_t^{(k)}$ belonging to $\tilde{M}_1, \ldots, \tilde{M}_I$ satisfy the rules expressed in Equation (7), where the antecedents are the current performance $x_t^{(k)}$ and last cognition $m_{t-1}^{(k)}$ and the consequent is the effects of $R2_{i,j}^{(k)}$.

$$\begin{aligned} R2_{i,j}^{(k)}\text{: } & \textbf{if } m_{t-1}^{(k)} \in \tilde{M}_i \text{ with probability } mc_{t-1,i}^{(k)} \\ & \textbf{and } x_t^{(k)} \in \tilde{S}_j \text{ with probability } ms_{t,j}^{(k)}, \\ & \textbf{then} \text{ the effect on } m_t^{(k)} \in \tilde{M}_1 \text{ is } w_{1,i,j} * r_{i,j}, \\ & \textbf{and} \ldots \ \textbf{and} \text{ the effect on } m_t^{(k)} \in \tilde{M}_I \text{ is } w_{I,i,j} * r_{i,j}, \end{aligned} \tag{7}$$

where $r_{i,j}$ is obtained according to Equation (2). The fuzzy rule node $RN_{i,j}$ indicates the combination of factors A and B, where the former is $m_{t-1}^{(k)} \in \tilde{M}_i$ and the latter is $x_t^{(k)} \in \tilde{S}_j$. $m_t^{(k)}$ and $m_{t-1}^{(k)}$ denote the cognition of k at time steps t and $t-1$, respectively. $x_t^{(k)}$ denotes the continuous score of k. $\tilde{M}_i$ and $\tilde{S}_j$ represent the i-th fuzzy cognition set and the j-th fuzzy score set, respectively.

Then, the current fuzzy cognition is obtained from the fuzzy rule nodes, satisfying rule $\boldsymbol{R}3 = \{\boldsymbol{R}3^{(1)}, \boldsymbol{R}3^{(2)}, \ldots, \boldsymbol{R}3^{(K)}\}$. $\boldsymbol{R}3^{(k)} = \{R3_1^{(k)}, \ldots, R3_I^{(k)}\}(k \in \{1, 2, \ldots, K\})$, where $R3_u^{(k)}(u \in 1, 2, \ldots, I)$ satisfies Equation (8).

$$\begin{aligned} R3_u^{(k)}\text{: } & \textbf{if} \text{ the effect of } RN_{1,1} \text{ on } m_t^{(k)} \in \tilde{M}_u \text{ is } w_{u,1,1} * r_{1,1}, \\ & \textbf{and} \ldots \\ & \textbf{and} \text{ the effect of } RN_{I,J} \text{ on } m_t^{(k)} \in \tilde{M}_u \text{ is } w_{u,I,J} * r_{I,J}, \\ & \textbf{then} \text{ the probability of } m_t^{(k)} \in \tilde{M}_u \text{ is} \\ & \sum_{i=1}^{I} \sum_{j=1}^{J} w_{u,i,j} * r_{i,j}, \end{aligned} \tag{8}$$

where the antecedents are the effects of the fuzzy rule nodes, and the consequent is the current cognition $m_t^{(k)} \in \tilde{M}_u$. This FRM in FDKT can be considered as a dynamic fuzzy inference system because its input contains a memory term for storing the past fuzzy cognition using the feedback unit [63].

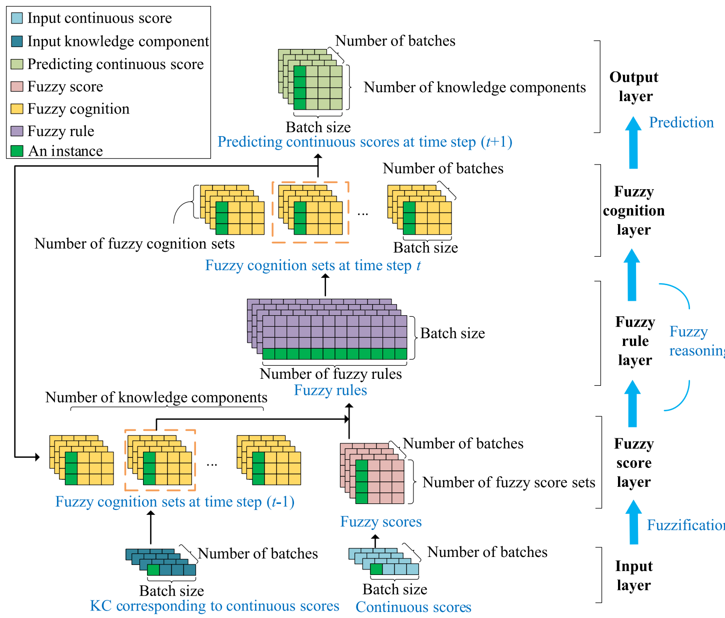

Fig. 6. The architecture of FDKT (taking the t-th time step as an example) is presented to explain the hidden semantics from the input to the output of the FDKT.

4.4 Layered Operation of FDKT

The layered operation of the proposed model FDKT is detailed in this subsection to describe the integration of the three modules, as shown in Figure 6. We denote $i^{(p)}(t)$ and $o^{(p)}(t)$ as the input and output in the p-th layer ($p \in \{1,2,3,4,5\}$ in the architecture) at time step t.

For the fuzzy cognition of the knowledge component at time step t, the operation for the 1, 2, 3, 4-th layers is shown in Equations (9)–(15). For the knowledge component that is not at time step t, the operation before the 5-th layer is shown in (16).

The input in the $(p+1)$-th layer equals the output in the p-th layer for the layer without memory terms ($p \in \{1,3,4\}$). In other words, $\boldsymbol{i}^{(p+1)}(t) = \boldsymbol{o}^{(p)}(t), p \in \{1,3,4\}$. In the 1-th layer, the output equals the input given by Equation (9).

$$o^{(1)}(t) = i^{(1)}(t). \tag{9}$$

$$\boldsymbol{o}^{(2)}(t) = (o_1^{(2)}(t), o_2^{(2)}(t), \ldots, o_J^{(2)}(t)), \tag{10}$$

where $o_j^{(2)}(t)$ satisfies Equation (11).

$$o_j^{(2)}(t) = \exp\left\{-\frac{(i^{(2)}(t)-\mu_j)^2}{\sigma_j^2}\right\}, j \in \{1,2,\ldots,J\}, \tag{11}$$

where J is the number of fuzzy score sets. μ_j and σ_j denote the mean and std in the membership function of the fuzzy score set $\tilde{S}_j$, respectively.

In the 3-th layer, that is, the fuzzy rule layer with the memory terms, its input contains two aspects $\boldsymbol{is}^{(3)}(t) = (is_1^{(3)}(t), is_2^{(3)}(t), \ldots, is_J^{(3)}(t))$ and $\boldsymbol{ic}^{(3)}(t) = (ic_1^{(3)}(t), ic_2^{(3)}(t), \ldots, ic_I^{(3)}(t))$. J and I

are the numbers of fuzzy score sets and fuzzy cognition sets, respectively.

$$\boldsymbol{o}^{(3)}(t) = (o^{(3)}_{1,1}(t), \dots, o^{(2)}_{I,J}(t)), \tag{12}$$

where $o^{(3)}_{i,j}(t)$ satisfies Equation (13).

$$o^{(3)}_{i,j}(t) = is^{(3)}_j(t) \prod ic^{(3)}_i(t), \tag{13}$$

where $is^{(3)}_j(t) = o^{(2)}(t)$ and $ic^{(3)}_i(t) = o^{(4)}(t-1)$.

$$\boldsymbol{o}^{(4)}(t) = (o^{(4)}_1(t), \dots, o^{(4)}_I(t)), \tag{14}$$

where $o^{(4)}_u(t), u \in \{1, 2, \dots, I\}$ satisfies Equation (15).

$$o^{(4)}_u(t) = \sum_{i=1}^{I} \sum_{j=1}^{J} w_{u,i,j} * i^{(4)}_{i,j}(t), \tag{15}$$

where $w_{u,i,j} \in \{w_{1,i,j}, w_{2,i,j}, \dots, w_{I,i,j}\}$ is an adjustable parameter.

For the fuzzy cognition of the knowledge components not at time step t, the operation for the 1, 2, 3, 4-th layers is given by Equation (16).

$$\boldsymbol{o}^{(4)}_u(t) = \boldsymbol{o}^{(4)}_u(t-1). \tag{16}$$

Thus, the fuzzy cognition of all the knowledge components $\boldsymbol{o}^{(4)}(t) = \{\boldsymbol{o}^{(4)}_1(t), \boldsymbol{o}^{(4)}_2(t), \dots, \boldsymbol{o}^{(4)}_K(t)\}$ is obtained, where K is the number of knowledge components. $\boldsymbol{o}^{(4)}_k(t), k \in \{1, 2, \dots, K\}$ is obtained using Equation (14) if k is the conducted knowledge component or Equation (16) otherwise. Subsequently, the input and output of the 5-th layer are given by Equation (17).

$$\boldsymbol{o}^{(5)}(t) = (o^{(5)}_1(t), \dots, o^{(5)}_K(t)), \tag{17}$$

X where $o^{(5)}_k(t), k \in \{1, 2, \dots, K\}$ satisfies Equation (18).

$$o^{(5)}_k(t) = f(\boldsymbol{i}^{(5)}_k(t)), \tag{18}$$

where $f(\cdot)$ is a linear function. $\boldsymbol{i}^{(5)}_k(t) = \boldsymbol{o}^{(4)}_k(t)$ and $\boldsymbol{o}^{(4)}_k(t) = (o^{(4)}_1(t), \dots, o^{(4)}_I(t))$.

4.5 Time Complexity

The time complexity of FDKT is analyzed as follows. The FDKT algorithm is presented in Algorithm 1, invoking Algorithm 2. The time complexity of Algorithm 2 is $O(I^2J)$, where I and J denote the numbers of fuzzy cognition sets and fuzzy score sets, respectively. Algorithm 2 is in the loop with respect to the time steps of Algorithm 1. Therefore, the time complexity of FDKT (Algorithm 1) for an epoch is $O(TI^2J)$, where T denotes the total number of time steps.

We also analyze the complexities for other KT models (shown in Table 1), which are detailed in Appendix. The time complexity of the proposed FDKT model can be observed to be lower than that of the neural network-based KT models because the values of the fuzzy cognitive and fuzzy score sets are set to integers less than 10, whereas the general representation dimension is set to tens or hundreds. Noteworthily, at lower time complexity, FDKT performs better than the general neural network-based KT models in continuous scenarios, with more convenient parameter optimization than traditional non-neural-network-based KT models.

5 INTRINSIC INTERPRETABILITY OF FDKT

As mentioned previously, there are two types of interpretability: intrinsic and post hoc. In this section, FDKT is explained using rules and hidden semantics to demonstrate its intrinsic interpretability [15], in other words, to answer the question, *how does the model work* (shown in Figure 2(b)). Then, an example is considered to clearly demonstrate the process followed by FDKT. The interpretability of FDKT is also illustrated from the post hoc aspect via experiments (Section 6.3).

5.1 Explanation by Rules

In this subsection, FDKT is explained with the help of the rules, where the rules may be the most powerfully explanatory model [64].

From the input to the output of FDKT, the current cognition of the students is deduced using the rules ($\boldsymbol{R}1, \boldsymbol{R}2$, and $\boldsymbol{R}3$) expressed in Section 4.3, and subsequently, the future performance is predicted according to the current cognition using Equation (5). Specifically, after the fuzzification of the input continuous score, $\boldsymbol{R}1$ is to select the fuzzy cognition of the knowledge component to be updated. The selected fuzzy cognition and the fuzzy score are both fed into the FRM, and they will first meet $\boldsymbol{R}2$. The number of the rules in $\boldsymbol{R}2$ depends on the combinations of the fuzzy cognition and the fuzzy score. Each fuzzy rule node generates its effect on the current fuzzy cognition, according to its corresponding rule in $\boldsymbol{R}2$. Subsequently, each rule in $\boldsymbol{R}3$ obtains the probability of the current fuzzy cognition by summing the effects of all the fuzzy rule nodes.

In the proposed model, the network structure is constructed based on the fuzzy rules, which rely on prior knowledge. This demonstrates that FDKT has intrinsic interpretability.

5.2 Explanation by Hidden Semantics

Based on common knowledge of this field, we make sense of the semantics of the hidden layers and parameters in the model. This makes the process significantly easier to understand.

5.2.1 Semantics of Hidden Layers. The input nodes are fed into the fuzzification module to obtain the fuzzy scores. Then, the fuzzy cognition nodes are obtained from fuzzy scores through fuzzy rule nodes using the FRM. Finally, the output prediction performance is obtained from the fuzzy cognition using the prediction module.

5.2.2 Semantics of Parameters. In the **fuzzification module**, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_J)$, and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_J)$ (Equation (1)) denote the mean and std of each fuzzy score set, respectively.

In the **FRM**, $r_{i,j}$ (Equation (2)) denotes the probability of the fuzzy rule node $RN_{i,j}$, specifically, the probability that the last cognition $m_{t-1}^{(k)} \in \tilde{M}_i$ and the current fuzzy score $x_t^{(k)} \in \tilde{S}_j$. $mc_{t,u}^{(k)}$ (Equation (3)) represents the probability that the current cognition $m_t^{(k)} \in \tilde{M}_u$. $w_{u,i,j}$ represents the contribution of the fuzzy rule node $RN_{i,j}$ to each fuzzy cognition of $\tilde{M}_u$. When $r_{i,j}$ increases by 0.1, $mc_{t,u}^{(k)}$ increases by $0.1 * w_{u,i,j}(w_{u,i,j} \in \boldsymbol{w}_2)$ (Equation (2)).

In the **prediction module**, $\boldsymbol{w}_1$ (Equation (5)) represents the contribution of the fuzzy cognition to the predicting performance. If $mc_{t,u}^{(k)}$ increases by 0.1, the prediction performance on k increases by $0.1 * w_{1,k,u}(w_{1,k,u} \in \boldsymbol{w}_1)$.

5.3 Example

By reviewing Figure 2, the interpretable model explains why the model can obtain this prediction. Based on this, a simple example is considered to discuss the interpretability of FDKT. As can be understood from Figure 7, FDKT can not only obtain the prediction results but also explain them.

Five fuzzy cognition sets (i.e., the first to fifth bin from the best to the worst) and four fuzzy score sets (i.e., poor, medium, good, and excellent) are defined in the example. It is worth noting that,

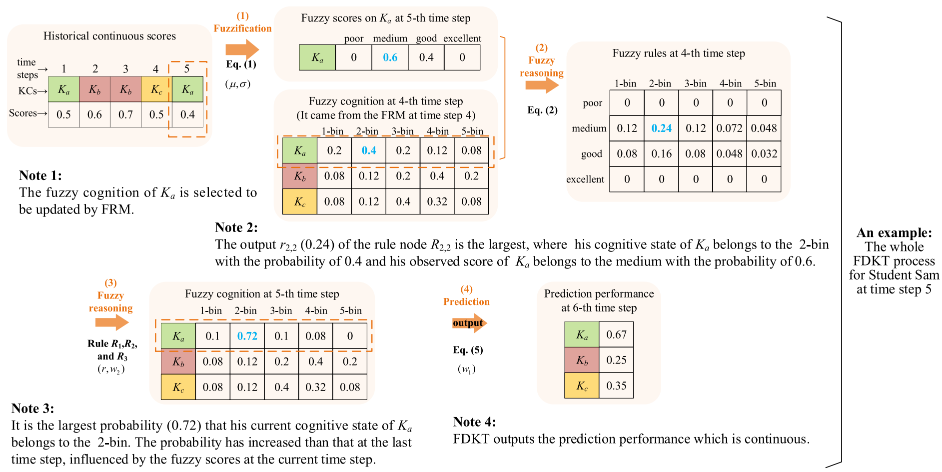

Fig. 7. An illustrative example of the whole FDKT process for an individual at a time step is provided. Five fuzzy cognition sets and four fuzzy score sets were defined. FDKT outputs the prediction performance for Sam. More importantly, it explains that he may obtain the highest score on K_a as he has a good level of mastery (2-bin) on it, where his current cognition was updated via the fuzzy reasoning.

after calculating the membership degrees of an individual to different fuzzy sets, we performed probability normalization on these membership degree values to ensure their sum is equal to 1. This is for the standardized degree distributions and the convenience of neural network computations [65–67]. Suppose that Sam performed some exercises related to three knowledge components K_a, K_b, and K_c, and the last fuzzy cognition of them is given. At time step t, Sam received a score of 0.4 when conducting an exercise related to K_a. $(0, 0.6, 0.4, 0)$ is obtained through the **fuzzification module**, representing that there is a great possibility that the score of 0.4 is indicative of mediocre performance.

Then, **FRM** infers the current fuzzy cognition of K_a through different combinations of the last fuzzy cognition and the current fuzzy score of K_a. The maximum possibility of the rule is $0.4*0.6 = 0.24$ of Rule 6, in which the last fuzzy cognition on K_a is in 2-bin with a probability of 0.4, and the current fuzzy performance is in the medium range with a probability of 0.6. Therefore, the effect of Rule 6 on the five probabilities of the current fuzzy cognition is the largest, compared with the other 19 rules.

Finally, the prediction $(0.67, 0.25, 0.35)$ on K_a, K_b, and K_c are obtained through the **prediction module**. FDKT outputs the prediction performance for Sam, as it explains that he may obtain the highest score on K_a because he has achieved a good level of mastery (2-bin) on it.

6 EXPERIMENTS

(1) How does FDKT perform in continuous score scenarios (Section 6.2) and (2) how FDKT does interpretation (Section 6.3). The parameters in FDKT are analyzed in Section 6.4.

6.1 Setup

The setup is introduced, including the datasets, baselines, and evaluation index.

6.1.1 Datasets. Four well-known datasets were used in the experiments: Algebra05, Algebra06, Bridge06 [68], and ASSISTments (https://sites.google.com/site/assistmentsdata/). To evaluate the performance of the models in continuous-score scenarios, the datasets were preprocessed as in

Table 3. Description of Data Sets

Name	Students	Exercises	Skills	Logs
Algebra05	514	172,758	435	605,051
Algebra06	1,247	549,165	1,701	1,805,754
Bridge06	1,100	129,186	564	1,816,138
ASSISTments	4,163	17,751	149	283,105

[29], according to [69, 70]. We filtered the logs of students who practiced less than 10 exercises [26, 71]. After preprocessing, the size of the datasets is listed in Table 3.

6.1.2 Baselines. To demonstrate the prediction performance of FDKT, it was compared with the following deep learning-based KT models: DKT[1] [21], DKVMN[2] [24], DeepIRT[3] [25], SAKT[4] [34], KQN[5] [27], AKT[6] [35], CKT[7] [36], DMN[8] [38], ADMN[8] [38], IADMN[8] [38], DKT+[9] [39], CL4KT[10] [42], GIKT[11] [47], and APGKT[12] [72]. The models are introduced in Section 2. We treated partially correct responses as wrong if the scores for the compared models were less than 0.5 due of their inapplicability to continuous scenarios [28].

In this paper, the Bayesian-based KT models were not included in the baselines, because they must mark the relationship between exercises and knowledge components and classify the exercises with the same knowledge components (detailed in Section 1). In this way, the temporal properties of the learning sequences would be altered after adopting this pretreatment. Thus, it has less reference value when compared between them with deep-learning-based KT models. Moreover, as mentioned previously, KT can be regarded as a dynamic cognitive diagnosis task. The cognitive diagnosis models require a students' interactive matrix with the same exercises, for example, a matrix with the size of 3000*20 where there are 3,000 students and 20 exercises. Note that there is no temporal relationship between these 20 exercises. However, the three data sets used in the experiments do not satisfy such an input. Students have different lengths of interaction sequences with the temporal relationship. For example, some students only have 10 interactions, while some have more than 3,000 interactions. Therefore, the cognitive diagnosis models were also not included in the baselines.

6.1.3 Evaluation. KT in continuous-score scenarios can be regarded as a regression task. Thus, two regression metrics, **RMSE** and **MAE**, were selected to quantify the prediction performance of the models [28].

The parameters used in FDKT are listed in Table 4. The batch size of the datasets was set to 128. The experiments were conducted using the **five-fold cross-validation** method to obtain stable results. All the experiments were implemented using the PyTorch public toolbox on a standard Ubuntu 16.04.7 LTS with TU102 USB Type-C UCSI Controller GPUs and 512 GB memory size.

[1] https://github.com/lingochamp/tensorflow-dkt [21]
[2] https://github.com/jennyzhang0215/DKVMN [24]
[3] https://github.com/ckyeungac/DeepIRT [25]
[4] https://github.com/TianHongZXY/pytorch-SAKT [34]
[5] https://github.com/JSLBen/Knowledge-Query-Network-for-Knowledge-Tracing [27]
[6] https://github.com/arghosh/AKT [35]
[7] https://github.com/bigdata-ustc/Convolutional-Knowledge-Tracing [36]
[8] https://github.com/nathan-f-elazar/Distributed-Memory-Networks [38]
[9] https://github.com/ckyeungac/deep-knowledge-tracing-plus [39]
[10] https://github.com/UpstageAI/cl4kt [42]
[11] https://github.com/ApexEDM/GIKT [47]
[12] https://github.com/DMiC-Lab-HFUT/APGKT-PRICAI2022 [72]

Table 4. Parameter Setting of FDKT

Parameter	Fuzzy score sets	Fuzzy cognition sets	Epoch
Value	6	6	100
Parameter	Optimizer	Weight decay	Learning rate
Value	Adam	0.001	0.02

Table 5. Comparison of Prediction Performance in Continuous Score Scenarios[4]

Datasets	Metrics	FDKT (Ours)	DKT [21]	DKVMN [24]	DeepIRT [25]	SAKT [34]	KQN [27]
Algebra05	MAE	**0.1700**	0.1882	0.2111	0.2122	0.1995	0.1920
	RMSE	**0.2130**	0.2473	0.3002	0.2889	0.2516	0.2622
Algebra06	MAE	**0.1590**	0.1646	0.1989	0.1954	0.2040	0.1928
	RMSE	<u>0.2075</u>	0.2162	0.2703	0.2600	0.2564	0.2555
Bridge06	MAE	**0.1450**	0.1621	0.2062	0.2064	0.2065	0.1921
	RMSE	**0.1850**	0.2060	0.2672	0.2641	0.2581	0.251
ASSISTments	MAE	<u>0.1874</u>	0.2508	0.2002	0.1593	**0.0832**	0.1644
	RMSE	**0.2588**	0.3238	0.3102	0.3459	0.2877	0.2679
Datasets	Metrics	AKT [35]	CKT [36]	DMN [38]	ADMN [38]	IADMN [38]	DKT+ [39]
Algebra05	MAE	0.2110	0.1932	0.1997	0.1992	0.1984	0.2048
	RMSE	0.2856	0.2575	0.2670	0.2661	0.2642	0.2784
Algebra06	MAE	0.1950	0.1892	0.1944	0.1933	0.1933	0.1923
	RMSE	0.2554	0.2519	0.2577	0.2555	0.2532	0.2534
Bridge06	MAE	0.1900	0.1951	0.2062	0.2020	0.2017	0.1556
	RMSE	0.2492	0.2491	0.2603	0.2557	0.2575	0.1878
ASSISTments	MAE	0.1920	0.1669	0.1613	0.1608	<u>0.1602</u>	0.1624
	RMSE	0.3023	0.2679	<u>0.2656</u>	0.2668	0.2662	0.2721
Datasets	Metrics	CL4KT [42]	GIKT [47]	APGKT [72]			
Algebra05	MAE	<u>0.1732</u>	0.1849	0.1842			
	RMSE	<u>0.2182</u>	0.2250	0.2297			
Algebra06	MAE	0.1592	0.1593	0.1602			
	RMSE	**0.1995**	0.2077	0.2089			
Bridge06	MAE	<u>0.1483</u>	0.1476	0.1513			
	RMSE	<u>0.1851</u>	0.1905	0.1895			
ASSISTments	MAE	0.1975	0.2061	0.1906			
	RMSE	0.2842	0.2848	0.2850			

[4]The bold and underlined results refer to the first and second best values, respectively.

6.2 Comparison of Prediction Performance

This subsection describes the performance of FDKT in continuous score scenarios, as compared with the baselines (**question 1**). To ensure fairness, the parameters, epochs, optimizer, weight decay, and learning rate of the models to be compared were set to be the same as those in FDKT. Smaller values of RMSE and MAE indicate better performance.

Table 5 presents the MAE and RMSE results of the prediction performance, when FDKT was compared with the deep-learning-based KT models. The prediction performance of FDKT outperforms those of both DKT and the other compared models in most cases in continuous score scenarios. This is attributed to the mechanisms such as fuzzy processing and fuzzy rules in FDKT that effectively adapt to continuous scenarios.

The Nemenyi test [73] was conducted to present a comprehensive comparison between FDKT and the baselines. The results were statistically compared over multiple datasets, as shown in Figure 8. Lower ranks indicate better performance. There is no significant difference in the same crossline-connected models. FDKT was found to perform better in continuous score scenarios.

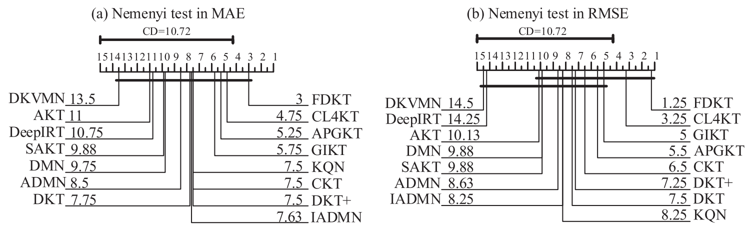

Fig. 8. Nemenyi test of the prediction performance in continuous score scenarios.

6.3 Illustration of Interpretability

This subsection shows the case studies to demonstrate how FDKT can do interpretation, including intrinsic and post hoc interpretability.

6.3.1 Case Study of Intrinsic Interpretability. As shown in Figure 9, to illustrate the working process of FDKT in a visually intuitive manner, we have selected three records of a student from the Algebra05 dataset, where a student has provided consecutive responses to a particular skill (No. 224). We will demonstrate how FDKT predicts the outcome for the third record. According to Table 4, in the experiment, we set the number of fuzzy score sets and fuzzy cognition sets to 6. Firstly, after the fuzzification module of FDKT, we obtained fuzzy scores and the last fuzzy cognition with a continuous score of 0.8. We can observe that the student's last fuzzy cognition has a higher probability (0.45) of belonging to the 4-bin, while the current fuzzy score has a higher probability (0.36) of belonging to the 5-bin. Based on the inference of fuzzy rules, we obtained the current fuzzy cognition. At this point, the student's fuzzy cognition for skill 224 has a higher probability (0.44) of belonging to the 5-bin. This indicates an improvement compared to the last fuzzy cognition. As a result, FDKT predicts a score of 0.86 for the student's performance on the skill-related exercises in the next time step. This predicted score represents an improvement compared to the score of 0.8. Furthermore, when comparing FDKT's predicted scores with the ground truth, we find that the predicted score trend (continuously increasing) aligns consistently with the actual scores. The above case study demonstrates the intrinsic interpretability of FDKT, that is, FDKT provides explanations for its corresponding prediction results.

6.3.2 Results of Post Hoc Interpretability. This subsection demonstrates the post hoc interpretability of FDKT (**question 2**), that is, to answer *what else FDKT tells us* (shown in Figure 2(b)).

According to the domain knowledge (the basic unit of interpretability [15]) in education data mining, better performance on exercises thanks to better knowledge mastery of students. Student cognition and prediction performance are the cause and effect for the KT task, respectively. In this subsection, the interpretability of FDKT is visualized from the following two aspects: (1) From the effect (performance) to cause (fuzzy cognition), and (2) From the cause (fuzzy cognition) to effect (performance).

From Performance to Cognition. The relation from the prediction scores to the fuzzy cognition is analyzed in this part. The fuzzy cognition includes its feature values and feature effects. The feature effects [64, 74] were obtained by multiplying the feature values with the weights in the optimized FDKT.

The feature value and effect distributions of the students with higher and lower prediction scores in Algebra05 are shown in Figure 10, where the weight in the optimized FDKT is 0.23, 0.50, 0.81, 0.60, 0.96, 1.38. Higher and lower scores denote normalized prediction scores higher than 0.8 and lower than 0.2, respectively. The fuzzy cognition from the 1 to 6 level denotes cognition from low to high. Comparing the fuzzy cognition of students between the higher (Figure 10(a-b)) and lower

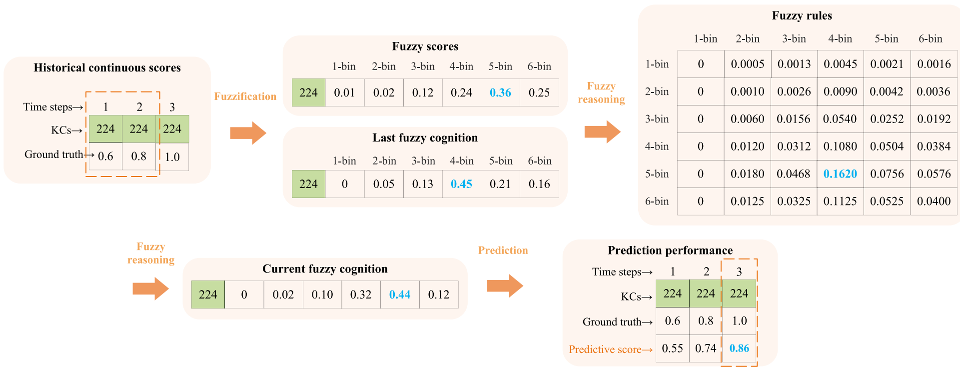

Fig. 9. Case study to demonstrate how FDKT can do intrinsic interpretation on Algebra05.

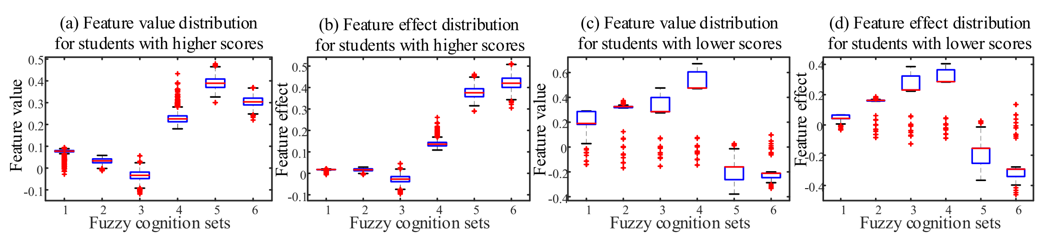

Fig. 10. Feature value and effect distributions of the students with higher (a-b) and lower (c-d) prediction scores in Algebra05. The fuzzy cognition sets from 1 to 6 denote the cognition from low to high. The better fuzzy cognition achieves high feature values and effects in (a-b), while those in (c-d) are on the contrary. This demonstrates that students with high prediction scores have high fuzzy cognition, which is in line with our initial understanding.

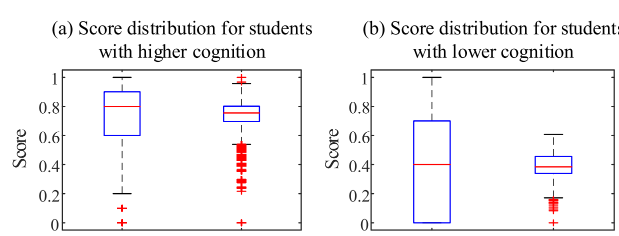

Fig. 11. Distributions of the target and prediction scores of Algebra05 for students with higher (a) and lower (b) cognition. The scores in (a) were statistically higher than those in (b). Both cases show excellent prediction performance. This demonstrates that the good prediction of FDKT is attributed to the good cognition estimate from the cause to effect.

prediction scores (Figure 10(c-d)), better fuzzy cognition achieves high feature values and effects in the former, while those in the latter are on the contrary. This demonstrates that students with high prediction scores have high fuzzy cognition, which is line with our common knowledge.

From Cognition to Performance. The relation from the fuzzy cognition to the prediction scores is analyzed as follows. The distributions of the target and prediction scores of Algebra05 for students with higher and lower cognition levels are shown in Figure 11. The probability that students belong to the top (4-6 levels) and bottom (1-3 levels) of the three fuzzy cognition sets are

Table 6. Comparison of Prediction Performance in Continuous Score Scenarios between Different Numbers of Fuzzy Sets[6]

Datasets	Metrics	Case A (2, 2)	Case B (2, 6)	Case C (6, 2)	Case D (6, 6)
Algebra05	MAE	0.1770	0.1768	0.1777	**0.1700**
	RMSE	0.2320	0.2300	0.2301	**0.2130**
Algebra06	MAE	0.1612	0.1610	0.1612	**0.1590**
	RMSE	0.2090	0.2090	0.2088	**0.2075**
Bridge06	MAE	0.1662	0.1474	0.1489	**0.1450**
	RMSE	0.2188	0.2004	0.2014	**0.1850**
ASSISTments	MAE	0.1944	**0.1864**	0.1936	0.1874
	RMSE	0.2710	0.2602	0.2596	**0.2588**

[6](i, j) indicate that the numbers of fuzzy cognition and fuzzy score sets are i and j, respectively.

denoted as p_{top} and p_{bottom}, respectively. Higher and lower cognition denote the cognition that $p_{top} > p_{bottom}$ and $p_{top} < p_{bottom}$, respectively. $|p_{top} - p_{bottom}| > \alpha(\alpha = 0.2)$ because cognition with a small probability difference cannot be arbitrarily defined.

Figure 11 is analyzed from the following two aspects. (1) Comparing two target distributions between the higher and lower cognition, the target scores of students with higher cognition are statistically higher than those with lower cognition, according to the student cognition obtained from the proposed model. This is consistent with our domain knowledge that students with greater cognitive ability can achieve higher grades. It also shows the evaluation of student cognition is reasonable in the proposed model. (2) Comparing the target and prediction for the same level of cognition, they show a relatively consistent distribution, in which the red transverse lines represent the median values. The median values of the target and prediction data, for higher and lower-cognition students, are around 0.8 and 0.4, respectively. This demonstrates the good prediction performance of FDKT is attributed to good cognition estimation from cause to effect.

6.4 Parameter Analysis

In this subsection, an analysis of the two hyper parameters, that is, the numbers of fuzzy cognition sets I and fuzzy score sets J, is presented as follows. As shown in Table 6, the results demonstrate that the FDKT performs best when both I and J are set to six. This illustrates the applicability of the FDKT to continuous scenarios because it is equivalent to considering that the input scores are only two sets (similar to the discrete scenario) when I is set to 2. For example, for the RMSE results of FDKT on the ASSISTments dataset, the best results were obtained in Case D (the number of both fuzzy sets is set to 6), with a 4.71% improvement over the results in Case A (the number of both fuzzy sets is set to 2).

The specific analysis is as follows. Defining a greater number of fuzzy sets (within a reasonable range) can effectively improve the accuracy of the FDKT in continuous scenarios and can be attributed to the following: (1) According to the definition of fuzzy rules in the FDKT (Section 4.2.2), the greater the number of fuzzy sets, the greater the number of fuzzy rules. (2) According to the fuzzy reasoning module in the FDKT framework (Figure 5(c)), the number of fuzzy rules is equal to that of hidden units in the FDKT fuzzy rule layer, which directly affects the network structure. (3) According to fuzzy theory, the higher the number of fuzzy rules, the more expert knowledge the model can incorporate [75]; moreover, according to the experience of neural networks, the higher the number of hidden units, the higher the number of network parameters and the stronger the representation capability of the model [76]. Therefore, one of the main reasons for the excellent performance of FDKT in continuous scenarios stems from the larger number of fuzzy sets.

In discrete scenarios, only two definite categories for the exercise scores exist (i.e., correct or incorrect answers, denoted by 1 and 0, respectively). That is, the score of a student on an exercise can be categorized under only two score sets (i.e., correct or incorrect set). Thus, the number of two fuzzy sets (i.e., fuzzy score sets and fuzzy cognitive sets) in FDKT is set to two to make the proposed FDKT more adaptable to discrete scenarios. Thus, the number of fuzzy rules in discrete scenarios is $2 * 2 = 4$ (according to Section 4.2.2), which is significantly smaller than that in continuous scenarios. Therefore, the FDKT has a smaller number of fuzzy rules in the discrete scenarios, which limits its accuracy according to the above analysis.

6.5 Discussion

The experiments answered the main questions in the experiments, which are summarized as follows: (1) From the prediction performance perspective, the proposed model outperforms the compared models in most cases, both on two regression metrics RMSE and MAE, in the continuous-score scenarios. (2) From the model interpretability perspective, the proposed model illustrates the post hoc interpretability both from cause to effect and effect to cause, respectively.

The reasons for the better performance of FDKT in continuous scenarios are analyzed as follows. (1) FDKT uses backpropagation to update the network parameters for improving the prediction performance by designing a reasonable loss function, similar to most neural network-based KT models. The direction of gradient descent guides the parameters to be updated in a better direction. (2) The process of updating the gradient-guided parameters is combined with domain-related expert experience through fuzzy rules for equalizing its prediction results with domain knowledge. In FDKT, educational expertise is combined in neural networks through fuzzy rules (detailed in Section 4.3). And we also demonstrated the consistency in the prediction results with expert knowledge through visualization (detailed in Section 6.3).

Compared with the existing neural-network-based KT model, the advantages of the proposed FDKT with an FNN are analyzed as follows. (1) FDKT improves the interpretability of the traditional neural-network-based KT model, both in intrinsic and post-hoc aspects. As for the intrinsic interpretability of the FDKT (detailed in Section 5), we designed a set of fuzzy rules regarding the fuzzy cognitive states and fuzzy performance scores, relying on prior knowledge. The post hoc interpretability of FDKT is illustrated through the experimental results (detailed in Section 6.3). (2) FDKT combines the advantages of both fuzzy theory and neural networks, i.e., the ability to combine language-based knowledge (e.g., expert experience) and the ease of training the model parameters (e.g., backpropagation). The FDKT combines neural networks and fuzzy theory, which solves the limitation that neural networks cannot receive linguistic knowledge, since fuzzy sets and rules are powerful tools for dealing with this type of linguistic data. (3) FDKT could address the uncertainty in the KT task, reflected in the following three aspects. (a) Performance of students on exercises is uncertain. For example, if a student scores 0.55, evaluating the score as high or low is not possible. (b) A student's knowledge of the knowledge component is uncertain. (c) Reasoning about the current time step cognition based on the performance of a student in an exercise and the previous time step cognition is uncertain. (4) FDKT extends the application scenarios of most neural network-based KT to make them suitable for continuous scenarios. Most existing KT models cannot directly handle continuous scoring scenarios. They must be fed into the network by binarizing the continuous scores and then encoding them as 0 or 1. Instead, FDKT extends KT to continuous scoring scenarios by representing continuous inputs as fuzzy sets after a fuzzy affiliation function.

Moreover, in other fields, the approaches for designing FNNs are as follows: Combining fuzzy systems with neural networks or deep learning in uncertain application scenarios is a powerful solution. We consider that this might include the following two approaches. (1) Converting the weights or inputs of neural networks into fuzzy sets. (2) Designing expert knowledge into fuzzy rules to be added between the input and output of the neural network.

7 CONCLUSION

Most deep learning-based KT models are less interpretable because of the difficulty in explaining the achievement of accurate predictions. To address this problem, a fuzzy knowledge tracing (FDKT) model is proposed with a fuzzy reasoning module that estimates student cognition of the knowledge components. The intrinsic and post-hoc interpretability of FDKT is demonstrated through rules, hidden semantics, and visualization experiments. In addition, FDKT performs better than the deep-learning-based KT models on continuous scores, broadening the application of KT. It should be pointed out that in discrete scenarios, students' practice scores have only two definite categories (i.e., correct or wrong answers), resulting in a much smaller number of fuzzy rules and hidden units than in continuous scenarios (see Section 6.4 for details). This limits the performance of the FDKT model in discrete scenarios. In the future, the authors plan to design reasonable mechanisms to further improve applicability of FDKT to discrete scenarios.

ACKNOWLEDGMENTS

The authors took advantage of the source codes of all the baselines for comparison. Our code can be available at https://github.com/DMiC-Lab-HFUT/FDKT. The computation is completed on the HPC Platform of Hefei University of Technology.

REFERENCES

[1] Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yu Yin, Shijin Wang, and Yu Su. 2022. NeuralCD: A general framework for cognitive diagnosis. *IEEE Transactions on Knowledge and Data Engineering* 35, 8 (2022), 8312–8327.

[2] Chenyang Bu, Fei Liu, Zhiyong Cao, Lei Li, Yuhong Zhang, Xuegang Hu, and Wenjian Luo. 2022. Cognitive diagnostic model made more practical by genetic algorithm. *IEEE Transactions on Emerging Topics in Computational Intelligence* 7, 2 (2022), 447–461.

[3] Zhenya Huang, Qi Liu, Yuying Chen, Le Wu, Keli Xiao, Enhong Chen, Haiping Ma, and Guoping Hu. 2020. Learning or forgetting? A dynamic approach for tracking the knowledge proficiency of students. *ACM Transactions on Information Systems (TOIS)* 38, 2 (2020). DOI: http://dx.doi.org/10.1145/3379507

[4] Yun-Wei Chu, Seyyedali Hosseinalipour, Elizabeth Tenorio, Laura M. Cruz Castro, Kerrie A. Douglas, Andrew Lan, and Christopher G. Brinton. 2022. Mitigating biases in student performance prediction via attention-based personalized federated learning. In *Proceedings of the ACM International Conference on Information & Knowledge Management (CIKM'22)*. ACM, 3033–3042.

[5] Haiping Ma, Jinwei Zhu, Shangshang Yang, Qi Liu, Haifeng Zhang, Xingyi Zhang, Yunbo Cao, and Xuemin Zhao. 2022. A prerequisite attention model for knowledge proficiency diagnosis of students. In *Proceedings of the ACM International Conference on Information & Knowledge Management (CIKM'22)*. ACM, 4304–4308.

[6] Xiangzhi Chen, Le Wu, Fei Liu, Lei Chen, Kun Zhang, Richang Hong, and Meng Wang. 2024. Disentangling cognitive diagnosis with limited exercise labels. *Proceedings of the 35th Advances in Neural Information Processing Systems* 36 (2024).

[7] Yingjie Liu, Tiancheng Zhang, Xuecen Wang, Ge Yu, and Tao Li. 2023. New development of cognitive diagnosis models. *Frontiers of Computer Science* 17, 1 (2023), 171604.

[8] Yuying Chen, Qi Liu, Zhenya Huang, Le Wu, Enhong Chen, Run-ze Wu, Yu Su, and Guoping Hu. 2017. Tracking knowledge proficiency of students with educational priors. In *Proceedings of the ACM on Conference on Information and Knowledge Management (CIKM'17)*. ACM, 989–998.

[9] Wonsung Lee, Jaeyoon Chun, Youngmin Lee, Kyoungsoo Park, and Sungrae Park. 2022. Contrastive learning for knowledge tracing. In *Proceedings of the ACM Web Conference*. ACM, 2330–2338.

[10] Fei Wang, Zhenya Huang, Qi Liu, Enhong Chen, Yu Yin, Jianhui Ma, and Shijin Wang. 2023. Dynamic cognitive diagnosis: An educational priors-enhanced deep knowledge tracing Perspective. *IEEE Transactions on Learning Technologies* 16, 3 (2023), 306–323. DOI: http://dx.doi.org/10.1109/TLT.2023.3254544

[11] B. Bakhshinategh, O. R. Zaiane, S. Elatia, and D. Ipperciel. 2017. Educational data mining applications and tasks: A survey of the last 10 years. *Education & Information Technologies* 2018, 23 (2017), 537–553.

[12] Koki Nagatani, Qian Zhang, Masahiro Sato, Yan-Ying Chen, Francine Chen, and Tomoko Ohkuma. 2019. Augmenting knowledge tracing by considering forgetting behavior. In *Proceedings of the 19th World Wide Web Conference (WWW'19)*. 3101–3107.

[13] Zhenya Huang, Qi Liu, Chengxiang Zhai, Yu Yin, Enhong Chen, Weibo Gao, and Guoping Hu. 2019. Exploring multi-objective exercise recommendations in online education systems. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 1261–1270.

[14] Fei Liu, Xuegang Hu, Shuochen Liu, Chenyang Bu, and Le Wu. 2023. Meta multi-agent exercise recommendation: A game application perspective. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'23)*. 1441–1452.

[15] Yu Zhang, Peter Tiňo, Aleš Leonardis, and Ke Tang. 2020. A survey on neural network interpretability. *arXiv preprint arXiv:2012.14261* (2020). https://arxiv.org/abs/2012.14261

[16] Finale Doshi-Velez and Been Kim. 2017. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608* (2017). https://arxiv.org/abs/1702.08608

[17] Diogo V. Carvalho, Eduardo M. Pereira, and Jaime S. Cardoso. 2019. Machine learning interpretability: A survey on methods and metrics. *Electronics* 8, 8 (2019), 832–866.

[18] Zachary Chase Lipton. 2016. The mythos of model interpretability. CoRR abs/1606.03490 (2016). *arXiv preprint arXiv:1606.03490* (2016). https://arxiv.org/abs/1606.03490

[19] Qi Liu, Shuanghong Shen, Zhenya Huang, Enhong Chen, and Yonghe Zheng. 2021. A survey of knowledge tracing. *arXiv preprint arXiv:2105.15106* (2021).

[20] Yu Lu, Deliang Wang, Qinggang Meng, and Penghe Chen. 2020. Towards interpretable deep learning models for knowledge tracing. In *Proceedings of the 23rd International Conference on Artificial Intelligence in Education (AIED'20)*. Springer, 185–190.

[21] Chris Piech, Jonathan Spencer, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas Guibas, and Jascha Sohl-Dickstein. 2015. Deep knowledge tracing. In *Proceedings of the 26th International Conference on Neural Information Processing Systems (NeurIPS'15)*. 505–513.

[22] Theophile Gervet, Ken Koedinger, Jeff Schneider, and Tom Mitchell. 2020. When is deep learning the best approach to knowledge tracing? *Journal of Educational Data Mining* 12, 3 (2020), 31–54.

[23] Xinyi Ding and Eric C. Larson. 2021. On the interpretability of deep learning based models for knowledge tracing. *arXiv preprint arXiv:2101.11335* (2021).

[24] Jiani Zhang, Xingjian Shi, Irwin King, and Dit Yan Yeung. 2017. Dynamic key-value memory networks for knowledge tracing. In *Proceedings of the 26th International Conference on World Wide Web (WWW'17)*. ACM, 765–774.

[25] Chun-Kit Yeung. 2019. Deep-IRT: Make deep learning based knowledge tracing explainable using item response theory. In *Proceedings of the 12th Educational Data Mining (EDM'19)*. 683–686.

[26] Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Hui Xiong, Yu Su, and Guoping Hu. 2019. EKT: Exercise-aware knowledge tracing for student performance prediction. *IEEE Transactions on Knowledge and Data Engineering (TKDE)* 33, 1 (2019), 100–115.

[27] J. Lee and D. Y. Yeung. 2019. Knowledge query network for knowledge tracing: How knowledge interacts with skills. In *Proceedings of the 9th International Conference on Learning Analytic & Knowledge (LAK'19)*. ACM, 491–500.

[28] Qi Liu, Runze Wu, Enhong Chen, Guandong Xu, Yu Su, Zhigang Chen, and Guoping Hu. 2018. Fuzzy cognitive diagnosis for modelling examinee performance. *ACM Transactions on Intelligent Systems and Technology (TIST)* 9, 4 (2018), 1–26.

[29] Fei Liu, Xuegang Hu, Chenyang Bu, and Kui Yu. 2021. Fuzzy Bayesian knowledge tracing. *IEEE Transactions on Fuzzy Systems (TFS)* 30, 7 (2021), 2412–2425.

[30] Albert T. Corbett and John R. Anderson. 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-adapted Interaction* 4, 4 (1994), 253–278.

[31] Xuegang Hu, Fei Liu, and Chenyang Bu. 2020. Research advances on knowledge tracing models in educational big data. *Journal of Computer Research and Development* 57, 12 (2020), 2523–2546.

[32] Jose Maria Alonso Moral, Ciro Castiello, Luis Magdalena, and Corrado Mencar. 2021. Designing interpretable fuzzy systems. In *Explainable Fuzzy Systems*. Springer, 119–168.

[33] Chin Teng Lin and C. S. George Lee. 1991. Neural-network-based fuzzy logic control and decision system. *IEEE Trans. Comput.* 40, 12 (1991), 1320–1336.

[34] P. Pandey and G. Karypis. 2019. A self attentive model for knowledge tracing. In *Proceedings of the 12th International Conference on Educational Data Mining (EDM'19)*. 384–389.

[35] Aritra Ghosh, Neil Heffernan, and Andrew S. Lan. 2020. Context-aware attentive knowledge tracing. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD'20)*. ACM, 2330–2339.

[36] Shuanghong Shen, Qi Liu, Enhong Chen, Han Wu, Zhenya Huang, Weihao Zhao, Yu Su, Haiping Ma, and Shijin Wang. 2020. Convolutional knowledge tracing: Modeling individualization in student learning process. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'20)*. 1857–1860.

[37] Jinze Wu, Zhenya Huang, Qi Liu, Defu Lian, Hao Wang, Enhong Chen, Haiping Ma, and Shijin Wang. 2021. Federated deep knowledge tracing. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining (WSDM'21)*. ACM, 662–670.

[38] N. F. Elazar. 2018. *Distributed Memory Networks for Knowledge Tracing*. Master's Thesis. Australian National University.

[39] Chun Kit Yeung and Dit Yan Yeung. 2018. Addressing two problems in deep knowledge tracing via prediction-consistent regularization. In *Proceedings of the 5th Annual ACM Conference on Learning at Scale*. ACM, 1–10.

[40] Sein Minn, Yi Yu, Michel C. Desmarais, Feida Zhu, and Jill-Jênn Vie. 2018. Deep knowledge tracing and dynamic student classification for knowledge tracing. In *Proceedings of the 18th IEEE International Conference on Data Mining (ICDM'18)*. IEEE, 1182–1187.

[41] Moyu Zhang, Xinning Zhu, Chunhong Zhang, Yang Ji, Feng Pan, and Changchuan Yin. 2021. Multi-factors aware dual-attentional knowledge tracing. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM'21)*, Gianluca Demartini, Guido Zuccon, J. Shane Culpepper, Zi Huang, and Hanghang Tong (Eds.). ACM, 2588–2597.

[42] Wonsung Lee, Jaeyoon Chun, Youngmin Lee, Kyoungsoo Park, and Sungrae Park. 2022. Contrastive learning for knowledge tracing. In *Proceedings of the 13th ACM Web Conference*. 2330–2338.

[43] Leyan Deng, Defu Lian, Chenwang Wu, and Enhong Chen. 2022. Graph convolution network based recommender systems: Learning guarantee and item mixture powered strategy. *Proceedings of the 33rd Advances in Neural Information Processing Systems* 35 (2022), 3900–3912.

[44] Yufei Zeng, Zhixin Li, Zhenbin Chen, and Huifang Ma. 2023. Aspect-level sentiment analysis based on semantic heterogeneous graph convolutional network. *Frontiers of Computer Science* 17, 6 (2023), 176340.

[45] Jinfu Liu, Xinshun Wang, Can Wang, Yuan Gao, and Mengyuan Liu. 2023. Temporal decoupling graph convolutional network for skeleton-based gesture recognition. *IEEE Transactions on Multimedia* (2023), 1–13.

[46] Xinshun Wang, Wanying Zhang, Can Wang, Yuan Gao, and Mengyuan Liu. 2023. Dynamic dense graph convolutional network for skeleton-based human motion prediction. *IEEE Transactions on Image Processing* 33 (2023), 1–15.

[47] Yang Yang, Jian Shen, Yanru Qu, Yunfei Liu, Kerong Wang, Yaoming Zhu, Weinan Zhang, and Yong Yu. 2021. GIKT: A graph-based interaction model for knowledge tracing. In *Machine Learning and Knowledge Discovery in Databases*, Frank Hutter, Kristian Kersting, Jefrey Lijffijt, and Isabel Valera (Eds.). Springer International Publishing, Cham, 299–315.

[48] Jimmy de la Torre. 2009. DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics* 34, 1 (2009), 115–130.

[49] Jeff Johns, Sridhar Mahadevan, and Beverly Woolf. 2006. Estimating student proficiency using an item response theory model. In *Proceedings of International Conference on Intelligent Tutoring Systems (ITS'06)*. Springer, 473–480.

[50] Philip I. Pavlik, Hao Cen, and Kenneth R. Koedinger. 2009. Performance factors analysis - A new alternative to knowledge tracing. In *Proceedings of the 14th International Conference on Artificial Intelligence in Education (AIED) (Frontiers in Artificial Intelligence and Applications)*, Vania Dimitrova, Riichiro Mizoguchi, Benedict du Boulay, and Arthur C. Graesser (Eds.). Vol. 200. IOS Press, 531–538.

[51] Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yuying Chen, Yu Yin, Zai Huang, and Shijin Wang. 2020. Neural cognitive diagnosis for intelligent education systems. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI'20)*, Vol. 34. 6153–6161.

[52] Song Cheng, Qi Liu, Enhong Chen, Zai Huang, Zhenya Huang, Yiying Chen, Haiping Ma, and Guoping Hu. 2019. DIRT: Deep learning enhanced item response theory for cognitive diagnosis. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM'19)*. 2397–2400.

[53] Jinjin Zhao, Shreyansh Bhatt, Candace Thille, Dawn Zimmaro, and Neelesh Gattani. 2020. Interpretable personalized knowledge tracing and next learning activity recommendation. In *Proceedings of the 7th ACM Conference on Learning @ Scale (L@S)*. 325–328.

[54] Shalini Pandey and Jaideep Srivastava. 2020. RKT: Relation-aware self-attention for knowledge tracing. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM'20)*. 1205–1214.

[55] Weibo Gao, Qi Liu, Zhenya Huang, Yu Yin, Haoyang Bi, Mu-Chun Wang, Jianhui Ma, Shijin Wang, and Yu Su. 2021. RCD: Relation map driven cognitive diagnosis for intelligent education systems. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 501–510.

[56] Hiromi Nakagawa, Yusuke Iwasawa, and Yutaka Matsuo. 2019. Graph-based knowledge tracing: Modeling student proficiency using graph neural network. In *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence (WI'19)*. IEEE, 156–163.

[57] Hanshuang Tong, Zhen Wang, Qi Liu, Yun Zhou, and Wenyuan Han. 2020. HGKT: Introducing hierarchical exercise graph for knowledge tracing. *arXiv preprint arXiv:2006.16915* (2020).

[58] Lotfi A. Zadeh. 1996. Fuzzy sets. In *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh*. World Scientific, 394–432.
[59] Shinichi Tamura, Seihaku Higuchi, and Kokichi Tanaka. 1971. Pattern classification based on fuzzy relations. *IEEE Transactions on Systems, Man, and Cybernetics* 1 (1971), 61–66.
[60] Lotfi A. Zadeh. 1965. Fuzzy sets. *Information and Control* 8, 3 (1965), 338–353.
[61] Madan M. Gupta and J. Qi. 1991. Theory of T-norms and fuzzy inference methods. *Fuzzy Sets and Systems* 40, 3 (1991), 431–450.
[62] Tzung-Pei Hong and Chai-Ying Lee. 1996. Induction of fuzzy rules and membership functions from training examples. *Fuzzy Sets and Systems* 84, 1 (1996), 33–47.
[63] Ching Hung Lee and Ching Cheng Teng. 2000. Identification and control of dynamic systems using recurrent fuzzy neural networks. *IEEE Transactions on Fuzzy Systems (TFS)* 8, 4 (2000), 349–366.
[64] Christoph Molnar. 2020. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Lulu.com.
[65] Tim Salimans and Durk P. Kingma. 2016. Weight normalization: A simple reparameterization to accelerate training of deep neural networks. *Advances in Neural Information Processing Systems* 29 (2016).
[66] Faisal Naeem, Gautam Srivastava, and Muhammad Tariq. 2020. A software defined network based fuzzy normalized neural adaptive multipath congestion control for the internet of things. *IEEE Transactions on Network Science and Engineering* 7, 4 (2020), 2155–2164.
[67] Qiying Feng, Long Chen, CL Philip Chen, and Li Guo. 2020. Deep fuzzy clustering–a representation learning approach. *IEEE Transactions on Fuzzy Systems* 28, 7 (2020), 1420–1433.
[68] J. Stamper. 2010. Algebra I 2005–2006, Algebra I 2006–2007, Bridge to Algebra 2006–2007. Challenge data Set from KDD Cup 2010 Educational Data Mining Challenge. [EB/OL]. (2010). https://pslcdatashop.web.cmu.edu/KDDCup/downloads.jsp
[69] Yutao Wang and Neil Heffernan. 2013. Extending knowledge tracing to allow partial credit: Using continuous versus binary nodes. In *Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED'13)*. Springer, 181–188.
[70] Korinn Ostrow, Christopher Donnelly, Seth Adjei, and Neil Heffernan. 2015. Improving student modeling through partial credit and problem difficulty. In *Proceedings of the 2nd ACM Conference on Learning @ Scale (L@S)*. ACM, 11–20.
[71] Kai Zhang and Yiyu Yao. 2018. A three learning states Bayesian knowledge tracing model. *Knowledge-Based Systems* 148 (2018), 189–201.
[72] Haotian Zhang, Chenyang Bu, Fei Liu, Shuochen Liu, Yuhong Zhang, and Xuegang Hu. 2022. APGKT: Exploiting associative path on skills graph for knowledge tracing. In *Proceedings of Pacific Rim International Conference on Artificial Intelligence (PRICAI'22)*, Sankalp Khanna, Jian Cao, Quan Bai, and Guandong Xu (Eds.). Springer, Cham, 353–365.
[73] Janez Demšar. 2006. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research* 7 (2006), 1–30.
[74] Joel Vaughan, Agus Sudjianto, Erind Brahimi, Jie Chen, and Vijayan N. Nair. 2018. Explainable neural networks based on additive index models. *arXiv preprint arXiv:1806.01933* (2018).
[75] Lotfi A. Zadeh. 1988. Fuzzy logic. *Computer* 21, 4 (1988), 83–93.
[76] Stephen I. Gallant. 1993. *Neural Network Learning and Expert Systems*. MIT Press.

Received 6 January 2023; revised 10 January 2024; accepted 20 March 2024