# Interpretable Cognitive State Prediction via Temporal Fuzzy Cognitive Map

Yuang Wei and Bo Jiang

***Abstract*—Understanding student cognitive states is essential for assessing human learning. The deep neural networks (DNN)-inspired cognitive state prediction method improved prediction performance significantly; however, the lack of explainability with DNNs and the unitary scoring approach fail to reveal the factors influencing human learning. Identifying and understanding these factors remain a challenge. Thus, this article proposes the temporal fuzzy cognitive map (tFCM) model, which combines the prediction power of DNNs with the interpretability of fuzzy cognitive maps. In the proposed tFCM model, cognitive states are modeled as fuzzy, multidimensional, and interrelated vectors, which are input to a long short-term memory network for prediction. This integration allows the proposed model to combine the exceptional ability of DNNs to uncover latent factors with the distinct benefits of fuzzy cognitive maps' ability to reveal potential correlations. A comparative experiment was designed and conducted on a large-scale dataset to assess the predictive performance and interpretability of the proposed tFCM model. The results demonstrate tFCM's superior performance and interpretability compared to existing models. The findings of this study contribute to the development of a multidimensional quantitative model to represent cognitive states and an interpretable model architecture for state prediction.**

***Index Terms*—Cognitive state prediction, deep neural networks (DNNs), fuzzy cognitive map, interpretable learner model.**

## List of Abbreviations

| | |
|---|---|
| CSK | Cognitive state of knowledge. |
| DINA | Deterministic inputs, noisy "and" gate. |
| DIRT | Deep item response theory. |
| DNN | Deep neural network. |
| FCM | Fuzzy cognitive mapping. |
| FLC | Fuzzy logic controller. |
| IRT | Item response theory. |
| LSTM | Long short-term memory. |
| MIRT | Multidimensional item response theory. |
| MIRT-C | Compensatory multidimensional item response theory. |
| RMSE | Root mean square error. |
| RNN | Recurrent neural network. |
| tFCM | Temporal fuzzy cognitive map. |

Manuscript received 16 April 2023; revised 27 July 2023; accepted 27 July 2023. Date of publication 22 August 2023; date of current version 4 January 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 61977058, in part the Natural Science Foundation of Shanghai under Grant 23ZR1418500, and in part by the Shanghai Science and Technology Innovation Action Plan under Grant 20511101600. *(Corresponding author: Bo Jiang.)*

Yuang Wei is with Shanghai Institute of AI for Education, East China Normal University, Shanghai 200062, China (e-mail: 2644595124@qq.com).

Bo Jiang is with the Department of Educational Information Technology, East China Normal University, Shanghai 200062, China (e-mail: bjiang.zh@gmail.com).

Digital Object Identifier 10.1109/TLT.2023.3307565

## I. Introduction

Human learning is a complicated cognitive process that is affected by various factors, e.g., the learner's environment, motivation, and prior knowledge [1]. The acquisition of knowledge involves complex interplay between internal cognitive processes and external stimuli, which makes it difficult to identify the factors that influence the cognitive process [2]. In addition, the complexity of the cognitive process also makes students' CSK fuzzy, which means that simply categorizing a student's CSK as "mastered" or "not mastered" is imprecise [3], [4]. For example, phrases like "He is moderate at this unit" or "She knows 70% of the chapter" are frequently used to describe a student's CSK. The fuzziness of CSK brings challenges in assessing students' cognitive states.

Assessing the CSK has received considerable attention in the fields of educational psychology and educational technologies. A prominent area of study is cognitive state diagnostic models, with IRT [5], [6] and DINA [7], [8], [9] being the most well-known cognitive state diagnostic models. Recently, some machine learning-inspired cognitive diagnostic algorithms have been proposed, e.g., the probabilistic matrix factorization framework [10] and factor models based on singular value decomposition [11]. Although current studies have significantly improved the predictive performance of cognitive diagnostic compared to the conventional IRT or DINA models, challenges remain in terms of handling the fuzziness of cognitive states effectively while providing sufficient interpretability and predictive capabilities.

To address this issue, this study attempts to model human learning as a fuzzy control process of the automatic control system. A typical automatic control system operates in a feedback loop by continuously adjusting relevant system parameters based on the feedback until the desired output is acquired [12]. This iterative learning process helps the control system adapt to changing conditions and uncertainties. Similarly, human learning involves an iterative process of receiving feedback, evaluating performance, making adjustments, and striving for improvement. Learners apply feedback to modify their learning strategies, revise their understanding, and enhance their abilities over time. Thus, the learning process can be analogized as a closed-loop control system [13]. In addition, considering the previously mentioned fuzziness, this closed-loop system evokes an efficient tool in industrial control known as fuzzy logic control (FLC) [14], which establishes connections between different system states using fuzzy probabilities, thereby enabling efficient handling of uncertainty in nonlinear state control systems [15], [16], [17], [18].

However, humans cannot be equated to control systems, which has motivated researchers to explore and create FLC specific to human cognition, referred to as fuzzy cognitive maps (FCM) [19]. By combining fuzzy logic with cognitive psychology theories, FCMs offer an effective solution to the fuzziness of the cognitive state [20], [21], [22]. Even in the era dominated by high-performance DNN models, FCM continues to be applied extensively in a wide range of domains due to its interpretability and reliable predictive performance [23]. For example, in the field of learning analytics, the inherent interpretability of FCM models due to their explicit interaction graph is particularly important. Thus, in this study, we utilize an FCM to model the cognitive state transition process of learners. To address this modeling challenge, we must understand the interconnections among states, establish the rules governing state influence, and design an intelligent model that excels in terms of both performance and interpretability.

Inspired by FLC and FCM, we propose a fuzzy method to predict students' cognitive states, named temporal FCM (tFCM). Rather than using a binary representation, the proposed method converts the learners' cognitive states into four fuzzy states. In addition, by combining a DNN and FCM, the proposed prediction model enhances the capacity to represent nonlinear relationships and improves prediction accuracy while maintaining sufficient interpretability. The performance of the proposed model was validated on large-scale educational datasets, and we elucidated the interpretability of two types of relationships, i.e., the connection between simultaneous states and the association between current and past state information. In summary, the proposed model effectively handles the nonlinearity and fuzziness of students' cognitive states during the learning process (where the nonlinearity captures the nonlinear variations in cognitive states, and the fuzziness reflects the probabilistic nature of students' cognitive states), which allows the proposed model to withstand the influence of confounding factors and align more closely with the evolution of the learning process.

The rest of this article is organized as follows. Related work is summarized in Section II. In Section III, we explain the structure of the temporal FCM for fuzzy cognitive state prediction. The performance and interpretability of the proposed method are assessed in Section IV, and we present a relevant discussion in Section V. Finally, Section VI concludes this article.

## II. Related Work

### A. Cognitive State Diagnosis

The IRT and DINA models are well-known cognitive diagnosis models in the educational psychometrics field. The IRT model relies on a 1-D variable (i.e., a latent trait) to evaluate learners' proficiency levels. This model employs an interaction function to estimate the probability of learners answering items correctly based on their latent traits. In contrast, the DINA model utilizes binary vectors to describe each learner's proficiency, where 1 indicates mastery, and 0 indicates nonmastery. These models have been expanded through further research. For example, MIRT [24] extended the capabilities of IRT to handle more complex and diverse latent attributes, and other extensions like MIRT-C [5] and MIRT-NC [25] have enabled the analysis of interactions between different levels of mastery. In addition, the dynamic IRT [26] model can track the entire learning performance sequence to uncover changes in cognitive states over time, and extended DINA models incorporate attribute and response time modeling by considering item responses and response times jointly [27]. Integrating DINA with higher order Markov processes allows us to better capture the variation process in students' skill levels [9]. Generally, these models offer good interpretability; however, their predictive performance is typically weaker compared to deep learning models.

In data mining, various models have demonstrated promising predictive performance with the support of large-scale datasets. For example,Thai-Nghe and Schmidt-Thieme [28] employed matrix factorization techniques in intelligent tutoring systems to evaluate students' cognitive states, and the authors in [29] and [30] employed nonnegative matrix factorization to infer the Q-matrix to mine students' cognitive states. In addition, Desmarais [29] performed an assessment to determine the fundamental skill factors and assumptions that impact learning achievements, and Gao et al. [31] focused on modeling student characteristics and mining cognitive states by considering the relationship between problems and skills. However, a limitation of these models is that the inferred latent factors are often uninterpretable, which introduces uncertainty factors.

### B. Fuzzy Cognitive States of Knowledge

The CSK may comprise various states; however, several previous studies only examined cognitive states as a single feature, which fails to encompass the uncertainty associated with learning. To address this issue, fuzzy logic techniques can be used, e.g., fuzzy recommendation of learning content based on student's learning styles and personal needs [32] or fuzzy cluster analysis of students' learning behavior. However, these methods lack dynamic changes and conceptual correlation processes [33]. In [34], a fuzzy mechanism was employed to represent the association between concepts and influence rules, which was then updated with the expansion of the map structure. The first condition for our research was to define a set of fuzzy logic relations that belong to an issue, and the key to fuzzy logic relations was the definition of their fuzzy sets. However, defining the cognitive state of learners is a complex task, and it is difficult to achieve accuracy using only a linear probability model. For example, stating that "Student A's knowledge mastery in mathematics discipline is 50% unclear," "Student B is ignorant of Chinese," or "Student C has 100% mastery of English" is insufficient. Instead, the state information can be conceptualized naturally using fuzzy sets. In [35], the knowledge state was divided into four grades, i.e., unknown (*Un*), insufficiently

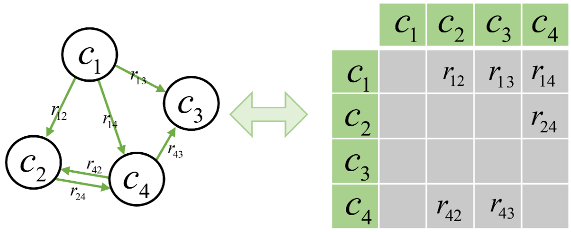



Fig. 1. Basic fuzzy cognitive map.

known (InK), known (*K*), and learned (*L*). These grades can represent both the knowledge state during the learning process and the stages of cognitive states. The membership function is composed of trapezoidal functions, as

Here, each function in the formula represents the fuzzy membership of different levels of states, which covers 100% of the states and exists symmetrically. Linear fuzzy membership functions emphasize linearity and complementarity; however, human cognitive processes are considered to be nonlinear and fuzzy. Therefore, this type of membership function in (1) shown at the bottom of this page, is not suitable for modeling human cognitive processes.

### C. Fuzzy Cognitive Map (FCM)

The basic FCM model proposed by [36] comprises *concepts*, *states*, and *relationships*. Here, the *concepts* represent a node in FCM, the *states* represent the state of the concept at a certain time, and the *relationships* represent the influence between concepts. The FCM is a directed graph with relational weight. For example, assume that the concept node is represented as $c_i$ and the relational weight as $r_{i,j}$. Thus, the graph structure and matrix can be employed to represent the FCM model, as shown in Fig. 1.

Generally, the state of $c_i$ can be represented by a set of dynamic time series $\{s_i^1, s_i^2, \ldots, s_i^T\}$, where $s_i^T$ denotes the state value of the $c_i$ concept at time *T*. The value of $s_i^{T+1}$ is affected by both $s_i^T$ and other concepts in the FCM that have an impact on $c_i$. Thus, the state of $c_i$ at each time point can be expressed as follows:

$$s_i^{(t+1)} = \varphi\left(s_i^{(t)} + \sum_{j \neq i} r_{ji} s_j^{(t)}\right). \tag{2}$$

FCM can construct a path in virtual space that may terminate at a fixed point or limit cycle in a simple model or at a chaotic attractor in a complex model, representing an uncontrollable model outcome [37]. To construct more complex models, Hagiwara [38] extended the linear relationship between concepts to a nonlinear relationship, while the authors in [39] and [40] introduced a fuzzy measure to quantify the impact of the causal relationship of conceptual node causality and proposed a probabilistic cognitive map model. Despite its remaining challenges, FCM attracts attention consistently due to its interpretability and is widely used in various fields, e.g., control system monitoring [41], intelligent decision-making methods [42], business modeling [43], medical atlases [44], and educational research [45]. FCM has been proven to be a very effective modeling method in the analysis of multidomain complex systems; however, its dependence on expert experience and knowledge remains a disadvantage. In addition, FCM, similar to other interpretable intelligent models, involves an inherent tradeoff between interpretability and performance.

## III. Proposed Model

### A. Nonlinear Fuzzy Cognitive States of Knowledge

To address the nonlinearity and uncertainty inherent in the learning process, linear membership functions are insufficient. Here, we propose a new nonlinear fuzzy cognitive state membership function for knowledge. The proposed function comprises nonlinear functions that account for the four CSKs. These states are defined as follows:

1) *Unknown (Un)*: The first state is defined as the unknown stage, which indicates that something (e.g., knowledge, skills, abilities, etc.) is ambiguous to the person or that they have only heard of it without fully comprehending.
2) *Insufficiently Known (InK)*: This state is defined as the accumulation stage, where the person has a certain basic understanding of something; however, there may be some blind spots in knowledge and opinions.
3) *Known (K)*: This state is defined as the bottleneck stage, where the person knows and can use something, and even knows the relevant knowledge.
4) *Learned (L)*: This state is defined as the breakthrough stage, in which the person has acquired complete mastery of a certain type of knowledge, skill, or ability, and can analyze and evaluate these knowledge, skills, or abilities critically. In addition, the person can propose innovative

$$\mu_{Un}(x) = \begin{cases} 1, & x \le 55 \\ 1-(x-55)/5, & 55 < x < 60 \\ 0, & x \ge 60 \end{cases} \qquad \mu_{\mathrm{InK}}(x) = \begin{cases} (x-55)/5, & 55 < x < 60 \\ 1, & 60 \le x \le 70 \\ 1-(x-70)/5, & 70 < x < 75 \\ 0, & x \le 55 \textit{ or } x \ge 75 \end{cases}$$

$$\mu_K(x) = \begin{cases} (x-70)/5, & 70 < x < 75 \\ 1, & 75 \le x \le 85 \\ 1-(x-85)/5, & 85 < x < 90 \\ 0, & x \le 70 \textit{ or } x \ge 90 \end{cases} \qquad \mu_L(x) = \begin{cases} (x-85)/5, & 85 < x < 90 \\ 1, & 90 \le x \le 100 \\ 0, & x \le 85. \end{cases} \tag{1}$$

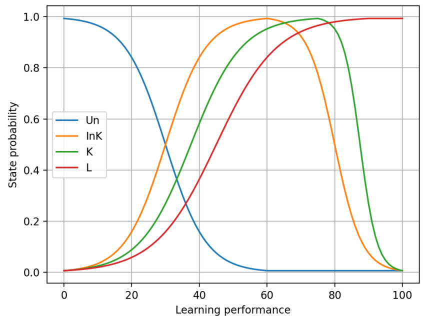



Fig. 2. Nonlinear membership function curve.

solutions and ideas based on their mastery of the given subject matter.

Then, to capture the nonlinearity and uncertainty, we developed the membership function according to the following theory. Initially, learning involves gathering diverse information, and memory involves encoding and storing that information. From a neuroscience perspective, the mechanism of "learning (acquiring information) to form memory and storage" relies on alterations in neuronal connections, which involves the creation and elimination of connections and the strengthening and weakening of existing connections. These modifications in connectivity are collectively referred to as synaptic plasticity [46]. The process of synaptic plasticity exhibits a gradual pace initially but undergoes rapid changes over time. Once it reaches a particular threshold, the changes cease, and the system stabilizes. Thus, we define the nonlinear change process of the CSK based on this pattern of variation. The membership function is given in (3) shown at the bottom of this page, and its function curve is shown Fig. 2.

The change curve shown in Fig. 2 depicts the nonlinear change rule of the CSK in the learning process. In the initial learning stage, the probability of learners being confused and ignorant is much higher than in other states. However, as the learning progresses and learning time accumulates, the probability of students being in the InK, *K*, and *L* states increases as the level of knowledge and skills increases. In this study, the CSK of students is represented as the probability set S = [$\mathrm{P}_{Un}$, $\mathrm{P}_{\mathrm{InK}}$, $\mathrm{P}_{K}$, $\mathrm{P}_{L}$]. For example, at a given moment, the probability of students being in the *Un* state is observed to be 82%, in InK to be 18%, in *K* to be 10%, and in *L* to be 6%; thus, the probability set S = [82%, 18%, 10%, 6%].

In addition, it is difficult to determine whether students who have passed the test have truly learned knowledge or just know it. Thus, InK, *K*, and *L* are set to similar probabilities to enhance the fuzziness of this stage, which can ensure the assessment results do not introduce greater bias when faced with other uncertain factors.

### B. Temporal Fuzzy Cognitive Map (tFCM) Model

The assessment of student's cognitive state performance is frequently dependent on analyzing their rate of correct responses to specific test problems. However, the accuracy rate of responses is a static evaluation criterion (typically encompassing question correctness and test scores), which poses difficulties in terms of satisfying the dynamic assessment requirements of learning. Such dynamic requirements may involve predicting future learning performance and other factors. To address this issue, the proposed tFCM evaluates and predicts the cognitive state of students, whose computational process can be illustrated, as shown in Fig. 3. First, the proposed model fuzzifies the students' responses using the membership function, and then it utilizes the state information to train the fuzzy state relation. Finally, the model evaluates and predicts the students' CSK dynamically by incorporating time factors. Fig. 4 shows the simplified unit structure of the proposed model.

*1) Nonlinear Fuzzy Cognitive Map:* The original FCM cannot address the nonlinear process of CSK change. However, neural networks have demonstrated excellent performance when handling nonlinear problems; thus, we integrated the nonlinear processing of neural networks into the FCM to enhance prediction performance while maintaining sufficient interpretability. Here, the students' cognitive state is represented as follows:

$$s_j^{(t+1)} = \varphi(f_j(s^{(t)})) \tag{4}$$

where $s^{(t)} = (s_1^t, s_2^t, \ldots, s_n^t)$ represents the probability value of all states associated with $s_j$ at time $t$, and function $f_j(\bullet)$ models the association relationship between states. By fusing these relationships, the state probability at time $t+1$ is obtained through $\varphi$ fuzzification. Note that the original FCM is a special case where $f_j(s^{(t)}) = s_i^{(t)} + \sum_{j\neq i} \omega_{ji} s_j^{(t)}$. To better fit the nonlinear problem and handle many uncertain factors in the learning process, we implement a feedforward neural network to construct $f_j(\bullet)$. Here, the ReLU activation function is selected as the neuron activation method to maintain the FCM at its original value of 0. The neuron calculation is expressed as follows:

$$\sum_{q=1}^{M} y_{(q,l)}^{(t)} = \sum_{q=1}^{M} \left( w_{(q,l+1)} \mathrm{ReLU} \left( \sum_{p=1}^{N} w_{(pq,l)} y_{(p,l-1)}^{(t)} \right) \right)$$

$$\mu_{Un}(x) = \begin{cases} \dfrac{1}{1+e^{\frac{1}{6}*(x-30)}} & x < 60 \\ \dfrac{1}{1+e^{5}} & x \geq 60 \end{cases} \qquad \mu_{\mathrm{InK}}(x) = \begin{cases} \dfrac{1}{1+e^{-\frac{1}{6}*(x-30)}} & 0 < x \leq 60 \\ \dfrac{1}{1+e^{\frac{1}{4}*(x-80)}} & 60 < x \leq 100 \end{cases}$$

$$\mu_{\mathrm{K}}(\mathrm{x}) = \begin{cases} \dfrac{1}{1+e^{-\frac{15}{2}*(x-\frac{75}{2})}} & 0 < x \leq 75 \\ \dfrac{1}{1+e^{\frac{5}{2}*(x-\frac{175}{2})}} & 75 < x \leq 100 \end{cases} \qquad \mu_{L}(x) = \frac{1}{1+e^{-\frac{1}{10}*(x-50)}} \quad 0 \leq x \leq 100. \tag{3}$$

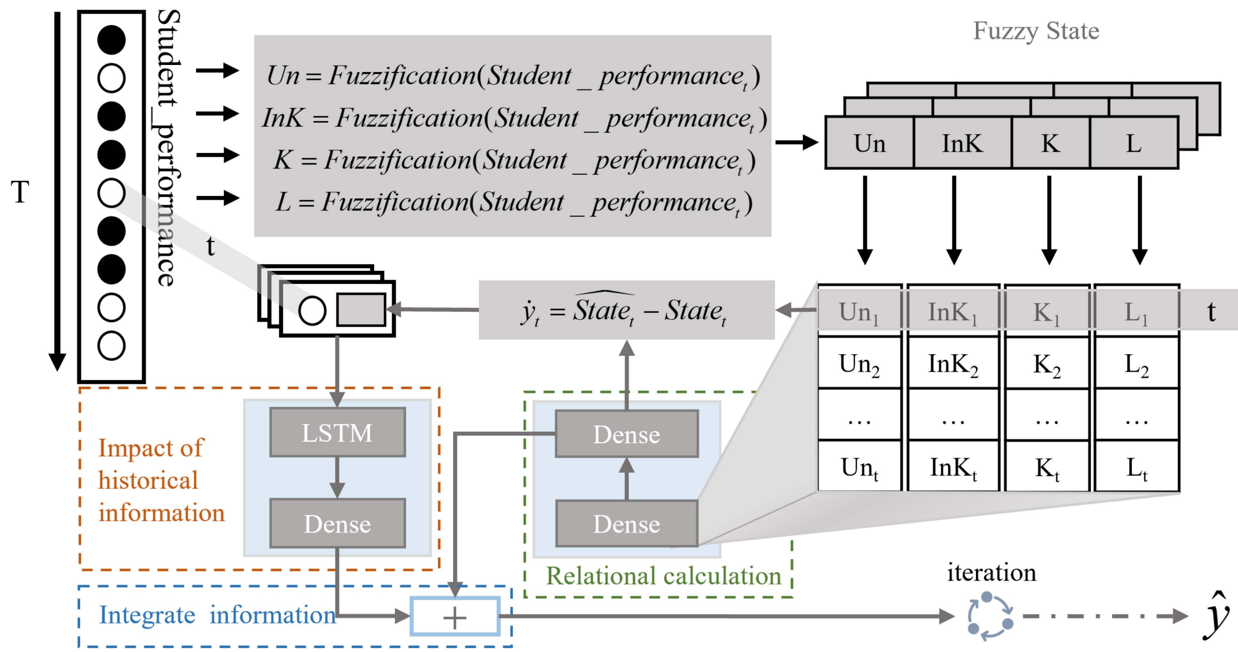



Fig. 3. Technology route of the proposed temporal fuzzy cognitive map encompasses a series of steps. Initially, the student test results undergo transformation into fuzzy vectors using fuzzy membership functions. Then, a fully-connected network is employed to establish associations between the states. In addition, an LSTM network is utilized to capture the impact of historical state information. Finally, the model generates prediction results based on these processes.

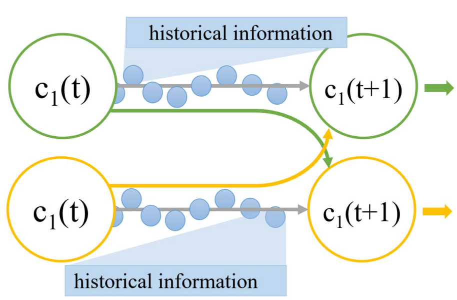



Fig. 4. Unit of the proposed temporal fuzzy cognitive map.

$$= f_j(s^{(t)}) \tag{5}$$

$$\mathrm{ReLU} = \begin{cases} z & z > 0 \\ 0 & z \leq 0 \end{cases} \tag{6}$$

where $l$ represents the number of fully connected network layers, $M$ represents the number of neurons in the layer $l$, and $y$ represents the output of layer $l$. In (5), $\sum_{q=1}^{M} y_{(q,l)}^{(t)}$ represents the neuron $q$ of layer $l-1$ that enters the $l$ layer after experiencing weight product and activation, and the neuron $q$ of layer $l$ directly outputs after experiencing weight product. Note that the activation function is not employed in the output layer of the final layer because using the ReLU activation function will limit the output result to the range of $[0, +\infty)$. According to the definition of the FCM model, the output of FCM falls within the range of $(-\infty, +\infty)$. If the model output is not within this range, the interpretability of the output will be weakened. Similarly, the original FCM is a special case of the proposed model.

*2) Fusion of Historical State Information:* Typically, the fuzzy CSKs are determined by both the current state and the previous state, where the former can be obtained from learning performance, and the latter is typically mixed with uncertainty factors, thereby making it complex and challenging to measure. Although the original FCM model can assess and predict the current state, it fails to integrate historical state information. To address this issue, previous studies have employed static constants to adjust the FCM model [47], [48]. However, time-series data provide dynamic information; thus, a structure that can accommodate dynamic input is required. RNN [49] possess robust capability to handle sequential information. In addition, information of different time lengths can be subject to forgetting; thus, we employ the LSTM network [50] to calculate and fuse historical state information

$$s_j^{(t+1)} = \varphi(h_j^{(t)} + f_j(s^{(t)})) \tag{7}$$

$$h_j^{(t)} = \mathrm{LSTM}(t, \mathrm{T}, h_j^{(t-1)}). \tag{8}$$

Here, $h_j^{(t)}$ represents the overall influence of various factors on the CSK, including historical state information, and parameter $T$ denotes the memory cycle, which reflects the periodic fluctuations in the student's CSK during the learning process, which may result from the impact of historical state information and other unknown factors. The historical state at time $t-1$, denoted $h_j^{(t-1)}$, contains various historical factors.

### C. Fuzzy State Relationship

Due to the uncertainty of the CSKs, it is challenging to identify the correlations between states. Nevertheless, by utilizing the powerful ability of an FCM to mine the conceptual relationship of complex systems, we can identify the relationship between CSKs after constructing the fuzzy model through the FCM model.

In the basic FCM model, the correlation coefficient $r_{ij}$ is defined as a measure of the correlation between two concepts $c_i$ and $c_j$. This coefficient can take three possible values. 1) When $r_{ij} > 0$, this indicates that the concept $c_i$ has a positive correlation with $c_j$, meaning that the growth of $c_j$ will have a positive effect on $c_i$. 2) When $r_{ij} < 0$, this represents a negative correlation between $c_i$ and $c_j$, which is opposite to the positive correlation. 3) When $r_{ij} = 0$, this means that $c_i$ and $c_j$ are not related, and it is difficult for them to interact with each other. In addition, the magnitude of $r_{ij}$ can be used to infer the strength of the relationship between the two concepts, where the closer $r_{ij}$ is to 0, the weaker the mutual influence between the two concepts, and the farther $r_{ij}$ is from 0, the stronger the influence.

In the original FCM, the manner in which the relationships between concepts change can be described as the degree to which one concept changes relative to another concept. In the proposed tFCM, as the state change is no longer defined by a linear relationship, we must utilize the limiting form to calculate the rate of change in the relationship as follows:

$$r_{ij}(s) = \lim_{\Delta s_i \to 0} \frac{f_j(s_i + \Delta s_i, s \not\subset s_i) - f_j(s_i, s \not\subset s_i)}{\Delta s_i} \tag{9}$$

where $s_i$ denotes the state value of the $i$ concept, and function $r_{ij}(s)$ expresses the degree of $f_j$'s increasing when $s_i$ increases $\Delta s_i$, given $s \not\subset s_i$ as a condition.

At the same time, learners' cognitive states are unlikely to change suddenly; thus, we can use frequency to approximate the probability. In addition, variations may occur in the state relationships among different individuals, and using probability representation can better reduce anomalies. For a state within a certain range, if there are $D$ samples within $[\alpha, \beta]$, then according to the law of large numbers and under the assumption that $s_i \in [\alpha, \beta]$, $\overline{r_{ij}}$ can be calculated approximately as follows:

$$\overline{r_{ij}} = \frac{1}{\mathrm{D}} \sum_{s_i \in [\alpha,\beta]} r_{ij}(s_i, s \not\subset s_i). \tag{10}$$

Note that utilizing a neural network as a tool for state calculation in the proposed model does not hinder the interpretability of the state relationship calculation. In reference to (4), $f_j(s^{(t)})$ represents the result of the output layer of the neural network. Here, we can make the following deduction:

$$\begin{aligned} \frac{\partial f_j}{\partial y(l)} &= \sum \frac{\partial f_j}{\partial y(l{+}1)} \frac{\partial y(l{+}1)}{\partial y(l)} \\ &= \sum \frac{\partial f_j}{\partial y(l{+}1)} \left( w_{(l+1)} \mathrm{ReLU}' \left( \sum w_{(l+1)} y(l) \right) \right) \end{aligned} \tag{11}$$

$$\frac{\partial f_j}{\partial s_j} = \sum \frac{\partial f_j}{\partial y(l+1)} \frac{\partial y(l{+}1)}{\partial y(l)} \cdots \frac{\partial y(1)}{\partial s_j} \tag{12}$$

where $l$ is the number of network layers, $y$ is the input of each layer, and $\omega$ is the weight of the input.

**Algorithm 1:** CSK Calculation.

1: **Input:** Training dataset $= \{\mathrm{t}, \mathrm{s}^{(t)}\}_{t=1}^{T}$
2: Initialize $\mathrm{y}_j^{(t)} = \mathrm{s}_j^{(1)}, \sigma_f$ and $\sigma_h = 0$
3: define loss $= \hat{y}_j^{(t)} - y_j^{(t)}$
4: **while** loss $\leq 0.015$
5: $\quad \sigma_f = \mathrm{nn}(f_j(\sigma_f), \{(s^{(t)}, y_j^{(t)}\}_{t=1}^{T})$
6: $\quad$ for t in$(1, T]$: $y_j^{(t)} = s_j^{(t)} - f_j(s^{(t)}|\omega_{f_j(s_j^{(t)} - h_j(t|\sigma_h))})$
7: $\quad \sigma_h = \mathrm{LSTM}(h_j(\sigma_h), \{(t, y_j^{(t)}\}_{t=1}^{T})$
8: $\quad$ for t in$(1, T]$: $y_j^{(t)} = s_j^{(t)} - h_j(t|\omega_{h_i(s_j^{(t)} - f_j(s_j^{(t)}|\sigma_f)})$
9: **return** $y_j^{(t)}$

### D. Estimation of Cognitive State of Knowledge (CSK)

Based on the derived formulas and structural design mentioned earlier, the historical state information is utilized in the training process, and the historical information is input at each timestamp to influence the CSK at the given timestamp. The CSK calculation process is described in Algorithm 1.

The algorithm uses the state value of the first timestamp as the initial value, and it utilizes the prediction error as the loss value. First, the algorithm employs a backpropagation neural network to estimate the states based on the data correlation between several states, and then it compares the estimated states with the real states to obtain the output correction. Then, the corrected output and time-series data are input into the LSTM network. Finally, the predicted values and historical state information are combined for evaluation purposes. Here, $\omega$ represents the network parameters.

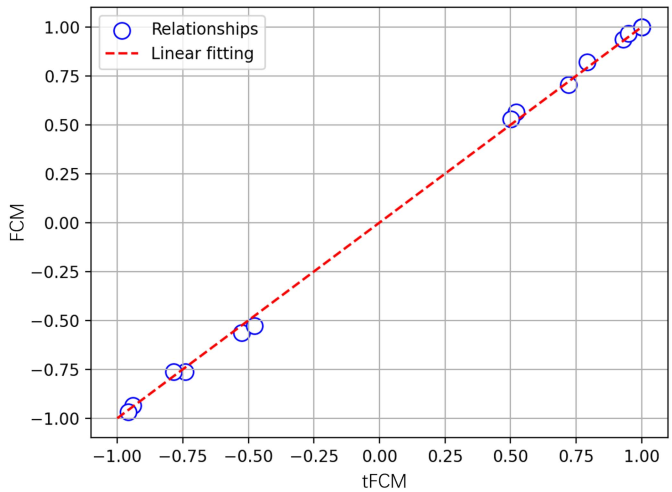
During the iterative process, the model parameters obtained in each iteration are used in the subsequent iteration. For example, the network parameter $\omega$ obtained in the previous iteration is used to estimate $s^{(t)}$, which can be expressed as $f_j(s^{(t)}|\omega_{f_j(s_j^{(t)} - h_j(t|\sigma_h))})$. In the training process, the final estimation result is output when the estimation error is less than 0.015.

## IV. Experimental Results

In this study, we used the ASSISTments2017[1] dataset to evaluate and predict the CSKs of random students. The ASSISTments2017 dataset is a widely used online education dataset containing more than 500 000 problem-solving records from greater than 10 000 students who used the ASSISTments system. This dataset includes data on students' responses, problem answers, problem metadata, and course metadata (among other types of data). Covering math and science subjects from fourth grade to high school, the dataset was collected between 2015

[1][Online]. Available: https://sites.google.com/view/assistmentsdatamining

(a)




(b)

Average state correlation of FCM model solution

| FCM | *Un** | *InK** | *K** | *L** |
|---|---|---|---|---|
| ***Un*** | 1 | -0.76 | -0.56 | -0.52 |
| ***InK*** | -0.76 | 1 | 0.56 | 0.52 |
| ***K*** | -0.93 | 0.93 | 1 | 0.7 |
| ***L*** | -0.96 | 0.96 | 0.82 | 1 |

(c)

Average state correlation of tFCM model solution

| tFCM | *Un** | *InK** | *K** | *L** |
|---|---|---|---|---|
| ***Un*** | 1 | -0.73 | -0.52 | -0.47 |
| ***InK*** | -0.78 | 1 | 0.52 | 0.5 |
| ***K*** | -0.94 | 0.93 | 1 | 0.72 |
| ***L*** | -0.95 | 0.94 | 0.79 | 1 |

Fig. 5. Comparison of fuzzy cognitive map and temporal fuzzy cognitive map relationship results. (a) The overall correlations between fuzzy cognitive map and temporal fuzzy cognitive map. (b) The average state correlation coefficients suggested by fuzzy cognitive map. (c) The average state correlation coefficients suggested by temporal fuzzy cognitive map.

and 2016, and the data are stored in CSV format. This dataset is primarily used to evaluate learning analytics methods in educational technology, e.g., predicting student performance, assessing learning progress, and identifying learning difficulties. The ASSISTments2017dataset has been widely utilized in various educational technology research and learning analytics competitions. In this experimental evaluation, we compared the predictive performance of the proposed tFCM with that of the conventional FCM [19] and LSTM [50] models. Note that the FCM, LSTM, and tFCM models had different inputs in this experiment. These inputs are described as follows.

*FCM:* The basic FCM model uses the fuzzy state data of each student as input, and it predicts the CSK using the basic FCM structure.

*LSTM:* The student learning result data with time series are used as the input to the RNN. The hidden layer of the LSTM model in this experiment contained 20 LSTM neurons. In addition, the LSTM output layer was equipped with a fully connected network to further process the output.

*Proposed tFCM:* The learning result data and fuzzy state data of students with time series were used as the input, and the CSK was estimated and predicted using Algorithm 1.

In this study, RMSE was used to evaluate each model's prediction. The RMSE is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum\nolimits_{t=1}^{T}\left(\hat{s}^{(t)} - s^{(t)}\right)^2}. \tag{13}$$

This formula expresses the comparative relationship between the estimated and actual values at a given time. Here, $\hat{s}^{(t)}$ is the predicted value at time $t$, and $s^{(t)}$ is the actual value. Note that a smaller RMSE value indicates better model performance. The experimental results shown in Table I demonstrate that the proposed tFCM model obtained good performance.

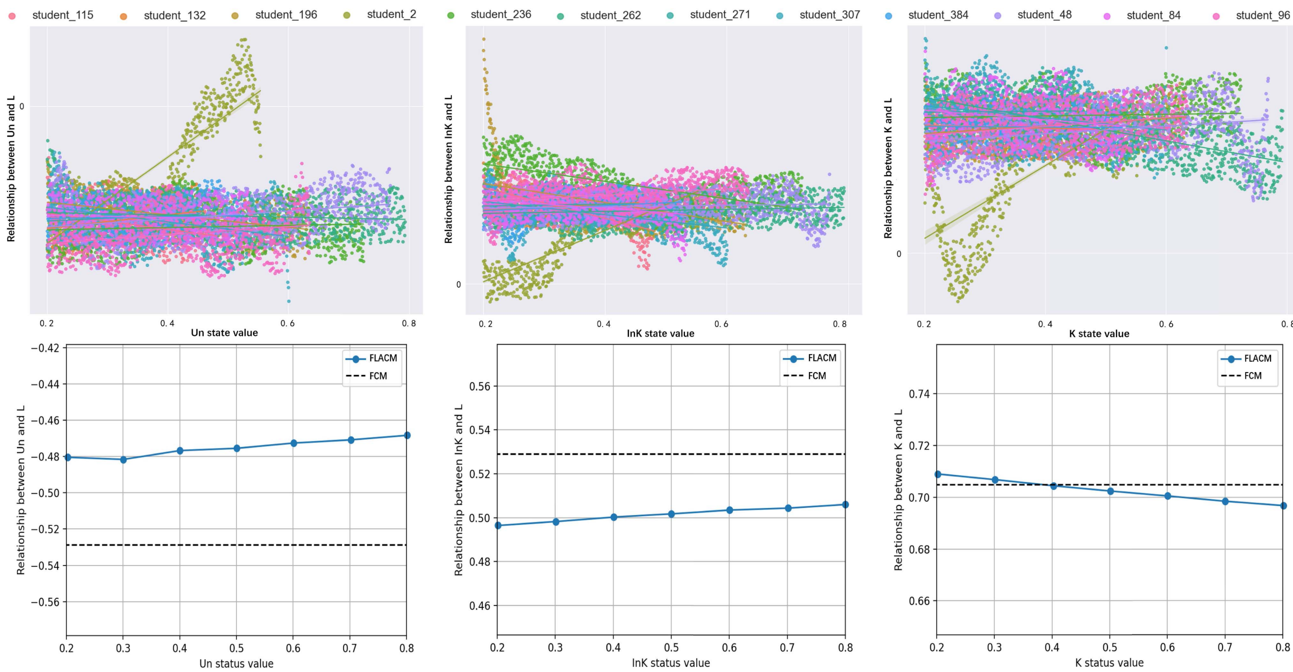
TABLE I
COMPARISON OF RMSE RESULTS

| Student_ID | State | tFCM | FCM | LSTM |
|---|---|---|---|---|
| Student_(1-40) | *Un* | 0.839536 | 0.786839 | **0.77099** |
| | InK | **0.06795** | 0.202212 | 0.266993 |
| | *K* | **0.04488** | 0.068582 | 0.130998 |
| | *L* | 0.151456 | **0.01504** | 0.087728 |
| Student_(41-150) | *Un* | 0.390927 | 0.261338 | **0.23411** |
| | InK | **0.68349** | 0.766356 | 0.812141 |
| | *K* | **0.3628** | 0.528568 | 0.552064 |
| | *L* | **0.15193** | 0.296341 | 0.343318 |
| Student_(151-250) | *Un* | 0.662251 | 0.519375 | **0.47229** |
| | InK | **0.31385** | 0.515987 | 0.497724 |
| | *K* | **0.13675** | 0.229049 | 0.275578 |
| | *L* | **0.10566** | 0.15233 | 0.150871 |
| Student_(251-350) | *Un* | 0.567508 | 0.466344 | **0.45469** |
| | InK | **0.28796** | 0.547831 | 0.580569 |
| | *K* | **0.29002** | 0.309423 | 0.341481 |
| | *L* | 0.297664 | **0.08136** | 0.18482 |
| Student_(351-450) | *Un* | 0.878534 | **0.79364** | 0.867248 |
| | InK | **0.14214** | 0.197902 | 0.179298 |
| | *K* | **0.01337** | 0.052746 | 0.038716 |
| | *L* | **0.07803** | 0.08915 | 0.111373 |
| Count | | **13** | 3 | 4 |

The bold font represents the better result.

## A. Analysis and Interpretability of Fuzzy State Relations

Changes in the CSK frequently result in differences in the final states, and these changes are influenced by various factors,




(a) (b) (c)

Fig. 6. State relation change diagram (the change of *Un*, *InK*, and *K* state values affects their relationship with L). (a)*Un*-*L*. (b) InK-*L*. (c) *K*-*L*.

e.g., the correlation between CSKs. Thus, it is essential to ensure that this relationship is interpretable. The proposed tFCM model inherits from the FCM model and should also be interpretable. To demonstrate sufficient interpretability, we randomly selected a single student as an example. Here, we define the relationship weight of FCM as $r$, the relationship weight of tFCM as $w$, and the values in Fig. 5(b) and (c) represent the directed relationship of the state. For example, $K \rightarrow \text{InK}^*$ is 0.93, and $\text{InK} \rightarrow K^*$ is 0.56, which indicates the influence relationship between $K$ and InK. It can be understood that if the learner is predicted to be in the $K$ state, there is also a high probability that the learner will also be in the InK state (0.93). However, if the learner is predicted to be in the InK state, there is a low probability that the learner will also be in the $K$ state (0.56).

As shown in Fig. 5(a), the relationship points calculated by both models are correlated linearly in the graph, which indicates that the relationships revealed by the FCM and tFCM models are highly similar, thereby suggesting that the proposed tFCM model incorporates FCM's interpretability in the relationship calculations. In addition, we can gain insights into this relationship by analyzing some of the learning processes. First, if the probability of a student being in the *Un* state is high, the probability of the student being in other states is relatively small, and the student may be unable to answer corresponding knowledge-related questions effectively. Second, we can consider the InK state as a cumulative stage of learning. During this stage, it is difficult to predict whether a student will enter the $K$ stage or proceed directly to the $L$ stage because we cannot determine if other factors will influence a student's deep understanding of certain knowledge. Thus, the correlation between InK and both the $K$ and $L$ stages should be at a similar level. However, becoming a master in a specific knowledge area is challenging; thus, the relationship between InK and $L$ is slightly weaker compared to the correlation with $K$. Finally, by examining the relationships between the $K$, $L$, and InK states, the association between $K$ and $L$ is higher compared to the association between InK and $L$, which implies that when student A is predicted to be in the $K$ and student B is predicted to be in the InK, student A has a greater probability of progressing quickly to the $L$ state. Therefore, we can conclude that when students reach the $L$ state, they are more likely to accumulate derived knowledge related to this knowledge than in the $K$ state because they have mastered the knowledge previously.

To verify the generalizability of these explanations to different groups of students, in the following, we demonstrate the process of the relationship between states changing with the change of state values using student groups. The experimental results are shown in Fig. 6, where the second column of figures indicates that the FCM calculates the relationships as static values, and the proposed tFCM handles time-series data, thereby enabling the exploration of dynamic change processes. Then, from the changing values and trends, we can observe the following.

1) In the relationship between *Un* and $L$, the correlation weakens as the probability of *Un* increases, which can be understood as follows. For unknown tasks, students become vague about whether they can complete the task as the degree of uncertainty increases.
2) If the student is very likely to be in the InK stage, they should have some understanding of the corresponding knowledge. Thus, as knowledge accumulates, the CSK may transition to a higher stage.
3) When students are in the $K$ stage, they have accumulated sufficient knowledge. Here, some breakthrough is required

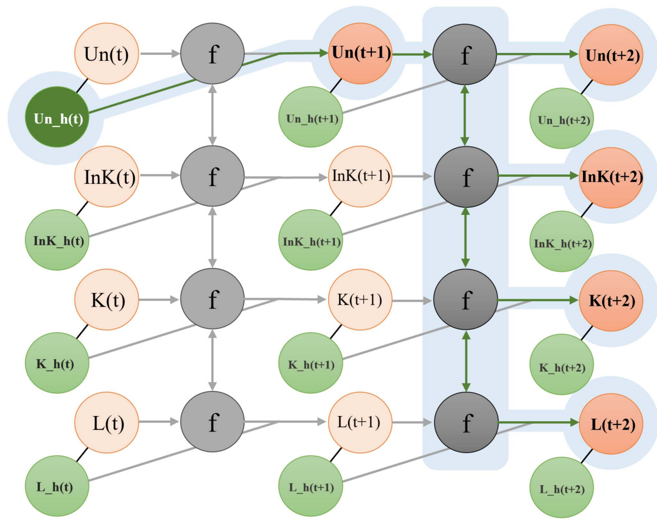



Fig. 7. Complex impact graph of historical state information.

to transform into *L*, similar to an epiphany or inspiration. In addition, the greater the probability of students being in *K*, the weaker the correlation between students and *L*. Here, it can be understood that thinking may solidify in the bottleneck period. The greater the probability of being in *K*, the more difficult it is to break through the bottleneck period. Thus, the correlation between *K* and *L* will be reduced accordingly.

### *B. Association Between Historical State Information and Cognitive State of Knowledge (CSKs)*

To make a reasonable assessment of students' CSK, it is important to consider the learning process as a whole. This includes combining historical state information with the current state to evaluate and predict the state of the student. In the proposed tFCM, as time progresses, the historical state information of a single CSK will also affect the probability of the given student being in other CSKs. The historical state information and current node state information will collectively influence the future state through the current node, which reflects the complexity of CSK changes. To represent this influence relationship visually, the green path in Fig. 7 highlights the impact of historical state information.

It is clear that historical state information is very important in terms of assessing students' CSK. By comparing and analyzing the historical state information and state changes with randomly selected students, we found that the impact of historical state information is not a simple one-to-one relationship. The corresponding experimental results are shown in Fig. 8 and provided in the Appendix. Note that the Appendix also provides a guide on how to interpret the experimental results shown in the figures. Taking Fig. 8 as an example, we first find that the most influential historical state information is *Un_h*, which affects the theme of state change. According to the definition of the fuzzy state, *Un* is negatively correlated with other states. Thus, in the initial stage, the decrease of *Un_h* will promote the decline of *Un* and the rise of InK, *K*, and *L*, which can be understood as students gradually accumulate knowledge from ignorance and gradually eliminate uncertainty. Then, as *Un_h* decreases, its influence gradually approaches *K_h*. Here, InK, *K*, and *L* will have a downward trend, which can be understood as students encountering difficulties in learning. Finally, with the sharp decline of *Un_h* and the increase of historical state information of other CSKs, students' cognitive states begin to rise again, which can be understood as students solving their learning problems through historical accumulation and maintaining a good state. Additional explanations and examples are provided in the Appendix.

By analyzing the historical state information and changes in students' CSK, we can identify the time points at which their state fluctuates during the learning process, which allows us to detect whether students are facing learning problems or have passed the period of fluctuation. By asking how students passed through the fluctuation period, we can gather insights into effective learning strategies and utilize them to accumulate teaching experience. This information can also be used to develop more effective adaptive learning methods, which can personalize the learning experience for individual students based on their unique cognitive state and learning history.

## V. Discussion

The goal of this study was to forecast the CSK and explore the interconnected dimensions to gain insights into cognitive states in the learning process. Our investigation has primarily analyzed the historical state data and examined the correlation between states. In exploring the link between historical state data and CSKs, we discovered a fascinating result (Fig. 8), i.e., the historical state data reflect the pattern of fluctuations in the students' CSK, potentially holding vital insights into their psychological and self-regulation mechanisms. Positive self-regulation enhances CSKs [51], [52], [53]. This finding verifies the connection between historical state data and CSKs. In comparison, the observed time points offer cues about the students' progression through various cognitive state stages.

In addition, by utilizing historical state data, we can identify the exact moments at which students encounter challenges and subsequently make informed judgments about their progress. Moreover, the interconnections between different stages of cognitive state allow educators or adaptive systems to gain understanding and provide effective guidance and advice. This empowers them to aid students by modifying the targeted learning paths, enhancing dedicated study periods, delivering psychological support, and extending various forms of assistance. Finally, researchers can perform further analyses of cognitive states in conjunction with other studies to achieve more precise individual evaluations [54], [55], [56]. Previous studies have shown that integrating multidimensional historical state information significantly improves the extraction of valuable insights [28]. This information includes various factors, e.g., the student's knowledge mastery, psychological influences, and environmental factors. Utilizing the informative nature of historical state

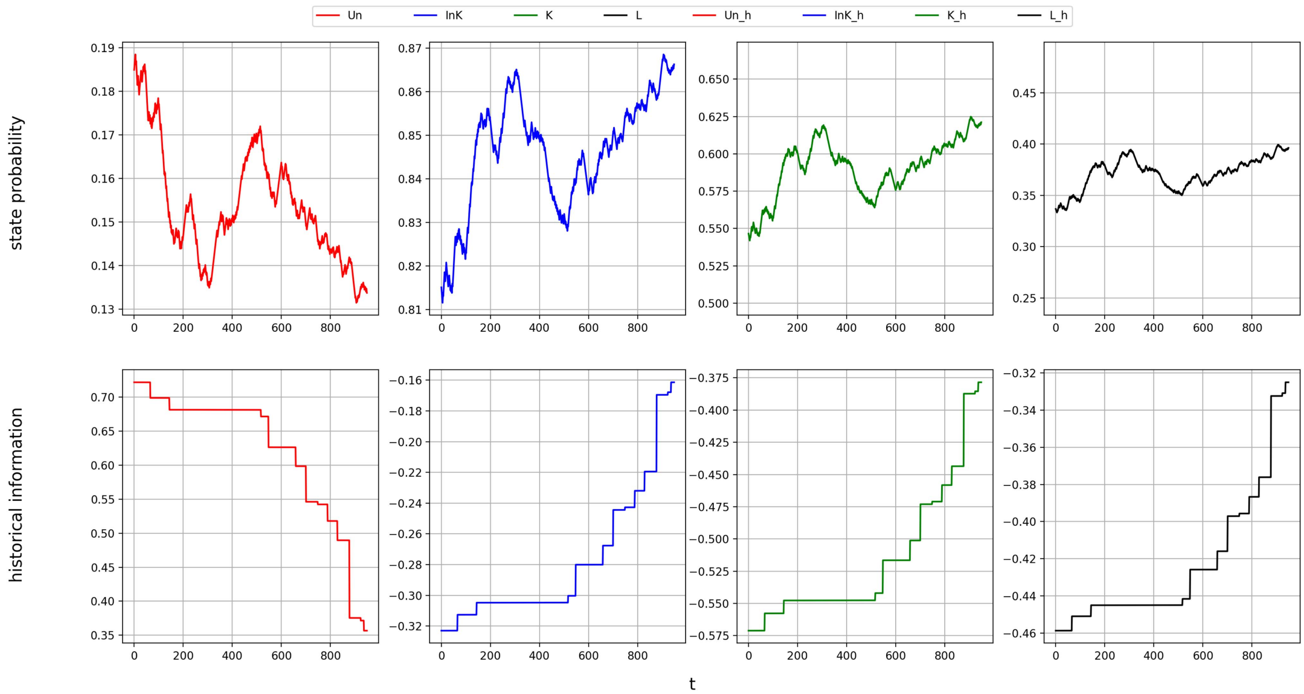



Fig. 8. Influence of historical state information on cognitive states of knowledge.

information enables us to effectively pinpoint the underlying causes of problems, predict the probability of future issues, and recognize the particular states in which such issues are likely to occur. These abilities are crucial for educators because they guide the learning journey and set a strong foundation for future exploration of students' individual characteristics and psychological factors.

In addition, by understanding how cognitive skills are transferred, we can determine if students are facing difficulties in their learning process, and we can also ascertain whether students have already achieved proficiency in the given subject. This knowledge empowers teachers and adaptive learning systems to customize their educational methods and provide advanced courses that align with the student's current learning level. Moreover, analyzing the patterns of cognitive skill development enables us to implement specific interventions to promote continual learning progress.

The fuzzy states employed in this study provide valuable insights into the learning process; however, an improved approach could be to utilize fuzzy states to model students' mastery of specific knowledge components. This approach may offer more precise and targeted insights, thereby facilitating more effective learning. However, individual knowledge component datasets frequently have limited samples. In this context, if datasets with sufficient data volume for personal knowledge component learning results are available, more accurate model construction will become increasingly feasible. In addition, incorporating additional factors, e.g., psychological [57], [58] and environmental factors [59], [60], into the prediction process will enhance the model's predictive performance and provide additional insights into the impact of various cognitive states.

## VI. Conclusion

This article has presented a framework that employs nonlinear fuzzy logic to predict students' fuzzy CSK. This approach captures both the nonlinear and uncertain nature of learning, and the framework defines four probability states to represent students' cognitive states. The proposed tFCM model, which integrates historical state information with the FCM model, is introduced to predict this fuzzy state. The proposed tFCM model offers advantages in terms of solving nonlinear problems while preserving FCM's interpretability. The interpretability of the proposed model was validated by examining the correlation between CSKs and the influence of historical state information on states. In addition, we have demonstrated that the proposed tFCM model outperforms similar existing models (FCM and LSTM). We believe that our findings highlight the importance of utilizing past data to detect crucial time points when students might face challenges and require extra assistance. In addition, our investigation has uncovered and examined plausible factors by evaluating the comprehensible connections between fuzzy states and historical states. We expect that the findings of this study will contribute to the expanding realm of uncertainty reasoning in adaptive learning research.

In the future, our goal is for students to improve their abilities rather than simply focus on acquiring knowledge. However, the complexity of competency-based education and the uncertainty it entails present various research challenges. Thus, in future studies, we plan to experiment with models that utilize data on competency development, not solely limited to extracting knowledge-related aspects. This will involve predicting the emotional value derived from learning, extracting factors that

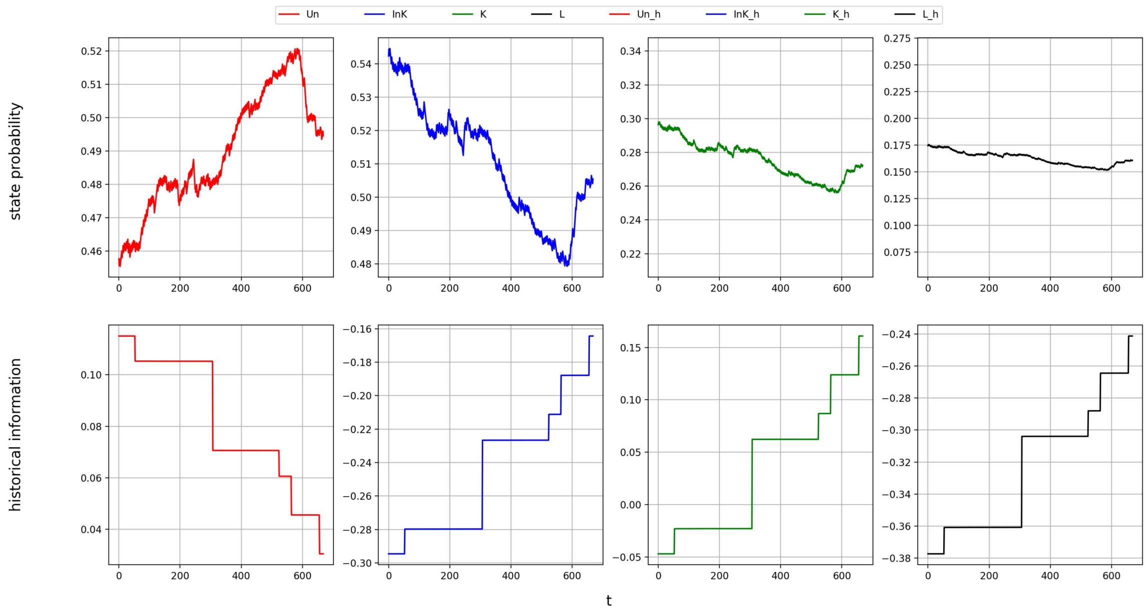



Fig. 9. Influence of students' historical state information on their cognitive states of knowledge (a).

enhance confidence, and exploring combined variables to calculate comprehensive abilities, among other things. In addition, we plan to integrate the proposed model into our personalized learning system, which was specifically designed for elementary school mathematics courses, to gather valuable user feedback and empirically assess the effectiveness of the model in a practical environment.

## Appendix
## Influence of Students' Historical State on Their CSK

Here, we present an example of the influence of students' historical state information on their CSK. First, there is a rule regarding the influence of historical state information on CSKs, as illustrated in Table II.

TABLE II
IMPACT OF HISTORICAL STATE VALUES AND TRENDS ON CSKs

| | Value and Trend | *Un* | *InK* | *K* | *L* |
|---|---|---|---|---|---|
| ***Un_h*** | P-P | ↑ | ↓ | ↓ | ↓ |
| | P-N | ↓ | ↑ | ↑ | ↑ |
| | N-P | ↓ | ↑ | ↑ | ↑ |
| | N-N | ↓ | ↑ | ↑ | ↑ |
| ***InK_h,K_h,L_h*** | P-P | ↓ | ↑ | ↑ | ↑ |
| | P-N | ↑ | ↓ | ↓ | ↓ |
| | N-P | ↑ | ↓ | ↓ | ↓ |
| | N-N | ↑ | ↓ | ↓ | ↓ |

1. P: Positive (indicates that value or trend is a positive number).
2. N: Negative (indicates that the value or trend is a negative number).
3. Arrow: the trend of change for the given state.

1) When *Un_h* is the most influential characteristic. Here, if *Un_h* is positive, and its change trend is positive, it is believed that the students' cognition of things is developing in the direction of chaos. At this time, the students' good state will deteriorate.
2) When the most influential feature is *L_h* or InK_*h* or *K_h*. For example, assume that *K_h* is the most influential at a given time. If *K_h* is positive and the trend of change is positive, we believe that students are in a state of being encouraged and that their CSK will improve. If the value or trend of *K_h* is negative, there is a problem in the learning process of students, which will lead to a decline in the students' state.

In Fig. 9, the most influential historical state information is *K_h*; thus, *K_h* will be the initial factor that affects the change in the CSK. In the initial stage, the value of *K_h* is negative, and the trend of change is positive, which leads to a decline in InK, *K*, and *L*, which indicates that the students may encounter learning difficulties. As *L_h*, InK_*h*, and *K_h* increase and *Un_h* becomes smaller, the historical state information has a gradual positive impact on the state. At this point, InK, *K*, and *L* exhibit an upward trend, which suggests that the students are attempting to solve problems or have already solved the problems. After resolving the problem successfully, the student's state has improved.

As shown in Fig. 10, the most influential historical state information at the initial stage is *L_h*. Thus, *L_h* becomes the initial factor affecting the change of CSK. Initially, the value of *L_h* is positive, and the trend of change is negative. As a result, InK, *K*, and *L* values decline, thereby indicating that students may be facing difficulties in learning. As *K_h* increases, *L_h* becomes smaller, and the influence of *K_h* on the state gradually becomes more dominant. At this point, *K_h* is positive and its trend is also positive; thus, the values for InK, *K*, and *L* show an upward trend, which indicates that the students have found a way to solve problems. Once the problem is solved, the students' states begin to improve rapidly.

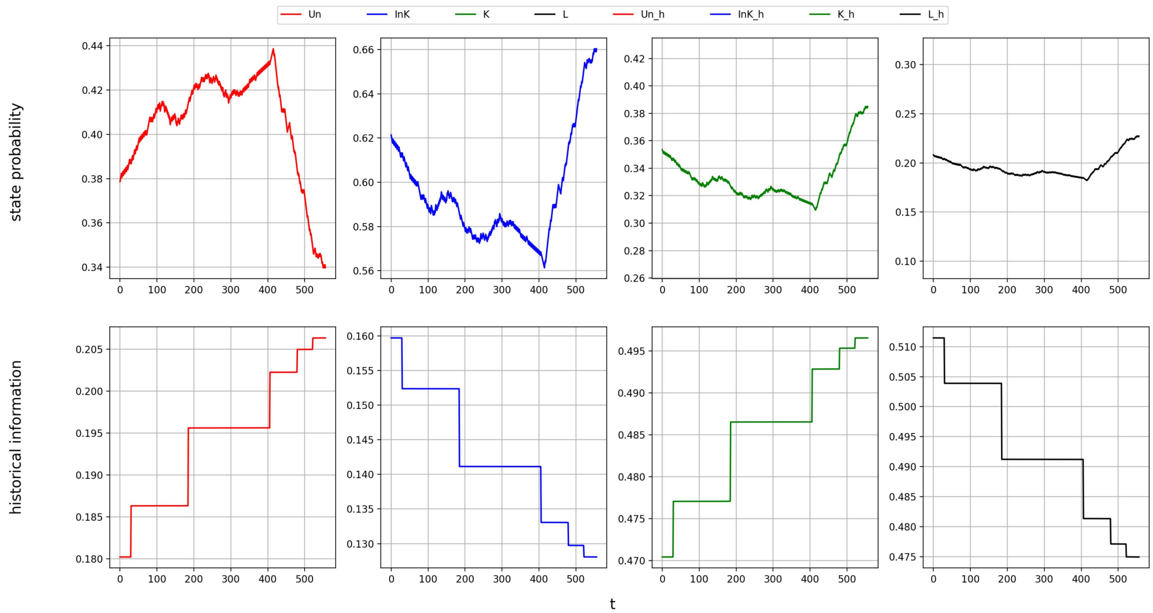



Fig. 10. Influence of students' historical state information on their cognitive states of knowledge (b).

## REFERENCES


[1] N. R. Council, *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC, USA: Nat. Academies Press, 2000.

[2] J. R. Anderson, *The Architecture of Cognition*. Cambridge, MA, USA: Harvard Univ. Press, 1983.

[3] A. Gisolfi, A. Dattolo, and W. Balzano, "A fuzzy approach to student modeling," *Comput. Educ.*, vol. 19, no. 4, pp. 329–334, 1992.

[4] K. Chrysafiadi and M. Virvou, *Modeling Student's Knowledge on Programming Using Fuzzy Techniques* (Intelligent Interactive Multimedia Systems and Services Series), vol. 6, Berlin, Germany:Springer, 2010, pp. 23–32.

[5] M. D. Reckase and R. L. McKinley, "The discriminating power of items that measure more than one dimension," *Appl. Psychol. Meas.*, vol. 15, no. 4, pp. 361–373, 1991.

[6] S. Cheng et al., "Dirt: Deep learning enhanced item response theory for cognitive diagnosis," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, 2019, pp. 2397–2400.

[7] J. De La Torre, "The generalized DINA model framework," *Psychometrika*, vol. 76, pp. 179–199, 2011.

[8] B. W. Junker and K. Sijtsma, "Cognitive assessment models with few assumptions, and connections with nonparametric item response theory," *Appl. Psychol. Meas.*, vol. 25, no. 3, pp. 258–272, 2001.

[9] S. Wang, Y. Yang, S. A. Culpepper, and J. A. Douglas, "Tracking skill acquisition with cognitive diagnosis models: A higher-order, hidden Markov model with covariates," *J. Educ. Behav. Statist.*, vol. 43, no. 1, pp. 57–87, 2018.

[10] Y. Chen et al., "Tracking knowledge proficiency of students with educational priors," in *Proc. ACM Conf. Inf. Knowl. Manage.*, 2017, pp. 989–998.

[11] A. Toscher and M. Jahrer, "Collaborative filtering applied to educational data mining," *J. Mach. Learn. Res.*, 2010.

[12] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. Chelmsford, MA, USA:Courier Corporation, 2013.

[13] K. Smith and M. Smith, *Cybernetic Principles of Learning and Educational Design*. Austin, TX, USA: Holt, Rinehart and Winston, 1966. [Online]. Available: https://books.google.com/books?id=67JRAQAAIAAJ

[14] C. Lee, "Fuzzy logic in control systems: Fuzzy logic controller. I," *IEEE Trans. Syst., Man, Cybern.*, vol. 20, no. 2, pp. 404–418, Mar./Apr. 1990.

[15] S. Tong and Y. Li, "Adaptive fuzzy output feedback tracking backstepping control of strict-feedback nonlinear systems with unknown dead zones," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 168–180, Feb. 2011.

[16] S. Tong and Y. Li, "Adaptive fuzzy output feedback control of MIMO nonlinear systems with unknown dead-zone inputs," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 1, pp. 134–146, Feb. 2013.

[17] S. Tong, T. Wang, Y. Li, and B. Chen, "A combined backstepping and stochastic small-gain approach to robust adaptive fuzzy output feedback control," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 2, pp. 314–327, Apr. 2013.

[18] P. Melin, J. Soto, O. Castillo, and J. Soria, "A new approach for time series prediction using ensembles of ANFIS models," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 3494–3506, 2012.

[19] W. Stach, L. Kurgan, W. Pedrycz, and M. Reformat, "Genetic learning of fuzzy cognitive maps," *Fuzzy Sets Syst.*, vol. 153, no. 3, pp. 371–401, 2005.

[20] E. I. Papageorgiou, "Review study on fuzzy cognitive maps and their applications during the last decade," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 2011, pp. 828–835.

[21] E. I. Papageorgiou and D. K. Iakovidis, "Intuitionistic fuzzy cognitive maps," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 2, pp. 342–354, Apr. 2013.

[22] K. Chrysafiadi and M. Virvou, "Using fuzzy cognitive maps for the domain knowledge representation of an adaptive e-learning system," in *Proc. Int. Joint Conf. Knowl.-Based Softw. Eng.*, 2012, pp. 257–265.

[23] E. I. Papageorgiou and J. L. Salmeron, "A review of fuzzy cognitive maps research during the last decade," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 1, pp. 66–79, Feb. 2013.

[24] T. A. Ackerman, "Multidimensional item response theory models," in John Wiley & Sons, Ltd., 2014. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118445112.stat06401

[25] J. B. Sympson, "A model for testing with multidimensional items," in *Proc. Computerized Adaptive Testing Conf.*, vol. 14, 1978, pp. 82–98.

[26] X. Wang, J. O. Berger, and D. S. Burdick, "Bayesian analysis of dynamic item response models in educational testing," *Ann. Appl. Statist.*, vol. 7, no. 1, pp. 126–153, 2013.

[27] P. Zhan, H. Jiao, and D. Liao, "Cognitive diagnosis modelling incorporating item response times," *Brit. J. Math. Stat. Psychol.*, vol. 71, no. 2, pp. 262–286, 2018.

[28] N. Thai-Nghe and L. Schmidt-Thieme, "Multi-relational factorization models for student modeling in intelligent tutoring systems," in *Proc. 7th Int. Conf. Knowl. Syst. Eng.*, 2015, pp. 61–66.

[29] M. C. Desmarais, "Mapping question items to skills with non-negative matrix factorization," *ACM Sigkdd Explorations Newslett.*, vol. 13, no. 2, pp. 30–36, 2012.

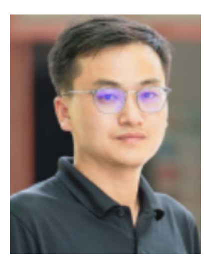
[30] Y. Sun, S. Ye, S. Inoue, and Y. Sun, "Alternating recursive method for q-matrix learning," in *Proc. Int. Conf. Educ. Data Mining*, 2014, pp. 14–20.

[31] L. Gao, Z. Zhao, C. Li, J. Zhao, and Q. Zeng, "Deep cognitive diagnosis model for predicting students' performance," *Future Gener. Comput. Syst.*, vol. 126, pp. 252–262, 2022.

[32] P. Alves, L. Amaral, and J. Pires, "Case-based reasoning approach to adaptive web-based educational systems," in *Proc. IEEE 8th Int. Conf. Adv. Learn. Technol.*, 2008, pp. 260–261.
[33] J. Chen, K. Huang, F. Wang, and H. Wang, "E-learning behavior analysis based on fuzzy clustering," in *Proc. 3rd Int. Conf. Genet. Evol. Comput.*, 2009, pp. 863–866.
[34] A. Kavcic, "Fuzzy user modeling for adaptation in educational hypermedia," *IEEE Trans. Syst., Man, Cybern., Part C (Appl. Rev.)*, vol. 34, no. 4, pp. 439–449, Nov. 2004.
[35] K. Chrysafiadi and M. Virvou, "Fuzzy logic for adaptive instruction in an e-learning environment for computer programming," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 1, pp. 164–177, Feb. 2015.
[36] B. Kosko, "Fuzzy cognitive maps," *Int. J. Man-Mach. Stud.*, vol. 24, no. 1, pp. 65–75, 1986.
[37] L. Chunmei, "Model method and application study of fuzzy cognitive map," Ph.D. dissertation, College Inform. Sci. Technol., Dong Hua Univ., Shanghai, China, Nov. 2006.
[38] M. Hagiwara, "Extended fuzzy cognitive maps," *IEEE Trans. Electron., Inf. Syst.*, vol. 114, no. 3, pp. 367–372, 1994.
[39] L. Xiangfeng, G. Jun, and Z. Xudong, "Probabilistic fuzzy cognitive map based on belief knowledge database," *Comput. Res. Develop.*, vol. 40, no. 7, pp. 925–933, 2003.
[40] L. Xiangfeng and G. Juan, "Probability fuzzy cognitive map," *J. China Univ. Sci. Technol.*, vol. 33, no. 1, pp. 26–33, 2003.
[41] C. D. Stylios and P. P. Groumpos, "Fuzzy cognitive maps: A model for intelligent supervisory control systems," *Comput. Ind.*, vol. 39, no. 3, pp. 229–238, 1999.
[42] H.-Q. Min, J.-X. Hui, Y.-S. Lu, and J.-Z. Jiang, "Probability fuzzy cognitive map for decision-making in soccer robotics," in *Proc. IEEE//WIC/ACM Int. Conf. Intell. Agent Technol.*, 2006, pp. 321–325.
[43] D. Kardaras and G. Mentzas, "Using fuzzy cognitive maps to model and analyze business performance assessment," *Adv. Ind. Eng. Appl. Pract. II*, vol. 2, pp. 63–68, 1997.
[44] J. L. Salmeron and E. I. Papageorgiou, "A fuzzy grey cognitive maps-based decision support system for radiotherapy treatment planning," *Knowl.-Based Syst.*, vol. 30, pp. 151–160, 2012.
[45] J. R. Cole and K. A. Persichitte, "Fuzzy cognitive mapping: Applications in education," *Int. J. Intell. Syst.*, vol. 15, no. 1, pp. 1–25, 2000.
[46] C. Herry and J. P. Johansen, "Encoding of fear learning and memory in distributed neuronal circuits," *Nature Neurosci.*, vol. 17, no. 12, pp. 1644–1654, 2014.
[47] B. F. Hobbs, S. A. Ludsin, R. L. Knight, P. A. Ryan, J. Biberhofer, and J. J. Ciborowski, "Fuzzy cognitive mapping as a tool to define management objectives for complex ecosystems," *Ecological Appl.*, vol. 12, no. 5, pp. 1548–1565, 2002.
[48] O. Osoba and B. Kosko, "Causal modeling with feedback fuzzy cognitive maps," in John Wiley & Sons, Ltd., ch. 25, p. 587615, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119485001.ch25
[49] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
[50] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[51] W. Eom and R. A. Reiser, "The effects of self-regulation and instructional control on performance and motivation in computer-based instruction," *Int. J. Instructional Media*, vol. 27, no. 3, 2000, Art. no. 247.
[52] V.-A. Romero-Zaldivar, A. Pardo, D. Burgos, and C. D. Kloos, "Monitoring student progress using virtual appliances: A case study," *Comput. Educ.*, vol. 58, no. 4, pp. 1058–1067, 2012.
[53] S. Joksimović, D. Gašević, V. Kovanović, B. E. Riecke, and M. Hatala, "Social presence in online discussions as a process predictor of academic performance," *J. Comput. Assist. Learn.*, vol. 31, no. 6, pp. 638–654, 2015.
[54] R. K. Verma and S. Kumari, "Effect of self-confidence on academic achievement of children at elementary stage," *Indian J. Res.*, vol. 5, no. 1, pp. 181–183, 2016.
[55] F. Fatma, "A study of anxiety of adolescents in relation to their gender, locality and academic achievement," *Int. J. Appl. Res.*, vol. 1, no. 12, pp. 407–411, 2015.
[56] A. Karimi, "The relationship between self-confidence with achievement based on academic motivation," *Kuwait Chapter Arabian J. Bus. Manage. Rev.*, vol. 4, no. 1, pp. 210–215, 2014.
[57] Z. A. Pardos, R. S. Baker, M. O. San Pedro, S. M. Gowda, and S. M. Gowda, "Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes," *J. Learn. Anal.*, vol. 1, no. 1, pp. 107–128, 2014.
[58] M. M. T. Rodrigo et al., "Affective and behavioral predictors of novice programmer achievement," in *Proc. 14th Annu. ACM SIGCSE Conf. Innov. Technol. Comput. Sci. Educ.*, 2009, pp. 156–160.
[59] R. S. Baker, "Modeling and understanding students' off-task behavior in intelligent tutoring systems," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2007, pp. 1059–1068.
[60] M. Cocea, A. Hershkovitz, and R. S. Baker, "The impact of off-task and gaming behaviors on learning: Immediate or aggregate?," in *Proc. Artif. Intell. Educ.*, 2009, pp. 507–514.

**Yuang Wei** received the master's degree in control science and engineering from the North China University of Technology, Beijing, China, in 2022. He is currently working toward the Ph.D. in computer science and technology (specializing in intelligent education) with the East China Normal University, Shanghai, China.

His research largely focuses on knowledge tracing and his current research interest includes the construction of knowledge maps and natural language processing models in education.



**Bo Jiang** received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2014.

He was an Associate Professor with the Department of Educational Information Technology, Zhejiang University of Technology. He is currently an Associate Professor with the Department of Educational Information Technology, East China Normal University, Shanghai, China. His research interests include learner modeling, learning analytics, and computer science education.

Dr. Jiang was the recipient of the 2021 Asia-Pacific Society on Computers in Education Early Career Research Award. He is currently serves on the Editorial Board member of IEEE Transactions on Learning Technologies and Research & Practice in Technology Enhanced Learning. He also serves as the Executive Committee member of the Asia-Pacific Society on Computers in Education.