Personalized Early Warning of Learning Performance for College Students: A Multilevel Approach via Cognitive Ability and Learning State Modeling

Hua Ma, *Senior Member, IEEE*, Wen Zhao, Yuqi Tang, Peiji Huang, Haibin Zhu, *Senior Member, IEEE*, Wensheng Tang, and Keqin Li, *Fellow, IEEE*

***Abstract*—To prevent students from learning risks and improve teachers' teaching quality, it is of great significance to provide accurate early warning of learning performance to students by analyzing their interactions through an e-learning system. In existing research, the correlations between learning risks and students' changing cognitive abilities or learning states are still underexplored, and the personalized early warning is unavailable for students at different levels. To accurately identify the possible learning risks faced by students at different levels, this article proposes a personalized early warning approach to learning performance for college students via cognitive ability and learning state modeling. In this approach, students' learning process data and historical performance data are analyzed to track students' cognitive abilities in the whole learning process, and model their learning states from four dimensions, i.e., learning quality, learning engagement, latent learning state, and historical learning state. Then, the Adaboost algorithm is used to predict students' learning performance, and an evaluation rule with five levels is designed to dynamically provide multilevel personalized early warning to students. Finally, the comparative experiments based on real-world datasets demonstrate that the proposed approach could effectively predict all students' learning performance, and provide accurate early warning services to them.**

***Index Terms*—Cognitive ability modeling, learning performance prediction, learning state modeling, multilevel early warning (EW), personalized EW.**

Manuscript received 2 November 2023; revised 20 January 2024 and 25 February 2024; accepted 21 March 2024. Date of publication 27 March 2024; date of current version 23 April 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62077014 and Grant 71971221, in part by the Key Project of Scientific Research Fund of Hunan Provincial Education Department under Grant 23A0061, in part by the Key Project of General Universities Teaching Reform Research Fund of Hunan Province under Grant HNJG-20230213, and in part by the Comprehensive Reform Project of Undergraduate Education and Teaching of Hunan Normal University. *(Corresponding author: Wen Zhao.)*

Hua Ma, Wen Zhao, Yuqi Tang, Peiji Huang, and Wensheng Tang are with the Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410081, China, and also with the College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China (e-mail: huama@hunnu.edu.cn; zhaowen@hunnu.edu.cn; 202170293851@hunnu.edu.cn; peiji_huang@hunnu.edu.cn; tangws@hunnu.edu.cn).

Haibin Zhu is with the Department of Computer Science and Mathematics, Nipissing University, North Bay, ON P1B 8L7, Canada (e-mail: haibinz@nipissingu.ca).

Keqin Li is with the Department of Computer Science, State University of New York, New Paltz, NY 12561 USA (e-mail: lik@newpaltz.edu).

Digital Object Identifier 10.1109/TLT.2024.3382217

Nomenclature

CDM	Cognitive diagnosis model [1].
IRT	Item response theory [2].
DINA	Deterministic inputs, noisy and-gate [3].
FuzzyCDF	Fuzzy cognitive diagnosis framework [4].
FC-CDF	Fuzzy cloud cognitive diagnosis framework [5].
BKT	Bayesian knowledge tracing [6].
DBKT	Dynamic Bayesian knowledge tracing [7].
LFA	Learning factors analysis [8].
DKT	Deep knowledge tracing [9].
LPKT	Learning process-consistent knowledge tracing [10].

I. Introduction

A. Motivations

By mining and analyzing the data related to students' learning processes, early warning (EW) of learning performance aims to identify the students who might be at risk for their learning in the future [11]. EW could not only urge students to make timely self-correction but also assist teachers in adapting their instructional strategies, and is of great significance in promoting students' academic success and improving teachers' instructional quality [12]. Moreover, in recent years, China is vigorously promoting the Engineering Education Accreditation.[1] It is necessary to establish an effective quality monitoring mechanism of the teaching process to ensure that students achieve the overall teaching objectives of a course and the training objectives of a major. Therefore, it has become a hotspot to provide dynamic and accurate EW to students with different characteristics on the basis of continuous monitoring of their whole learning process [13], [14].

Despite the efforts of previous research in EW of learning performance, there are still the following limitations [15], [16], [17].

1) EW is usually modeled as a classification problem of learning performance, and mainly focuses on identifying students at risk of dropout [17], failure [18], or delayed graduation [19]. In practical teaching process, we should not only pay more attention to the underachieving students, but also pay attention to all students' fluctuations

[1] [Online]. Available: https://www.ceeaa.org.cn/jyzyrzxh/index/index.html

in academic performance. For example, both George and Bob rank around 50% in the exams in a course. George's performance has remained around the top 50% of the class in previous courses, while Bob's ranking has consistently declined from the top 10% to 50%. Therefore, it is urgent to warn George about the fluctuations and send prompts to teachers to provide appropriate intervention and personalized instruction to George. However, existing research lacks the ability to analyze the different learning characteristics of students at different levels [20], failing to provide personalized EW to all students.

2) Existing research usually predicts a student who might be at risk for their learning in the future based on the student's demographic information and test scores. Nevertheless, it is difficult to reveal the correlation between learning risks and students' changing cognitive abilities or learning states. For example, suppose that George and Bob with similar background have an average score of about 70 in a phase exam of a multiphase course. However, their learning process data recorded in an e-learning system could demonstrate some significant differences between George and Bob. George has mastered all key knowledge concepts and actively participated in the daily learning process. However, as a result of carelessness, George has poor performance in this exam. Bob has insufficient mastery on knowledge concepts and has not actively engaged in daily learning activities. Obviously, since George and Bob face different levels of learning risks, we should send different EW signals to them. However, the existing research has failed to accurately evaluate the students' changing cognitive abilities and learning states. Thus, it is not conducive to achieve accurate and personalized EW [20].

Aiming at the limitations of existing research, this article proposes a personalized EW approach to learning performance via cognitive ability and learning state modeling. In this approach, the students' changing cognitive states are dynamically analyzed based on their learning process data. Then, the learning process data and historical performance data from an e-learning system are mixed to model the students' learning features from multiple dimensions. Based on these features, the students' learning performance in the next phase is predicted. By comprehensively analyzing the historical and predicted data about the students' learning performance, a five-level evaluation rule for EW is designed to provide multilevel and differentiated early warning results to all students. The experiments based on real-world datasets demonstrate that the proposed approach is effective to predict all students' learning performance, and provide accurate early warning services.

B. Our Contributions

The main contributions of this article are as follows.

1) To explore the correlation between learning risks and students' changing cognitive abilities or learning states, this article proposes a comprehensive approach to model students' cognitive abilities and learning states. The students' cognitive abilities are measured with a DKT model, and the students' learning states are modeled from four dimensions, i.e., learning quality (LQ), learning engagement (LE), latent learning state (LLS), and historical learning state (HLS). The modeling results of students' cognitive abilities and learning states serve as a crucial decision-making foundation for personalized early warning.

2) To accurately identify the earning risks of students at different levels, this article innovatively designs a personalized early warning approach to learning performance from the perspective of multiclassification. The learning performance of students is predicted via an Adaboost algorithm and divided into five grades, i.e., excellent, good, medium, pass, and fail. Then, the five-level evaluation rules are designed to provide differentiated EW to students. The proposed approach is conducive to preventing students from encountering learning risks and helping teachers improve their teaching for achieving the teaching objectives of a curriculum.

II. Related Work

This section reviews the existing research on cognitive ability modeling and EW approaches of learning performance. The related work on cognitive ability modeling is divided into two categories, i.e., cognitive diagnosis and knowledge tracing. The related work on EW is presented from three aspects, i.e., traditional EW, EW via cognitive ability analysis, and multilevel EW.

A. Cognitive Ability Modeling

In the domain of smart education, the students' mastery of knowledge concepts is employed to represent their cognitive abilities [5]. The modeling of cognitive ability helps us understand a student's knowledge concepts mastery based on their response behaviors and outcome data. In recent years, researchers have proposed various approaches for cognitive ability modeling, such as cognitive diagnosis, knowledge tracing, coverage models, differential models, and deviation models [21]. Among them, cognitive diagnosis and knowledge tracing are currently the two most mainstream methods.

1) Cognitive Diagnosis: In the domain of psychological and educational measurement, cognitive diagnosis refers to the diagnostic assessment of a student's cognitive process and knowledge concepts [22]. CDMs are typically classified into discrete CDMs and continuous ones.

IRT [2] is one of the most representative continuous CDMs. IRT assumes that each student possesses a unique latent trait, representing their potential ability. It models students' cognitive abilities with a single value by analyzing students' responses to exercises. However, a single value cannot reflect student's abilities in distinct knowledge concepts. Besides, IRT has some limitations, e.g., the complex calculations and the difficulty in satisfying the assumption of a single dimension. In contrast, the DINA [3] model establishes the link between students' cognitive ability and knowledge concepts via the

Q-matrix [23], which indicates the association between exercises and knowledge concepts. With the incorporation of Q-matrix and students' historical learning data, each student's cognitive ability is represented by a binary multidimensional vector.

To solve the problem of the DINA model's inability to diagnose subjective exercise, FuzzyCDF [4] utilizes fuzzy intersection and fuzzy union operations to model the cognitive response patterns of objective and subjective questions, respectively. Thus, students' cognitive ability is expressed as the membership degree of a fuzzy set, i.e., a real number in a range from 0 to 1. To further improve the diagnostic efficiency of FuzzyCDF, FC-CDF [5] utilizes the cognitive cloud to analyze students' skill proficiency with three numerical features, i.e., the expectation, degree of variation, and variation frequency. The experiments indicate that FC-CDF has significantly enhanced the execution time of cognitive diagnosis. To improve the execution efficiency of cognitive diagnosis, we presented a neutrosophic cognitive diagnosis approach [24] with consideration of the complexity of knowledge concepts and the influence of forgetting factors. This approach employs the neutrosophic set theory to comprehensively evaluate students' cognitive ability on knowledge concepts from three features, i.e., understanding level, degree of misunderstanding, and uncertainty.

2) Knowledge Tracing: The knowledge tracing is a dynamic approach to evaluate the students' mastery of specific knowledge concepts by analyzing their interaction data with the exercises [25]. It is applied to many downstream tasks, such as cognitive ability modeling and performance prediction [21]. The classical knowledge tracing models could be divided into three categories as follows.

1) *Probabilistic models:* It assumes that the student's learning process follows a Markov process and that the outcomes are statistically interpretable. For example, BKT model, proposed by Corbett and Anderson [6], is a specialized instance of the hidden Markov model. It models students' learning processes with Bayesian networks, and diagnoses students' cognitive abilities by incorporating both slip and guess factors. However, BKT models each knowledge concept independently, but ignores the prerequisite or similar relationship within concepts. To address this limitation, Käser et al. [7] put forward a DBKT model. This model utilizes a dynamic Bayesian network to model the hierarchical relationships among knowledge concepts. Consequently, it could simultaneously model students' cognitive abilities on different knowledge concepts within a single model.
2) *Logistic model:* It estimates the parameters about students' learning abilities and practices the parameters about test difficulty and test differentiation through mathematical functions to predict the probability of students answering exercises correctly. For example, the learning factors analysis (LFA) [8] model assumes that learning is an ongoing and dynamic process, and it models students' cognitive abilities using variables, such as their participation rate and the number of attempts at exercises. By using them, the LFA model would predict the probability of students answering exercises correctly.
3) *Deep learning-based models:* It utilizes the robust feature extraction capabilities of deep learning algorithms to effectively model the intricate cognitive processes of students. Considering the fact that knowledge tracing is a sequential prediction task, Piech et al. [9] put forward a DKT model by introducing a recurrent neural network (RNN). However, DKT fails to capture students' learning gain, which limits its capacity of modeling students' real learning processes. To tackle it, Shen et al. [10] proposed an LPKT model. It integrates the effect of learning gain and forgetting to model students' learning processes. The experiments show that LPKT is more effective in predicting students' future performance.

B. *EW of Learning Performance*

EW refers to analyzing students' learning data (e.g., learning background, learning behaviors, test scores) using certain criteria, and sending prompt signals to teachers and students based on the analysis results [26].

1) Traditional EW: Traditional research on EW typically relies on the analysis of students' scores and learning behavioral data for detecting students' potential learning risks in advance, such as dropout, failure, and delayed graduation [27].

Dropout prediction is designed to predict whether a student will quit a course before it ends, and the outcomes are usually binary [28]. For example, by collecting the data from students' logs and demographic information to analyze the correlation between dropouts from different courses, Feng et al. [17] proposed a context-aware feature interaction network to predict online students' dropout behavior in an e-learning system. Based on textual data (i.e., papers, emails, and tasks) collected from students' learning processes, Phan et al. [29] put forward a decision support framework based on logistic regression (LR) for identifying students at risk of dropout in early phases of a course.

Besides, there is a significant focus on identifying learning risks, such as failure and delayed graduation, in traditional research on EW. Hu et al. [18] employed a decision tree (DT) and AdaBoost algorithm to predict whether students pass the final exam by collecting their learning behavior data from online courses. Asif et al. [19] conducted a case study by collecting learning data from students at a college to identify those students who are at risk of delayed graduation in the early years of college. The results of the study indicate a correlation between students' final graduation performance and their learning performance in certain courses in the early years. Therefore, it is significant to monitor students' academic grades in real time for predicting their learning performance and achieving accurate EW. By analyzing students' historical data, Polyzou and Karypis [30] predicted whether students would pass the course exams before the start of the semester. Alshanqiti and Namoun [31] proposed a hybrid regression model to optimize the prediction accuracy of student academic performance. To evaluate whether students would pass the course exams, Osmanbegovic and Suljic [32] investigated the impact of students' socio-demographic variables, achieved results from high school and from the entrance exam, and attitudes toward studying.

2) EW via Cognitive Ability Analysis: In traditional research on EW, it is difficult to accurately model students' changing cognitive states by analyzing their grades or learning behavioral data. Therefore, the EW approach via cognitive ability analysis is gradually receiving attention from researchers.

There has been some progress in the existing research. Wang et al. [33] proposed an EW model from the perspective of knowledge concepts. They identified key knowledge concepts with Bayesian networks and sensitivity analysis, and the students' overall mastery on these knowledge concepts is predicted for EW based on their learning assessment data. Zou et al. [34] presented a new approach to analyze students' learning behavior data according to two novel features, i.e., knowledge points and question types. Based on the analysis results, the students' final performances are predicted by neural networks. To improve the accuracy of prediction, Alcaraz et al. [35] proposed that students' learning performance could be predicted by analyzing their weekly assignments. Besides, Okubo et al. [36] put forward an integrated review system via cognitive ability to improve students' learning. The system designs a review dashboard based on students' cognitive abilities, and then provides students with appropriate learning materials to support their reviews.

There have been efforts to explore the impact of students' cognitive abilities on their learning performance. However, the existing research usually ignored the fluctuations of students' cognitive abilities in the learning process, failing to accurately capture the correlation between students' learning performance and their changing cognitive abilities. Our approach devotes to address this issue.

3) Multilevel EW: Researchers have proposed some multilevel EW approaches based on the students' learning performance. The learning performance refers to students' situation related to learning process, including their completion of learning tasks, grades, and LE [37]. The LE is usually measured by the scores and the frequency of interactions [38].

The existing research on learning performance prediction usually applies the classification algorithms or network models to predict the students' final exam grades or learning performance (i.e., excellent, good, and poor) by analyzing their learning data [39]. Pandey and Taruna [40] proposed a multilevel classification model of learning performance based on a DT algorithm. In this model, the students' learning performance are classified into four levels, i.e., A, B, C, and F. Guo et al. [41] put forward a learning performance prediction model via the artificial neural networks to classify the students' learning performance into five levels, i.e., O, A, B, C, and D. Arnold and Pistilli [42] used traffic lights to provide EW to students based on their academic performance. Romero et al. [43] compared different data mining methods and techniques for classifying students based on their Moodle usage data and the final marks obtained in their courses. Jain and Solanki [44] analyzed the performance of four machine learning algorithms on educational dataset used for the early prediction of student performance, and used a multiclass classification in which students are divided into three classes, namely, poor-, average-, and good-performing students. Hua [45] analyzed the multilevel EW from a theoretical perspective by dividing the EW into five levels, i.e., severe, moderate, light, normal, and best. This mechanism provides an EW service to students by analyzing qualitative or quantitative data about the mind state, learning performance, and activity participation. However, there is no research on the algorithm implementation and experimental validation of multilevel personalized EW.

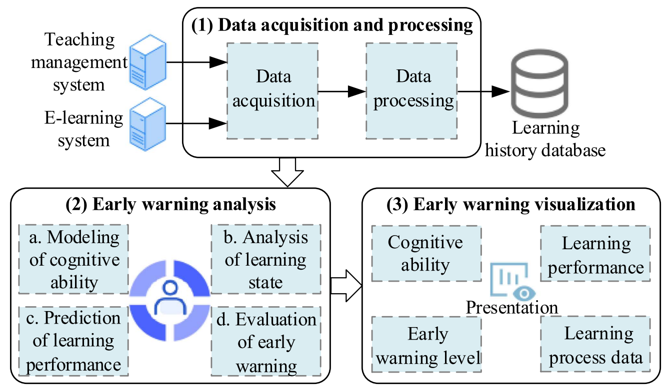


Fig. 1. Framework of the system.

The above approaches do not differentiate the learning characteristics of students at different levels, failing to provide personalized EW to all students. In this case, teachers cannot provide targeted guidance to those students at risk of learning.

III. Description of EW System

Aiming at the limitations of existing work and the practical teaching requirements, a personalized EW system has been developed, and applied to the practical teaching process for college students. An overview of the system's framework is shown in Fig. 1.

The core modules of the system are as follows.

1) *Data acquisition and processing:* In this module, the data related to students' learning process are collected from an e-learning system, whereas historical performance data are collected from a teaching management system. After the data are cleansed, the preprocessed data are stored in the learning history database.
2) *EW analysis:* By utilizing students' learning data from the learning history database, this module provides personalized EW analysis to students. The main functions are as follows.
   a) Dynamically model the students' skill proficiency required by exercises via cognitive ability modeling approach.
   b) Analyze the students' learning state features from four dimensions based on their skill proficiency.
   c) Predict the students' learning performance in the current phase based on their learning state features.
   d) Evaluate the students' EW level based on the predicted learning performance in the current phase.
3) *EW visualization:* To illustrate EW results in a visual manner, students' cognitive abilities, learning performance, EW level, and their learning process data are presented by graphs, e.g., radar charts, pie charts, bar charts, and

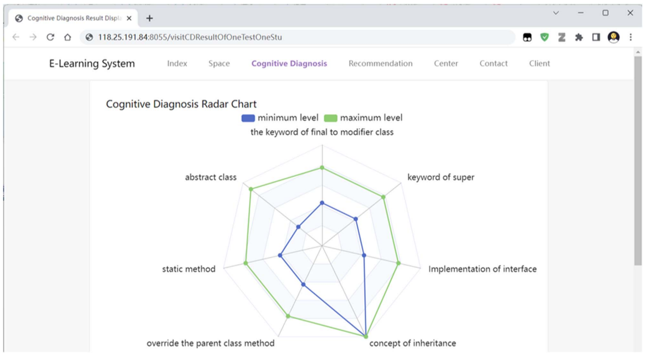


Fig. 2. Diagram of cognitive ability diagnosis.

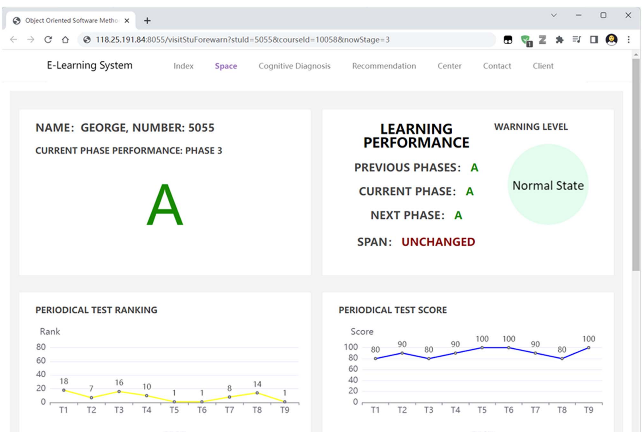


Fig. 3. Diagram of early warning of learning performance.

line graphs. Two examples about the implementation of cognitive ability diagnosis and EW are shown in Figs. 2 and 3.

IV. Research Design

To further explore the research design of the personalized EW system introduced in Section III, we will explain in detail the problem definition, the personalized EW framework of learning performance, and the personalized EW approach to learning performance in this section.

A. Problem Statement

The problem scenario of personalized EW is described as follows. In the learning process of a course, a teacher assigns learning tasks to students in the e-learning system according to the order of chapters or knowledge concepts. Then, the system records all the learning data generated by students as they complete the tasks. Considering that the learning period of a course is usually as long as 3–4 months, we could analyze the students' data relevant to this course's different learning phases corresponding to different learning tasks. Based on the analysis result, we could accurately model students' cognitive abilities and learning states for predicting their learning performance in the future. Finally, the differentiated EW results are provided dynamically and accurately to all students. To further explicit the problem, three definitions are given as follows.

Definition 1. Learning phase: A learning phase refers to a period in a course's learning process, and corresponds to a certain learning task. Usually, a learning phase lasts for 1–3 weeks. If a course reaches the learning phase #i, which is defined as the current phase, we name the learning phases from #1 to #(i-1) as previous phases and phase #(i+1) as the next phase.

Definition 2. Learning process data: It refers to the data related to students' participation in learning activities from learning phase #1 to phase #i of the course. For example, after George participates in an exam, the learning process data, including the time to complete the exam, scores, and response records. The data about response records reveal whether the student answered correctly for each exercise.

Definition 3. Historical performance data: It refers to the student's grades in the prerequisite courses or previous phases of the course.

Then, a multilevel personalized EW problem is defined as follows. Given student u's learning process data in phases #1–#i and historical performance data, how to predict u's learning performance in phase #(i+1) and evaluate his EW level. This problem is shown in Fig. 4.

From Fig. 4, George is in the learning phase #i of the course, and has taken a periodical test. Then, George's learning process data about this test are recorded by the e-learning system. We assume that this test includes j exercises, denoted as e_1, e_2, …, e_j. The system will record the related data of answering each exercise, e.g., answer record, score, and answer time. After George finishes all learning tasks for each phase, there are two types of tests for George to complete. To assess George's mastery of the knowledge concepts learned in the phase, a periodical test is organized for him to complete within a specified time. In addition, to assist George in consolidating and reinforcing his mastery of knowledge concepts, the self-tests would be automatically generated by the system based on the diagnosis results of his cognitive ability. George could decide whether to complete these tests after class according to his actual needs. Meanwhile, George's historical performance data are extracted from the teaching management system. As shown in Fig. 4, George's historical performance data include the grades of four prerequisite courses, i.e., c_1–c_4. In addition, it is necessary to use the Q-matrix for accurately assessing George's cognitive ability. The Q-matrix is annotated by experts to describe the correlation between exercises and knowledge concepts, and the element in Q-matrix denotes whether a knowledge concept is required to answer an exercise correctly. Specifically, the value of 1 represents that the knowledge concept is required to answer the exercise correctly, otherwise, the value is 0.

In summary, this article needs to address two key issues as follows: how to model student George's cognitive ability and learning state with the input data consisting of his learning process data produced in the current phase, historical performance data, and the Q-matrix? Based on this, how to predict George's learning performance in the next phase and finally give a personalized EW result that could accurately reflect the fluctuation of his learning situation?

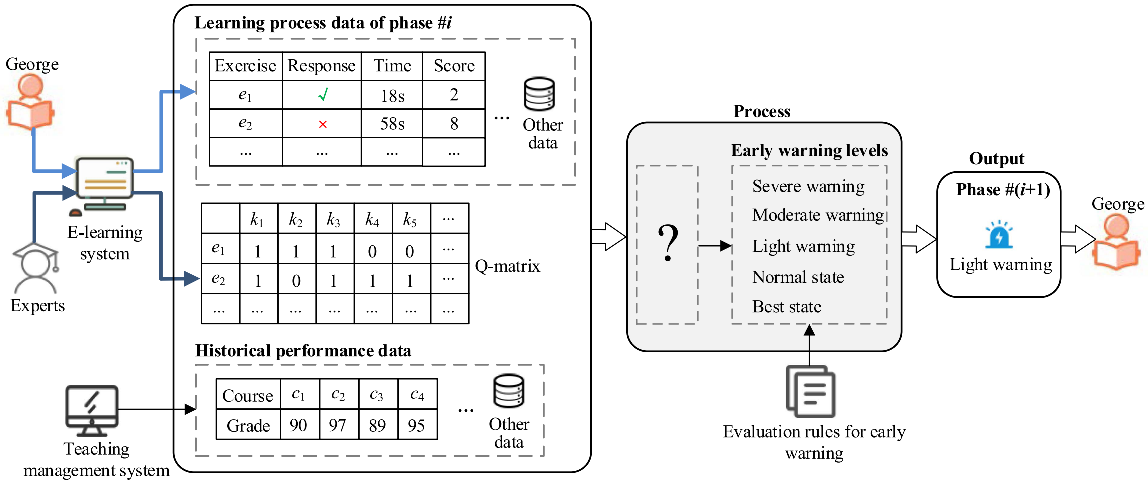


Fig. 4. Multilevel personalized early warning problem.

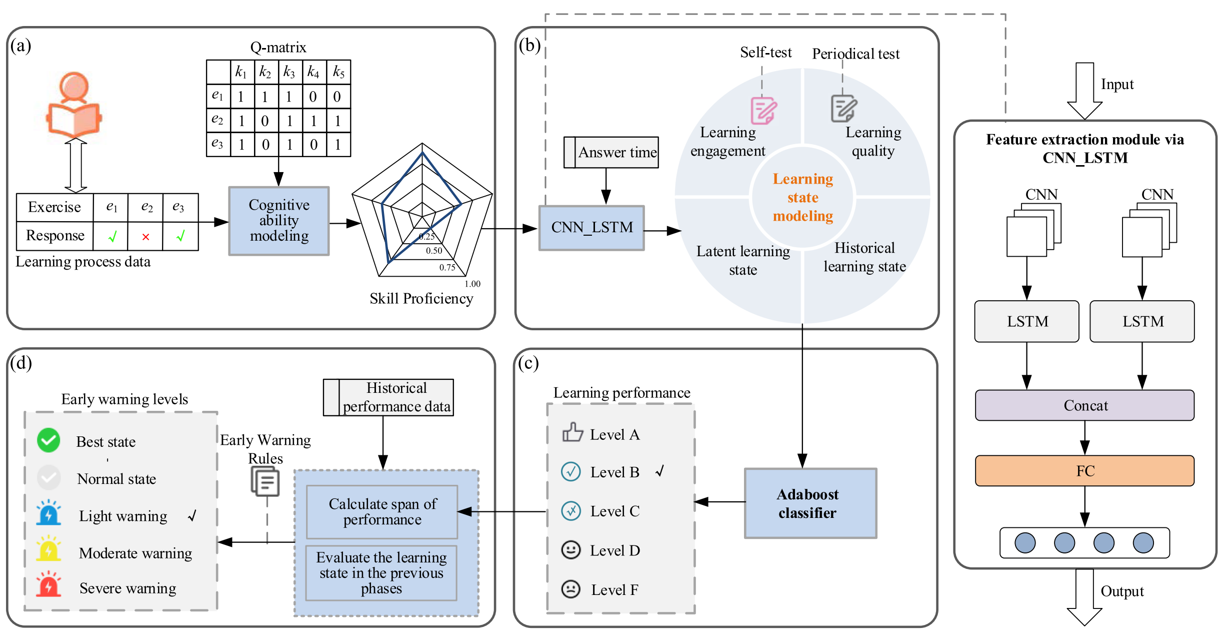


Fig. 5. Framework of PerLEW2LP. (a) Model cognitive ability. (b) Analyze learning state. (c) Predict learning performance. (d) Evaluate early warning level.

B. Personalized EW Framework

This article proposes a *p*erson*al*ized *e*arly *w*arning approach to *l*earning *p*erformance based on cognitive ability and learning state modeling, denoted as PerLEW2LP. By tracking the learning performance of all students, PerLEW2LP identifies those who might be at risk for their learning, providing precise EW services to all students. The framework of PerLEW2LP is shown in Fig. 5.

The process of PerLEW2LP consists of the following four steps.

1) Diagnose the students' mastery of knowledge concepts (i.e., skill proficiency) via a cognitive ability modeling approach.
2) Model comprehensively the learning state features by combining students' skill proficiency from four dimensions, i.e., LQ, LE, LLS, and HLS.
3) Employ the Adaboost algorithm to predict students' learning performance in the next phase based on their learning state features.
4) Provide the multilevel personalized EW to students by analyzing their performance in the current phase and HLS.

C. Personalized EW Approach

1) Model Student's Cognitive Ability: To assess students' LQ and predict their learning performance, it is essential to model their cognitive abilities. Cognitive diagnosis and knowledge

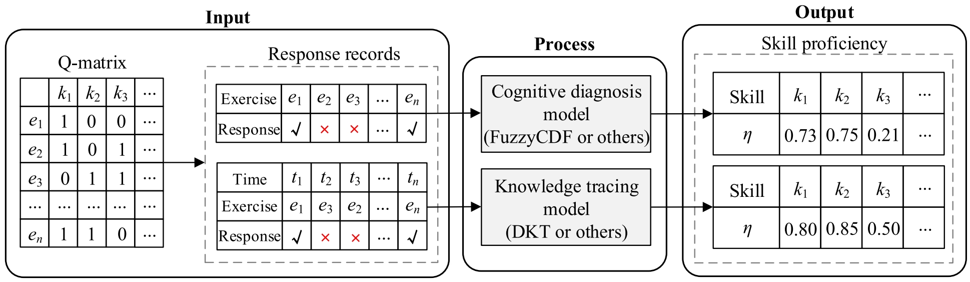


Fig. 6. Process of cognitive ability modeling.

tracing are currently the two most common approaches to modeling cognitive ability. They are suitable for different scenarios [21]. Cognitive diagnosis employs the limited data about the students' answer results of exercises to evaluate their cognitive states in the current phase. In contrast, knowledge tracing uses more data about students' answer sequences from the progressive learning process to analyze their cognitive abilities. The modeling flow of the two approaches is shown in Fig. 6.

a) Cognitive diagnosis: Cognitive diagnosis typically employs probability formulas to establish the static model of students' cognitive abilities from two aspects, i.e., exercises and knowledge concepts. Here, knowledge concepts are synonymous with skills. Suppose that George has finished n exercises in the learning process, and these exercises are related to k skills, his responses to n exercises are represented as an answer sequence $x_1, x_2, x_3, \ldots, x_n$. Then, by introducing Q-matrix, George's cognitive ability is modeled as a proficiency vector over k skills, denoted as $\eta = \{\eta_1, \eta_2, \ldots, \eta_k\}$.

However, different forms of cognitive diagnosis results would be generated by different approaches. Specifically, in the DINA [3] model, a student's mastery of skill i, denoted as η_i, is represented as either 0 or 1. In FuzzyCDF [4], η_i is represented as a fuzzy real number between 0 and 1. In the neutrosophic CD approach [24], η_i is represented as a fuzzy interval with three values, e.g., [0.50, 0.75, 0.46], indicating the degree of understanding, misunderstanding, and uncertainty.

By analyzing the results of students' responses to exercises over a continuous timeslot, knowledge tracing aims to predict their skill proficiency by applying sequential prediction techniques. Suppose Bob answers n exercises in the timeslot from t_1 to t_n of the learning process, his answer records on n exercises are denoted as an answer sequence $x_1, x_2, x_3, \ldots, x_n$. Then, by introducing Q-matrix, Bob's mastery of skills examined in n exercises is modeled to predict his probability of answering next exercise correctly. Since each exercise is usually related to a single skill in knowledge tracing, the probability of a student answering an exercise correctly corresponds to the student's skill proficiency.

b) Knowledge tracing: To model students' cognitive abilities more precisely, knowledge tracing considers not only students' response records but also the order of their answer sequences. In addition, with the development of deep neural networks, the input data used in knowledge tracing changed from 1-D data to multidimensional data. For example, considering the forgetting factors in a learning process, LPKT [10] model further analyzes students' answer time data to diagnose their cognitive abilities.

In PerLEW2LP, we would apply DKT [9] model to diagnose students' cognitive abilities. First, referring to Q-matrix, the response records of Bob would be converted into input vectors consisting of three types of data, i.e., exercise id, skill id, and response result, which denotes whether an exercise is answered correctly. The input vector is then fed into the DKT model, after the transformation of hidden states, the probability vector y of Bob's correctly answering these exercises is output. The probability of Bob's correctly answering an exercise denotes his skill proficiency in the DKT model. The main equations of DKT are shown as follows:

$$h_t = \tanh(W_{\mathrm{hx}} x_t + W_{\mathrm{hh}} h_{t-1} + b_{\mathrm{h}}) \quad (1)$$

$$y_t = \sigma(W_{\mathrm{yh}} h_t + b_{\mathrm{y}}) \quad (2)$$

where W_{hx} denotes the weight matrix of input data, W_{hh} denotes the weight matrix of the long short-term memory (LSTM) network, b_{h} and b_{y} denote the bias of the hidden state and output layer, respectively, σ denotes the sigmoid function, and h_t denotes the student's hidden cognitive state at timeslot #t.

2) Analyze Student's LS: After assessing students' cognitive abilities, we further model their learning state features from four aspects, i.e., LQ, LE, LLS, and HLS. Take student u as an example, u's learning state is described as a vector $f_u = \{q_u, e_u, l_u, s_u\}$, where q_u, e_u, l_u, and s_u are four vectors, representing u's LQ, LE, LLS, and HLS, respectively. Next, we will introduce each aspect of the learning state features in detail.

a) Learning quality: It is evident that students' LQ is easily influenced by their learning states. Therefore, we reversely model students' learning states by analyzing their LQ in skills [46]. Since a periodical test is uniformly organized by a teacher for all the students, the score and skill proficiency related to this test directly reflect students' actual learning effects. Based on this, we select three metrics to measure students' LQ, i.e., completion rate, average accuracy, and average skill proficiency of the periodical test, denoted as z^{c}, z^{α}, and z^{η}. The three metrics are given by

$$z^{\mathrm{c}} = n^{\mathrm{z}} / n^{\mathrm{a}} \quad (3)$$

$$z^{\alpha} = n_{\mathrm{z}}^{\mathrm{y}} / n_{\mathrm{z}}^{\mathrm{q}} \quad (4)$$

$$z^{\eta} = \sum_{k \in K_{\mathrm{z}}} \eta_{\mathrm{z}}^{k} \Big/ |K_{\mathrm{z}}| \tag{5}$$

where n^{z} is the number of periodical tests completed by one student, n^{a} denotes the total number of periodical tests, $n_{\mathrm{z}}^{\mathrm{y}}$ is the number of exercises correctly answered by the student in the completed periodical test, $n_{\mathrm{z}}^{\mathrm{q}}$ is the total number of exercises completed by the student, η_{z}^{k} denotes the proficiency of the student on skill *k*, and K_{z} refers to the set of skills involved in periodical tests completed by this student, $k \in K_{\mathrm{z}}$.

Thus, the *u*'s LQ is defined by $q_u = \{z_u^{\mathrm{c}}, z_u^{\alpha}, z_u^{\eta}\}$.

b) Learning engagement: LE is a crucial factor for measuring students' learning outcomes [47]. The self-tests are automatically generated by the e-learning system according to students' skill proficiency, which aims to assist them in consolidating and reinforcing poorly mastered skills. Students could decide whether to complete these tests after class according to their actual needs. In this case, the completion of these tests could serve as an indicator of their LE. Based on this, we select three metrics to measure students' LE, including times of completing self-tests, average accuracy rate of self-tests, and average skill proficiency of self-tests. They are denoted as κ^{n}, κ^{α}, and κ^{η}, respectively. κ^{α} and κ^{η} are calculated by

$$\kappa^{\alpha} = n_{\kappa}^{\mathrm{y}} / n_{\kappa}^{\mathrm{q}} \tag{6}$$

$$\kappa^{\eta} = \sum_{k \in K} \eta_{\kappa}^{k} \Big/ |K_{\kappa}| \tag{7}$$

where n_{κ}^{y} is the number of exercises answered correctly by one student in the self-tests, n_{κ}^{q} is the total number of exercises completed by the student, η_{κ}^{k} is the student's proficiency on skill *k*, and K_{κ} is the set of skills involved in self-tests, $k \in K_{\kappa}$.

Thus, *u*'s LE is given by $e_u = \{\kappa_u^{\mathrm{n}}, \kappa_u^{\alpha}, \kappa_u^{\eta}\}$.

c) Latent learning state: The LLS of students influences their learning performance. It is challenging to observe directly. The skill proficiency and answer time, which are generated by students in the learning process, could reflect their academic effectiveness and answering patterns. In this case, we would explore students' LLS from the two types of data with a joint network combining a convolutional neural network (CNN) and LSTM network.

The specific details of the structure are given as follows.

1) *Input layer:* Suppose that student *u* completes J periodical tests, each test examining M skills. The answer time matrix of *u* is defined as $X_u = [a_{j,m}^{u}]_{\mathrm{J}\times\mathrm{M}}$, where $a_{j,m}^{u}$ denotes the answer time that *u* has spent on exercise *m* in test *j*. The matrix of skill proficiency is represented as $Y_u = [\eta_{j,k}^{u}]_{\mathrm{J}\times\mathrm{K}}$, where $\eta_{j,k}^{u}$ denotes *u*'s proficiency on skill *k* involved in test *j*. The missing values in X_u and Y_u are filled with zero.
2) *CNN layer:* The matrix of answer time and skill proficiency of student *u* are fed into CNN, respectively. Then, a nonlinear mapping is performed on this matrix via the rectified linear unit (ReLU) activation function, and the matrices c_u^{X} and c_u^{Y} are generated. Taking X_u as an example, the convolution layer is defined as

$$c_u^{X} = \sigma(W_c X_u + b_c) \tag{8}$$

where W_{c} is the weight matrix of the convolutional layer. We specifically choose a convolutional kernel of $1\times N$ to extract horizontal features of the data. The value of *N* is set as $\lfloor \mathrm{M}/2 \rfloor$, and M denotes the size of a column in matrix X_u, b_{c} denotes the bias of CNN, and σ denotes the ReLU activation function.
3) *LSTM layer:* To further explore the LLS of student *u*, the outputs c_u^{X} and c_u^{Y} from the convolutional layer are input to the LSTM to extract the hidden state features of two types of data. The hidden state features are denoted as h_u^{X} and h_u^{Y}, respectively. Taking c_u^{X} as an example, c_u^{X} is divided into J rowwise vectors based on the number of periodical tests, and these vectors are represented as $c_u^{\mathrm{X}} = \{c_u^{\mathrm{X}_1}, c_u^{\mathrm{X}_2}, \ldots, c_u^{\mathrm{X}_{\mathrm{J}}}\}$, where $c_u^{\mathrm{X}_j}$ denotes the vector of answer time for *u* on periodical test *j*. The hidden state feature is obtained by

$$h_u^{\mathrm{X}_j} = \sigma(W_{\mathrm{L}}^{\mathrm{c}} c_u^{\mathrm{X}_j} + W_{\mathrm{L}}^{\mathrm{h}} h_u^{\mathrm{X}_{j-1}} + b_{\mathrm{L}}) \quad \forall j{=}1, 2, \ldots, \mathrm{J} \tag{9}$$

where $W_{\mathrm{L}}^{\mathrm{c}}$ and $W_{\mathrm{L}}^{\mathrm{h}}$ denote the weight matrices of the LSTM layer, b_{L} denotes the bias of the LSTM layer, $h_u^{\mathrm{X}_{j-1}}$ denotes the hidden state feature obtained by inputting the (*j*-1)th vector, and σ denotes the ReLU activation function.
4) *Output layer:* The hidden state features $h_u^{\mathrm{X}_{\mathrm{J}}}$ and $h_u^{\mathrm{Y}_{\mathrm{J}}}$ are inputted into the full connection (FC) layer. Then, *u*'s LLS is obtained by

$$l_u = FC(h_u^{\mathrm{X}_{\mathrm{J}}} || h_u^{\mathrm{Y}_{\mathrm{J}}}) \tag{10}$$

where $||$ denotes the concatenation operation.

d) Historical learning state: The students' academic performance in the previous phases could not only serve as a significant predictor for their learning outcomes in the next phase, but also reflect the students' learning state in previous phases [48]. Since the difficulty is different for different courses, the assessment standards of courses are also different. Accordingly, we classify the students' HLS into five levels based on their scores and rankings, denoted as A, B, C, D, and F. The five levels represent excellent, good, moderate, pass, and failure, respectively. To evaluate students' HLS, we first transform their scores and rankings into a five-level format. In this article, the scores above 95 are viewed as level A, the scores between 85 and 95 as B, the scores between 75 and 85 as C, the scores between 65 and 75 as D, and the scores below 65 as F. Similarly, the rankings are transformed into five levels. The students ranked in the top 15% get level A, those from 15% to 30% get B, those from 30% to 50% get C, those from 50% to 80% get D, and those within the bottom 20% get F.

The HLS is evaluated according to the learning states in the current and previous phases, denoted as s^{L} and s^{O}, respectively. The evaluation vector of student *u* on the HLS is defined as $s_u = \{s_u^{\mathrm{L}}, s_u^{\mathrm{O}}\}$. To calculate s_u^{L} and s_u^{O}, the different data need to be used in the following two cases.

1) In the first learning phase of a course, the students' historical performance data of all prerequisite courses are extracted from the teaching management system. The

student *u*'s grades from the most recently completed prerequisite courses are used to calculate s_u^{L}. The *u*'s average grade from all the prerequisite courses is used to calculate s_u^{O}.

2) In the second or subsequent phase of the course, the students' historical performance data in the previous phases are extracted from the e-learning system. The *u*'s grades in the last phase are used to calculate s_u^{L}. The *u*'s average grade in all previous phases is used to calculate s_u^{O}.

Based on the above data, both s_u^{L} and s_u^{O} are calculated in the same way. Take the calculation of s_u^{L} as an example

$$s_u^{\mathrm{L}} = \left\lfloor g_u^{\mathrm{I}} * w^{\mathrm{g}} + r_u^{\mathrm{I}} * (1 - w^{\mathrm{g}}) \right\rfloor \tag{11}$$

where g_u^{I} and r_u^{I} denote the level of *u*'s score and ranking, both represented by integers ranging from 1 to 5 corresponding to levels from A to F. For example, $g_u^{\mathrm{I}} = 1$ denotes that *u* scored in the range of 95–100, $r_u^{\mathrm{I}} = 2$ denotes that *u*'s ranking is in the range of the top 15%–30%, and w^{g} is the weight coefficient and is set as 0.5 by default. s_u^{L} is an integer ranging from 1 to 5, representing the levels of *u*'s performance. The integers 1–5 correspond to levels A, B, C, D, and F, respectively.

3) Predict Student's Learning Performance: At the end of the learning tasks released by the teacher at the current phase, students' learning state features, including the LQ, LE, LLS, and HLS, are fed into the classifier as input data to predict their learning performance in the next phase.

In multiclassification tasks, the problem of imbalanced sample sizes is easily lead to misclassification, consequently reducing the accuracy of the results. There are some approaches proposed in existing research for dealing with the multiclassification problems with unbalanced data, e.g., the under sampling approach, oversampling approach, or ensemble learning algorithm [49]. Based on careful experimental comparisons, we select the Adaboost algorithm with the best performance to assess students' learning performance. Details about experimental results are presented in Section V.

As one of ensemble learning algorithms, Adaboost improves the overall classification effect by solving the problem that minority samples are too difficult to classify. For each iteration, Adaboost first calculates the weights of each sample in the training data based on the classification results and overall accuracy of the last iteration. In this case, the base classifier is trained by adjusting data distribution and sample weights [49]. Then, the modified dataset with updated weights is subsequently fed into the base classifier for training. Finally, the classifiers obtained from each iteration are combined into an ultimate decision classifier. This classifier performs excellently in handling imbalanced data in multiclassification tasks, and is widely adopted in practical applications.

After inputting student *u*'s learning state (f_u) into the Adaboost classifier, the *u*'s predicted learning performance in the next phase is generated, denoted as v_u^{p}. v_u^{p} is an integer between 1 and 5, representing the five levels of learning performance, respectively, i.e., A, B, C, D, and F.

4) Evaluate Student's EW Level: Referring to the classification criteria from Taiwan's EW system [45], we classify the EW results into five levels, i.e., severe warning, moderate warning, light warning, normal state, and best state. Inspired by Hua [45], we design the five-level EW evaluation rules for accurately identifying the changes of students' learning performance and providing personalized EW to students. The EW rules are given in Table I.

TABLE I
EARLY WARNING EVALUATION RULES

Learning performance in the next phase	Learning state in previous phases	Span	Early warning level
A	A, B, C, D, F	ε=0	Normal state
		ε<0	Best state
B	A, B, C, D, F	ε=1	Light warning
		ε=0	Normal state
		ε<0	Best state
C, D	A, B	ε≥1	Moderate warning
		ε=0	Light warning
		ε<0	Normal state
	C, D, F	ε≥2	Moderate warning
		ε=1	Light warning
		ε=0	Normal state
		ε<0	Best state
F	A, B, C, D	——	Severe warning
	F		Moderate warning

Specifically, three factors are involved in the rule, i.e., student *u*'s predicted learning performance in the next phase, learning states in previous phases, and the span between learning performance in the current and next phases. The span between *u*'s learning performance is given by

$$\varepsilon_u = v_u^{\mathrm{p}} - s_u^{\mathrm{L}} \tag{12}$$

where v_u^{p} denotes *u*'s learning performance in the next phase, and s_u^{L} denotes *u*'s learning performance in the current phase, which also refers to *u*'s learning state in the current phase. $\varepsilon<0$ denotes an improvement in students' learning performance. $\varepsilon = 0$ denotes stability in students' learning performance. $\varepsilon = 1$ denotes a slight decline in students' learning performance. $\varepsilon\geq 2$ denotes a serious decline in students' learning performance.

In summary, the *u*'s EW level is evaluated based on *u*'s learning performance in the next phase, learning state in the previous phases, and span between learning performance. For example, suppose that two students, u_1 and u_2, are taking the same course. Based on the PerLEW2LP, the analysis process of their EW levels is shown as follows. 1) If u_1 gets level A for $s_{u_1}^{\mathrm{O}}$, B for $s_{u_1}^{\mathrm{L}}$, and C for $v_{u_1}^{\mathrm{p}}$, thenε_{u_1} is equal to 1. It is evident that u_1's learning state has been consistently declining. Based on the warning rule, u_1 will receive a moderate warning notification. 2) If u_2 gets level D for $s_{u_2}^{\mathrm{O}}$, D for $s_{u_2}^{\mathrm{L}}$, and F for $v_{u_2}^{\mathrm{p}}$, then ε_{u_2} is equal to 1. Since the u_2's learning state has declined to level F, a severe warning notification will send to u_2 according to the warning rules.

V. EXPERIMENTS

Experiments are implemented by Python 3.6 run in a Linux server with a 2.3 GHz Inter Xeon CPU. The CNN_LSTM model and Adaboost algorithm are implemented by PyTorch and Sklearn library, respectively.

TABLE II
DATASET INFORMATION

Type	Quantity	
	Course #1	Course #2
Students	130	130
Chapters	12	11
Periodical tests	20	15
Questions	500	271
Skills	253	188
Answer records	1950	1795

A. Experiment Setup

1) Datasets: To verify the effectiveness of PerLEW2LP, we conduct experiments on a real dataset collected from the e-learning system of Hunan Normal University. The dataset consists of learning data of students majoring in software engineering for two courses, i.e., Java Programming and Principles of Database Systems, denoted as Course #1 and Course #2, respectively. Both courses are instructed by the same teacher across four consecutive semesters, i.e., fall 2021, spring 2022, fall 2022, and spring 2023. Detailed learning data for each student are recorded in the database. Data are collected by class at the end of the course. The dataset information is as given in Table II. The dataset is already available online via a URL.[2]

2) Parameter Settings: For all approaches, the common hyperparameters are listed as follows: the convolution unit is set as 1-D, the batch size is fixed as 20, the epoch is set as 100, the loss function is the cross-entropy error, the Adam optimizer is used to update all parameters, and the learning rate is fixed as 0.001. To prevent the model from overfitting, we added dropout regularization with the probability of 0.2.

3) Evaluation Metrics: We select *precision*, *recall,* and $F1$ as evaluation metrics, and the parameter is weighted. These metrics are defined by

$$precision = \frac{TP}{TP + FP} \tag{13}$$

$$recall = \frac{TP}{TP + FN} \tag{14}$$

$$F1 = \sum_{i=1}^{C} \frac{a_i}{n} * \frac{2 * precision_i * recall_i}{precision_i + recall_i} \tag{15}$$

where TP and TN denote the number of positive and negative samples that are correctly classified, respectively, FP and FN denote the number of misclassified negative and positive samples, respectively, C is the number of categories, n is the number of samples, anda_i is the number of samples in category i.

To further evaluate the accuracy of PerLEW2LP, we select the *Kappa* coefficient [50] as the evaluation metric. It is widely used for consistency testing in statistics, which could also be applied as an indicator for evaluating classification accuracy. The *Kappa* coefficient is given by

$$K = (\omega_{\mathrm{o}} - \omega_{\mathrm{e}})/(1 - \omega) \tag{16}$$

[2][Online]. Available: https://github.com/ZW-301/early-learning-warning.git

$$\omega_e = \sum_{i=1}^{\mathrm{C}} a_i * b_i \Big/ n^2 \tag{17}$$

where ω_{o} denotes the precision, C denotes the total number of categories, n denotes the total number of samples, a_i denotes the actual number of samples in category i, and b_i denotes the predicted number of samples in category i.

B. Comparative Experiments

Our approach differs from existing approaches. Consequently, we conducted comparative experiments based on different algorithms to achieve the best results. The details of PerLEW2LP are as follows.

1) The DKT model is applied to model students' cognitive abilities.
2) The CNN_LSTM structure is employed to mine students' LLS features.
3) The Adaboost algorithm is used to predict students' learning performance. To verify the effectiveness of PerLEW2LP, we collect the students' final exam grades on two courses as true labels for calculating their learning performance and EW level. In experiments, a fivefold cross-validation is conducted, and the dataset is divided into two subdatasets, i.e., Course #1 and Course #2.

1) Comparison of Different Neural Network Structures: In PerLEW2LP, the CNN_LSTM structure is used to extract the LLS features of students. To verify the effectiveness of EW approaches to learning performance based on CNN_LSTM structure, we replace the CNN_LSTM structure with two other neural network structures, i.e., the joint network combining a CNN and RNN, and the joint network combining a CNN and gated recurrent unit (GRU), denoted as CNN_RNN and CNN_GRU, respectively. Besides, other modules remain the same, with Adaboost as the classification algorithm and the DKT model as the cognitive ability modeling approach. The two variant approaches are denoted as EW_CNN_RNN and EW_CNN_GRU. The experimental results are given in Table III.

The experimental results show that PerLEW2LP outperforms EW_CNN_RNN and EW_CNN_GRU in prediction accuracy and stability. The results are analyzed as follows: 1) due to the problem of gradient disappearance and gradient explosion, RNN has trouble dealing with long time sequences. Therefore, EW_CNN_RNN performs worse than PerLEW2LP in all evaluation metrics and 2) since the GRU structure has fewer parameters, it is typically faster than LSTM in training speed. However, because of the limited data scale, there is no significant difference between PerLEW2LP and EW_CNN_GRU in training speed. Moreover, PerLEW2LP even performs better in prediction accuracy and stability.

2) Comparison of Different Classification Algorithms: To verify the effectiveness of EW approaches to learning performance based on the Adaboost algorithm, similar to [51], we select seven baseline approaches for comparison, including linear support vector machine (SVM) denoted as LSVM, SVM with radial basis function (RBF) kernel denoted as SVM_RBF, LR, DT, random forest (RF), gradient boosting decision tree (GBDT),

TABLE III
COMPARISON OF NEURAL NETWORK STRUCTURES

Approach	Course #1				Course #2			
	Precision	*Recall*	*F1*	*Kappa*	*Precision*	*Recall*	*F1*	*Kappa*
EW_CNN_RNN	0.812	0.625	0.694	0.505	0.862	0.690	0.758	0.562
EW_CNN_GRU	0.845	0.683	0.727	0.542	0.880	0.702	0.776	0.576
PerLEW2LP	**0.875**	**0.720**	**0.784**	**0.572**	**0.895**	**0.745**	**0.780**	**0.595**

TABLE IV
COMPARISON OF CLASSIFICATION ALGORITHMS

Approach	Course #1				Course #2			
	Precision	*Recall*	*F1*	*Kappa*	*Precision*	*Recall*	*F1*	*Kappa*
EW_LSVM	0.725	0.580	0.650	0.445	0.750	0.585	0.687	0.468
EW_SVM_RBF	0.760	0.605	0.652	0.432	0.762	0.605	0.655	0.423
EW_LR	0.734	0.63	0.670	0.376	0.715	0.602	0.692	0.302
EW_DT	0.690	0.704	0.700	0.495	0.766	0.704	0.690	0.470
EW_RF	0.802	0.683	0.728	0.508	0.833	0.689	0.725	0.554
EW_GBDT	0.855	0.710	0.765	0.542	0.865	0.722	0.766	0.525
EW_DANN	0.651	0.548	0.582	0.376	0.667	0.580	0.601	0.359
PerLEW2LP	**0.875**	**0.720**	**0.784**	**0.572**	**0.895**	**0.745**	**0.780**	**0.595**

TABLE V
COMPARISON OF COGNITIVE ABILITY MODELING

Approach	Course #1				Course #2			
	Precision	*Recall*	*F1*	*Kappa*	*Precision*	*Recall*	*F1*	*Kappa*
EW_DINA	0.842	0.690	0.725	0.520	0.878	0.702	0.743	0.576
PerLEW2LP	**0.875**	**0.720**	**0.784**	**0.572**	**0.895**	**0.745**	**0.780**	**0.595**

and deep artificial neural network (DANN). Specifically, we replace the Adaboost algorithm with each of the above six different classification algorithms. Meanwhile, other modules use the same process workflow as PerLEW2LP, utilizing the CNN_LSTM as neural network structure and employing the DKT model for cognitive ability modeling. These variant approaches are denoted as EW_LSVM, EW_SVM_RBF, EW_LR, EW_DT, EW_RF, EW_GBDT, and EW_DANN, respectively.

As given in Table IV, the experimental results demonstrate that PerLEW2LP performs better than other EW approaches. Specifically, the Kappa values of PerLEW2LP in the two courses are 0.572 and 0.595, respectively, which indicate better stability. The reason is that the diversity of learning resources and tasks in an e-learning system leads to the various learning states of students. However, ensemble learning algorithms, such as Adaboost, RF, and GBDT, have advantages in identifying complex learning states. Thus, EW_RF, EW_GBDT, and PerLEW2LP are more effective compared with EW_SVM_LK, EW_SVM_RBF, EW_LR, EW_DT, and EW_DANN. Moreover, since the data are unbalanced, the experimental results display higher precision but lower recall. However, EW_DT has reduced its precision as a result of overfitting.

3) Comparison of Different Cognitive Ability Models: Since PerLEW2LP uses DKT to model students' cognitive abilities, we conduct a comparison experiment to verify the effectiveness of the EW approach based on the DKT model. Specifically, we replace the DKT model with the DINA model and the new approach is denoted as EW_DINA. In addition, the remaining modules of EW_DINA are kept the same as PerLEW2LP. Adaboost is used as the classification algorithm and CNN_LSTM as the neural network structure. The results are presented in Table V.

From Table V, PerLEW2LP performs better than EW_DINA in all metrics. The dataset used in the experiments includes two courses with a learning period of 2–3 months. In the long-term learning period, the DKT model, as a dynamic cognitive modeling approach, could better predict students' knowledge acquisition process and timely identify their changing cognitive abilities, which is suitable for the dataset. Therefore, PerLEW2LP could more accurately predict changes of students' learning performance, and provide better EW services to students.

4) Error Analysis of EW Approaches: To further evaluate the accuracy of PerLEW2LP, we employ a confusion matrix, which is widely used to evaluate the accuracy of classification models, to analyze the errors of all the afore-mentioned PerLEW2LP variant approaches. The confusion matrix of PerLEW2LP on Course #1 and Course #2 is given in Table VI.

The classification results of students' EW include five levels, i.e., the best state, normal state, light warning, moderate warning, and severe warning, denoted as B, N, L, M, and S, respectively. From Table VI, the superscripts P and A denote the predicted result and actual result, respectively. For example, B^P and B^A correspond to the predicted and actual results of the best state, respectively. Our experiments show that the actual results of EW for 130 students in Course #1 are as follows: 55 students should receive the best state notification, 33 a normal state notification, 21 a light warning one, 16 a moderate one, and 5 a severe one. According to Table VI, the predicted results of PerLEW2LP for Course #1 are as follows: the best state is sent to 46 students, normal one to 21, light warning to 14, moderate one to 11, and severe one to 2.

TABLE VI
CONFUSION MATRIX OF PERLEW2LP

Category	Course #1					Course #2				
	B^P	N^P	L^P	M^P	S^P	B^P	N^P	L^P	M^P	S^P
B^A	46	7	2	0	0	13	5	1	0	0
N^A	1	21	5	6	0	6	54	8	2	0
L^A	1	2	14	3	1	1	3	13	2	0
M^A	1	0	2	11	2	1	0	1	8	2
S^A	0	0	0	3	2	1	0	1	3	6

TABLE VII
ERROR ANALYSIS

Module	Approach	Course #1		Course #2	
		Type I errors	Type II errors	Type I errors	Type II errors
Neural network	EW_CNN_RNN	0.183	0.133	0.184	0.146
	EW_CNN_GRU	0.123	0.162	0.177	0.138
	PerLEW2LP	**0.076**	**0.200**	**0.131**	**0.153**
Classification algorithm	EW_LSVM	0.208	0.192	0.231	0.154
	EW_SVM_RBF	0.223	0.192	0.200	0.169
	EW_SVM_LR	0.223	0.185	0.292	0.154
	EW_DT	0.138	0.185	0.185	0.162
	EW_RF	0.146	0.162	0.169	0.162
	EW_GBDT	0.138	0.162	0.154	0.146
	EW_DANN	0.308	0.277	0.285	0.285
	PerLEW2LP	**0.076**	**0.200**	**0.131**	**0.153**
Cognitive ability modeling	EW_DINA	0.138	0.162	0.169	0.162
	PerLEW2LP	**0.076**	**0.200**	**0.131**	**0.153**

TABLE VIII
ABLATION EXPERIMENTS

Approach	Course #1				Course #2			
	Precision	*Recall*	*F1*	*Kappa*	*Precision*	*Recall*	*F1*	*Kappa*
PerLEW2LP-LQ	0.800	0.603	0.647	0.472	0.802	0.641	0.677	0.503
PerLEW2LP-LLS	0.841	0.715	0.752	0.550	0.868	0.729	0.762	0.548
PerLEW2LP-LE	0.818	0.690	0.737	0.538	0.877	0.706	0.748	0.580
PerLEW2LP-HLS	0.795	0.656	0.714	0.522	0.843	0.680	0.722	0.541
PerLEW2LP	**0.875**	**0.720**	**0.784**	**0.572**	**0.895**	**0.745**	**0.780**	**0.595**

Based on the confusion matrix, we could calculate total errors consisting of two types of errors, i.e., Types I and II. A Type I error means that a student's actual result is worse than the predicted one. In contrast, a Type II error means that a student's actual result is better than the predicted one. Obviously, in an EW system, the negative effect of mistaking poor students for excellent ones is much more serious than mistaking excellent students for poor ones. The reason is that the poor students urgently need more teachers' attention and timely instruction. Therefore, the fewer Type I errors in the prediction results obtained by an EW approach, the better its prediction effect.

The two types of errors are analyzed for all variants of PerLEW2LP in the two courses, and results are given by Table VII. The mean value of PerLEW2LP is 10.35% for Type I errors and 17.65% for II. From the experimental results, PerLEW2LP has the best EW performance since it has the minimum Type I errors and the minimum total number of errors.

5) Ablation Experiments: To analyze the validity and importance of different types of learning state features on prediction results, we conducted ablation experiments. Four learning state features including LQ, LE, LLS, and HLS, are abandoned from PerLEW2LP, respectively. The four variant approaches are denoted as PerLEW2LP-LQ, PerLEW2LP-LE, PerLEW2LP-LLS, and PerLEW2LP-HLS, respectively.

From Table VIII, it is clear that the impact of four features on students' learning performance is quite different. Among them, LQ has the most significant impact on the EW effect compared with other features. After removing the LQ, the recall values of Course #1 and Course #2 decrease by 0.117 and 0.104, F1 decreases by 0.137 and 0.103, and Kappa coefficient decreases by 0.100 and 0.092, respectively. The accuracy and stability of PerLEW2LP decrease significantly. It is shown that students' learning effectiveness is significantly influenced by their LQ.

VI. CONCLUSION

To meet the EW needs of college students at different levels, we propose a personalized EW approach to learning performance via cognitive ability and learning state modeling. First, we employ a cognitive ability modeling approach to track students' changing cognitive abilities. Then, based on students' learning processes data and historical performance data, we comprehensively model students' learning states from four dimensions, which provide differentiated learning performance prediction and personalized EW services to all students. Finally, the extensive experiments on real-world datasets demonstrate the accuracy and effectiveness of our approach.

This article tries to provide multilevel and differentiated EW results to all college students. Nevertheless, the proposed personalized EW approach has a lot of room for improvement. In the future, we will first explore to evaluate more comprehensively the students' learning performance by collecting richer and more diverse learning process data (e.g., the number and duration of learning resources learned) from the e-learning system. Second, the data used in the experiments show a significant imbalance among students at different EW levels. The imbalance data may reduce the accuracy of our approach in predicting these levels. We will apply the data-level algorithms, e.g., oversampling and undersampling techniques, to improve the prediction performance. Finally, to further verify the generalizability and feasibility of our approach in different educational scenarios, we will promote the widespread application of the EW system. We will also further explore effective early warning approaches of teaching oriented to the achievement of curriculum objectives.

References

[1] P. D. Nichols and K. Joldersma, "Cognitive diagnostic assessment for education: Theory and applications," *J. Educ. Meas.*, vol. 45, pp. 407–411, 2008.

[2] X. Fan, "Item response theory and classical test theory: An empirical comparison of their item/person statistics," *Educ. Psychol. Meas.*, vol. 58, no. 3, pp. 357–381, 1998.

[3] J. De La Torre, "DINA model and parameter estimation: A didactic," *J. Educ. Behav. Statist.*, vol. 34, no. 1, pp. 115–130, 2009.

[4] Q. Liu et al., "Fuzzy cognitive diagnosis for modelling examinee performance," *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 4, pp. 1–26, 2018.

[5] H. Ma, Z. Huang, H. Zhu, W. Tang, H. Zhang, and K. Li, "Predicting examinee performance based on a fuzzy cloud cognitive diagnosis framework in e-learning environment," *Soft Comput.*, vol. 27, no. 24, pp. 18949–18969, 2023.

[6] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," *User Model. User-Adapted Interaction*, vol. 4, no. 4, pp. 253–278, 1994.

[7] T. Käser, S. Klingler, A. G. Schwing, and M. Gross, "Dynamic Bayesian networks for student modeling," *IEEE Trans. Learn. Technol.*, vol. 10, no. 4, pp. 450–462, Oct.–Dec. 2017.

[8] H. Cen, K. Koedinger, and B. Junker, "Learning factors analysis - a general method for cognitive model evaluation and improvement," in *Proc. 6th Int. Conf. Intell. Tuoring Syst.*, 2006, pp. 164–175.

[9] C. Piech et al., "Deep knowledge tracing," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 505–513.

[10] S. Shen et al., "Learning process-consistent knowledge tracing," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discov. Data Mining*, 2021, pp. 1452–1460.

[11] J. E. Raffaghelli, M. E. Rodríguez, A.-E. Guerrero-Roldán, and D. Baneres, "Applying the UTAUT model to explain the students' acceptance of an early warning system in higher education," *Comput. Educ.*, vol. 182, 2022, Art. no. 104468.

[12] H. Zeineddine, U. Braendle, and A. Farah, "Enhancing prediction of student success: Automated machine learning approach," *Comput. Elect. Eng.*, vol. 89, 2021, Art. no. 106903.

[13] S. Batool, J. Rashid, M. W. Nisar, J. Kim, H.-Y. Kwon, and A. Hussain, "Educational data mining to predict students' academic performance: A survey study," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 905–971, 2023.

[14] L. Zhang, J. D. Basham, and S. Yang, "Understanding the implementation of personalized learning: A research synthesis," *Educ. Res. Rev.*, vol. 31, 2020, Art. no. 100339.

[15] A. Bayazit, N. Apaydin, and I. Gonullu, "Predicting at-risk students in an online flipped anatomy course using learning analytics," *Educ. Sci.*, vol. 12, no. 9, 2022, Art. no. 581.

[16] V. Christou et al., "Performance and early drop prediction for higher education students using machine learning," *Expert Syst. Appl.*, vol. 225, 2023, Art. no. 120079.

[17] W. Feng, J. Tang, and T. X. Liu, "Understanding dropouts in MOOCs," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 517–524.

[18] Y.-H. Hu, C.-L. Lo, and S.-P. Shih, "Developing early warning systems to predict students' online learning performance," *Comput. Hum. Behav.*, vol. 36, pp. 469–478, 2014.

[19] R. Asif, A. Merceron, and M. K. Pathan, "Predicting student academic performance at degree level: A case study," *Int. J. Intell. Syst. Appl.*, vol. 7, no. 1, pp. 49–61, 2014.

[20] Z. Zhu, Z. Li, Y. Liu, and Y. Zou, "Research process review on learning early warning," *Modern Educ. Technol.*, vol. 30, no. 6, pp. 39–46, 2020.

[21] J. Dai, X. Gu, and B. Jiang, "All roads lead to Rome: Cognitive diagnosis and knowledge tracing–A comparison of two modeling methods for mainstream learners' knowledge state," *Modern Educ. Technol.*, vol. 32, no. 4, pp. 88–98, 2022.

[22] R. Henson and J. Douglas, "Test construction for cognitive diagnosis," *Appl. Psychol. Meas.*, vol. 29, no. 4, pp. 262–277, 2005.

[23] K. K. Tatsuoka, "Rule space: An approach for dealing with misconceptions based on item response theory," *J. Educ. Meas.*, vol. 20, pp. 345–354, 1983.

[24] H. Ma, Z. Huang, W. Tang, H. Zhu, H. Zhang, and J. Li, "Predicting student performance in future exams via neutrosophic cognitive diagnosis in personalized e-learning environment," *IEEE Trans. Learn. Technol.*, vol. 16, no. 5, pp. 680–693, Oct. 2023.

[25] S. Mao, J. Zhan, Y. Wang, and Y. Jiang, "Improving knowledge tracing via considering two types of actual differences from exercises and prior knowledge," *IEEE Trans. Learn. Technol.*, vol. 16, no. 3, pp. 324–338, Jun. 2023.

[26] W. Xiao, C. Ni, and R. Li, "A literature review of international studies of early-warning for learning (in Chinese)," *Chin. J. Distance Educ.*, no. 2, pp. 70–78, 2018.

[27] A. Cano and J. D. Leonard, "Interpretable multiview early warning system adapted to underrepresented student populations," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 198–211, Apr.–Jun. 2019.

[28] X. Jin, Q. Jiang, and W. Zhao, "Research on the academic early warning and intervention in online learning space: Situation, key issues and innovative application," *Modern Educ. Technol.*, vol. 31, no. 6, pp. 79–87, 2021.

[29] M. Phan, A. De Caigny, and K. Coussement, "A decision support framework to incorporate textual data for early student dropout prediction in higher education," *Decis. Support Syst.*, vol. 168, 2023, Art. no. 113940.

[30] A. Polyzou and G. Karypis, "Feature extraction for next-term prediction of poor student performance," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 237–248, Apr.–Jun. 2019.

[31] A. Alshanqiti and A. Namoun, "Predicting student performance and its influential factors using hybrid regression and multi-label classification," *IEEE Access*, vol. 8, pp. 203827–203844, 2020.

[32] E. Osmanbegovic and M. Suljic, "Data mining approach for predicting student performance," *Econ. Rev.: J. Econ. Bus.*, vol. 10, no. 1, pp. 3–12, 2012.

[33] J. Wang, Z. Yu, Z. Mou, and Y. Li, "Knowledge-based learning early warning modeling and simulation based on learning and evaluation Big Data," *Modern Distance Educ.*, no. 4, pp. 28–37, 2019.

[34] Y. Zou, Z. Zhu, Y. Liu, and Z. Li, "A novel learning early-warning model based on knowledge points and question types," in *Proc. 9th Int. Conf. Inf. Educ. Technol.*, 2021, pp. 68–72.

[35] R. Alcaraz, A. Martínez-Rodrigo, R. Zangróniz, and J. J. Rieta, "Early prediction of students at risk of failing a face-to-face course in power electronic systems," *IEEE Trans. Learn. Technol.*, vol. 14, no. 5, pp. 590–603, Oct. 2021.

[36] F. Okubo, T. Shiino, T. Minematsu, Y. Taniguchi, and A. Shimada, "Adaptive learning support system based on automatic recommendation of personalized review materials," *IEEE Trans. Learn. Technol.*, vol. 16, no. 1, pp. 92–105, Feb. 2023.

[37] J. Lu, C.-S. Yu, and C. Liu, "Learning style, learning patterns, and learning performance in a WebCT-based MIS course," *Inf. Manage.*, vol. 40, no. 6, pp. 497–507, 2003.

[38] Y. Jung and J. Lee, "Learning engagement and persistence in massive open online courses (MOOCS)," *Comput. Educ.*, vol. 122, pp. 9–22, 2018.

[39] X. Wang, Y. Zhao, C. Li, and P. Ren, "ProbSAP: A comprehensive and high-performance system for student academic performance prediction," *Pattern Recognit.*, vol. 137, 2023, Art. no. 109309.

[40] M. Pandey and S. Taruna, "A multi-level classification model pertaining to the student's academic performance prediction," *Int. J. Adv. Eng. Technol.*, vol. 7, no. 4, 2014, Art. no. 1329.

[41] B. Guo, R. Zhang, G. Xu, C. Shi, and L. Yang, "Predicting students performance in educational data mining," in *Proc. Int. Symp. Educ. Technol.*, 2015, pp. 125–128.
[42] K. E. Arnold and M. D. Pistilli, "Course signals at Purdue: Using learning analytics to increase student success," in *Proc. 2nd Int. Conf. Learn. Analytics Knowl.*, 2012, pp. 267–270.
[43] C. Romero, S. Ventura, P. G. Espejo, and C. Hervás, "Data mining algorithms to classify students," in *Proc. 1st Int. Conf. Educ. Data Mining*, 2008, pp. 8–17.
[44] A. Jain and S. Solanki, "An efficient approach for multiclass student performance prediction based upon machine learning," in *Proc. Int. Conf. Commun. Electron. Syst.*, 2019, pp. 1457–1462.
[45] J. Hua, "The early warning system for studying in Taiwan's colleges and universities and its reference," *Jiangsu Higher Educ.*, no. 5, pp. 136–138, 2007.
[46] P. Zhou, P. Liu, D. Zhang, Z. Wang, and D. Wu, "Research on modeling of classroom teaching activities in learning cyber space," *China Educ. Technol.*, no. 4, pp. 44–51, 2019.
[47] F. Wu and T. Li, "The differentiation design of instructional objectives based on distance learner," *Modern Distance Educ. Res.*, no. 3, pp. 66–71, 2013.
[48] L. Wang, "Learner-platform interaction behavior mining and learning prediction model building," *Chin. J. Distance Educ.*, no. 5, pp. 62–67, 2021.
[49] L. Wang, M. Han, X. Li, N. Zhang, and H. Cheng, "Review of classification methods for unbalanced datasets," *Comput. Eng. Appl.*, vol. 57, no. 22, pp. 42–52, 2021.
[50] O. Allouche, A. Tsoar, and R. Kadmon, "Assessing the accuracy of species distribution models: Prevalence, Kappa and the true skill statistic (TSS)," *J. Appl. Ecol.*, vol. 43, no. 6, pp. 1223–1232, 2006.
[51] N. Wu, L. Zhang, Y. Gao, M. Zhang, X. Sun, and J. Feng, "CLMS-Net: Dropout prediction in MOOCs with deep learning," in *Proc. ACM Turing Celebration Conf.-China*, 2019, pp. 1–6.

Hua Ma (Senior Member, IEEE) received the B.S. degree in computer science and technology, the M.S. degree in computer application technology, and the Ph.D. degree in software engineering from Central South University, Changsha, China in 2003, 2006, and 2016, respectively.

He is currently a Full Professor with the College of Information Science and Engineering, Hunan Normal University, Changsha. His research interests include Big Data mining in education, personalized learning, and services computing.

Dr. Ma is a Senior Member of China Computer Federation.

Wen Zhao received the B.S. degree in computer science and technology from Hunan Women University, Changsha, China, in 2021. She is currently working toward the M.S. degree in software engineering with the College of Information Science and Engineering, Hunan Normal University, Changsha.

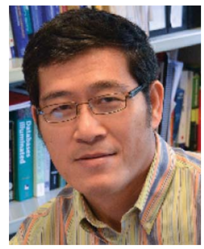
Her research interests include personalized learning and early warning of learning performance.

Yuqi Tang received the B.S. degree in information management and information system from the Hunan University of Chinese Medicine, Changsha, China, in 2021. She is currently working toward the M.S. degree in software engineering with the College of Information Science and Engineering, Hunan Normal University, Changsha.

Her research interests include personalized learning and cognitive diagnosis.

Peiji Huang received the B.S. degree in software engineering from Hunan Normal University, Changsha, China, in 2022, where he is currently working toward the M.S. degree in software engineering with the College of Information Science and Engineering.

His research focuses on personalized learning.



Haibin Zhu (Senior Member, IEEE) received the B.S. degree in computer engineering from the Institute of Engineering and Technology, Zhengzhou, China, in 1983, and the M.S. and Ph.D. degrees in computer science and technology from the National University of Defense Technology, Changsha, China, in 1988 and 1997, respectively.

He is a Full Professor with the Department of Computer Science and Mathematics and the Founding Director of the Collaborative Systems Laboratory, Nipissing University, North Bay, ON, Canada. He has accomplished more than 230 research works, including 40+ IEEE transactions articles, six books, five book chapters, three journal issues, and four conference proceedings.

Dr. Zhu is a Fellow of the International Institute of Cognitive Informatics and Cognitive Computing, a Senior Member of ACM, and a Full Member of Sigma Xi.

Wensheng Tang received the B.S. degree in electronic information science and technology from Hunan Normal University, Changsha, China in 1992, and the M.S. and Ph.D. degrees in computer science and technology from National University of Defense Technology, Changsha, in 1997 and 2009, respectively.

He is currently a Full Professor with the College of Information Science and Engineering, Hunan Normal University. His research interests include personalized learning, distributed computing, and cloud computing.

Keqin Li (Fellow, IEEE) received the B.S. degree in computer science from Tsinghua University, Beijing, China in 1985, and the Ph.D. degree in computer science from the University of Houston, Houston, TX, USA, in 1990.

He is a SUNY Distinguished Professor of computer science with the State University of New York, New Paltz, NY, USA. He has authored or coauthored more than 870 journal articles, book chapters, and refereed conference papers. He holds more than 70 patents announced or authorized by the Chinese National Intellectual Property Administration. His current research interests include cloud computing, fog computing and mobile edge computing, energy-efficient computing and communication, embedded systems and cyber-physical systems, heterogeneous computing systems, Big Data computing, high-performance computing, CPU-GPU hybrid and cooperative computing, computer architectures and systems, computer networking, machine learning, and intelligent and soft computing.

Dr. Li was the recipient of the several Best Paper Awards. He is among the world's top 5 most influential scientists in parallel and distributed computing in terms of both single-year impact and career-long impact based on a composite indicator of Scopus citation database. He has chaired many international conferences. He is currently an Associate Editor for *ACM Computing Surveys* and *CCF Transactions on High Performance Computing*. He was on the editorial boards of IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Computers, IEEE Transactions on Cloud Computing, IEEE Transactions on Services Computing, and IEEE Transactions on Sustainable Computing. He is an AAIA Fellow. He is also a Member of Academia Europaea (Academician of the Academy of Europe).