IFS/ACM Transactions on Machine Learning, vol. 3, no. 1, pp. 20-29, 2026

# Personalized Forgetting-aware Multi-view Graph Knowledge Tracing for Vocational Bachelor's Practical Training Platform Evaluating

Xintao Wang

*Shenzhen Polytechnic University, Shenzhen 518055 Guangdong, China*

ARTICLE INFO

*Keywords*:
vocational bachelor's practical training platforms
knowledge tracing
personalized forgetting-aware calibration

ABSTRACT

Vocational bachelor's practical training platforms play an important role in cultivating high-level technical talents, where accurately evaluating students' learning progress is essential for personalized guidance and adaptive training optimization. Knowledge tracing provides an effective way to infer students' evolving knowledge states from historical learning interactions. However, existing methods still suffer from two limitations: they usually fail to jointly model static knowledge structures and dynamic cognitive evolution, and they often ignore personalized forgetting behaviors caused by time intervals, skill characteristics, and repeated practice. To address these issues, this paper proposes a personalized forgetting-aware multi-view graph knowledge tracing method for student performance evaluation in vocational bachelor's practical training platforms. Specifically, a multi-view graph aggregation module is first constructed, where a heterogeneous graph captures high-order semantic relations among students, exercises, and skills, while a session graph models the temporal evolution of students' response-aware knowledge states. Then, a personalized forgetting-aware calibration module is designed to estimate individualized forgetting intensity by jointly considering hidden knowledge states, exercise representations, skill representations, time intervals, and historical practice frequency. Based on the calibrated knowledge state, an adaptive performance evaluation module integrates target exercise information, related skill representations, and historical learning experience for response prediction. Furthermore, a forgetting-aware adaptive loss is introduced to emphasize cognitively unstable and uncertain interactions during optimization. Experimental results on benchmark datasets show that the proposed method consistently outperforms representative knowledge tracing baselines in terms of AUC, ACC, and Recall.

## 1. Introduction

With the continuous deepening and structural adjustment of the higher education system, vocational bachelor's education has become an important field for cultivating high-level, interdisciplinary, and application-oriented technical talents, while also facing unprecedented opportunities and challenges [1, 2, 3]. Unlike general undergraduate education, which emphasizes the construction of theoretical knowledge systems, and unlike traditional higher vocational education, which focuses mainly on job-oriented skill training, vocational bachelor's education stresses the deep integration of theoretical knowledge, technical application, engineering practice, and professional competence [4, 5, 6]. Its fundamental goal is to cultivate high-quality technical talents who can adapt to industrial upgrading, technological innovation, and increasingly complex occupational requirements. In the context of a new round of scientific and technological revolution and industrial transformation, enterprises have placed higher demands on students' practical ability, comprehensive application ability, innovation ability, and continuous learning ability [7, 8]. Traditional practical training modes, which mainly rely on single-skill training, fixed operational procedures, and result-oriented assessment, are no longer sufficient to meet the advanced requirements of vocational bachelor's talent cultivation. Therefore, the experimental and practical training system of vocational bachelor's education is undergoing a profound transformation from traditional practical training to comprehensive experimental training. This transformation is reflected in the shift from isolated skill practice to integrated multi-skill training, from verification-oriented experiments to project-based, comprehensive, and open-ended practice, and from final-score-oriented evaluation to process-oriented assessment of learning development, competence formation, and continuous improvement [9].

In this process, accurately and effectively evaluating the training effectiveness of the experimental and practical training system has become a key issue for improving the quality of vocational bachelor's education [10]. Existing evaluation approaches mainly depend on course grades, practical training reports, teacher ratings, or stage-based assessments. Although these methods can reflect students' final performance to some extent, they are insufficient for revealing changes in knowledge mastery, skill development trajectories, competence weaknesses, and learning forgetting during the training process [11]. They also provide limited support for the precise adjustment and continuous optimization of the practical training system. Without dynamic perception and scientific evaluation of training effectiveness, the experimental and practical training system may remain at the static level of task completion and result recording, making it difficult to form a closed-loop mechanism of effect evaluation, problem diagnosis, system adjustment, and

DOI: https://doi.org/10.70891/TML.2026.040005
ISSN of IFS/ACM TML: 3078-5510

*Corresponding author
wangxintao@szpu.edu.cn (X. Wang)

quality improvement. Therefore, for vocational bachelor's experimental and practical training systems, it is necessary to develop an intelligent method that can characterize students' knowledge and skill mastery, reflect changes in the training process, and support precise evaluation and optimization decisions, thereby providing effective support for the continuous improvement of practical training systems and the enhancement of talent cultivation quality [12].

Knowledge tracing, as a dynamic learner-state modeling method in educational data mining and intelligent education, provides a feasible technical foundation for evaluating the effectiveness of vocational bachelor's practical training platforms [12, 13]. Its essential goal is to analyze students' historical interaction sequences in learning platforms or training systems, infer their mastery levels of knowledge concepts, skill units, or practical competence components, and predict their future performance in subsequent tasks, projects, or assessments. In the context of vocational bachelor's practical training, knowledge tracing should not be limited to estimating whether a student has mastered a theoretical concept. Instead, it can be extended to the continuous evaluation of operational skills, project implementation ability, engineering thinking, and comprehensive practical competence. Specifically, students' experimental procedures, project submissions, code revisions, device operations, stage tests, error feedback, and repeated training behaviors can all be regarded as educational interaction data [14]. By sequentializing, structuring, and graph-modeling these data, knowledge tracing models can construct dynamic profiles of students' competence states, thereby providing a fine-grained, process-oriented, and interpretable basis for evaluating the training effectiveness of practical training systems [15].

Existing knowledge tracing methods can generally be divided into three categories. The first category is sequence-based knowledge tracing [16]. This type of method treats students' historical learning behaviors as temporally ordered interaction sequences and uses recurrent neural networks, long short-term memory networks, or attention mechanisms to capture the temporal evolution of students' knowledge states [17]. Sequence-based methods are useful for modeling how students progress from basic training tasks to comprehensive projects, because they can represent temporal dependencies in continuous learning processes [18]. However, they often compress knowledge states into unified latent vectors, which may overlook the structural relationships among different skills and the cognitive differences among individual students [19]. The second category is memory-network-based knowledge tracing. These methods introduce external memory matrices to store students' mastery levels of different knowledge concepts or skill units and update these states through reading and writing operations [20, 21]. Compared with purely sequential models, memory-network-based methods can provide a more explicit representation of concept-level mastery and can better describe students' competence changes across different practical training modules. Nevertheless, their state updating processes are usually implemented through implicit neural operations, making it difficult to provide sufficient explanations for forgetting, repeated practice, and personalized learning characteristics [22, 23]. The third category is graph-structure-based knowledge tracing [24]. This type of method constructs graph structures among students, tasks, projects, knowledge points, and skill units, and employs graph neural networks to mine dependency relations and high-order semantic information. Graph-based methods are particularly suitable for vocational bachelor's practical training platforms because practical tasks often involve prerequisite relations, skill combinations, and project-level dependencies [25, 26].

However, existing knowledge tracing methods still face two major limitations when applied to student performance evaluation in vocational bachelor's practical training platforms. First, most existing methods fail to jointly model static knowledge structures and dynamic cognitive evolution in a unified framework. Sequence-based models mainly focus on temporal response patterns, but they usually ignore the high-order semantic relations among students, exercises, and skills. In contrast, graph-based methods can capture structural dependencies, but they often insufficiently describe how students' knowledge states evolve along continuous practical training activities. This separation makes it difficult to comprehensively represent the complex learning process in practical training scenarios. Second, current methods usually model forgetting with a shared or simplified temporal decay function, while ignoring the personalized nature of knowledge retention. In real learning environments, forgetting is jointly affected by students' current knowledge states, exercise characteristics, skill properties, time intervals, and repeated practice frequency. Without explicitly modeling these individualized forgetting factors, existing methods may overestimate outdated historical interactions or underestimate the memory consolidation effect of repeated practice, leading to less reliable student performance prediction.

To address the above limitations, this paper proposes a personalized forgetting-aware multi-view graph knowledge tracing method for student performance evaluation in vocational bachelor's practical training platforms. Specifically, this paper first constructs a multi-view graph aggregation module to jointly model static and dynamic learning information. A heterogeneous graph is built to capture high-order semantic relations among students, exercises, and skills, while a session graph is introduced to characterize the temporal evolution of students' response-aware cognitive states. In this way, the model can simultaneously exploit structural knowledge dependencies and sequential learning transitions. Furthermore, this paper designs a personalized forgetting-aware calibration module to explicitly estimate the forgetting intensity of each student on specific skills. The module jointly considers hidden knowledge states, exercise representations, skill representations, time intervals, and historical practice frequency, thereby modeling both knowledge decay and memory consolidation caused by repeated practice. Based on the calibrated knowledge state, an adaptive performance evaluation module is further developed to integrate target exercise information, related skill representations, and historical learning experience for response prediction. Finally, a forgetting-aware adaptive loss is introduced to assign larger optimization weights to cognitively unstable and uncertain interactions, improving the reliability of student performance evaluation.

The contributions of this paper are listed as follows:

- This paper proposes a personalized forgetting-aware multi-view graph knowledge tracing method for student performance evaluation in vocational bachelor's practical training platforms, which jointly models static knowledge relations, dynamic cognitive evolution, and individualized forgetting behaviors, to provide a more comprehensive solution for practical training evaluation.

- This paper designs a multi-view graph aggregation module and a personalized forgetting-aware calibration module where the heterogeneous graph captures high-order semantic relations among students, exercises, and skills, while the session graph models response-aware cognitive evolution. Meanwhile, the personalized forgetting mechanism estimates knowledge decay by considering knowledge states, exercise and skill representations, time intervals, and repeated practice frequency.

- This paper conducts extensive experiments on benchmark datasets to evaluate the effectiveness of the proposed method. Experiment results show that the proposed method consistently outperforms representative knowledge tracing baselines in terms of AUC, ACC, and Recall.

## 2. Related Works

### 2.1. Researches on Knowledge Tracing

Early researches on knowledge tracing is rooted in educational psychology and cognitive science, and primarily relied on conceptual models to quantify the evolution of learners' knowledge states. Bayesian Knowledge Tracing models the discrete transition of knowledge states within a hidden Markov framework, while Item Response Theory predicts learner performance by characterizing the interaction between student ability and item difficulty [27, 28, 29]. Although these approaches offer strong interpretability, their reliance on strong Markov assumptions and limited parametric capacity makes it difficult to capture long-range dependencies in learning sequences and fine-grained cognitive heterogeneity among learners. With the rapid development of deep learning, knowledge tracing has gradually evolved along three major technical paradigms: sequential models, memory-augmented models, and graph-based models. Sequential models exploit autoregressive architectures to capture temporal dependencies in learning interactions. For example, DKT [4] first introduced LSTM into knowledge tracing, enabling end-to-end mapping from interaction sequences to latent knowledge states. Subsequent Transformer-based models, such as SAKT [24], AKT [11], and DTransformer [12], further leverage self-attention mechanisms to model long-range dependencies, while DIMKT [13] and CL4KT [14] enrich sequential representations through difficulty-aware embeddings or contrastive learning. However, these methods typically compress knowledge states into unified latent vectors or black-box representations, thereby overlooking the intrinsic mechanisms of cognitive structures and the fine-grained modeling of individual differences. Memory-augmented models introduce external memory modules to enhance the representational capacity and interpretability of internal states. A representative example is DKVMN [9], which maps knowledge concepts to external key-value memory slots and dynamically reads from and writes to them to track concept mastery. Nevertheless, the modeling of learning regularities, such as forgetting, still largely depends on implicit matrix updates and lacks explicit cognitive or pedagogical meaning. Graph-based models focus on explicitly capturing topological dependencies among knowledge concepts. Models such as GKT [10], SGKT [15], and GIKT [16] construct graph structures over learners, questions, and knowledge concepts, and employ graph neural networks for information propagation. However, static graph topology is insufficient to reflect the temporal evolution of knowledge structures. Although HawkesKT [22], IEKT [21], and GIKT-Dyn [23] further incorporate dynamic mechanisms into graph-based knowledge tracing, the complexity of graph propagation still makes it challenging to develop a lightweight and unified framework that can jointly handle individual cognitive heterogeneity and multifactorial knowledge evolution.

### 2.2. Researches on Forgetting Behavior

Inspired by the Ebbinghaus forgetting curve, recent studies have increasingly investigated how factors such as time intervals and repeated practice affect knowledge forgetting. Traditional BKT and its variants primarily focus on the process of knowledge acquisition, while modeling forgetting only in a relatively coarse-grained manner. Although Yudelson et al. [25] incorporated individual differences into knowledge tracing, they did not systematically characterize the dynamic nature of forgetting. The emergence of deep learning has provided new perspectives for modeling forgetting behavior. Nagatani et al.[26] were among the first to introduce a time-decay function into deep knowledge tracing, but their approach considers only a single temporal factor. Pandey et al. [17] incorporated learning time intervals in RKT, yet did not account for the effect of repeated practice. Ghosh et al. [20] enabled dynamic adjustment of forgetting rates in AKT, thereby improving contextual adaptability. Bai et al. [18] proposed LefoKT, which disentangles forgetting from question relevance through relative forgetting attention; however, its parameter design still relies on local windows. Kuling et al. [19] introduced Hebbian decay in KUL-KT to simulate biologically inspired forgetting, but did not consider the moderating effect of concept difficulty. Overall, although existing research on forgetting modeling has gradually shifted from single-factor assumptions toward multi-factor formulations, it still suffers from two common limitations: insufficient integration of multiple forgetting-related factors and inadequate modeling of individual-specific forgetting patterns.

## 3. Method

### 3.1. Problem Definition

To systematically evaluate learning progresses of students in vocational bachelor's practical training platform evaluating, this paper first defines the student performance prediction task and its underlying relational structures.

**Student performance prediction.** Given a student's historical exercise sequence from time step 1 to $t-1$ in a digital teaching platform, denoted as

$$\mathcal{E}_{1:t-1} = \{(e_1, a_1), (e_2, a_2), \dots, (e_{t-1}, a_{t-1})\}, \tag{1}$$

where $e_i$ denotes the exercise at time step $i$, and $a_i$ denotes the corresponding response result. The goal of student performance prediction is to estimate the probability $p_t$ that the student will correctly answer the next exercise at time step $t$.

**Heterogeneous graph.** The static correlations among students, exercises, and skills are modeled as a heterogeneous graph

$$G_h = \{V_h, A_h\}, \tag{2}$$

where $V_h$ is the set of nodes consisting of three types of entities, namely students, exercises, and skills. $A_h = \{r_A, r_C\}$ denotes the set of relation types, where $r_A$ represents the answering relation and $r_C$ represents the containment relation. Three meta-relation paths are designed to construct neighbors in the heterogeneous graph: exercises answered by the same student, different exercises involving the same skill, and skills co-occurring in the same exercise.

**Session graph.** To characterize the dynamic evolution of students' cognitive states, a session graph is constructed as

$$G_s = \{V_s, A_s\}, \tag{3}$$

where $V_s$ is the node set. Each node $e_i^{a_i} \in V_s$ is constructed from the pair $(e_i, a_i)$, and $A_s$ denotes the partial-order relation matrix reflecting the chronological order of the student's exercise sequence.

Based on the above definitions, this paper proposes a personalized forgetting-aware multi-view graph knowledge tracing method for vocational bachelor's practical training platform evaluating. As shown in Fig. 1, the overall framework consists of three major modules: a multi-view graph aggregation module, a personalized forgetting-aware calibration module, and an adaptive performance evaluation module.

### 3.2. Multi-view Graph Aggregation

In knowledge tracing, student performance is affected by both static relational attributes and dynamic sequential features. Therefore, this paper proposes a multi-view graph aggregation module. Specifically, a heterogeneous graph is used to model the static topological correlations between exercises and skills, while a session graph is employed to describe the dynamic evolution of students' cognitive states over time. This design effectively improves the discriminative capability of student, exercise, and skill representations.

The heterogeneous graph representation extraction adopts a wavelet graph convolution network to propagate information along predefined meta-paths, thereby extracting exercise and skill representations with domain-specific semantic consistency. Given the $i$-th node $\beta_i$ in the heterogeneous graph, which can represent either a skill state $k_i$ or an exercise state $e_i$, and its neighboring node set $\mathcal{N}_i$, the operation of the $l$-th wavelet graph convolutional layer is defined as

$$\beta_i^l = \mathrm{ReLU}\left(\frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i\cup\{i\}} \psi_s \theta^l \psi_s^{-1} \beta_j^{l-1}\right), \tag{4}$$

where $\beta_i^l$ denotes the representation of the $i$-th node output by the $l$-th wavelet graph convolutional layer, $\theta^l = \mathrm{diag}(\theta^l)$ is the learnable spectral-domain filter in the $l$-th layer, $\mathrm{ReLU}(\cdot)$ is the activation function, and $\psi_s$ is the heat-kernel wavelet basis matrix. The matrix $\psi_s$ transforms graph signals from the spatial domain to the wavelet domain and captures the local structural characteristics of nodes at different scales $s$. After multi-layer propagation on the heterogeneous graph $G_h = \{V_h, A_h\}$, the exercise representation $\widetilde{e}$ and the skill representation $\widetilde{k}$ are obtained.

The session graph representation extraction fuses exercise representations with response feedback to construct session-flow nodes, and then evolves students' hidden knowledge states through a gated mechanism on the session graph. First, the exercise and its corresponding response are combined as a node through information propagation:

$$c_i = e_i \oplus a_i, \tag{5}$$

where $e_i$ denotes the $i$-th exercise in the session and $\oplus$ denotes element-wise addition. In the gated graph neural network, the input of the gated recurrent unit is denoted as $\alpha_i$, which contains information from the neighboring nodes of the current node $c_i$ and represents the influence of the previous exercise on the next exercise in the answering session. $\alpha_i$ is formulated as

$$\begin{aligned}\alpha_i = \mathrm{Concat}\Big( & M_i^I([c_1, c_2, \dots, c_n]W^I + b^I),\\ & M_i^O([c_1, c_2, \dots, c_n]W^O + b^O)\Big),\end{aligned} \tag{6}$$

where $W^I$ and $W^O$ are parameter matrices, $b^I$ and $b^O$ are bias vectors, $[c_1, c_2, \dots, c_n]$ denotes the processed exercise-response list, $M_i^I$ and $M_i^O$ denote the $i$-th rows of the corresponding matrices, and $\mathrm{Concat}(\cdot)$ denotes the concatenation

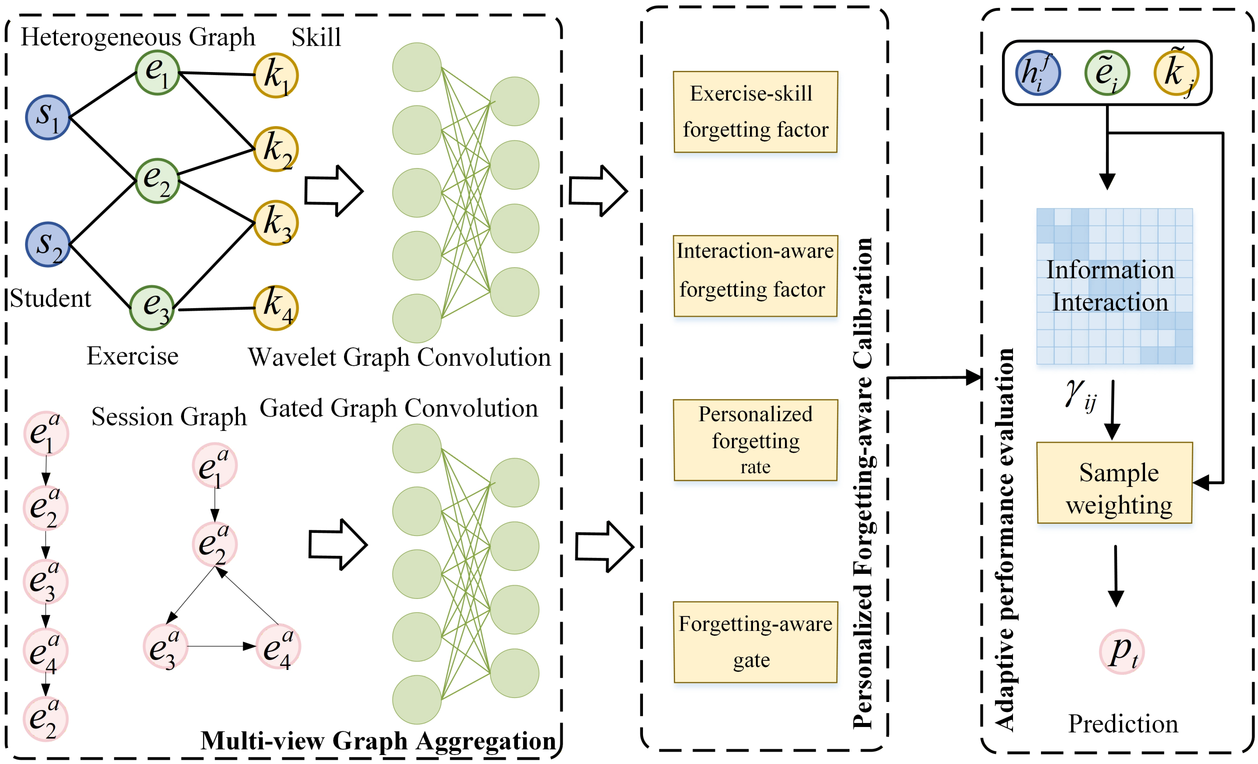


**Figure 1:** The overall framework of the proposed method.

operation. Then, based on the gated recurrent unit inputs $[\alpha_1, \alpha_2, \dots, \alpha_n]$, the hidden knowledge states of the corresponding student are learned and denoted as $[h_1, h_2, \dots, h_n]$. The update process of $h_i$ can be formalized as follows:

$$c_i^t = c_i \oplus \alpha_i, \tag{7}$$

$$\omega_i^t = \sigma(U_\omega c_i^t + V_\omega h_i^{t-1}), \tag{8}$$

$$r_i^t = \sigma(U_r c_i^t + V_r h_i^{t-1}), \tag{9}$$

$$\widetilde{h}_i^t = \tanh(U_h c_i^t + V_h (r_i^t \odot h_i^{t-1})), \tag{10}$$

$$h_i^t = (1 - \omega_i^t) \odot h_i^{t-1} + \omega_i^t \odot \widetilde{h}_i^t, \tag{11}$$

where $U_*$ and $V_*$ are learnable parameters, $* \in \{\omega, r, h\}$, $\sigma(\cdot)$ denotes the sigmoid activation function, and $\odot$ denotes element-wise multiplication. The update gate $\omega_i^t$ controls how much historical information is updated, while the reset gate $r_i^t$ is used to filter information. The term $r_i^t \odot h_i^{t-1}$ indicates that the new information is generated from part of the previous information. The term $(1-\omega_i^t)\odot h_i^{t-1}$ represents the forgotten part of the previous information, and $\omega_i^t \odot \widetilde{h}_i^t$ represents the newly retained information.

### 3.3. Personalized Forgetting-aware Calibration

After obtaining the static representations of exercises and skills from the heterogeneous graph and the dynamic hidden knowledge states from the session graph, an essential issue is how to calibrate the student's historical knowledge state according to individualized forgetting patterns. In real learning scenarios, knowledge forgetting is not only determined by the time interval between learning interactions, but is also jointly affected by students' memory capacity, exercise attributes, skill characteristics, and repeated practice behaviors. Therefore, this paper designs a personalized forgetting-aware calibration module to model the retention and decay of students' knowledge states in a fine-grained manner. Different from conventional methods that employ a globally shared temporal decay function, the proposed module dynamically estimates the forgetting intensity of a student on a specific skill by jointly considering the exercise representation, skill representation, time interval, and historical practice frequency. In this way, the model can adaptively distinguish different forgetting patterns across students thereby providing a more accurate representation of students' evolving knowledge states.

Given the $i$-th learning interaction of student $s$, let $\widetilde{h}_i$ denote the hidden knowledge state generated by the session graph encoder, $\widetilde{e}_i$ denote the exercise representation, and $\widetilde{k}_i$ denote the representation of the skill associated with the current exercise. In addition, $\Delta\tau_i$ denotes the time interval between the current interaction and the previous relevant interaction, and $r_{s,k_i}$ denotes the number of historical practices that student $s$ has performed on skill $k_i$. The personalized forgetting rate is defined as

$$f_{s,i} = f_{\text{base}} \cdot f_i^{\text{item}} \cdot f_{s,i}^{\text{int}} \cdot \left(1 - m_{s,i} \cdot f_{s,i}^{\text{rep}}\right), \tag{12}$$

where $f_{\text{base}}$ is the base forgetting rate, $f_i^{\text{item}}$ denotes the forgetting property related to the current exercise and its associated skill, $f_{s,i}^{\text{int}}$ measures the instantaneous forgetting intensity under the current interaction context, $f_{s,i}^{\text{rep}}$ denotes the memory consolidation effect brought by repeated practice, and $m_{s,i}$ represents the state-dependent memory retention capacity inferred from the student's hidden knowledge state. Specifically, the exercise-skill forgetting factor is calculated as

$$f_i^{\text{item}} = \sigma\left(W_e[\widetilde{e}_i; \widetilde{k}_i] + b_e\right), \tag{13}$$

where $[\cdot;\cdot]$ denotes the concatenation operation, and $W_e$ and $b_e$ are learnable parameters. This factor reflects the intrinsic forgetting property of the current exercise and its associated skill. Exercises involving more complex or less frequently practiced skills may correspond to a higher forgetting tendency. The interaction-aware forgetting factor is formulated as

$$f_{s,i}^{\text{int}} = \sigma\left(W_{\text{int}}[\widetilde{h}_i; \widetilde{e}_i; \widetilde{k}_i; \Delta\tau_i] + b_{\text{int}}\right), \tag{14}$$

where $W_{\text{int}}$ and $b_{\text{int}}$ are learnable parameters. This component estimates the instantaneous forgetting intensity under the current learning context by jointly considering the student's current hidden knowledge state, the target exercise, the related skill, and the elapsed time interval. To further capture the student's current retention ability without introducing an additional student representation, the memory retention capacity is inferred from the hidden knowledge state:

$$m_{s,i} = \sigma\left(W_m \widetilde{h}_i + b_m\right), \tag{15}$$

where $W_m$ and $b_m$ are learnable parameters. Since $\widetilde{h}_i$ is generated from the student's historical response sequence, $m_{s,i}$ can reflect the student's current knowledge stability and memory retention ability in a dynamic manner.

Moreover, to explicitly model the memory consolidation effect of repeated practice, the repetition-aware factor is defined as

$$f_{s,i}^{\text{rep}} = \sigma\left(W_{\text{rep}}[r_{s,k_i}; \Delta\tau_i; \widetilde{k}_i] + b_{\text{rep}}\right), \tag{16}$$

where $r_{s,k_i}$ denotes the number of previous practices of student $s$ on skill $k_i$. A larger value of $f_{s,i}^{\text{rep}}$ indicates a stronger memory consolidation effect caused by repeated practice, which reduces the overall forgetting rate through the term $(1 - m_{s,i} \cdot f_{s,i}^{\text{rep}})$. Finally, the personalized forgetting rate is used to calibrate the student's hidden knowledge state through a forgetting-aware gate:

$$\widetilde{h}_i^f = (1 - f_{s,i}) \odot \widetilde{h}_i, \tag{17}$$

where $\odot$ denotes element-wise multiplication, and $\widetilde{h}_i^f$ is the forgetting-calibrated knowledge state. This operation adaptively controls the retention degree of historical knowledge according to the current knowledge state, exercise characteristics, skill properties, time interval, and repeated practice frequency.

The proposed personalized forgetting-aware calibration mechanism provides a fine-grained way to model knowledge retention and decay. When a student has repeatedly practiced a certain skill, $f_{s,i}^{\text{rep}}$ becomes larger, thereby reducing the overall forgetting rate and preserving more historical knowledge. In contrast, when the time interval is long, the current exercise or skill has a stronger forgetting tendency, or the student's hidden state indicates weak knowledge stability, the estimated forgetting rate increases accordingly. As a result, the model can more accurately simulate the individualized process of knowledge consolidation and decay without introducing an additional student representation.

### 3.4. Adaptive Performance Evaluation

After obtaining the forgetting-calibrated knowledge state from the personalized forgetting-aware calibration module, this subsection further evaluates the student's future performance by adaptively integrating the calibrated cognitive state, the target exercise representation, and the related skill representations. Different from conventional prediction modules that directly use the original hidden state for response prediction, the proposed adaptive performance evaluation module takes the forgetting-calibrated state as the core student knowledge representation, thereby enabling the final prediction to reflect both knowledge acquisition and knowledge decay.

Given the target exercise $e_t$, let $\widetilde{e}_t$ denote its exercise representation, and let $\widetilde{K}_{e_t}$ denote the set of neighboring skill representations associated with $e_t$. For each related skill, $\widetilde{k}_j \in \widetilde{K}_{e_t}$ represents the $j$-th skill representation. In addition, $\widetilde{E}_t$ denotes the student's historical exercise experience set, and $\widetilde{h}_t^f$ denotes the forgetting-calibrated knowledge state generated by the personalized forgetting-aware calibration module. The module constructs multi-dimensional interaction features by considering four types of relationships: the interaction between the target exercise and the calibrated knowledge state, the interaction between related skills and the calibrated knowledge state, the influence of historical exercises on the target exercise, and the influence of historical exercises on the related skills. Specifically, the pair $\langle \widetilde{e}_t, \widetilde{h}_t^f \rangle$ represents the student's current mastery of the target exercise after considering forgetting effects. The pair $\langle \widetilde{k}_j, \widetilde{h}_t^f \rangle$ represents the student's mastery of the $j$-th skill associated with the target exercise. The pair $\langle \widetilde{e}_t, \widetilde{E}_t \rangle$ characterizes the influence of the student's historical exercise experience on the target exercise, while $\langle \widetilde{k}_j, \widetilde{E}_t \rangle$ describes the influence of historical exercises on the skills related to the target exercise. These interaction terms jointly provide a comprehensive basis for evaluating the student's future response performance. To adaptively fuse different interaction factors, this paper computes the importance weight of each interaction term through a similarity-based attention function:

$$\gamma_{ij} = \text{Softmax}\left(g(z_i, z_j)\right), \tag{18}$$

$$p_t = \sigma\left(\sum_{z_i \in \{\widetilde{E}_t, \widetilde{h}_t^f\}} \sum_{z_j \in \{\widetilde{K}_{e_t}, \widetilde{e}_t\}} \gamma_{ij}\left(W^T[z_i, z_j] + b^T\right)\right), \tag{19}$$

where $g(\cdot)$ denotes the dot-product operation, $\gamma_{ij}$ is the adaptive fusion weight, $W^T$ is a learnable parameter matrix, $b^T$ is a bias vector, $[\cdot,\cdot]$ denotes the concatenation operation, and $\sigma(\cdot)$ is the sigmoid activation function. The predicted probability $p_t$ indicates the probability that the student correctly answers the target exercise $e_t$ at time step $t$. The basic prediction objective is the binary cross-entropy loss between the predicted probability and the ground-truth response:

$$\mathcal{L} = -\sum_t \left[ y_t \log p_t + (1 - y_t) \log(1 - p_t) \right], \tag{20}$$

where $y_t \in \{0, 1\}$ denotes the ground-truth response label. However, in knowledge tracing scenarios, different learning interactions do not contribute equally to model optimization. Samples with higher forgetting intensity or weaker knowledge stability are usually more difficult to predict and contain more informative signals for evaluating the student's real cognitive state. Therefore, this paper further introduces a forgetting-aware adaptive loss to dynamically adjust the contribution of each training sample:

$$\mathcal{L} = -\sum_t \omega_t \left[ y_t \log p_t + (1 - y_t) \log(1 - p_t) \right], \tag{21}$$

where $\omega_t$ is the adaptive sample weight. It is computed according to the personalized forgetting rate and prediction uncertainty:

$$\omega_t = 1 + \lambda_f f_{s,t} + \lambda_u \left(1 - |2p_t - 1|\right), \tag{22}$$

where $f_{s,t}$ is the personalized forgetting rate estimated by the previous module, and $1 - |2p_t - 1|$ measures the prediction uncertainty. When $p_t$ is close to 0.5, the model is uncertain about the student's response, and the corresponding sample receives a larger weight. $\lambda_f$ and $\lambda_u$ are hyperparameters that control the contributions of forgetting intensity and prediction uncertainty, respectively.

Through this design, the adaptive performance evaluation module not only predicts student responses by integrating exercise, skill, and historical learning information, but also assigns greater optimization attention to cognitively unstable and high-forgetting interactions. As a result, the model can better distinguish whether an incorrect response is caused by insufficient mastery, knowledge decay, or uncertainty in the current learning state, thereby improving the reliability and sensitivity of student performance evaluation.

# 4. Experimental Results

## 4.1. Experimental Settings

To evaluate the effectiveness and superiority of the proposed method, two representative student performance prediction datasets are adopted as benchmark datasets, namely ASSIST09 and ASSIST12. Both datasets are collected from the ASSISTments online learning platform and have been widely used in knowledge tracing research. ASSIST09 is collected during the 2009–2010 academic year. It contains interaction records from 3852 students, covering 123 skills and 17737 exercises. ASSIST12 is collected during the 2012–2013 academic year and is substantially larger in scale, including 27485 students, 265 skills, and 2709436 exercises. These two datasets differ in terms of data scale, skill coverage, and interaction density, which enables a more comprehensive evaluation of the proposed model under different learning scenarios. For both datasets, the data are divided into training and testing sets with a ratio of 8:2.

The proposed method is implemented based on TensorFlow and trained on a Linux server equipped with an NVIDIA GTX A100 GPU. In all experiments, the embedding dimensions of exercises, skills, and student knowledge states are set to 100. The number of graph encoder layers is set to 3. The learning rate is fixed at 0.00025, and the dropout rate is set to 0.8 to alleviate overfitting. The maximum number of training epochs is set to 500, and the batch sizeis set to 6. To comprehensively assess the prediction performance of the model, three widely used evaluation metrics are employed: AUC, ACC, and recall. For all three metrics, higher values indicate better predictive performance.

The proposed method is compared with eight representative baseline models, including LefoKT [18] DSN-KT [30], Bi-CLKT [31], HD-KT [32], LBKT [33], JKT [34], and DMKT [35].

## 4.2. Result Comparison

To evaluate the effectiveness of the proposed method, this paper conducts comparative experiments on two widely used educational datasets, ASSIST09 and ASSIST12. Following common practice in knowledge tracing, and model performance is evaluated using AUC, ACC, and Recall. This paper compares the proposed method with several representative knowledge tracing baselines, including sequence-based, graph-based, behavior-aware, and cognitive-representation-based models. These baselines cover different technical paradigms, such as deep sequential modeling, joint graph convolution, dual-view cognitive representation, and learning-behavior-oriented tracing, thereby providing a comprehensive benchmark for assessing the proposed multi-view graph fusion framework.

As shown in Table 1, the proposed method consistently achieves the best overall performance on both datasets across all three metrics. This demonstrates that the proposed model can more accurately infer students' evolving knowledge states and predict their future responses than existing baselines. Compared with sequence-oriented methods, the proposed model does not rely solely on chronological response patterns, but further exploits the static semantic dependencies among students, exercises, and skills through heterogeneous graph aggregation. Compared with graph-based methods, the proposed model additionally constructs session graphs to capture the dynamic evolution of student cognition, which enables the learned representations to reflect both structural knowledge relations and temporal learning transitions. Therefore, the proposed model achieves more robust performance under different learning scenarios.

**Table 1**
The average results with standard deviation over five random runs are reported on two datasets.

| Methods | ASSIST09 | | | ASSIST12 | | |
|---|---|---|---|---|---|---|
| | AUC | ACC | Recall | AUC | ACC | Recall |
| DSN-KT | $56.47 \pm 1.13$ | $55.91 \pm 0.86$ | $56.82 \pm 1.42$ | $57.71 \pm 1.08$ | $56.04 \pm 1.67$ | $57.33 \pm 0.94$ |
| Bi-CLKT | $55.38 \pm 1.55$ | $54.76 \pm 1.21$ | $56.02 \pm 0.79$ | $56.69 \pm 1.36$ | $55.74 \pm 0.98$ | $56.35 \pm 1.74$ |
| JKT | $65.21 \pm 0.91$ | $63.08 \pm 1.48$ | $66.94 \pm 1.26$ | $67.63 \pm 1.63$ | $65.75 \pm 0.87$ | $66.91 \pm 1.19$ |
| DMKT | $58.16 \pm 1.72$ | $56.31 \pm 0.95$ | $59.72 \pm 1.31$ | $65.27 \pm 1.44$ | $62.84 \pm 1.08$ | $63.02 \pm 1.69$ |
| HD-KT | $70.89 \pm 1.06$ | $67.91 \pm 1.52$ | $73.81 \pm 0.88$ | $72.94 \pm 1.27$ | $68.46 \pm 1.81$ | $70.77 \pm 1.15$ |
| LBKT | $63.76 \pm 1.39$ | $63.42 \pm 0.83$ | $66.27 \pm 1.76$ | $66.86 \pm 1.12$ | $66.03 \pm 1.46$ | $68.36 \pm 0.97$ |
| LefoKT | $72.35 \pm 1.21$ | $69.21 \pm 1.09$ | $74.43 \pm 1.37$ | $74.16 \pm 1.42$ | $69.98 \pm 1.16$ | $75.21 \pm 0.88$ |
| Ours | **75.28 ± 0.74** | **70.81 ± 1.18** | **75.86 ± 0.93** | **76.05 ± 1.05** | **72.04 ± 0.89** | **77.18 ± 1.33** |

The performance advantage mainly comes from the complementary design of three modules. First, the multi-view graph aggregation module integrates heterogeneous graph modeling and session graph modeling. The heterogeneous graph captures high-order relations among exercises and skills through wavelet graph convolution, which helps alleviate the sparsity of student-exercise interactions and improves the discriminability of exercise and skill representations. Meanwhile, the session graph encodes response-aware learning trajectories through gated graph propagation, allowing the model to track the dynamic transition of students' hidden knowledge states. This static-dynamic collaboration provides a richer representation foundation than methods that model either graph structure or temporal sequence alone. Second, the personalized forgetting-aware calibration module further improves knowledge state estimation by explicitly modeling individualized knowledge retention and decay. Unlike conventional forgetting mechanisms that mainly depend on a shared temporal decay function, the proposed model estimates the forgetting intensity of a student on a specific interaction by jointly considering the hidden knowledge state, exercise representation, skill representation, time interval, and historical practice frequency. In this way, the model can distinguish whether historical knowledge should be retained or weakened according to the current learning context. Repeated practice contributes to memory consolidation, while longer intervals or unstable knowledge states increase the estimated forgetting intensity. This mechanism reduces the negative influence of outdated or unreliable historical interactions and makes the predicted knowledge state more consistent with realistic learning behavior. Third, the adaptive performance evaluation module strengthens the final prediction by integrating the forgetting-calibrated knowledge state with the target exercise, related skills, and historical exercise experience. The similarity-based adaptive fusion mechanism assigns different importance weights to multiple interaction terms, enabling the model to focus on the most relevant evidence for the target prediction. Moreover, the forgetting-aware adaptive loss gives higher training weights to samples with stronger forgetting intensity or higher prediction uncertainty.

**Table 2**
Ablation study on ASSIST09.

| Methods | AUC | ACC | Recall |
|---|---|---|---|
| w/o Heterogeneous Graph | 71.40 | 66.15 | 71.32 |
| w/o Session Graph | 72.57 | 67.26 | 72.48 |
| w/o Personalized Forgetting | 73.59 | 67.52 | 73.40 |
| w/o Repetition Factor | 74.56 | 68.31 | 74.37 |
| w/o Adaptive Fusion | 74.22 | 68.18 | 73.05 |
| w/o Forgetting-aware Loss | 75.87 | 69.60 | 74.68 |
| Ours | **75.28** | **70.81** | **75.86** |

This design encourages the model to learn more from cognitively unstable and difficult interactions, thereby improving its sensitivity to ambiguous student states and enhancing overall prediction reliability.

Overall, the comparison results confirm that the proposed model benefits from the joint modeling of structural dependency, temporal cognitive evolution, personalized forgetting, and adaptive optimization. The consistent superiority over all baselines indicates that simply enhancing sequential modeling, graph representation, or behavioral features is insufficient for accurate knowledge tracing. By calibrating multi-view representations with personalized forgetting dynamics, the proposed framework provides a more comprehensive and cognitively plausible solution for student performance prediction in vocational bachelor's practical training platform evaluation.

## 4.3. Ablation Result

To further verify the effectiveness of each key component in the proposed method, this paper conducts ablation experiments on ASSIST09 and ASSIST12. Specifically, this paper compares the full model with the following six variants:

- **w/o Heterogeneous Graph** removes the heterogeneous graph aggregation branch and only uses the session graph to model student learning sequences.
- **w/o Session Graph** removes the session graph encoder and only uses the heterogeneous graph to represent static exercise-skill relations.

- **w/o Personalized Forgetting** removes the personalized forgetting-aware calibration module and directly uses the original hidden knowledge state for prediction.

- **w/o Repetition Factor** removes the repetition-aware memory consolidation term from the forgetting estimation module.

- **w/o Adaptive Fusion** replaces the similarity-based adaptive interaction fusion with a simple concatenation-based prediction layer.

- **w/o Forgetting-aware Loss** replaces the forgetting-aware adaptive loss with the standard binary cross-entropy loss.

The ablation results are reported in Table 2. All variants are evaluated using the same experimental settings as the main comparison experiments. AUC, ACC, and Recall are used as evaluation metrics. As shown in Table 2, removing any key component leads to performance degradation, which verifies the necessity of each module in the proposed framework.

First, the variant without the heterogeneous graph suffers the most significant performance drop. This indicates that static relational structures among students, exercises, and skills are essential for learning discriminative representations. Without the heterogeneous graph, the model cannot sufficiently exploit high-order exercise-skill dependencies, making it more difficult to infer student mastery in sparse interaction scenarios. Second, removing the session graph also causes a clear decline in performance. This demonstrates that static graph information alone is insufficient for knowledge tracing, since student learning is inherently sequential and evolves over time. The session graph captures the dynamic transition patterns of response-aware learning behaviors, enabling the model to update students' hidden knowledge states according to their historical practice trajectories. Therefore, the combination of heterogeneous graph and session graph is critical for jointly modeling static knowledge structures and dynamic cognitive evolution.Third, the performance decreases when the personalized forgetting-aware calibration module is removed. This result confirms that directly using the original hidden state cannot fully reflect the realistic process of knowledge retention and decay. By estimating personalized forgetting rates according to the hidden knowledge state, exercise representation, skill representation, time interval, and historical practice frequency, the proposed module can adaptively calibrate outdated or unstable historical information. This helps the model distinguish whether a student's incorrect response is caused by insufficient mastery or knowledge decay. Fourth, removing the repetition factor weakens the model performance, although the degradation is smaller than that caused by removing the whole personalized forgetting module. This suggests that repeated practice provides useful evidence for memory consolidation. When a student repeatedly practices a skill, the model should retain more related historical knowledge rather than treating all past interactions with the same decay pattern. The repetition-aware factor therefore improves the model's ability to capture the positive effect of practice frequency on knowledge stability. Fifth, the variant without adaptive fusion performs worse than the full model. This shows that simple feature concatenation cannot sufficiently model the complex interactions among the forgetting-calibrated knowledge state, target exercise, related skills, and historical exercise experience. In contrast, the adaptive fusion mechanism assigns different importance weights to different interaction terms, allowing the model to focus on the most informative evidence for the current prediction. Finally, replacing the forgetting-aware adaptive loss with the standard binary cross-entropy loss also leads to a performance decrease. This indicates that not all learning interactions contribute equally to model optimization. Samples with higher forgetting intensity or higher prediction uncertainty are usually more informative for learning robust student representations. By assigning larger weights to these difficult and cognitively unstable samples, the forgetting-aware loss improves the sensitivity and reliability of the prediction model.

## 5. Conclusion

In this paper, this paper proposed a personalized forgetting-aware multi-view graph knowledge tracing method for student performance evaluation in vocational bachelor's practical training platforms. The proposed method jointly models static knowledge relations, dynamic cognitive evolution, and individualized forgetting behaviors. Specifically, a heterogeneous graph is constructed to capture high-order semantic relations among students, exercises, and skills, while a session graph is used to model the temporal evolution of students' response-aware knowledge states. Furthermore, a personalized forgetting-aware calibration module estimates forgetting intensity by considering knowledge states, exercise and skill representations, time intervals, and historical practice frequency. Based on the calibrated knowledge state, the adaptive performance evaluation module integrates target exercise information, related skills, and historical learning experience for prediction. Experimental results demonstrate that the proposed method outperforms representative baselines. In future work, more practical behavior features will be incorporated to support personalized feedback and adaptive training evaluation.

## Acknowledgement

This work was supported by Shenzhen Educational Science 14th Five-Year Plan 2024 Research Project: Innovative Construction and Empirical Study of the Vocational Undergraduate Experimental and Practical Training System from an Interdisciplinary Perspective. The author declares that there is no conflict of interest.

# References

[1] I. Farran, I. Nunez, Converging pathways: New approaches to integrate vocational education training and higher education, Journal of Vocational Education & Training, 2025, vol. 77, no. 4, pp. 1147–1165.

[2] H. M. Naveen, NEP, 2020: General education embedded with skill and vocational education, International Journal of Scientific Research in Science, Engineering and Technology, 2022, vol. 9, no. 1, pp. 65–75.

[3] P. Li, J. Gao, J. Zhang, S. Jin, Z. Chen, Deep reinforcement clustering, IEEE Transactions on Multimedia, 2023, vol. 25, pp. 8183–8193.

[4] C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas, J. Sohl-Dickstein, Deep knowledge tracing, Proceedings of the 2015 Annual Conference on Neural Information Processing Systems, Montreal, Quebec, Canada, 2015, pp. 505–513.

[5] J. Gao, M. Liu, P. Li, J. Zhang, Z. Chen, Deep multiview adaptive clustering with semantic invariance, IEEE Transactions on Neural Networks and Learning Systems, 2024, vol. 35, no. 9, pp. 12965–12978.

[6] J. Gao, M. Liu, P. Li, A. A. Laghari, A. R. Javed, N. Victor, T. R. Gadekallu, Deep incomplete multiview clustering via information bottleneck for pattern mining of data in extreme-environment IoT, IEEE Internet of Things Journal, 2024, vol. 11, no. 16, pp. 26700–26712.

[7] Z. Zhan, X. Mao, H. Liu, S. Yu, STGL: Self-supervised spatio-temporal graph learning for traffic forecasting, Journal of Artificial Intelligence Research, 2025, vol. 2, no. 1, pp. 1–8.

[8] W. Liu, Channel reorganization for few-shot segmentation, Journal of Artificial Intelligence Research, 2024, vol. 1, no. 1, pp. 36–40.

[9] J. Zhang, X. Shi, I. King, D. Yeung, Dynamic key-value memory networks for knowledge tracing, Proceedings of the 26th International Conference on World Wide Web, ACM, Perth, Australia, 2017, pp. 765–774.

[10] H. Nakagawa, Y. Iwasawa, Y. Matsuo, Graph-based knowledge tracing: Modeling student proficiency using graph neural network, Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, ACM, Thessaloniki, Greece, 2019, pp. 156–163.

[11] X. Guo, Z. Huang, J. Gao, M. Shang, M. Shu, J. Sun, Enhancing knowledge tracing via adversarial training, Proceedings of the 29th ACM International Conference on Multimedia, ACM, Virtual Event, China, 2021, pp. 367–375.

[12] Y. Yin, L. Dai, Z. Huang, S. Shen, F. Wang, Q. Liu, E. Chen, X. Li, Tracing knowledge instead of patterns: Stable knowledge tracing with diagnostic transformer, Proceedings of the ACM Web Conference 2023, ACM, Austin, TX, USA, 2023, pp. 855–864.

[13] S. Shen, Z. Huang, Q. Liu, Y. Su, S. Wang, E. Chen, Assessing student's dynamic knowledge state by exploring the question difficulty effect, Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, Madrid, Spain, 2022, pp. 427–437.

[14] W. Lee, J. Chun, Y. Lee, K. Park, S. Park, Contrastive learning for knowledge tracing, Proceedings of the ACM Web Conference 2022, ACM, Virtual Event, Lyon, France, 2022, pp. 2330–2338.

[15] Z. Wu, L. Huang, Q. Huang, C. Huang, Y. Tang, SGKT: Session graph-based knowledge tracing for student performance prediction, Expert Systems with Applications, 2022, vol. 206, p. 117681.

[16] Y. Yang, J. Shen, Y. Qu, Y. Liu, K. Wang, Y. Zhu, W. Zhang, Y. Yu, GIKT: A graph-based interaction model for knowledge tracing, Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, Ghent, Belgium, 2020, pp. 299–315.

[17] S. Pandey, J. Srivastava, RKT: Relation-aware self-attention for knowledge tracing, Proceedings of the 29th ACM International Conference on Information and Knowledge Management, ACM, Virtual Event, Ireland, 2020, pp. 1205–1214.

[18] Y. Bai, X. Li, Z. Liu, Y. Huang, M. Tian, W. Luo, Rethinking and improving student learning and forgetting processes for attention based knowledge tracing models, Proceedings of the 39th AAAI Conference on Artificial Intelligence, AAAI Press, Philadelphia, PA, USA, 2025, pp. 27822–27830.

[19] G. Kuling, M. Zitnik, Ken utilization layer: Hebbian replay within a student's ken for adaptive exercise recommendation, 2025. arXiv:2507.00032.

[20] A. Ghosh, N. Heffernan, A. S. Lan, Context-aware attentive knowledge tracing, Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Virtual Event, CA, USA, 2020, pp. 2330–2339.

[21] T. Long, Y. Liu, J. Shen, W. Zhang, Y. Yu, Tracing knowledge state with individual cognition and acquisition estimation, Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, Virtual Event, Canada, 2021, pp. 173–182.

[22] C. Wang, W. Ma, M. Zhang, C. Lv, F. Wan, H. Lin, T. Tang, Y. Liu, S. Ma, Temporal cross-effects in knowledge tracing, Proceedings of the 14th ACM International Conference on Web Search and Data Mining, ACM, Virtual Event, Israel, 2021, pp. 517–525.

[23] L. Tu, X. Zhu, Y. Ji, Graph-based dynamic interactive knowledge tracing, Proceedings of the 2023 8th International Conference on Distance Education and Learning, ACM, Beijing, China, 2023, pp. 49–56.

[24] S. Pandey, G. Karypis, A self attentive model for knowledge tracing, Proceedings of the 12th International Conference on Educational Data Mining, International Educational Data Mining Society (IEDMS), Montréal, Canada, 2019.

[25] M. V. Yudelson, K. R. Koedinger, G. J. Gordon, Individualized bayesian knowledge tracing models, Proceedings of the 16th International Conference on Artificial Intelligence in Education, Springer, Memphis, TN, USA, 2013, pp. 171–180.

[26] K. Nagatani, Q. Zhang, M. Sato, Y.-Y. Chen, F. Chen, T. Ohkuma, Augmenting knowledge tracing by considering forgetting behavior, Proceedings of the World Wide Web Conference, ACM, San Francisco, CA, USA, 2019, pp. 3101–3107.

[27] Y. Zhou, M. Liu, S. Chen, Z. Chen, Deep time series contrastive clustering with cross-view reliable cluster diffusion, Expert Systems with Applications, 2026, p. 133028.

[28] Z. Chen, M. Liu, Y. Li, H. Zheng, Z. Liu, H. Zhang, L. Zhao, Double-incomplete multi-view clustering with self-induced semantic label diffusion, Information Science, 2026, vol. 753, p. 123626.

[29] M. Liu, Y. Li, Z. Chen, Y. Lin, Information-driven complementarity and consistency mining for multi-view clustering, IEEE Signal Processing Letters, 2026, vol. 33, pp. 216–220.

[30] R. Gao, Student classroom knowledge tracking based on deep semantics-robust network, J Appl Sci Eng, 2024, vol. 28, pp. 2051–2058.

[31] X. Song, J. Li, Q. Lei, W. Zhao, Y. Chen, A. Mian, Bi-CLKT: Bi-graph contrastive learning based knowledge tracing, Knowledge-Based Systems, 2022, vol. 241, p. 108274.

[32] H. Ma, Y. Yang, C. Qin, X. Yu, S. Yang, X. Zhang, H. Zhu, HD-KT: Advancing robust knowledge tracing via anomalous learning interaction detection, Proceedings of the ACM Web Conference 2024, ACM, Singapore, 2024, pp. 4479–4488.

[33] B. Xu, Z. Huang, J. Liu, S. Shen, Q. Liu, E. Chen, J. Wu, S. Wang, Learning behavior-oriented knowledge tracing, Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, ACM, Long Beach, CA, USA, 2023, pp. 2789–2800.

[34] X. Song, J. Li, Y. Tang, T. Zhao, Y. Chen, Z. Guan, JKT: A joint graph convolutional network based deep knowledge tracing, Information Sciences, 2021, vol. 580, pp. 510–523.

[35] Q. Li, X. Yuan, J. Yue, X. Shen, R. Liang, S. Liu, Z. Yan, Dual-view multi-scale cognitive representation for deep knowledge tracing, Knowledge-Based Systems, 2025, vol. 310, p. 113010.