Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: https://www.science-gate.com/IJAAS.html

A knowledge-structured and interpretable machine learning framework for knowledge tracing and learning outcomes

Bagabold Gendensuren, Byambasuren Ivanov *, Munkhtsetseg Namsraidorj, Enkhtuul Bukhsuren

Department of Information and Computer Sciences, School of Information Technology and Electronics, National University of Mongolia, Ulaanbaatar, Mongolia

ARTICLE INFO

Article history:
Received 5 March 2026
Received in revised form
17 July 2026
Accepted 9 August 2026

Keywords:
Educational data mining
Graph diffusion
Knowledge graph
Knowledge tracing
Machine learning

ABSTRACT

Knowledge Tracing (KT) models are widely used to estimate learners' knowledge states. Recent approaches often rely on deep neural architectures, which can be difficult to interpret, computationally expensive, and challenging to deploy in real-world educational settings. In addition, explicitly modeling relationships between topics remains a key challenge. This study proposes a knowledge-structured framework that automatically identifies latent relationships between topics from real learning data and integrates them with machine learning–based prediction. Topic similarity is estimated using cosine similarity to construct a knowledge graph, and learners' knowledge states are updated through a graph-based propagation process. This approach enables related topics to influence each other, providing a more structured view of learning dynamics. Experimental results show that the proposed method produces more stable and consistent knowledge estimates compared with approaches based solely on raw performance data. The model achieves a Recall of 0.92, an AUC of 0.61, and an RMSE of 0.55, indicating strong sensitivity to learning risk while maintaining stable predictions.

1. Introduction

Accurately monitoring how learners' knowledge develops over time and anticipating their ability to understand upcoming topics remains a central concern in education. When such information is available, instructors can adapt learning materials to individual needs, revisit concepts that have not been fully understood, and provide targeted support to reduce learning gaps.

A common approach for modeling this process is Knowledge Tracing (KT) (Corbett and Anderson, 1994). KT models describe learners' knowledge states over time by estimating whether a given concept has been mastered. In its classical form, KT relies on four parameters: the probability that a learner initially knows an idea, the likelihood of acquiring it through practice, the chance of guessing correctly without mastery, and the probability of making an error despite knowing the concept. While this formulation is transparent and straightforward, it treats concepts independently and overlooks the relationships between topics.

To address sequential learning patterns, Deep Knowledge Tracing (DKT) was later proposed by Piech et al. (2015), using Long Short-Term Memory (LSTM) networks. Although DKT improves the modeling of temporal dependencies, it shares a key limitation with traditional KT: neither approach explicitly represents how topics are related or how prior learning across concepts influences future performance. An alternative perspective was introduced by Pavlik et al. (2009) through Performance Factors Analysis (PFA), which applies logistic regression to capture the effects of accumulated correct and incorrect responses. PFA offers improved interpretability, yet it still lacks an explicit representation of conceptual structure and temporal ordering.

These limitations have motivated the exploration of structured representations that capture topic–subtopic relationships. Recent studies have incorporated Knowledge Graphs (KGs) and Graph Neural Networks (GNNs) into KT models to better reflect prerequisite structures and learning progression. For instance, Graph-based Interaction Knowledge Tracing (GIKT) (Yang et al., 2021) models interactions between questions and skills

* Corresponding Author.
Email Address: byambaa@num.edu.mn (B. Ivanov)
https://doi.org/10.21833/ijaas.2026.08.005
Corresponding author's ORCID profile:
https://orcid.org/0000-0001-5236-7857

using graph structures, allowing knowledge transfer to be represented more explicitly. Similarly, Graph-EKLN (Liu et al., 2021) encodes learner–exercise–concept relations as nodes and edges, with semantic similarity reflected through weighted connections. Graph-based Knowledge Tracing models knowledge concepts as nodes and their relationships as edges, and updates learners' knowledge states through graph neural networks. It also supports learning latent graph structures directly from student interaction data (Nakagawa et al., 2019).

Despite these advances, many graph-based KT models depend on deep, multi-layer neural architectures. Such designs often require substantial computational resources, large training datasets, and provide limited insight into internal decision processes. As a result, recent research has increasingly emphasized the development of interpretable KT models that make learning dynamics more transparent.

In this study, we propose an enhanced KT framework that constructs a knowledge graph based on topic–subtopic relationships and employs cosine similarity together with propagation algorithms to estimate future topic mastery using machine learning techniques. By modeling conceptual relationships directly at the graph level, the proposed approach enables a more precise analysis of knowledge evolution and cross-topic influence, while maintaining interpretability and practical applicability.

2. Related work

Research on Knowledge Tracing (KT) has progressed through several related, though not always convergent, directions. Initial approaches framed learner knowledge using probabilistic models with deliberately simplified assumptions. Subsequent studies have increasingly adopted graph-based representations to capture dependencies between related topics that earlier models overlooked. At the same time, increasing attention has been paid to model transparency, while a separate line of research has examined how machine learning techniques can be used to assess and validate KT predictions. This section reviews representative contributions from these strands and outlines the limitations that motivate the framework proposed in this study.

2.1. Traditional knowledge tracing models

One of the earliest and most influential models in KT research is Bayesian Knowledge Tracing (BKT), which represents a learner's knowledge state as a binary variable indicating whether a concept has been mastered (Corbett and Anderson, 1994). The model updates this state over time using four parameters—learning, guess, slip, and forget—estimated from observed responses. Because of its simplicity and intuitive probabilistic formulation, BKT has been widely adopted in early intelligent tutoring systems. However, the model relies entirely on binary correctness signals, which limits its ability to reflect partial or evolving understanding. In addition, concepts are treated as independent, leaving semantic relationships between topics unmodeled.

To incorporate temporal learning patterns more explicitly, DKT introduced recurrent neural networks, particularly Long Short-Term Memory (LSTM) architectures, to process sequences of learner responses (Piech et al., 2015). While this approach improves the modeling of time-dependent behavior, it offers little transparency regarding how predictions are formed. Performance Factors Analysis (PFA) employs a different approach, utilizing logistic regression to model the cumulative impact of prior correct and incorrect responses, yielding a model that is more interpretable than its neural alternatives (Pavlik et al., 2009). Nevertheless, PFA still does not explicitly represent sequence-level temporal structure or conceptual dependencies across topics.

Taken together, these traditional models provide an important foundation for KT research. At the same time, they offer only a limited perspective on how knowledge develops across related concepts and do not explicitly describe how understanding transfers between topics.

2.2. Graph-based knowledge tracing: Topic–subtopic dependency modeling

As the limitations of independent concept modeling became clearer, graph-based approaches emerged as a natural extension of KT. These methods aim to represent prerequisite relations, semantic proximity, and knowledge transitions directly within a structured graph (Namsraidorj et al., 2025).

GIKT models interactions between questions and skills through graph structures, enabling knowledge transfer to be expressed more explicitly than in purely sequential models (Yang et al., 2021). Graph-based approaches represent learners, exercises, and knowledge concepts as connected elements, making it possible to capture the relationships among them. Building on these ideas, Cui et al. (2024) introduced a dual-graph ensemble learning approach for knowledge tracing that captures heterogeneous relationships among learning interactions and combines the resulting graph models through online knowledge distillation.

More recent work has explored advanced graph neural architectures. For example, GCN-based KT models have shown improved ability to capture hierarchical topic structures, resulting in higher AUC scores compared to traditional DKT approaches. Word2Vec has been widely used to generate semantic vector representations of concepts, whereas Graph Attention Networks (GAT) allow attention-based weighting of graph edges and neighboring nodes. Together, these techniques provide an effective foundation for modeling topic–

subtopic relationships in knowledge graphs (Mikolov et al., 2013; Veličković et al., 2017).

Additionally, dual-graph-based knowledge tracing approaches integrate concept graphs with learner interaction sequences, offering a broader view of learning dynamics (Nakagawa et al., 2019). Although these models capture latent dependencies and semantic proximity more effectively, many of them rely on deep architectures that increase computational cost and further complicate interpretation

2.3. Interpretable and explainable knowledge tracing models

Despite improvements in predictive accuracy, interpretability remains a persistent challenge in KT research. Concerns about the black-box nature of deep neural models have motivated increasing interest in Interpretable Knowledge Tracing (IKT) approaches (Piech et al., 2015; Yeung and Yeung, 2018).

Linear and logistic models, such as AFM and PFA, allow parameters to be interpreted directly, while extensions that incorporate Item Response Theory (IRT) seek to jointly model learner ability and task difficulty (Yudelson et al., 2013). More recent approaches employ attention mechanisms to make the influence of specific topics on predictions more visible (Gantulga et al., 2020). Although these methods improve transparency, they generally do not integrate explicit semantic relationships between topics or model knowledge propagation at the graph level

2.4. Hybrid KT and machine learning validation

Another line of research focuses on validating KT outputs using conventional supervised machine learning methods. Baker and Inventado (2014) highlighted the importance of evaluating KT models with standard ML metrics such as Accuracy, AUC, and RMSE to assess predictive reliability. Knowledge Tracing representations can also be used with traditional machine learning methods to improve prediction and evaluate model performance.

These approaches aim to improve prediction while keeping the model useful for practical educational applications. However, fully integrated pipelines that combine graph-based semantic similarity, explicit knowledge propagation, KT modeling, and machine learning validation remain relatively rare (Nakagawa et al., 2019; Veličković et al., 2017). This limitation directly motivates the unified framework introduced in this study.

3. Proposed methodology and framework

This section presents a hybrid Knowledge Tracing framework designed to address the common limitations of existing models. Previous graph-based KT models have used relationships between concepts and information propagation to model how knowledge concepts influence one another (Tong et al., 2020). Following this idea, our framework uses a knowledge graph, a simple propagation mechanism, and supervised machine learning to keep the model understandable and practical to use. A review of previous Knowledge Tracing (KT) studies highlights several open issues that motivate this work. First, recent graph-based KT models attempt to represent relationships between topics, but most of them rely on deep neural architectures (Nakagawa et al., 2019; Veličković et al., 2017). While these models can be effective, their internal reasoning is difficult to comprehend, making it challenging to interpret concept-level relationships from a human perspective (Piech et al., 2015; Yeung and Yeung, 2018). This creates a need for a clearer and more explainable way to represent semantic links between topics.

Second, knowledge propagation across topics is often handled implicitly within neural networks. Although such approaches can capture complex patterns, they provide limited insight into how mastery in one topic affects related topics over time. A more explicit and mathematically transparent propagation mechanism would allow this process to be analyzed more clearly.

Third, many graph-based KT models assess their performance mainly within the same modeling framework. Independent validation using supervised machine learning models and standard evaluation metrics remains relatively limited, making it more challenging to assess their practical reliability.

Finally, deep KT models are often computationally expensive to train and deploy. This limits their applicability in real Learning Management Systems (LMS), especially in settings with limited computational resources. A lightweight and efficient architecture is therefore important for real-world use.

Taken together, these limitations motivate a hybrid framework that combines interpretable graph construction, explicit knowledge propagation, and machine learning–based validation in a practical and efficient manner.

3.1. Framework overview

The proposed framework is built around four key elements. First, a knowledge graph is created to describe the relationships between topics, with similarity between topics measured using cosine similarity (Mikolov et al., 2013). Second, a diffusion-based propagation process is applied to update learners' knowledge states by allowing information to flow between related topics (Wang and Zhang, 2021). Third, the resulting knowledge estimates are examined using supervised machine learning models to assess their predictive value (Baker and Inventado, 2014). Finally, the framework presents the results in a clear and interpretable form to support instructional decision-making.

By combining these elements, the framework provides a structured approach to modeling knowledge development, thereby eliminating the need for complex, deep neural architectures (Piech et al., 2015).

3.2. Knowledge graph construction

A Knowledge Graph (KG) is constructed to represent semantic relationships among topics and subtopics. The graph is defined as:

$$G = (V, E, W) \quad (1)$$

where, $\boldsymbol{V} = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ denotes the set of topic nodes, $\boldsymbol{E}$ represents edges between topics, and $\boldsymbol{W} = [\boldsymbol{w_{ij}}]$ is a weighted adjacency matrix.

Each topic is represented as a feature vector derived from its learning content. Semantic similarity between topics $\boldsymbol{v_i}$ and $\boldsymbol{v_j}$ is computed using cosine similarity:

$$CosineSim(v_i, v_j) = \frac{(v_i \cdot v_j)}{(||v_i|| \times ||v_j||)} \quad (2)$$

The edge weight between two topics is defined as:

$$w_{ij} = \begin{cases} CosineSim(v_i, v_j), if CosineSim(v_i, v_j) > \theta \\ 0, \quad otherwise \end{cases} \quad (3)$$

where, $\boldsymbol{\theta}$ is a threshold that removes weak or irrelevant connections.

By filtering out weak connections, the resulting graph focuses on semantically relevant relationships and remains straightforward to interpret.

3.3. Graph diffusion–based knowledge propagation

To model the dynamic influence of related topics, a graph diffusion–based propagation mechanism is applied. Let X^t denote the learner's topic-level mastery state at iteration t, and let X^0 represent the initial mastery state derived from observed performance. The propagation process is defined as:

$$X^{t+1} = \alpha \tilde{W} X^t + (1 - \alpha) X^0 \quad (4)$$

where, $\tilde{W}$ is the normalized adjacency matrix and $\alpha \in [0, 1]$ controls the strength of propagation.

This formulation allows mastery in one topic to influence related topics based on their connections. The process converges when $X^{t+1} = X^t$, yielding the stabilized representation:

$$X^{prop} = (I - \alpha\tilde{W})^{-1} (1 - \alpha) X^0 \quad (5)$$

where, I is the identity matrix. The resulting matrix X^{prop} represents the propagated mastery values across all topics.

Machine learning–based validation: The propagated knowledge state X^{prop} is used as the input feature matrix for supervised machine learning validation. The objective is to estimate the learner's actual performance outcome as:

$$\hat{y} = f(X^{prop}) \approx \text{y\_actual} \quad (6)$$

where, f(·) represents the prediction model used in this study.

Evaluation metrics: The performance evaluation metrics used in this study are summarized in Table 1.

Table 1: Performance evaluation metrics

Metrics	Formula
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
AUC (ROC curve)	—
RMSE	$\sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2}$

These metrics provide complementary measures of the model's predictive performance.

3.4. Implementation details and algorithm

The propagation step follows Eq. (4), where α controls the balance between information propagated from related topics and the initial mastery state X^{prop}. To make the implementation easier to follow, the main parameter settings used in this study are described below. The dataset includes item-level attempt records, item–topic mappings, and topic-level relationships derived from learning sequences (Nakagawa et al., 2019; Piech et al., 2015). The threshold value was set to θ = 0.60 and used to convert continuous performance scores into binary learning outcomes. This value was chosen based on the overall distribution of scores, where many topic-level values are relatively low, making it useful to separate stronger and weaker performance (Baker and Inventado, 2014; Fawcett, 2006).

The propagation coefficient was set to α = 0.50. This allows the model to balance the original performance information with the structural information coming from related topics. A window size of 2 was used to construct the topic transition matrix, capturing immediate transitions between topics observed in the learning process. The transition matrix $\tilde{W}$ was normalized using row normalization so that the outgoing weights for each topic are comparable. Knowledge propagation was applied for a single step. Given the relatively small dataset, additional iterations were not used, as they tend to smooth the values too much and reduce meaningful differences between topics. The main parameter settings are summarized in Table 2.

Algorithm 1 presents the overall workflow of the proposed method, including data preparation, construction of the Student–Topic matrix, topic transition modeling, knowledge propagation, and prediction.

Algorithm 1. Knowledge-structured learning with propagation

Input: item-level data, topic mapping, θ, α, window size
Output: predicted success probabilities
1. Prepare and clean the item-level data.
2. Construct the Student–Topic matrix (X) based on topic-level performance.
3. Extract topic sequences from learner activity.
4. Build the topic transition matrix ($\tilde{W}$) using a sliding window.
5. Normalize $\tilde{W}$ using row normalization.
6. Apply knowledge propagation to obtain X^{prop}.
7. Convert performance scores into binary labels using θ.
8. Train the prediction model using X^{prop}.
9. Generate success probabilities for each student and topic.

Table 2: Implementation parameters

Parameter	Description	Value
θ (theta)	Threshold for binary learning outcome	0.60
α (alpha)	Propagation coefficient	0.50
Window size	Topic transition window	2
Normalization method	Normalization of $\tilde{W}$	Row normalization
Propagation step	Number of propagation updates	1
Convergence rule	Stopping condition	Fixed one-step propagation (to avoid oversmoothing)

The framework is designed so that its results are easy to understand. The weights of the graph edges indicate the degree of relatedness between topics, the propagation process illustrates how knowledge in one topic influences others, and the machine learning model highlights which topics contribute most significantly to the predictions.

Because the framework is lightweight and uses a small number of parameters, it can be applied in real-time learning management systems. In practice, it helps lecturers identify topics that learners struggle with, recognize potential learning risks early, and make more informed teaching decisions based on data. Within the scope of this study, the following hypotheses are proposed:

H1: The incorporation of knowledge propagation leads to improved predictive performance compared to using raw performance data alone, as reflected in higher AUC and lower RMSE.
H2: The knowledge-structured model more effectively identifies low-performing topics compared to raw performance–based estimates, resulting in clearer differentiation of learning risk.
H3: The integration of topic structure and propagation produces model outputs that are more interpretable, as evidenced by consistent alignment between student–topic representations and topic-level summaries.
H4: The proposed framework maintains stable and conservative probability estimates under limited data conditions, avoiding extreme or overconfident predictions.

3.5. Experimental setup and system architecture

In this study, the proposed knowledge-structured machine learning framework was implemented and evaluated using real learning data collected from eight programming olympiad training teams, each consisting of three students. This setup follows the standard team structure used in programming contests. The primary objective of the implementation is to assess whether the framework can be applied in a reproducible, lightweight, and decision-support-oriented manner. The overall process follows the system architecture illustrated in Fig. 1. The proposed architecture consists of five layers:

- Data layer: Item-level data collected from the training environment were preprocessed by standardizing identifiers, handling missing values, and normalizing scores. This ensures a consistent and reliable input for subsequent analysis.
- Feature construction layer: Item-level performance is aggregated with difficulty-based weighting to build the Student - Topic matrix X.
- Knowledge graph layer: Based on learners' observed topic sequences, a window-based sequence method is used to construct the topic transition matrix $\tilde{W}$. The graph structure is learned directly from data.
- Propagation layer: The matrices X and $\tilde{W}$ are combined to compute the propagated matrix X^{prop}, allowing prerequisite and nearby topic effects to be reflected in the student–topic state.
- Prediction and output layer: Using raw states, propagated states, and learner similarity information, the model estimates topic-level success probabilities. The output file (predictions.csv) provides Student × Topic probabilities that can be used for instructional decisions.

Overall, the architecture eschews deep neural networks and is designed to remain transparent, interpretable, and straightforward to deploy in real-world systems.

- Data and Preprocessing: We used item-level data collected from a real training environment. During preprocessing, Student and item IDs were standardized, missing values were handled, and score values were mapped to a consistent scale. This step provides a stable input for all subsequent computations.
- Student–Topic Matrix (X): The Student–Topic matrix was built from item-level performance. Item difficulty was used as a weight so that performance on more difficult items contributes more to the topic-level estimate. The resulting matrix X represents the initial learning state at the topic level.
- Topic Graph ($\tilde{W}$) and Knowledge Propagation: To learn topic dependencies from real learning behavior, we applied a window-based sequence

analysis and constructed $\tilde{W}$. The resulting matrix is typically dense and contains many non-zero transitions, reflecting the actual flow of learning across topics.

Next, we applied propagation over the topic graph to compute X^{prop}. This step makes topic estimates more stable, especially when data are sparse or uneven across topics.

- Success Probability Prediction: Based on the propagated topic states, we estimated the probability that each learner will successfully learn each topic. In implementation, each topic was modeled separately, and a fallback strategy was used when class separation was insufficient. As a result, predicted probabilities remain within [0, 1] and do not collapse into extreme values.

4. Results and discussion

4.1. Results of student–topic matrix

In the initial stage of the study, a Student–Topic matrix (X) was constructed using item-level performance data, with question difficulty applied as a weighting factor. The results indicate that learners' levels of topic mastery are unevenly distributed across topics. While mastery appears relatively stable for some topics, others exhibit greater variability and dispersion among learners. Fig. 2 presents the Student × Topic matrix, which represents the initial knowledge state of each learner for each topic. As observed from the matrix, several topics demonstrate consistently higher mastery levels across most learners. In contrast, others—most notably Topic L—show generally low mastery values for the majority of learners. This pattern suggests that topic mastery varies substantially depending on both the learner and the topic under consideration.

The color gradient used in the visualization provides an intuitive representation of mastery levels. Yellow and light green shades correspond to high mastery values (close to 1), green–blue tones indicate moderate mastery, and blue to dark purple colors represent low or near-zero mastery levels. This visual encoding allows for a direct comparison of mastery differences across learners and topics.

Overall, these observations suggest that mastery estimates based solely on raw performance data can partially reflect learners' knowledge states but fail to capture the influence of structural relationships between topics and the effects of learning order. Consequently, it becomes necessary in the subsequent stage to infer inter-topic dependencies from actual learning trajectories and to incorporate this structural information to refine mastery estimation.

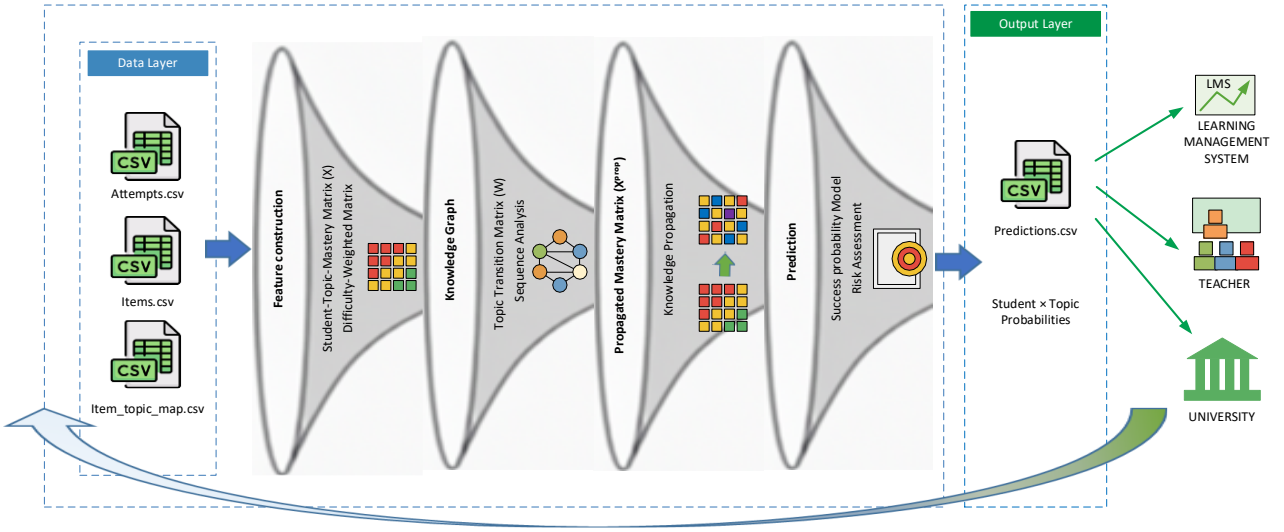

Fig. 1: Overall architecture of the proposed knowledge-structured framework

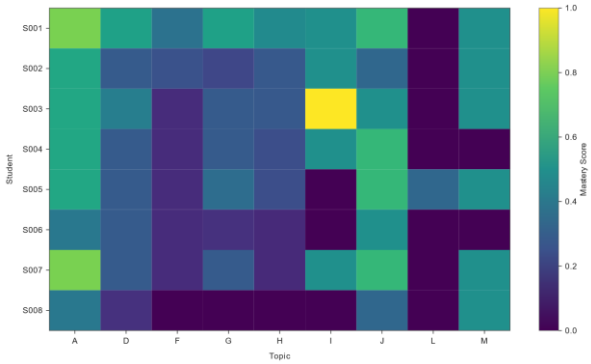

Fig. 2: Student–topic mastery matrix (X)

4.2. Results of the topic transition matrix ($\tilde{W}$)

Based on learners' actual learning sequences, a topic transition matrix ($\tilde{W}$) was constructed using a window-based sequence analysis approach. This matrix captures the tendency of transitions from one topic to subsequent topics and reflects structural information about the learning process learned directly from data, rather than being manually defined. The topic transition matrix ($\tilde{W}$) derived from real learning trajectories is illustrated as a heatmap in Fig. 3. In this representation, rows correspond to source topics, columns denote

subsequent issues, and the matrix values indicate the relative transition strengths from one topic to another. As observed from the heatmap, the $\tilde{W}$ matrix exhibits a dense structure with multiple non-zero entries for most topics. This pattern suggests that learning content is not acquired in isolation, but instead follows a layered and interconnected progression.

Importantly, these dependencies are inferred directly from observed learning behavior, demonstrating the feasibility of data-driven discovery of topic relationships. Distinct transition patterns are evident for specific topics. In particular, transitions toward Topics G and H show relatively higher weights, indicating that these topics function as core or foundational nodes for multiple other topics. Notably, the strong transition from Topic I to Topic H suggests an explicit sequential dependency, consistent with a prerequisite relationship between these topics. Conversely, some topics exhibit fewer and more concentrated transitions, which may indicate that they correspond to later stages of instruction or more specialized concepts. Such structural characteristics are critical for subsequent knowledge propagation, as they determine how strongly mastery in preceding topics influences related topics.

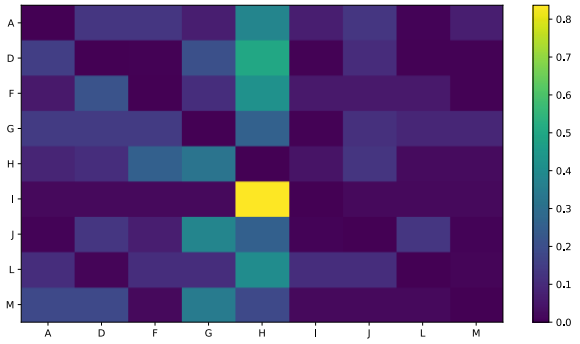

Fig. 3: Topic transition matrix ($\tilde{W}$)

Overall, the analysis confirms that the topic transition matrix ($\tilde{W}$) effectively captures meaningful structural relationships among topics based on learners' actual learning paths. These findings provide both theoretical justification and practical support for applying knowledge propagation at the Student–Topic mastery level in the subsequent stage of the framework.

4.3. Results of knowledge propagation

Based on the topic transition matrix ($\tilde{W}$), a knowledge propagation process was applied to the Student–Topic mastery matrix, resulting in the propagated mastery matrix (X^{prop}). This propagation mechanism allows prior mastery of earlier topics to influence subsequent related topics, thereby enriching raw mastery estimates with structural information derived from the learning sequence. The results indicate that the propagated matrix X^{prop} exhibits a more balanced distribution than the original matrix X.

In particular, topics affected by data sparsity or weak discrimination in the raw estimates show more stable and consistent values after propagation. This observation suggests that knowledge propagation successfully transfers structural information from the learning process into topic-level mastery estimates.

As shown in Table 3, the impact of knowledge propagation varies across topics. Topics with relatively low initial mastery levels, such as H, G, and L, exhibit substantial increases in their propagated mean values, indicating that prerequisite and neighboring topics make meaningful contributions to their updated mastery estimates. The pronounced increase observed for Topic H suggests that it is strongly connected to other topics within the learned structure. In contrast, topics such as A, I, and J, which initially showed higher or more isolated mastery patterns, experience decreases in propagated values. This behavior reflects the smoothing effect of propagation, whereby mastery estimates are redistributed in a structured manner according to topic dependencies.

Table 3: Topic-level effects of knowledge propagation

Topic	Mean X (Raw)	Mean X^{prop} (Propagated)	Δ (X^{prop} − X)
A	0.6000	0.3220	−0.2780
D	0.3214	0.3234	+0.0019
F	0.1562	0.2110	+0.0548
G	0.2679	0.4611	+0.1932
H	0.2200	0.8106	+0.5906
I	0.3750	0.1688	−0.2062
J	0.5417	0.3040	−0.2376
L	0.0417	0.1046	+0.0630
M	0.3750	0.1729	−0.2021

Δ: Denotes the change in average topic-level mastery induced by knowledge propagation.

Building on these propagated mastery estimates, the next stage focuses on predicting the probability of successful learning for each topic at the learner level. Fig. 4 presents a heatmap of the predicted probabilities of successful learning at the learner–topic level. This visualization provides a concise overview of how success probabilities are distributed across learners and topics, allowing for the observation of overall patterns, structural differences, and potential risk areas. As shown in the heatmap, several topics exhibit consistently high and stable probabilities across most learners, whereas other topics display lower and more uneven probability distributions. To quantitatively support the patterns observed in Fig. 4, Table 4 summarizes the mean, variability, and range of predicted success probabilities for each topic. The results in Table 4 are consistent with the patterns shown in Fig. 4. Topics D and F have high average probabilities and very small variation across learners, which suggests that these topics are learned in a stable and consistent manner. By contrast, Topics H and L have noticeably lower average probabilities. In particular, Topic L shows similarly low values for all learners, indicating that it represents a clear learning risk.

Taken together, the heatmap visualization and topic-level statistics demonstrate that the proposed framework can effectively distinguish between well-learned topics and those that remain problematic. This separation is useful in practice, as it allows instructors to identify topics that require additional support and to focus instructional efforts where learning difficulties are most likely to occur.

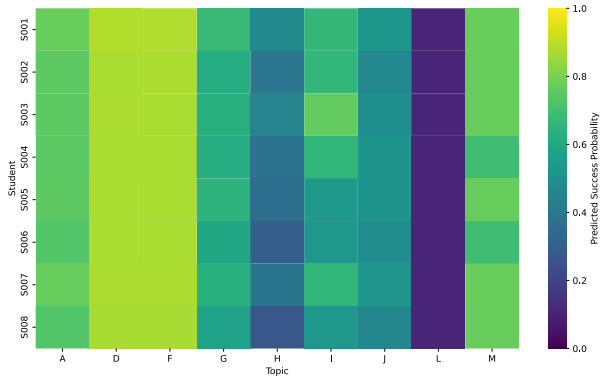

Fig. 4: Predicted success probabilities across learners and topics

Table 4: Topic-level summary of predicted success probabilities

Topic	Mean P(success)	Std	Min	Max	Interpretation
D	0.8750	0.0044	0.8683	0.8834	Well mastered
F	0.8750	0.0036	0.8690	0.8818	Well mastered
A	0.7500	0.0165	0.7275	0.7727	Stable
M	0.7500	0.0349	0.6925	0.7705	Stable
G	0.6250	0.0308	0.5773	0.6792	Moderate
I	0.6250	0.0825	0.5318	0.7604	Moderate (heterogeneous)
J	0.5000	0.0236	0.4631	0.5249	Borderline
H	0.3750	0.0690	0.2705	0.4734	At risk
L	0.1046	0.0000	0.1046	0.1046	High risk

Mean denotes the average predicted success probability for each topic, while Std reflects variability across learners

4.4. Validation of predictions and performance evaluation

This section examines the reliability and discriminative performance of the predicted success probabilities using quantitative metrics derived from real learning data. At the student–topic level, observed learning outcomes were aligned with predicted probabilities through item–topic mappings. Both classification-based and probability-based evaluation criteria were employed to obtain a comprehensive assessment. Classification performance was measured using Accuracy, Precision, Recall, and the Area Under the ROC Curve (AUC), while the agreement between predicted probabilities and observed outcomes was assessed using the Root Mean Square Error (RMSE). This combined evaluation framework captures not only predictive performance but also the stability and interpretability of probability estimates. The proposed model achieved a Recall of 0.92, indicating strong sensitivity in identifying unsuccessful learning cases. Such behavior is particularly important in educational contexts, where early detection of learning risk supports timely and targeted instructional intervention. An AUC value of 0.61 further indicates that the model distinguishes learning outcomes more effectively than random classification. The RMSE value of 0.55 suggests a moderate level of agreement between predicted and observed outcomes, while also reflecting stable probability estimates under limited data conditions. The comparatively lower Precision indicates a conservative prediction tendency, favoring risk detection over strict classification accuracy.

Overall, these results demonstrate that the proposed model produces predictions that are statistically consistent with observed learning data and effective for identifying topics associated with learning risk. Nevertheless, performance metrics alone do not fully explain the mechanisms underlying these outcomes. Accordingly, the following section provides a detailed discussion from both theoretical and practical perspectives, analyzing the influence of topic structure and knowledge propagation, and outlining the strengths and limitations of the proposed knowledge-structured approach.

The primary objective of this study was to examine whether automatically uncovering knowledge structures from real learning data and integrating them with machine learning–based prediction can provide a more interpretable and practically useful solution to the Knowledge Tracing problem. The results show that combining graph-based structure, knowledge propagation, and interpretable modeling offers clear advantages over approaches based solely on raw performance data.

With respect to H1, the findings indicate that knowledge propagation updates learner states in a structured manner based on topic relationships, rather than simply increasing raw scores. This leads to more balanced and realistic mastery estimates and is supported by improved AUC and lower RMSE compared to the baseline model.

For H2, the model consistently identifies topics with low mastery, forming stable learning risk patterns across learners. These patterns are less visible in raw performance data, where noise and sparsity can obscure meaningful relationships.

Regarding H3, the results show that the model outputs are interpretable and useful for decision support. Changes in topic mastery after propagation can be explained by the underlying topic structure, and the observed patterns align with instructional practices in the training environment.

Finally, for H4, the model produces stable probability estimates even under limited data conditions. The predictions remain balanced and avoid extreme values, making the approach suitable for real-world educational settings.

4.5. Limitations and future work

Despite its strengths, this study has several limitations. First, the dataset was relatively small and drawn from a specific educational context, which may limit generalizability. Second, topic

definitions and mappings were fixed in advance; different curricular designs may yield alternative structures. Third, while the framework avoids deep architectures to preserve interpretability, this choice may limit predictive performance in very large-scale settings.

Future work will focus on validating the framework across larger and more diverse datasets, exploring adaptive topic representations, and integrating temporal dynamics without sacrificing transparency. These directions provide opportunities to further strengthen the balance between interpretability, robustness, and predictive power.

In summary, the hypothesis-driven evaluation and explanatory case analysis jointly demonstrate that the proposed hybrid knowledge-structured framework provides an effective, interpretable, and practically deployable alternative to complex deep Knowledge Tracing models, particularly in real-world educational settings.

To support reproducibility, the implementation details of the proposed framework are fully described in the manuscript. Due to the nature of the dataset, anonymized data can be shared upon reasonable request, subject to institutional and privacy constraints.

5. Conclusion

This study proposes a knowledge-structured Knowledge Tracing framework that combines graph-based representations, knowledge propagation, and interpretable machine learning. The results show that this approach produces more stable and realistic estimates of learner knowledge than methods based only on raw performance data.

The findings indicate that knowledge propagation helps organize learner states in a more structured way, improves the identification of at-risk topics, and produces outputs that are easier to interpret for instructional decision-making. In addition, the model maintains stable probability estimates even when data are limited, making it suitable for real-world learning environments.

Overall, the proposed framework provides a transparent and computationally efficient alternative to more complex deep learning–based knowledge tracing models. Future work will focus on testing the approach on larger and more diverse datasets, incorporating temporal dynamics, and further improving its practical applicability.

List of abbreviations

AFM	Additive factors model
AUC	Area under the ROC curve
BKT	Bayesian knowledge tracing
DKT	Deep knowledge tracing
GAT	Graph attention networks
GCN	Graph convolutional network
GIKT	Graph-based interaction knowledge tracing
GNNs	Graph neural networks
ID	Identifier
IKT	Interpretable knowledge tracing
IRT	Item response theory
KG	Knowledge graph
KT	Knowledge tracing
LMS	Learning management systems
LSTM	Long short-term memory
ML	Machine learning
PFA	Performance factors analysis
RMSE	Root mean square error
ROC	Receiver operating characteristic

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Baker RS and Inventado PS (2014). Educational data mining and learning analytics. In: Larusson JA and White B (Eds.), Learning analytics: From research to practice: 61–75. Springer, New York, USA. https://doi.org/10.1007/978-1-4614-3305-7_4 **PMCid:PMC4355053**

Corbett AT and Anderson JR (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-adapted Interaction, 4: 253-278. https://doi.org/10.1007/BF01099821

Cui C, Yao Y, Zhang C, Ma H, Ma Y, Ren Z, Zhang C, and Ko J (2024). DGEKT: A dual graph ensemble learning method for knowledge tracing. ACM Transactions on Information Systems, 42(3): 78. https://doi.org/10.1145/3638350

Fawcett T (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8): 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

Gantulga B, Munkhtsetseg N, Garmaa D, and Batbayar S (2020). Using five principles of object-oriented design in the transmission network management information. In: Pan JS, Li J, Tsai PW, and Jain L (Eds.), Advances in intelligent information hiding and multimedia signal processing: 325–333. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-13-9714-1_36

Liu M, Shao P, Zhang K (2021). Graph-based exercise- and knowledge-aware learning network for student performance prediction. In: Fang L, Chen Y, Zhai G, Wang J, Wang R, and Dong W (Eds.), Artificial Intelligence: First CAAI International Conference, CICAI 2021, Hangzhou, China, June 5–6, 2021, Proceedings, Part I 1: 27–38. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-030-93046-2_3

Mikolov T, Chen K, Corrado G, and Dean J (2013). Efficient estimation of word representations in vector space. Arxiv Preprint Arxiv:1301.3781. https://doi.org/10.48550/arXiv.1301.3781

Nakagawa H, Iwasawa Y, and Matsuo Y (2019). Graph-based knowledge tracing: Modeling student proficiency using graph neural network. In the IEEE/WIC/ACM International Conference on Web Intelligence, ACM, Thessaloniki, Greece: 156-163. https://doi.org/10.1145/3350546.3352513 **PMid:31029866**

Namsraidorj M, Namsraidorj B, and Enkhtur A (2025). Ontology-based recommendation system using GitHub Classroom and Bookwidget. TEM Journal, 14(1): 861. https://doi.org/10.18421/TEM141-76

Pavlik PI, Cen H, and Koedinger KR (2009). Performance factors analysis: A new alternative to knowledge tracing. In the Proceedings of the 14th International Conference on Artificial Intelligence in Education, IOS Press, Amsterdam, The Netherlands: 531–538. https://doi.org/10.3233/978-1-60750-028-5-531

Piech C, Bassen J, Huang J, Ganguli S, Sahami M, Guibas LJ, and Sohl-Dickstein J (2015). Deep knowledge tracing. In the Proceedings of the Advances in Neural Information Processing Systems, Montreal, Canada: 505–513.

Tong S, Liu Q, Huang W, Hunag Z, Chen E, Liu C, Ma H, and Wang S (2020). Structure-based knowledge tracing: An influence propagation view. In the IEEE International Conference on Data Mining (ICDM), IEEE, Sorrento, Italy: 541-550. https://doi.org/10.1109/ICDM50108.2020.00063

Veličković P, Cucurull G, Casanova A, Romero A, Lio P, and Bengio Y (2017). Graph attention networks. ArXiv Preprint ArXiv:1710.10903. https://doi.org/10.48550/arXiv.1710.10903

Wang J and Zhang L (2021). Discriminant mutual information for text feature selection. In: Jensen CS et al. (Eds.), Database systems for advanced applications: 136–151. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-030-73197-7_9

Yang Y, Shen J, Qu Y, Liu Y, Wang K, Zhu Y, Zhang W, and Yu Y (2021). GIKT: A graph-based interaction model for knowledge tracing. In: Hutter F, Kersting K, Lijffijt J, and Valera I (Eds.), Machine learning and knowledge discovery in databases: 299–315. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-030-67658-2_18

Yeung CK and Yeung DY (2018). Addressing two problems in deep knowledge tracing via prediction-consistent regularization. In the Proceedings of the Fifth Annual ACM Conference on Learning at Scale, ACM, London, UK: 1-10. https://doi.org/10.1145/3231644.3231647

Yudelson MV, Koedinger KR, and Gordon GJ (2013). Individualized Bayesian knowledge tracing models. In: Lane HC, Yacef K, Mostow J, and Pavlik P (Eds.), Artificial intelligence in education: 171–180. Springer, Berlin, Germany. https://doi.org/10.1007/978-3-642-39112-5_18