Malaysian Journal of Computing, 10 (2): 2262-2272, 2025

eISSN: 2600-8238

TOWARDS A UNIFIED FRAMEWORK FOR KNOWLEDGE TRACING WITH GRAPH CONVOLUTIONAL AND NEURAL ARCHITECTURES

Yaxi Su[1], Mohamad Yusof Darus[2*], Norizan Mat Diah[3], Jinpeng Huang[4], and Azlin Ramli[5]

[1,2*,3,5]*Faculty of Computer and Mathematical Sciences.Universiti Teknologi MARA (UiTM), 40450, Shah Alam, Selangor*
[4]*Cloud Computing Technology and Application Department, Shandong Institute of Commerce and Technology, Shandong, China, 250103*
[1]suyaxi0301@163.com, [2]*yusof_darus@uitm.edu.my, [3]norizan289@uitm.edu.my, [4]hbaiwan2025@163.com, [5]azlin.ramli.study@gmail.com

ABSTRACT

This paper sets out to propose a unified theoretical framework for knowledge tracing (KT) that combines graph convolutional networks (GCNs) with neural sequence architectures in intelligent tutoring systems. While existing methods have achieved some success, they face limitations in modelling relational dependencies among concepts and the temporal progression of learner behaviour. Building on socio-constructivist views of knowledge as a network of relations and connectionist accounts of learning as adaptation over time, the framework integrates graph-based relational reasoning with sequence-based temporal modelling. The argument advanced here is that the integration offers interpretable representations of knowledge states while preserving predictive performance. The paper draws together recent developments in graph-enhanced KT and attention-based models and outlines design heuristics for scalable deployment. Key issues are identified, including computational cost, data sparsity, and explainability for classroom use. It is anticipated that the framework will inform the design of more systems and provide a tractable agenda for empirical validation across multiple domains and learner populations.

Keywords: *Adaptive Learning, Graph Convolutional Networks, Intelligent Tutoring Systems, Knowledge Tracing, Neural Sequence Models, Socio-Constructivist models.*

Received for review: 24-07-2025; Accepted: 26-09-2025; Published: 01-10-2025
DOI: 10.24191/mjoc.v10i2.7915

1. Introduction

Progress in intelligent tutoring system (ITS) technologies has highlighted the fundamental nature of knowledge tracing (KT) - modelling and predicting students' learning of knowledge components over time, as the basis of personalised instruction. KT facilitates adaptive interventions by inferring the unmeasured state of knowledge from observable learner interactions (Corbett & Anderson, 1994). Early KT models, such as Bayesian Knowledge Tracing (BKT), employed probabilistic methods to account for the temporal dynamics, but were based on the simplifying assumptions of independence among concepts and persistent learning rates (Khajah et al., 2016). More recently, algorithms based on deep learning and neural networks, such as recurrent neural networks (RNN) and attention mechanisms, capture sequential dependencies more flexibly, leading to better prediction performance (Pandey & Karypis, 2019). However, most such models model knowledge components in isolation and do not consider delicate dependencies between concepts, which characterise typical educational domains.

The paper is conceptual in that it addresses a fundamental need in the KT literature; there is currently no framework capable of modelling relational dependencies between knowledge components and student learning over time. For modelling learners' existing works, such as GCNs (graph convolutional networks), have exhibited effectiveness in using the topological structure of knowledge graphs (Ghosh et al., 2020). However, these are not yet integrated as part of a unified framework with (neural) sequence models. Such a disconnection hampers the potential of ITS to take full advantage of the structural and sequential properties of learning tasks.

This work is informed by socio-constructivist learning theories that view knowledge as relational and dynamically constructed through experience (Vygotsky, 1978), alongside connectionist principles in neural computation. Guided by socio-constructivist accounts of knowledge as an organised, relational structure and by connectionist views of learning as adaptive computation, we conceptualise learning as movement through a structured knowledge space. This movement is constrained by the space's topology, prior knowledge, and the temporal progression of learner interactions.

This paper presents a framework that unifies GCNs and neural architectures in KT, unifying graph-based relational reasoning with sequence-based temporal modelling. The main claim is that such a combination provides theoretically sound and methodologically sound learner models, allowing better learning interventions that are more accurate, interpretable and learner-targeted. By so doing, the current paper sets out to add to the current discussion on adaptive education technologies and serves to direct future empirical research on KT model development and validation.

Graph convolutional networks have an exceptional merit in this special case. By taking advantage of the structural information of knowledge graphs (concepts and relations), GCNs can improve the modelling of student knowledge states, thus predicting the learning effect well. Interfacing knowledge graph embeddings with neural architectures has been recently explored, offering a significant performance improvement, but particularly within educational scenarios (Li & Wu, 2023). Furthermore, the use of attention mechanisms in these systems can customise learning paths by incorporating expert knowledge into learning models (Tato & Nkambou, 2022).

The rapid development of artificial intelligence (AI) and machine learning (ML) technologies has created a vast opportunity for researchers to address various social, economic, and security issues more effectively. For instance, agriculture is a promising field about ML use to optimise crop production with better prediction by scale models. Fashoto et al. (2021) have also been able to construct a multiple linear regression model with high accuracy rates to forecast maize crop yields in Eswatini, thus assisting the country's food planning (Fashoto et al., 2021). On the other hand, the ML approach has also been applied to social issues like divorce among women of Malaysia by Aimran et al. (2022), where the Decision Tree (C5.0) was the most successful approach with an accuracy of nearly 78% (Aimran et al., 2022). The greatest predictors for divorce in this study included the wife's occupation, nature of marriage and ethnicity. From the perspective of cybersecurity, Sunardi et al. (2023) used RAT and data mining analysis to characterise victims and offenders of cyber fraud in Indonesia (Sunardi et al., 2023). This study revealed that female users with low security awareness are most at risk of becoming victims. Overall, these three studies demonstrate how the integration of ML and sociological theory can be applied to understand complex phenomena, improve policy effectiveness, and have a positive impact on society.

2. Literature Review

Knowledge tracing (KT) is a classical problem in educational technology, also known as Bayesian knowledge tracing (BKT), a statistical model first proposed by Corbett and Anderson (1994). BKT models student knowledge as an unobserved binary variable that evolves with observed performance. Though BKT is very intuitive and understandable, it has a strong assumption (independent components of knowledge, a constant learning probability, etc.), which makes it less flexible in modelling complex learning processes. Some of these limitations were addressed by extensions, but BKT still is not enough to describe the subtle, dynamic nature of learners in the real world (Khajah et al., 2016). This restriction also requires more powerful models that might encode temporal dependence and relational structure.

Expanding on these criticisms, Piech et al. (2015) proposed DKT, which uses RNNs to model the sequence of student responses. Their finding showed that the RNN-based models performed much better than the BKT in prediction accuracy. DKT was a watershed model in that it framed KT as a sequence modelling problem, which let it tap into the representational power of neural networks to model complex interactions observed in student response data. Later research revealed significant limitations of DKT, such as poor interpretability as well as ignoring domain knowledge, in which the hierarchy and association among concepts will be considered (Yeung & Yeung, 2018). This gap indicates that a combined model of sequential dynamics and the structure of knowledge is needed.

A graph-based method appeared to overcome the lack of relationship information. Ghosh et al. (2020) have presented Context-Aware Attentive Knowledge Tracing (CAKT) for practical purposes by extending the relationship of concepts included in KG to include the writing process. CAKT demonstrated that graph structures capture the learning spaces for predicting the next knowledge states of learners more accurately and representing the knowledge states more readably. This aligns with socio-constructivist models of knowledge as the interconnected network of concepts (Vygotsky, 1978). However, CAKT still views graph reasoning and sequence modelling as two disconnected parts, which are heuristic instead of principled. The absence of theoretical unification renders such models hard to extrapolate.

Similarly, Nakagawa et al. (2021) proposed Graph Enhanced Knowledge Tracing (GKT), an approach that utilises a GCN to propagate information over an existing knowledge graph, aiming to capture structural dependencies between concepts to generalise beyond the training set. GKT makes the embeddings more interpretable and is consistent with the structure of the domain; however, it does not capture the temporal nature of the learning sequence straightforwardly and is not as suitable when the order of interactions has a strong impact on the outcomes (Shen et al., 2024). On the other hand, neural sequence models generally focus on the temporal structure but tend to under-represent concept–concept relations (Ke et al., 2024). This also reinforces the intuition that the learning trajectories are too complicated to be adequately modelled by either graph-based or purely neural sequential models, calling for a joint relational and predictive design (Wu et al., 2022; Shen et al., 2024).

Hybrid approaches have also been the focus of more recent studies. Zhang et al. (2022) introduced Sequence Graph Knowledge Tracing (SGKT) to combine these two types of knowledge (where a Graph Convolutional Network (GCN) would model knowledge structure and a Transformer higher-order feature dependencies in sequential knowledge productions). Their results indicate that a joint utilisation of graph reasoning and sequence modelling can be more effective than using either alone. But they have little theoretical backing other than empirical observations, very little in the way of theoretical modelling suggesting how or why these different aspects should, in theory, come together. Because of this, it constitutes an important step towards integration, but it raises questions about its generalisation, scalability, and theoretical support.

Combined, these studies yield precious information but demonstrate important deficiencies. BKT and its extensions emphasize interpretability but concern only simple dependencies (Corbett & Anderson, 1994; Khajah et al., 2016). DKT and its descendants (e.g., Pandey & Karypis, 2019) optimize for predictive power but not for structure and transparency. Graph-based models (Ghosh et al., 2020) take domain knowledge into consideration; however, they are not capable of capturing sequential dynamics in their entirety. There are some promising hybrid models (Wu et al., 2022), which are nevertheless under-theorised. The absence of a coherent theoretical or methodological structure to organize the relational and the temporal dimensions of learning, is a bottleneck for advancing knowledge in the field.

This is a conceptual piece that is rooted in the theoretical grounding of social-constructivist learning (Vygotsky, 1978) and connectionist principles to support an integrated framework that formalises the interplay between graph convolutional and neural architectures for KT. This framework aims to inspire the creation of more adaptive, interpretable, and effective ITSs by articulating a principled synthesis of graph-based relational reasoning and sequence-based temporal modelling.

Table 1. Comparative Summary of Existing KT Models

Model	Theoretical Basis	Strengths	Limitations	Key References
Bayesian Knowledge Tracing (BKT)	Probabilistic	Simple, interpretable	Assumes independence, fixed rates	Corbett & Anderson (1994)
Deep Knowledge Tracing (DKT)	Connectionist	Captures temporal dynamics	Poor interpretability, ignores structure	Piech et al. (2015)
Context-Aware Attentive KT (CAKT)	Socio-constructivist + Attention	Incorporates graph	Heuristic integration, weak theory	Ghosh et al. (2020)
Sequence Graph KT (SGKT)	Hybrid	Combined strengths	Under-theorised	Wu et al. (2022)

Table 1 summarises the key characteristics, strengths, and limitations of representative KT models discussed in the literature, highlighting the lack of a unified approach that integrates relational and temporal learning dimensions.

Recent developments in knowledge tracing (KT) have considered incorporating more complex relational and behavioural cues in predictive student learning state models. Hiromi et al. (2021) introduced a graph-based knowledge tracing model with graph neural networks (GNNs) to represent the latent graph structures of student coursework and reformulate KT into a time-series node classification problem (Hiromi et al., 2021). Their approach outperformed the state of the art in terms of prediction accuracy and interpretability on two popular benchmark datasets while not assuming the presence of explicit knowledge graphs. Similarly, Qiang et al. (2022) proposed a neural Turing machine-based skill-aware KT (NSKT) model that models the latent relevance between conjunctive skills in questions (Qiang et al., 2022). Empirical results on three real datasets showed that NSKT-based models achieved better prediction and interpretation performance compared to previous deep KT models and discovered semantically meaningful conditional dependencies over concepts. Consistent with these studies, Zhang et al. (2022) proposed an attention-augmented encoder-decoder KT model that incorporated three types of learning processes, end and interval, using both multiheaded and channel attention mechanisms (Zhang et al., 2022). Their model improves proneness and learnability, and acquisitions of aspects are effectively verified and shaped through an intelligent learning system plugged into an offline courseware in computer and English. Cumulatively, these studies point to the prospect of integrating relational, skill-relevance, and multi-behavioural information via sophisticated neural architectures to enhance KT efficacy and generalizability in education.

3. Theoretical and Conceptual Framework

The unified KT framework in ITS was proposed to consist of the GCNs and Neural Sequence Architectures, which were underpinned by two learning theories, social constructivist learning and connectionist learning. Collectively, these theories provide a basis for understanding learning as a process that is both dynamic, patterned and contextually dependent.

According to socio-constructivist theory (Vygotsky, 1978), knowledge is created in the encounter between the individual and the social and material context – it is relational from the outset. In KT, this view stresses the need for modelling and capturing the interrelationships among knowledge components (skills or concepts), which usually form a hierarchy or a network with nodes and edges, respectively, in a curriculum. This relational orientation also corresponds to knowledge graphs as representations of domain knowledge, where GCNs benefit from spreading information through interlinked nodes to characterise the development of learners' mastery over a network of knowledge (Ghosh et al., 2020).

In contrast, connectionist learning theory (the foundation of multilayer neural networks) characterises learning as the continuous modification of distributed representations following exposure to streams of stimuli (Rogers et al., 2014). This view motivates the development of RNNs and Transformer-based models in KT (Pandey & Karypis, 2019), in which the history of student responses to questions is used to predict future performance. Neural sequence models are effective at capturing the temporal dynamics of learning behaviours but are less capable of encoding explicit relational knowledge.

The conceptual framework proposed in this paper synthesises these complementary perspectives into a unified KT model. GCNs serve to model the relational structure of the knowledge domain, informed by socio-constructivist principles, while neural sequence models capture the temporal learning dynamics, aligned with connectionist principles. The interaction between these two components enables the learner model to reason over both the graph-structured knowledge space and the sequential trajectory of student interactions.

Table 2: Theoretical Alignment of Framework Components

Framework Component	Learning Theory	Computational Method
Relational Knowledge	Socio-constructivist (Vygotsky, 1978)	Graph Convolutional Networks
Sequential Behaviour	Connectionist (Rogers et al., 2014)	RNNs / Transformers

Table 2 delineates the alignment between the theoretical foundations and the computational components of the proposed framework, illustrating how each learning theory informs specific modelling strategies.

Application to the Research Problem

Existing KT models often focus either on sequential dynamics (e.g., DKT) or relational structure (e.g., GKT), but rarely both in a theoretically principled manner (Wu et al., 2022). This disjoint treatment leads to models that fail to fully exploit the rich, structured, and temporal nature of learning. The unified framework conceptualises learning as a process of navigating and updating a structured knowledge graph over time, enabling ITS to generate more accurate, interpretable, and adaptive recommendations. Grounded in existing learning theories, it not only provides methodological consistency but also theory-driven rigour that fills current gaps and steers future empirical studies of integrated KT models.

Unified Theoretical-Conceptual Framework

Below is a schematic representation of the proposed framework.

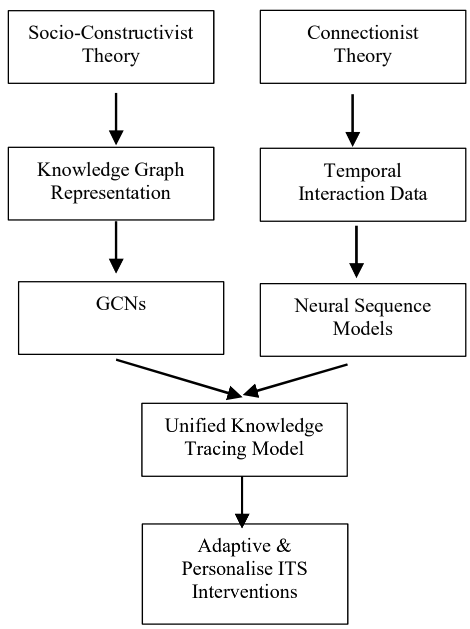

Figure 1. Proposed Framework.

The figure illustrates the framework integrating socio-constructivist (relational) and connectionist (temporal) principles: a GCN component encodes concept relations; a neural sequence component models learner trajectories; the fusion layer supports interpretable, personalised interventions.

4. Discussion and Implications

The conceptualisation frame for embedding Graph Convolutional Networks (GCN) and neural sequence architecture into the knowledge tracing (KT) contributes to empirical and applied understandings in learning theory and ITS design. Theoretical implications of the framework, its strengths and weaknesses, and possible educational applications are then considered.

Theoretical Contributions

The model extends the theoretical foundations of KT by integrating two traditions in the science of learning: the socio-constructivist conception of knowledge as an organised, interconnected network (Vygotsky, 1978) and the connectionist view of learning as a dynamic, adaptive process (Rogers et al., 2014). Current models often emphasise one view at the expense of the other. Graph-based KT effectively uses relational dependencies among concepts but tends to under-represent the temporal dynamics of learning (Ghosh et al., 2020; Wu et al., 2022). By contrast, neural sequence models capture the dynamics of the learning process **while ignoring the relations between concepts** (Ke et al., 2024). By explicitly unifying these standpoints, the proposed framework characterises learning as a dynamic walk over a structured concept graph through time, yielding representations that are both predictive and interpretable (Shen et al., 2024; Cheng et al., 2024). This view aligns with contemporary educational psychology, which emphasises the structure and development of learner knowledge (Shute & Rahimi, 2021). The integration provides a theoretical base for the findings by providing a plausible account for how learners could process both the conceptual and temporal relations in parallel when learning. It also addresses the need for a KT model that is both predictive and interpretable.

Strengths of the Framework

The framework exploits the synergies between graph neural networks and neural sequence models. Graph components capture prerequisite and co-occurrence relations of concepts, leading to better structural awareness and interpretable feedback. Sequence features describe order and timing and are used as temporal indicators of learner behaviours during and within sessions. The argument advanced here is that a stripped-down fusion can provide valid predictions and classroom-ready explanations. This design aligns with recent findings on session-graph KT, hierarchical transformers, and attention-based interpretability, and it can be extended to dynamic graphs when relations evolve (Wu et al., 2022).

The framework of this section gives a theoretical motivation to hybrid KT models, which have recently been proposed in the empirical literature, such as Wu et al. (2022). Furthermore, by detailing the theoretical basis for integration, it encourages methodological consistency and lays the groundwork for more conceptually driven model development and application in future studies.

Limitations and Challenges

This framework has strengths and constraints. A first constraint is computational cost. Training graph neural networks on large or dense concept graphs is memory-hungry because message passing expands neighbourhoods layer by layer, and intermediate activations and edge indices must be stored on the device. Full-graph or poorly sampled batches can exceed a single GPU's memory; dynamic-graph updates add further overhead for time encoders and evolving edges (Gao et al., 2024; Cheng et al., 2024). For the sequence component, standard self-attention scales quadratically with sequence length, which increases both memory traffic and run-time; more efficient kernels (e.g., Flash Attention) reduce memory reads/writes and help stabilise training at longer lengths (Tay et al., 2022; Dao et al., 2022). In practice, mini-batch subgraph or neighbour sampling and adaptive sampling strategies can control peak memory while preserving accuracy (Shen et al., 2024; Younesian et al., 2023).

A second constraint concerns graph availability and quality. Interpretability improves when domain knowledge is encoded as concept graphs, yet high-quality educational knowledge graphs are uneven across subjects and costly to curate; recent work shows value in programming education but highlights the dependency on graph coverage (Qu et al., 2024).

A third constraint is external validity. Logical coherence does not guarantee field efficacy. The framework requires verification in live ITS across multiple subjects and learner profiles, with reports on compute, latency, and explanation usefulness for teachers (Shen et al., 2024).

Taken together, these limitations suggest specific mitigations: (i) neighbour/subgraph sampling or dynamic graph training to bound memory; (ii) efficient attention for long sequences; and (iii) staged pilots that track both predictive gains and operational costs in classroom settings (Gao et al., 2024; Dao et al., 2022).

Potential Applications

There are several attractive applications of the unified framework. ITS developers can take advantage of the framework to build systems that predict learner performance more precisely and, simultaneously, deliver more relevant, context-based feedback and intervention. For instance, by pinpointing not only which concepts a learner is struggling with but also how those concepts relate to each other and how the learner's performance has changed over time, instructors and systems can develop more targeted remediation (Rahimi & Shute, 2021).

Furthermore, the framework can support the design of learning analytics dashboards that visualise the relational and temporal aspects of learners' progress, providing actionable feedback for teachers and students. Beyond individual learning, the framework's capacity to model collective learning trajectories may also support collaborative learning environments where group dynamics and shared knowledge structures play a critical role.

The general nature of the proposed approach makes an important contribution to KT theory by integrating relational with temporal modelling in a logical theoretical framework. It takes the dual advantage of GCNs and neural sequence models to overcome some constraints of current approaches, including interpretability, adaptability, and predictive power. Despite continuing challenges, related to complexity, data quality, and empirical testing, the framework provides a strong basis for furthering ITS design and KT research. The framework should be further developed, empirically tested and validated in real educational systems for a diverse group of learners.

Future Work

Empirical validation in the wild.

The next step is **empirical validation in real ITS deployments, across multiple domains and learner populations**. Studies should compare the unified model against leading sequence-only and graph-only baselines on accuracy, calibration, and robustness, and should assess explanation usefulness for teachers (Wu et al., 2022; Ke et al., 2024; Shen et al. 2024).

Efficiency and scaling.

Large concept graphs and long histories raise computational and memory load. Use neighbour/subgraph sampling for GNNs, sequence truncation or efficient attention for long inputs, and knowledge distillation for compact deployments. Report a simple resource envelope (GPU/CPU, peak memory, throughput) to make trade-offs explicit (Cheng et al., 2024; Dao et al., 2022; Younesian et al., 2023).

Adaptive knowledge graphs.

Move beyond static, expert-defined graphs. Incorporate **dynamic graph updates** driven by live interactions and curriculum change, with periodic pruning to control sparsity and drift. Where domain KGs are scarce, semi-automatic construction pipelines can bootstrap workable graphs (Cheng et al., 2024; Qu et al., 2024).

Explainability and usability.

Deliver teacher-facing views that attribute predictions to both **relational** (concept nodes/edges) and **temporal** (events/sessions) signals. Evaluate whether these explanations improve intervention decisions without harming performance (Lu et al., 2023).

Generalisability and equity.

Test across subjects (e.g., mathematics, programming, language), institutions, and demographic groups. Report subgroup performance and fairness diagnostics to ensure benefits are broadly shared (Shen et al., 2024).

Acknowledgement

The authors gratefully acknowledge the research support provided by the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia

Funding

The author(s) received no specific funding for this work.

Author Contribution

Author 1, Author 2 and Author 5 collaborated on crafting the literature review and supervising the article writing process. For the research methodology, Author 1, Author 2, Author 3, Author 4 and Author 5 collectively contributed. The analysis and interpretation of results were undertaken by Author 1, Author 2 and Author 5.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

Aimran, N., Rambli, A., Afthanorhan, A., Mahmud, A., Sapri, A., & Aireen, A. (2022). Prediction of Malaysian women divorce using machine learning techniques. *Malaysian Journal of Computing, 7*(2), 1067–1081. https://doi.org/10.24191/mjoc.v7i2.17077

Cheng, K., Peng, L., Wang, P., Ye, J., Sun, L., & Du, B. (2024). Dygkt: dynamic graph learning for knowledge tracing. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 409-420. https://doi.org/10.1145/3637528.3671773

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*(4), 253–278. https://doi.org/10.1007/BF01099821

Dao, T., Fu, D. Y., Ermon, S., Rudra, A., & Ré, C. (2022). Flashattention: fast and memory-efficient exact attention with io-awareness.. https://doi.org/10.48550/arxiv.2205.14135

Fashoto, S. G., Mbunge, E., Ogunleye, G., & Van den Burg, J. (2021). Implementation of machine learning for predicting maize crop yields using multiple linear regression and backward elimination. *Malaysian Journal of Computing, 6*(1), 679–697. https://mjoc.uitm.edu.my/main/images/journal/vol6-1-2021/4-Fashoto-et-al-Vol-61.pdf

Gao, S., Li, Y., Shen, Y., Shao, Y., & Chen, L. (2024). ETC: Efficient Training of Temporal Graph Neural Networks over Large-Scale Dynamic Graphs. *PVLDB, 17*(5), 1060–1072. https://doi.org/10.14778/3641204.3641215

Ghosh, S., Heffernan, N., & Lan, A. S. (2020). Context-aware attentive knowledge tracing. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2330–2339). ACM. https://doi.org/10.1145/3394486.3403307

Hiromi, N., Iwasawa, Y., & Matsuo, Y. (2021). Graph-based knowledge tracing: modeling student proficiency using graph neural networks. Web Intelligence, 19(1-2), 87-102. https://doi.org/10.3233/web-210458

Ke, F., Wang, W., Tan, W., Du, L., Jin, Y., Huang, Y., … & Yin, H. (2024). Hitskt: a hierarchical transformer model for session-aware knowledge tracing. Knowledge-Based Systems, 284, 111300. https://doi.org/10.1016/j.knosys.2023.111300

Khajah, M., Lindsey, R., & Mozer, M. C. (2016). How deep is knowledge tracing? https://doi.org/10.48550/arxiv.1604.02416

Li, W. and Wu, G. (2023). One-shot based knowledge graph embedded neural architecture search algorithm. Frontiers in Computing and Intelligent Systems, 3(3), 1-5. https://doi.org/10.54097/fcis.v3i3.7982

Lu, Y., Tong, L., & Cheng, A. (2023). Advanced knowledge tracing: incorporating process data and curricula information via an attention-based framework for accuracy and interpretability. https://doi.org/10.31219/osf.io/3k7z2

Nakagawa, H., Iwasawa, Y., & Matsuo, Y. (2021). Graph-Based Knowledge Tracing: Modelling Student Proficiency Using Graph Neural Networks. Web Intelligence, 19(1–2), 87–102. https://doi.org/10.3233/WEB-210458

Pandey, S., & Karypis, G. (2019). A self-attentive model for knowledge tracing. *arXiv*. https://arxiv.org/abs/1907.06837

Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. In *Advances in Neural Information Processing Systems* (Vol. 28). Curran Associates. https://papers.nips.cc/paper/2015/hash/bac9162b47c56fc8a4d2a519803d51b3-Abstract.html

Qu, K., Li, K. C., Wong, B. T. M., Wu, M. M., & Liu, M. (2024). A survey of knowledge graph approaches and applications in education. Electronics, 13(13), 2537. https://doi.org/10.3390/electronics13132537

Qiang, H., Su, W., Sun, Y., Huang, T., & Shi, J. (2022). Ntm-based skill-aware knowledge tracing for conjunctive skills. Computational Intelligence and Neuroscience, 2022, 1-16. https://doi.org/10.1155/2022/9153697

Rahimi, S., & Shute, V. J. (2021). Learning Analytics Dashboards in Educational Games. In S. Ifenthaler & M. Khalil (Eds.), *Visualizations and Dashboards for Learning Analytics* (pp. 527–546). Springer. https://doi.org/10.1007/978-3-030-81222-5_24

Rogers, T. T. and McClelland, J. L. (2014). Parallel distributed processing at 25: further explorations in the microstructure of cognition. Cognitive Science, 38(6), 1024-1077. https://doi.org/10.1111/cogs.12148

S. Shen *et al*., "A Survey of Knowledge Tracing: Models, Variants, and Applications," in *IEEE Transactions on Learning Technologies*, vol. 17, pp. 1858-1879, 2024, http://doi.org/10.1109/TLT.2024.3383325

Sunardi, S., Effendy, M., & Fadhillah, A. (2023). Cyber fraud profiling with routine activity theory using data mining techniques. *Malaysian Journal of Computing, 8*(1), 529–546. https://doi.org/10.24191/mjoc.v8i1.20296

Tay, Y., Dehghani, M., Bahri, D., & Metzler, D. (2022). Efficient transformers: a survey. ACM Computing Surveys, 55(6), 1-28. https://doi.org/10.1145/3530811

Tato, A. and Nkambou, R. (2022). Infusing expert knowledge into a deep neural network using attention mechanism for personalized learning environments. Frontiers in Artificial Intelligence, 5. https://doi.org/10.3389/frai.2022.921476

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wu, Z., Huang, L., Huang, Q., Huang, C., & Tang, Y. (2022). SGKT: Session Graph-Based Knowledge Tracing for Student Performance Prediction. *Expert Systems with Applications, 206*, 117681. https://doi.org/10.1016/j.eswa.2022.117681

Yeung, C.-K., & Yeung, D.-Y. (2018). Addressing Two Problems in Deep Knowledge Tracing via Prediction-Consistent Regularization. *arXiv:1806.02180*. https://arxiv.org/abs/1806.02180

Younesian, T., Daza, D., van Krieken, E., Thanapalasingam, T., & Bloem, P. (2023). Grapes: Learning to sample graphs for scalable graph neural networks. *arXiv preprint arXiv:2310.03399*. https://doi.org/10.48550/arXiv.2310.03399

Zhang, K., Qin, Z., & Kuang, Y. (2022). Knowledge Tracing via Attention-Enhanced Encoder-Decoder. *Computational Intelligence and Neuroscience, 2022*, 1552745. https://doi.org/10.1155/2022/1552745