IUSE: Advancing Deep Knowledge Tracing and Learning Path Recommendation to Enhance STEM Education at HBCUs

Afsana Nasrin, Prairie View A&M University
Christian Onyeke

Mr. Christian Onyeke is a Graduate Research Assistant and Ph.D. student in Electrical and Computer Engineering at Prairie View A&M University, Prairie View, Texas. My research focuses on Deep Knowledge Tracing (DKT), where I develop interpretable and trustworthy AI models to improve personalized learning and student performance predictions. Christian holds a Master's Degree in Data Analytics from the University of Houston, Downtown, a Master's Degree in Petroleum Engineering from the University of Louisiana at Lafayette, and a Bachelor's Degree in Mechanical Engineering from the University of Nigeria, Nsukka (UNN). With a strong interdisciplinary background, I combine data science, engineering, and AI to design adaptive learning systems that promote fairness, explainability, and effective educational interventions.

Uchechukwu Okechukwu, Prairie View A&M University
Dr. Yujian Fu P.E., Alabama A&M University

Dr. Yujian Fu is an associate professor of computer science department at Alabama A&M University. Her research interests fall in formal verification of cyber physical systems, behavioral analysis of mobile security, software architecture and design analys

Lijun Qian
xiangfang Li, Prairie View A&M University

Dr. Xiangfang Li is an associate professor in the Department of Electrical and Computer Engineering at Prairie View A&M University. Her research interests encompass computational biology, computer networking and communications, and machine learning and AI.

Prof. Xishuang Dong, Prairie View A&M University

Xishuang Dong is Associate Professor of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. His research interests include AI for education, natural language processing, computer systems biology, and Internet of Things.

IUSE: Advancing Deep Knowledge Tracing and Learning Path Recommendation to Enhance STEM Education at HBCUs

Abstract

Historically Black Colleges and Universities (HBCUs) play a vital role in expanding access to STEM education for underrepresented students but face challenges such as low retention and graduation rates that call for data-driven interventions. This project advances Deep Knowledge Tracing (DKT) research by improving performance in low-resource environments, interpretability, and real-world applicability to personalize learning at HBCUs. Through semi-supervised and ensemble DKT models, unlabeled student data are effectively leveraged, while tabular generative AI augments datasets to enhance prediction accuracy. To address the "black-box" issue of DKT, SHAP (Shapley Additive Explanations) and large language models (LLMs) are integrated to provide transparent, human-understandable explanations of learning predictions. Complementing this, ensemble learning with Item Response Theory (IRT) improves interpretability by linking model outputs to student ability and item difficulty. Furthermore, learning path recommendation is reframed as a multi-task sequence prediction problem, enabling shared learning representations and non-redundant, individualized study trajectories. Together, these innovations strengthen transparency and performance in STEM education at HBCUs, laying a foundation for AI-driven systems that equitably support student success, with support from the NSF Improving Undergraduate STEM Education (IUSE) program.

Introduction

Historically Black Colleges and Universities (HBCUs) play a pivotal role in broadening access to higher education for underrepresented students in STEM disciplines. However, challenges such as lower retention and graduation rates underscore the need for advanced, data-driven educational interventions [1]. These outcomes are largely associated with systemic, non-unique factors, such as higher proportions of first-generation and low-income students, and limited access to personalized academic support, that are also observed at many minority-serving and public institutions. As a result, solutions developed in the HBCU context are broadly applicable to other institutions facing similar student success challenges. Deep Knowledge Tracing (DKT) has emerged as a powerful approach for modeling students' knowledge states and predicting learning outcomes, thereby enabling timely and targeted support [2]. The study of this project advances the state of DKT research by addressing critical limitations in interpretability, low-resource environments, and real-world applicability, while also enhancing learning path recommendations to personalize STEM education at HBCUs.

HBCUs often face constraints in collecting and annotating large-scale educational data, which can hinder the performance of data-hungry deep models. To the best of our knowledge, we build the first dataset for knowledge tracing and learning path recommendation at HBCUs with labels indicating students' likelihood of passing (1) or failing (0) courses [15, 16]. To further mitigate the data-hungry issues, research on semi-supervised learning and ensemble DKT models demonstrates how unlabeled student interactions can be effectively leveraged to enhance model robustness in low-resource settings [3]. In addition, the use of tabular generative AI to create synthetic student data has been shown to augment real datasets, improving learning outcome predictions [4]. These innovations underscore the potential of generative and semi-supervised strategies to empower institutions with limited resources, ensuring that predictive models remain both effective and equitable. Moreover, a major barrier to deploying DKT in practice is its “black-box” nature, which limits trust among educators and students. To address this, we integrate SHAP (Shapley Additive Explanations) with DKT to quantify the influence of individual features, while leveraging large language models (LLMs) to generate natural-language explanations of model predictions. Another complementary effort combines ensemble learning with Item Response Theory (IRT) to improve interpretability in tracing student learning outcomes [5]. These approaches not only improve predictive accuracy but also provide transparent, actionable insights into student abilities, item difficulty, and performance trends, bridging the gap between advanced analytics and classroom usability.

Beyond tracing student outcomes through advanced DKT, this study also advances learning path recommendation by framing it as a multi-task sequence prediction problem [6]. By sharing representations between DKT and recommendation tasks, and by introducing a non-repeat penalty to avoid redundant recommendations, the proposed multi-task model produces more effective, individualized learning trajectories.

The following sections present these studies from two perspectives: deep knowledge tracing and learning path recommendation, both of which have strong potential to significantly improve the lower retention and graduation rates at HBCUs.

Deep Knowledge Tracing

1. Semi-supervised Learning

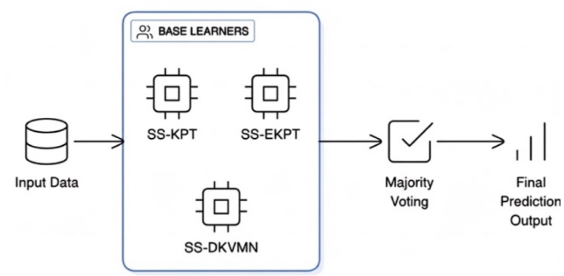

Figure 1. Diagram of the proposed semi-supervised DKT through introducing ensemble learning with majority voting

Due to strict privacy protections in educational applications, collecting large-scale, high-quality student interaction data from real-world learning environments is highly limited, especially for underserved learners in low-resource institutions. For example, the widely used ASSISTments dataset contains records from only about 4,000 students, making it insufficient for training DKT models without risking overfitting. To address the limitations of low-resource DKT, this study integrates semi-supervised learning with ensemble deep learning [3] to develop a novel semi-supervised DKT framework (Figure 1). The method first trains multiple semi-supervised (SS) DKT variants, including Knowledge Proficiency Tracing (KPT) [7], Exercise-Correlated KPT (EKPT), and Dynamic Key-Value Memory Networks (DKVMN) [8], using both the limited labeled interactions and a large pool of unlabeled data. These models including SS-KPT, SS-EKPT, and SS-DKVMN, are subsequently ensembled through majority voting, effectively combining the strengths of semi-supervised training and the bagging strategy. The proposed approach significantly outperforms baselines across all settings. It achieves the highest AUC (0.8601), Accuracy (0.8512), Precision (0.8294), and Recall (0.8595). These results highlight the effectiveness of integrating semi-supervised learning and ensemble strategies in enhancing low-resource deep knowledge tracing.

2. Data Augmentation

This study leverages generative AI to further advance DKT by addressing the persistent challenge of limited student interaction data. Generative models, widely used in fields such as natural language processing and computer vision, are capable of producing high-quality synthetic data and have proven effective in mitigating data scarcity. Building on this capability, the study applies TabDDPM [9], a diffusion-based generative model, to synthesize realistic educational interaction records and augment the training set for DKT framework. Extensive experiments on the ASSISTments datasets demonstrate the effectiveness of this approach in improving DKT performance.

TABLE I: Comparing performance between DKT and the proposed method (DKT + TabDDPM) [4].

Model	Accuracy (%)	AUC (%)	Precision (%)	Recall (%)
DKT	58.38±0.98	54.59±1.27	60.45±0.56	86.04±3.70
DKT + TabDDPM	**63.46±0.48**	**66.00±0.85**	**63.08±0.44**	**92.40±1.06**

For instance, Table I presents a performance comparison between the baseline DKT model and the proposed DKT augmented with TabDDPM. The proposed method outperforms the baseline across accuracy, AUC, precision, and recall. In other words, the synthetic samples generated by TabDDPM substantially boost AUC, recall, and accuracy, showing an especially strong effect in reducing false negatives. More results and discussions were published in [4].

3. Interpretable AI

3.1 Integrating Ensemble Learning with Item Response Theory to Improve the Interpretability of Student Learning Outcome Tracing

Student learning outcome (SLO) tracing seeks to monitor learners' progress by predicting their likelihood of passing or failing courses using DKT techniques. However, traditional DKT models generally lack interpretability, limiting their usefulness in educational environments where transparent and explainable decision-making is essential. To address this challenge, this study aims to improve both the interpretability and predictive performance of SLO tracing models. The approach integrates ensemble learning with Item Response Theory (IRT), a well-established statistical framework that relates a learner's latent ability to the probability of correctly answering an item [10]. In particular, the DeepIRT model incorporates IRT into deep neural networks, enabling the capture of temporal learning patterns while preserving interpretable parameters such as item difficulty and student ability [11]. In this work, as shown in Figure 2, multiple IRT-based interpretable DKT models were developed and combined using a bagging ensemble strategy to enhance generalization and provide a more holistic representation of students' knowledge states. Through aggregated predictions, the ensemble not only improves accuracy but also supports interpretability via visualizations of student ability estimates, item difficulty levels, and predicted performance probabilities [12]. The

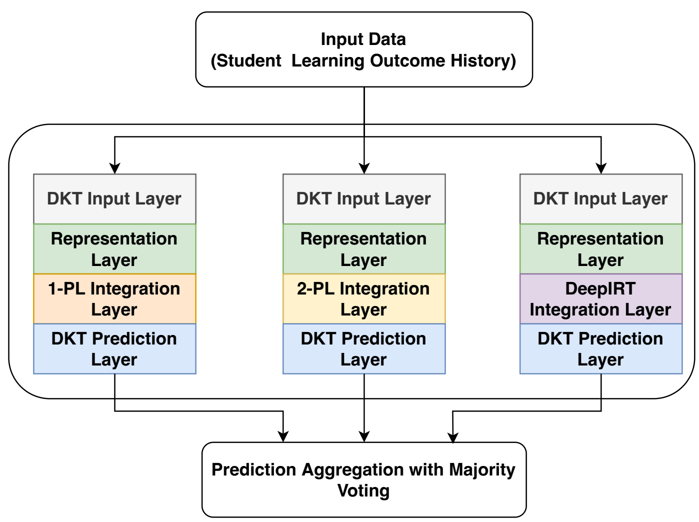

Figure 2. Diagram of the proposed method implementation.

proposed framework was evaluated using an SLO tracing dataset from Prairie View A&M University (PVAMU), a HBCU.

Experimental results show that the proposed ensemble IRT-based DKT method significantly improves interpretability and predictive performance compared with DeepIRT. Across metrics, the ensemble model maintains a more balanced performance, as observed in the Computer Science (CSC) department where it achieves a higher F1-score than DeepIRT (0.9084 vs. 0.9075). These improvements result from more accurate ability estimation and difficulty calibration, leading to smoother and more stable predicted probabilities. Overall, the ensemble framework provides clearer insights into learning patterns and more comprehensive knowledge-state representations than single-model approaches.

3.2 Integrating SHAP and LLMs to Enhance Interpretability of DKT

This study proposed a novel framework that enhances the interpretability of DKT by integrating SHapley Additive exPlanations (SHAP) [13] to quantify the contribution of individual features to model predictions. In addition, we leverage Large Language Models (LLMs) [14] to translate these quantified contributions into structured natural language explanations. These explanations offer educators and learners transparent insights into the features influencing predictions and the reasoning behind model decisions. To validate the effectiveness of the proposed approach, we conduct extensive experiments analyzing feature contributions in DKT using SHAP, alongside evaluations of the interpretability provided through LLM-generated explanations. Experimental results demonstrate that combining SHAP with LLMs produces context-aware and accessible explanations.

Learning Path Recommendation

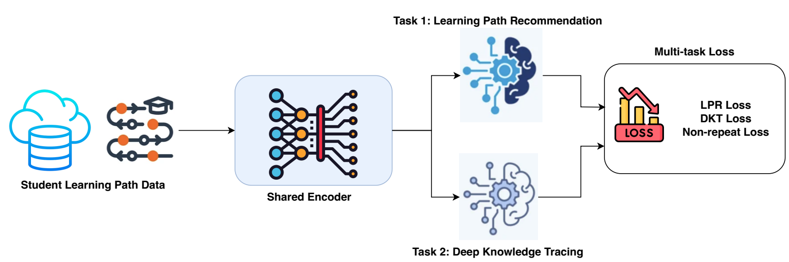

Figure 3. Diagram of the proposed multi-task learning path recommendation.

Personalized learning is a student-centered educational approach that adapts content, pace, and assessment to meet each learner's unique needs. As the key technique to implement the personalized learning, learning path recommendation sequentially recommends personalized learning items such as lectures and exercises. Advances in deep learning, particularly deep reinforcement learning, have made modeling such recommendations more practical and effective. This study proposes a multi-task LSTM model that enhances learning path recommendation, as shown in Figure 3, by leveraging shared information across tasks. The approach reframes learning path recommendation as a sequence-to-sequence (Seq2Seq) prediction problem, generating personalized learning paths from a learner's historical interactions. The model uses a shared LSTM layer to capture common features for both learning path recommendation and deep knowledge tracing, along with task-specific LSTM layers for each objective. To avoid redundant recommendations, a non-repeat loss penalizes repeated items within the recommended learning path. The proposed method achieves the highest performance across all evaluation metrics: an Accuracy of 0.3489, F1 Score of 0.3241, Precision of 0.3202, and Recall of 0.3489. This consistent improvement indicates that the proposed approach introduces significant enhancements that enable more effective modeling of student learning paths compared to traditional RNN, LSTM, and even Seq2Seq-based models.

Conclusion and Future Work

This project shows that advancing Deep Knowledge Tracing and learning path recommendation can help address persistent retention and graduation challenges at HBCUs by improving both predictive accuracy and personalization. Through semi-supervised learning, ensemble modeling, and tabular generative AI, the approach effectively overcomes data scarcity in low-resource environments. Interpretability is strengthened by integrating SHAP, large language models, and IRT-based ensemble learning, which provide educators with transparent insights into student abilities, item difficulty, and the factors shaping predictions. These interpretable mechanisms build trust in AI-driven systems and support more informed decision-making. Additionally, the proposed multi-task learning path recommendation model generates individualized, non-redundant learning trajectories. Collectively, these innovations create a scalable, equitable foundation for AI-enabled educational support that enhances STEM student success at HBCUs.

In the future, it will focus on two key directions for AI-enhanced STEM education for HBCUs. First, uncertainty-aware, LLM-based learning path recommendation will leverage the powerful reasoning capabilities of LLMs to generate more robust learning paths while quantifying recommendation uncertainty, enabling learners and instructors to make more confident and trustworthy decisions. Second, trustworthy knowledge tracing will be advanced by integrating LLMs with interpretable post-processing methods, as well as prompt engineering, allowing users, such as students and educators, not only to monitor learning outcomes but also to understand the evidence and rationale behind the tracing process. These studies will produce powerful tools that help improve the historically low retention and graduation rates in HBCU STEM education. Finally, while the current study focuses on model development and offline validation, the proposed DKT and learning path recommendation system is designed for near-term deployment in undergraduate STEM courses at HBCUs, where it will be integrated with existing learning management systems. It will include classroom deployment and IRB-approved studies to evaluate its impact on student learning outcomes, course performance, retention, and graduation rates.

Acknowledgement

This research work is supported by NSF under award number 2235731. Any opinions, findings, and conclusions or recommendations expressed in this work are those of the author(s) and do not necessarily reflect the views of NSF.

References

1. Kuo, M. M., Li, X., Qian, L., Obiomon, P., and Dong, X. (2024). Deep Knowledge Tracing for Personalized Adaptive Learning at Historically Black Colleges and Universities. arXiv preprint arXiv:2410.13876.
2. Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L. J., and Sohl-Dickstein, J. (2015). Deep knowledge tracing. Advances in neural information processing systems, 28.
3. Dong, X., Li, L., Okechukwu, U., and Qian, L. (2025). Integrating Semi-Supervised Learning and Ensemble Deep Learning for Low-resource Deep Knowledge Tracing. International Joint Conference on Neural Networks (IJCNN).
4. Dong, X., Fu, Y., Kuo, M. M., Sarker, S., Qian, L., and Li, X. (2024, June). Board 262: Enhancing Deep

Knowledge Tracing via Diffusion Models for Personalized Adaptive Learning. In 2024 ASEE Annual Conference & Exposition.
5. Onyeke, C., Qian, L., Obiomon, P., and Dong, X. (2025). Integrating Ensemble Learning with Item Response Theory to Improve the Interpretability of Student Learning Outcome Tracing. Applied Sciences, 15(23), 12594. https://doi.org/10.3390/app152312594
6. Nasrin, A., Qian, L., Obiomon, P., and Dong, X. (2025). Enhancing Learning Path Recommendation via Multi-task Learning. AIxHEART 2025 (International Conference on AI x Humanities, Education, and Art).
7. Huang, Z., et al. (2020). Learning or forgetting? A dynamic approach for tracking the knowledge proficiency of students. ACM Transactions on Information Systems (TOIS), 38(2), 1-33.
8. Zhang, J., Shi, X., King, I., and Yeung, D. Y. (2017, April). Dynamic key-value memory networks for knowledge tracing. In Proceedings of the 26th international conference on World Wide Web (pp. 765-774).
9. A. Kotelnikov, D. Baranchuk, I. Rubachev, and A. Babenko, "Tabddpm: Modelling tabular data with diffusion models," in International Conference on Machine Learning. PMLR, 2023, pp. 17 564–17 579.
10. Cai, L., Choi, K., Hansen, M., and Harrell, L. (2016). Item response theory. Annual Review of Statistics and Its Application, 3(1), 297-321.
11. Tsutsumi, E., Nishio, T., and Ueno, M. (2024, July). Deep-IRT with a Temporal Convolutional Network for Reflecting Students' Long-Term History of Ability Data. In International Conference on Artificial Intelligence in Education (pp. 250-264). Cham: Springer Nature Switzerland.
12. Gu, W., Liu, Z., and Liu, S. (2023, October). Interpretable Deep Knowledge Tracing with Graph Relationship Information. In 2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology (ICCASIT) (pp. 290-295). IEEE.
13. Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.
14. Kumar, P. (2024). Large language models (LLMs): survey, technical frameworks, and future challenges. Artificial Intelligence Review, 57(10), 260.
15. Kuo, M. M., Li, X., Obiomon, P., Qian, L., and Dong, X. (2025). Tracing Student Learning Outcome at Historically Black Colleges and Universities via Deep Knowledge Tracing. IEEE Access.
16. Kuo, M. M., Li, X., Obiomon, P., Qian, L., & Dong, X. (2025). Improving Student Learning Outcome Tracing at HBCUs Using Tabular Generative AI and Deep Knowledge Tracing. *IEEE Access*.