The Future of
Engineering Education
2024 Annual Conference & Exposition

Oregon Convention Center
Portland, OR . June 23 - 26, 2024

ASEE

Paper ID #41658

Board 262: Enhancing Deep Knowledge Tracing via Diffusion Models for Personalized Adaptive Learning

Prof. Xishuang Dong, Prairie View A&M University

Xishuang Dong is Assistant Professor of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. His research interests include foundation AI, deep learning, object detection, natural language processing, computer systems biology, and Internet of Things.

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

Ming-Mu Kuo, Prairie View A&M University

Ming-Mu Kuo is a doctoral student of Electrical and Computer Engineering Department, Senior Data Analyst of Office of Institutional Research & Effectiveness, Prairie View A&M University. His research interests include online education, computer assisted language learning, knowledge tracing, and deep learning.

Shouvon Sarker, Prairie View A&M University
Lijun Qian
Dr. 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.

Enhancing Deep Knowledge Tracing via Diffusion Models for Personalized Adaptive Learning

Ming-Mu Kuo[1], Shouvon Sarker[1], Xiangfang Li[1], Yujian Fu[2], Lijun Qian[1], Xishuang Dong[1]
[1]Electrical and Computer Engineering Department
Roy G. Perry College of Engineering
Prairie View A&M University
[2]Electrical Engineering and Computer Science Department
Alabama A&M University

Abstract

In contrast to pedagogies like evidence-based teaching, personalized adaptive learning (PAL) distinguishes itself by closely monitoring the progress of individual students and tailoring the learning path to their unique knowledge and requirements. A crucial technique for effective PAL implementation is knowledge tracing, which models students' evolving knowledge to predict their future performance. Based on these predictions, personalized recommendations for resources and learning paths can be made to meet individual needs. Recent advancements in deep learning have successfully enhanced knowledge tracking through Deep Knowledge Tracing (DKT). This paper introduces generative AI models to further enhance DKT. Generative AI models, rooted in deep learning, are trained to generate synthetic data, addressing data scarcity challenges in various applications across fields such as natural language processing (NLP) and computer vision (CV). This study aims to tackle data shortage issues in student learning records to enhance DKT performance for PAL. Specifically, it employs TabDDPM, a diffusion model, to generate synthetic educational records to augment training data for enhancing DKT. The proposed method's effectiveness is validated through extensive experiments on ASSISTments datasets. The experimental results demonstrate that the AI-generated data by TabDDPM significantly improves DKT performance, particularly in scenarios with small data for training and large data for testing.

1. Introduction

Unlike various pedagogies such as evidence-based teaching, personalized adaptive learning (PAL) [1] stands out by continuously monitoring each student's progress and adjusting the educational journey according to their distinct knowledge and needs. Instead of a one-size-fits-all approach, PAL tailors the learning journey for each learner. A pivotal tool in making PAL effective is knowledge tracing (KT) [2], which tracks the progression of students' understanding to anticipate their upcoming performance. Using these insights, tailored suggestions for study materials and pathways can be provided to cater to individual requirements. Additionally, content that might be deemed too elementary or overly advanced can be bypassed or postponed.

Recent progress in deep learning has notably improved KT with the introduction of Deep Knowledge Tracing (DKT) [3]. DKT is designed to leverage advanced deep learning methods to understand the temporal patterns in sequences of student interactions, predicting a student's response to a new question. Previous work has demonstrated that DKT surpasses conventional KT models across various benchmark datasets. The evolution of DKT has been presented by diverse approaches, including memory architectures [4], attention mechanisms [5], graph-based learning [6], textual features [7], and forgetting-based approach [8]. Furthermore, some research work [9], [10], [11] have delved into data augmentation strategies to boost DKT's efficiency. Nevertheless,

there's limited exploration into harnessing Generative AI methods [12] to bolster DKT by generating synthetic data.

Generative AI (GAI) models [12] leverage deep learning techniques to create synthetic data, effectively addressing data scarcity challenges in diverse applications like natural language processing (NLP) and computer vision (CV). Key players in the realm of GAI models include ChatGPT and DALL-E-2[13], both rooted in transformer architecture, featuring both an encoder and a decoder [14]. These pretrained GAI models extend their capabilities beyond initial training objectives by excelling in downstream tasks. A notable advancement in the field is the emergence of diffusion models [15], [16] as the latest state-of-the-art GAI models. This family of models has disrupted the longstanding dominance of generative adversarial networks (GANs) [17] in the intricate task of image synthesis [18]. Moreover, diffusion models exhibit promise across various domains, such as computer vision [19], [20] and natural language processing [21].

This paper introduces diffusion models as a solution to address data scarcity challenges in student learning records, aiming to enhance the performance of DKT for PAL. In particular, the task involves small data for training and large data for testing. In detail, the proposed approach utilizes TabDDPM [22], a diffusion model capable of generating synthetic educational records. TabDDPM exhibits universality, applicable to any tabular dataset and capable of handling various feature types. Notably, it outperforms alternative methods designed for tabular data, including GAN-based and VAE-based approaches. The proposed method's effectiveness is validated through extensive experiments on ASSISTments datasets. The experimental results demonstrate that the AI-generated data by TabDDPM significantly improves DKT performance, particularly in scenarios with small data for training and large data for testing.

The contributions in this paper can be summaries as:

- To the best of our knowledge, this study represents the pioneering exploration of resolving data shortage issues on DKT through GAI models. The approach leverages diffusion models to generate educational records, effectively augmenting the original training datasets used to train DKT models.
- The experimental results, evaluated using various metrics, unequivocally show a significant improvement in DKT performance through the proposed method. Additionally, there is a consistent increase in DKT performance as the amount of AI-generated training data increases. This finding aligns with previous research on data augmentation through GAI techniques, further validating the effectiveness of leveraging artificial intelligence for enhancing the training datasets in educational contexts.

2. Related Work
2.1. Knowledge Tracing
Knowledge Tracing (KT) is designed to effectively monitor a student's learning progress as they interact with teaching materials [2]. In essence, KT aims to observe, represent, and quantify a student's knowledge state, including their mastery level of the skills underlying the teaching materials. It can be broadly categorized into traditional knowledge tracing and deep learning-based knowledge tracing. Traditionally, two prominent works for knowledge tracing are Bayesian Knowledge Tracing [26] and Factor Analysis Models [27]. Conversely, recent advancements in deep learning have inspired researchers to apply deep learning techniques to knowledge tracing. In this context, the knowledge tracing task is commonly framed as a sequence prediction task with machine learning. Recurrent Neural Network (RNN) based techniques have significantly

advanced the development of deep knowledge (DK) in the context of knowledge tracing [28]. Despite these advancements, a persistent challenge in enhancing KT for real-world applications is the limited availability of education data. The scarcity of such data poses an ongoing obstacle to further improving the effectiveness of knowledge tracing methodologies in practical educational settings.

2.2. Generative Models for Tabular Data Generation

The exploration of generative models for tabular data generation is currently a vibrant area of research due to the increasing demand for high-quality synthetic data in various tabular tasks. This demand arises for several reasons. Firstly, tabular datasets, especially in the education domain, are often limited in size compared to datasets in vision or natural language processing (NLP) problems, which can leverage massive amounts of "extra" data available from the Internet. Secondly, creating synthetic datasets that do not include actual user data allows for public sharing without compromising anonymity. In addition, unlike unstructured images or natural texts, tabular data is typically structured, and the interpretability of individual features raises questions about the necessity for complex "deep" architectures. Therefore, simple interpolation techniques such as SMOTE [29] (originally designed for addressing class imbalance) can serve as effective solutions.

Recent research has introduced various models for tabular data generation, including tabular Variational Autoencoders (VAEs) [30], GAN-based approaches [31], and diffusion model-based approaches [22]. These models contribute to addressing the challenges associated with limited tabular data size and the need for synthetic datasets in a privacy-preserving manner.

3. Methodology

This paper aims to address the data scarcity issue in Deep Knowledge Tracing (DKT) by generating synthetic data using diffusion models. The approach encompasses the implementation of DKT models and the utilization of specific diffusion models for synthetic data generation.

3.1. Deep Knowledge Tracing (DKT)

Inspired by the groundbreaking advancements in deep learning [23], there has been swift progress in the development of knowledge tracing models. A notable contribution in this domain is Deep Knowledge Tracing (DKT) [3], which leverages Recurrent Neural Networks (RNNs) [24] to capture temporal dynamics within a sequence of interactions between a student's questions and answers. This enables the model to predict a student's response to a new question based on their historical interactions. A fundamental representation of a simple RNN network for DKT is defined as follows:

$$h_t = \tanh\left(W_{hx}x_t + W_{hh}h_{t-1} + b_t\right) \tag{1}$$
$$y_t = \delta\left(W_{yh}h_t + b_t\right) \tag{2}$$

where both $\tanh(\cdot)$ and $\delta(\cdot)$ are applied element-wise. The model is parameterized by an input weight matrix W_{hx}, a recurrent weight matrix W_{hh}, an initial state h_0, and an output weight matrix W_{yh}. Biases for the latent and output units are denoted by b_h and b_y. The inputs (x_t) to the dynamic network can be either one-hot encodings or compressed representations of a student's action, while the prediction (y_t) is a vector representing the probability of correctly answering each exercise in the dataset. Empirical results have demonstrated that DKT surpasses traditional KT models on various benchmark datasets. This endeavor underscores the potential of employing

deep learning models to effectively address the KT problem [2].

3.2. TabDDPM

Denoising diffusion probabilistic models (DDPM) [25] have garnered significant attention in the generative modeling community, often surpassing alternative methods in producing both realistic and diverse samples [16]. Kotelnikov et al. introduced TabDDPM [22] to explore the applicability of DDPM's versatility to general tabular challenges. Such challenges are prevalent in numerous industrial applications characterized by datasets comprising diverse and heterogeneous features. TabDDPM offers a streamlined DDPM variant tailored for tabular data, accommodating mixed data types, including numerical and categorical features. For instance, a tabular data sample is structured as $x = [x_{num}, x_{cat_1}, ..., x_{cat_C}]$, encompassing x_{num} numerical features and C categorical features denoted by x_{cat_i}. Notably, TabbDPM employs multinomial diffusion to represent categorical and binary features, while Gaussian diffusion is employed for numerical features. The reverse diffusion process within TabDDPM is characterized by a multi-layer neural network, producing outputs equivalent in dimensionality to x_0. During the network's training, the objective encompasses minimizing a combined loss function, comprising the mean-squared error L_t^{simple} for the Gaussian diffusion to process numerical attributes of tabular data and individual KL divergences L_t^i for each multinomial diffusion to process categorial attributes of tabular data. Furthermore, the cumulative loss for multinomial diffusions is adjusted by dividing it by the count of categorical features, ensuring comprehensive training dynamics.

$$L_t^{TabDDPM} = L_t^{simple} + \frac{\sum_{i \le C} L_t^i}{C} \tag{3}$$

3.3. Proposed Method

The flow of the proposed method is depicted in Figure 1. The process begins by randomly sampling data, serving as inputs to TabDDPM, which generates synthetic education data. Subsequently, this synthetic education data is integrated with the original training education data, forming a combined dataset used to train the DKT model for knowledge tracing. Finally, the trained DKT model is evaluated using testing education data, employing multiple evaluation metrics to assess its performance.
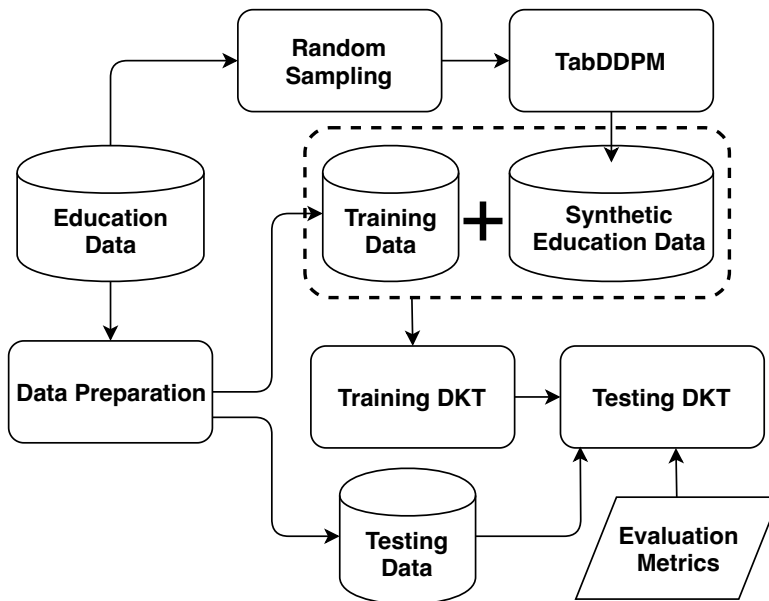

Figure 1. Flow of the proposed method

4. Experiment

4.1. Dataset

The ASSISTments datasets[1] consist of longitudinal data gathered from the ASSISTment platform,

---

[1] https://sites.google.com/site/assistmentsdata/

an online tutoring service. Widely recognized as the predominant datasets for evaluating KT models, they feature the highest number of questions overall. These datasets comprise math exercises from grade school, sourced from the Massachusetts Comprehensive Assessment System (MCAS), encompassing various question types like multiple choice, text, and open-ended questions. Multiple versions of the ASSISTments datasets exist, each collected during distinct time periods. For this study, ASSISTments2009 was utilized, collected during the 2009-2010 school year, initially containing 525, 535 interactions (i.e., student responses to questions in the dataset), including duplicates.

In real-world applications, the amount of training data often falls short compared to the testing data used by users to apply AI models for specific tasks. To address this need, this study splits the ASSISTments2009 dataset into training data (5%), validation data (20%), and testing data (75%). To generate synthetic educational data, TabDDPM [22] is applied to the ASSISTments2009 data, resulting in multiple education datasets. Only two categorical attributes, namely *skill_id* and *user_id*, along with one continuous attribute *overlap_time*, and the ground truth correct are utilized in this process.

4.2. Evaluation Metrics

Various evaluation metrics are employed to evaluate the performance of our proposed model, which includes Accuracy, precision, and recall. *Accuracy* is calculated by dividing the number of student answers identified correctly over the total number of student answers.

$$Accuracy = \frac{N_{correct}}{N_{total}} \tag{4}$$

Precision defines the capability of a model to represent only correct student answers while Recall computes the aptness to refer all corresponding correct student answers.

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

where *TP* (True Positive) tallies the total number of student answers that match the ground truth, *FP* (False Positive) quantifies the number of student answers that do not align with the ground truth but are incorrectly identified as correct. *FN* (False Negative) enumerates the instances where correct student answers are incorrectly classified as incorrect. Additionally, the study utilizes the Area Under the Curve (AUC) as a metric, which gauges the binary classifier's efficacy in distinguishing between classes. A higher AUC indicates superior performance in discerning between positive and negative classes.

4.3. Experimental results and discussion

TABLE 1: **Comparing performance between DKT and the proposed method (DKT + TabDDPM).**

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**

Table 1. illustrates a comparison of performance between the baseline model (DKT) and the proposed model. The proposed model surpasses DKT in terms of accuracy, AUC, precision, and recall. Notably, TabDDPM-generated samples contribute significantly to enhancing AUC, recall, and accuracy, with a more pronounced impact on reducing false negatives to increase the recall.

TABLE 2: **Comparing performance between various proposed method.**

Model	Accuracy (%)	AUC (%)	Precision (%)	Recall (%)
DKT + TabDDPM (1, 000)	59.30±0.43	57.29±1.14	60.50±0.16	90.10±2.50
DKT + TabDDPM (3, 000)	60.72±0.95	62.10±0.52	61.14±0.52	90.60±1.15
DKT + TabDDPM (5, 000)	60.22±2.22	60.02±4.48	61.51±0.78	87.55±6.14
DKT + TabDDPM (7, 000)	62.23±0.43	63.78±0.77	62.42±0.26	91.01±1.60
DKT + TabDDPM (10, 000)	62.23±0.43	63.78±0.77	62.42±0.26	91.01±1.60
DKT + TabDDPM (15, 000)	**63.62±0.32**	65.97±0.43	**63.45±0.33**	90.97±0.95
DKT + TabDDPM (30, 000)	63.46±0.48	**66.00±0.85**	63.08±0.44	92.40±1.06
DKT + TabDDPM (50, 000)	63.46±0.44	65.63±0.77	63.17±0.36	91.85±0.66
DKT + TabDDPM (100, 000)	63.44±0.39	65.63±0.71	62.95±0.25	**92.97±0.62**

Additionally, Table II provides a detailed breakdown of performance based on varying numbers of generated samples. Notably, even a small number of generated samples, such as 1,000 samples, contributes to improving knowledge tracing (KT) performance. As the number of generated samples increases, the proposed method exhibits increased stability, reflected in the standard deviation of accuracy, precision, and recall. This aligns with prior research suggesting that involving more data in the training process enhances model performance. Furthermore, Figure 2 visually depicts that an increased amount of generated data further enhances performance. This observation aligns with the findings presented in Table II, emphasizing the positive impact of augmented data on model performance.
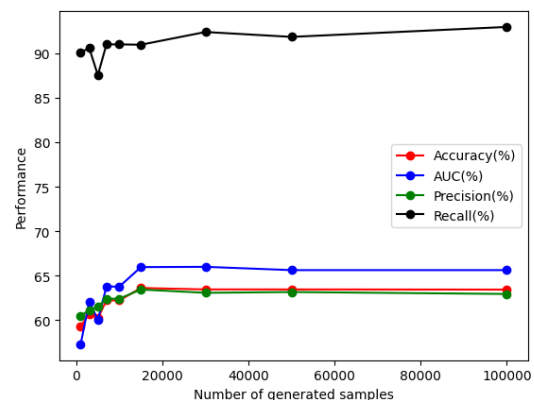

Figure 2. Performance Comparison

5. Conclusion

Knowledge tracing stands as a crucial technique for implementing personalized adaptive learning. In this paper, diffusion models were utilized to improve deep knowledge tracing, with a specific focus on applying TabDDPM to generate synthetic education data. This augmentation aims to address the limitations posed by the scarcity of training data for real-world knowledge tracing applications.

In the future, we envisions exploring two promising directions for future research. First, there is a plan to extend the generation of synthetic education data to encompass more aspects of knowledge tracing tasks. This expansion aims to enhance the diversity and richness of the generated data, contributing to a more comprehensive understanding of the learning process. Secondly, the proposed model will be validated in the context of personalized learning path recommendation to further advance personalized adaptive learning. This direction emphasizes tailoring educational experiences based on individual learning patterns, fostering a more effective and personalized approach to learning. The combination of these future directions is anticipated to contribute to the ongoing evolution of knowledge tracing and its applications in the personalized adaptive learning scenario.

References

1. H. Peng, S. Ma, and J. M. Spector, "Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment," Smart Learning Environments, vol. 6, no. 1, pp. 1–14, 2019.
2. G. Abdelrahman, Q. Wang, and B. Nunes, "Knowledge tracing: A survey," ACM Computing Surveys, vol. 55, no. 11, pp. 1–37, 2023.
3. C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas, and J. Sohl-Dickstein, "Deep knowledge tracing," Advances in neural information processing systems, vol. 28, 2015.
4. G. Abdelrahman and Q. Wang, "Knowledge tracing with sequential key- value memory networks," in Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval, 2019, pp. 175–184.
5. A. Ghosh, N. Heffernan, and A. S. Lan, "Context-aware attentive knowledge tracing," in Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, 2020, pp. 2330– 2339.
6. H. Nakagawa, Y. Iwasawa, and Y. Matsuo, "Graph-based knowledge tracing: modeling student proficiency using graph neural network," in IEEE/WIC/ACM International Conference on Web Intelligence, 2019, pp. 156–163.
7. Q. Liu, Z. Huang, Y. Yin, E. Chen, H. Xiong, Y. Su, and G. Hu, "Ekt: Exercise-aware knowledge tracing for student performance prediction," IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 1, pp. 100–115, 2019.
8. G. Abdelrahman and Q. Wang, "Deep graph memory networks for forgetting-robust knowledge tracing," IEEE Transactions on Knowledge and Data Engineering, 2022.
9. W. Lee, J. Chun, Y. Lee, K. Park, and S. Park, "Contrastive learning for knowledge tracing," in Proceedings of the ACM Web Conference 2022, 2022, pp. 2330–2338.
10. X. Song, J. Li, Q. Lei, W. Zhao, Y. Chen, and A. Mian, "Biclkt: Bigraph contrastive learning based knowledge tracing," Knowledge-Based Systems, vol. 241, p. 108274, 2022.
11. S. Lee, Y. Choi, J. Park, B. Kim, and J. Shin, "Consistency and monotonicity regularization for neural knowledge tracing," arXiv preprint arXiv:2105.00607, 2021.
12. Y. Cao, S. Li, Y. Liu, Z. Yan, Y. Dai, P. S. Yu, and L. Sun, "A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt," arXiv preprint arXiv:2303.04226, 2023.
13. A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, "Zero-shot text-to-image generation," in International Conference on Machine Learning. PMLR, 2021, pp. 8821–8831.
14. X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang, "Pre-trained models for natural language processing: A survey," Science China Technological Sciences, vol. 63, no. 10, pp. 1872–1897, 2020.
15. Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," Advances in neural information processing systems, vol. 32, 2019.
16. P. Dhariwal and A. Nichol, "Diffusion models beat gans on image synthesis," Advances in neural information processing systems, vol. 34, pp. 8780–8794, 2021.
17. D. Saxena and J. Cao, "Generative adversarial networks (gans) challenges, solutions, and future directions," ACM Computing Surveys (CSUR), vol. 54, no. 3, pp. 1–42, 2021.
18. Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," arXiv preprint arXiv:2011.13456, 2020.
19. R. Yang and S. Mandt, "Lossy image compression with conditional diffusion models," arXiv preprint arXiv:2209.06950, 2022.
20. R. Yang, P. Srivastava, and S. Mandt, "Diffusion probabilistic modeling for video generation," Entropy, vol. 25, no. 10, p. 1469, 2023.
21. X. Li, J. Thickstun, I. Gulrajani, P. S. Liang, and T. B. Hashimoto, "Diffusion-lm improves controllable text generation," Advances in Neural Information Processing Systems, vol. 35, pp. 4328–4343, 2022.
22. A. Kotelnikov, D. Baranchuk, I. Rubachev, and A. Babenko, "Tabddpm: Modelling tabular data with diffusion models," in International Confer- ence on Machine Learning. PMLR, 2023, pp. 17 564–17 579.
23. I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.
24. Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," arXiv preprint arXiv:1506.00019, 2015.
25. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," Advances in neural information processing systems, vol. 33, pp. 6840– 6851, 2020.
26. Y. Cui, M.-W. Chu, and F. Chen, "Analyzing student process data in game-based assessments with bayesian knowledge tracing and dynamic bayesian networks." Journal of Educational Data Mining, vol. 11, no. 1, pp. 80–

100, 2019.

27. C.-K. Yeung, "Deep-irt: Make deep learning based knowledge tracing explainable using item response theory," arXiv preprint arXiv:1904.11738, 2019.

28. X. Xiong, S. Zhao, E. G. Van Inwegen, and J. E. Beck, "Going deeper with deep knowledge tracing." International Educational Data Mining Society, 2016.

29. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," Journal of artificial intelligence research, vol. 16, pp. 321–357, 2002.

30. L. Xu, M. Skoularidou, A. Cuesta-Infante, and K. Veeramachaneni, "Modeling tabular data using conditional gan," Advances in neural information processing systems, vol. 32, 2019.

31. A. Torfi, E. A. Fox, and C. K. Reddy, "Differentially private synthetic medical data generation using convolutional gans," Information Sciences, vol. 586, pp. 485–500, 2022.

MING-MU KUO
Ming-Mu Kuo is a PhD student of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. His research interests are in data science and AI-enhanced education.

SHOUVON SARKER
Shouvon Sarker is a PhD student of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. His research interests are in generative AI and data mining.

XIANGFANG LI
Xiangfang Li is Associate Professor of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. Her research interests are in computational and systems biology, systems pharmacology, computer networking and communication, and artificial intelligence. She is the recipient of the Outstanding Researcher of the Year award from the Roy G. Perry College of Engineering of PVAMU in 2017.

YUJIAN FU
Yujian Fu is Professor of Electrical Engineering and Computer Science Department, Alabama A&M University. Her research interests are in the areas of formal specification, verification and validation, optimization, reinforcement learning, mobile security, behavior based malware detection, specification based intrusion detection, formal specification & verification of cyber physical systems, real time embedded software design and quality assurance, software testing, runtime verification, software architecture, service oriented architecture (SOA), software evolution.

LIJUN QIAN
Lijun Qian is Professor of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. His research interests are in the area of big data processing, artificial intelligence, wireless communications and mobile networks, network security and intrusion detection, and computational systems biology.

XISHUANG DONG
Xishuang Dong is Assistant Professor of Electrical and Computer Engineering Department, Roy G. Perry College of Engineering, Prairie View A&M University. His research interests include deep learning, object detection, natural language processing, computer systems biology, and Internet of Things.