A Closed-Loop Personalized Learning Agent Integrating Neural Cognitive Diagnosis, Bounded-Ability Adaptive Testing, and LLM-Driven Feedback

Zhifeng Wang
Faculty of Artificial Intelligence in Education
Central China Normal University
Wuhan 430079, China
zfwang@ccnu.edu.cn

Xinyue Zheng
Faculty of Artificial Intelligence in Education
Central China Normal University
Wuhan, China
zxy0624@foxmail.com

Chunyan Zeng
School of Electrical and Electronic Engineering
Hubei University of Technology
Wuhan 430068, China
cyzeng@hbut.edu.cn

***Abstract*—As information technology advances, education is moving from one-size-fits-all instruction toward personalized learning. However, most methods handle modeling, item selection, and feedback in isolation rather than as a closed loop. This leads to coarse or opaque student models, assumption-bound adaptivity that ignores diagnostic posteriors, and generic, non-actionable feedback. To address these limitations, this paper presents an end-to-end personalized learning agent, EduLoop-Agent, which integrates a Neural Cognitive Diagnosis model (NCD), a Bounded-Ability Estimation Computerized Adaptive Testing strategy (BECAT), and large language models (LLMs). The NCD module provides fine-grained estimates of students' mastery at the knowledge-point level; BECAT dynamically selects subsequent items to maximize relevance and learning efficiency; and LLMs convert diagnostic signals into structured, actionable feedback. Together, these components form a closed-loop framework of "Diagnosis–Recommendation–Feedback." Experiments on the ASSISTments dataset show that the NCD module achieves strong performance on response prediction while yielding interpretable mastery assessments. The adaptive recommendation strategy improves item relevance and personalization, and the LLM-based feedback offers targeted study guidance aligned with identified weaknesses. Overall, the results indicate that the proposed design is effective and practically deployable, providing a feasible pathway to generating individualized learning trajectories in intelligent education.**

***Index Terms*—Personalized Learning; Cognitive Diagnosis; Computerized Adaptive Testing; Large Language Models**

I. Introduction

As information technology advances, education is moving beyond one-size-fits-all instruction toward genuinely personalized learning [1]–[8]. Conventional classroom models, however, still struggle to accommodate heterogeneity in students' knowledge mastery, learning abilities, and study habits, limiting access to appropriate resources and guidance and, ultimately, dampening learning effectiveness [9]–[14].

Corresponding author: Chunyan Zeng, Email: cyzeng@hbut.edu.cn.

Cognitive Diagnosis Models (CDMs) and Computerized Adaptive Testing (CAT) offer principled tools to address this challenge by capturing individual differences and improving both assessment and instruction [15]–[18]. CDMs provide fine-grained, interpretable estimates of students' mastery at the knowledge-point level; recent deep-learning advances further enhance accuracy and diagnostic fidelity [12], [19], [20]. CAT dynamically selects items based on real-time ability estimates, with maximum (Fisher/global) information criteria serving as classical and effective selection principles [17], [20]. When combined, CDMs' granular posteriors and CAT's adaptive mechanisms can power targeted resource recommendations and learning interventions that move beyond the traditional "one test for all" paradigm [13], [21], [22].

Despite steady progress, much of the prior work optimizes only one stage of this pipeline—modeling, item selection, or feedback—rather than delivering an integrated, closed-loop solution. This fragmentation often yields (i) coarse or weakly interpretable student models that overlook knowledge-point mastery, (ii) adaptive policies tied to restrictive parametric assumptions that underutilize diagnostic posteriors, and (iii) generic feedback that is not evidence-grounded and thus less actionable for learners.

Large Language Models (LLMs) introduce a complementary capability: converting diagnostic and recommendation outputs into fluent, structured, and actionable feedback. Compared with traditional methods, LLM-generated guidance can articulate specific weaknesses and next steps, helping students close gaps and refine strategies toward truly individualized instruction [23]–[25].

To address these limitations, we propose and implement a personalized learning system that integrates Neural Cognitive Diagnosis (NeuralCD, NCD), a Bounded-Ability Estimation CAT strategy (BECAT), and LLM-based intelligent feedback. The NCD module performs fine-grained cognitive diagnosis at the knowledge-point level; BECAT uses bounded-ability estimation to select informative, level-appropriate items in real time; and the LLM component transforms diagnostic signals and recommendations into interpretable, evidence-grounded study guidance. Together, these components form a closed-loop framework of "Diagnosis–Recommendation–Feedback," aligning modeling, measurement, and intervention within a single system.

This study makes three primary contributions:

1) We design an integrated architecture agent that unifies cognitive diagnosis, adaptive item selection, and feedback generation.
2) We develop a bounded-ability CAT policy that leverages diagnostic posteriors for stable, relevant recommendations.
3) We introduce an LLM-based feedback module with a prompt-integration mechanism to deliver targeted, actionable guidance grounded in model evidence. We validate the system empirically on a public benchmark and analyze its effectiveness for personalized learning.

The remainder of the paper is organized as follows. Section II reviews related research on CDMs, adaptive testing and recommendation strategies, and educational applications of LLMs. Section III details the overall system workflow, the NCD model, the BECAT strategy, and the prompt-integration mechanism. Section IV presents implementation details, datasets, and experimental results. Section V concludes with limitations and future directions.

II. Related Work

This section reviews three strands that underpin personalized learning: (i) cognitive diagnosis models, (ii) adaptive testing and recommendation strategies, and (iii) educational applications of large language models. We trace their evolution and synthesize common limitations—manual assumptions, heuristic or assumption-bound adaptivity, and stage-wise fragmentation—thereby motivating an integrated, closed-loop approach that aligns diagnosis, item selection, and feedback.

A. Cognitive Diagnosis Models

Since Rasch introduced Item Response Theory (IRT) in the 1960s [26], educational measurement has steadily advanced to better represent learners' multidimensional knowledge and evolving skills. Key developments include Multidimensional IRT (MIRT) [27]; cognitive diagnosis models (CDMs) grounded in Q-matrix specifications—such as DINA and G-DINA [28]; and Bayesian Knowledge Tracing (BKT) and its variants [29]. While influential, these approaches often depend on manually specified item–skill mappings, distributional assumptions, and parameter calibration, which can limit scalability and reduce fidelity in complex learning environments.

To mitigate these limitations based on deep learning [6], [30]–[47], Wang et al. proposed Neural Cognitive Diagnosis [48], which leverages neural networks to learn student–item interactions while enforcing an attribute monotonicity assumption. NCD yields interpretable, fine-grained mastery estimates at the knowledge-point level and scales to large datasets. Beyond improved diagnostic accuracy, its skill-level posteriors naturally support downstream personalization, including targeted item recommendation to address weak prerequisites and the construction of coherent, prerequisite-aware learning paths.

B. Adaptive Testing and Recommendation Strategies

Adaptive testing and recommendation strategies are central to personalized learning systems: they aim to select, in real time, the most informative or pedagogically valuable next item based on a learner's current state, thereby improving measurement efficiency and instructional precision. Early approaches largely relied on heuristic or parametric rules—e.g., maximum information criteria in IRT [26] or mastery probabilities in BKT [29]. These methods typically require manual parameterization and Q-matrix specifications and can struggle with complex skill structures and cross-concept dependencies.

Recent work leverages deep learning [49]–[77] to overcome some of these limitations. Sequence models and attention mechanisms automatically extract latent features while capturing temporal learning behaviors and inter-skill relationships, leading to more dynamic and personalized recommendations [78], [79]. Graph-based approaches further incorporate knowledge graphs to guide item selection, promoting coherent learning paths and facilitating knowledge transfer [80]. Overall, the field is moving toward more intelligent and fine-grained adaptive systems; however, many solutions still optimize item selection in isolation, without tightly integrating fine-grained diagnostic posteriors or closing the loop with evidence-grounded feedback—gaps addressed by the system proposed in this work.

C. Applications of Large Language Models in Education

Large language models (LLMs) have accelerated progress in intelligent education by enabling scalable, high-quality language generation. Transformer-based systems such as the GPT family can synthesize personalized learning materials and exercises aligned with learner profiles and trajectories, thereby supporting adaptive instruction and peer-like study resources. As interactive tutors, LLMs provide real-time formative feedback and clarification, improving the flexibility and accessibility of learning experiences [24]. In assessment, LLMs assist with automatic grading, item and hint generation, and learning analytics, which helps reduce instructor workload while maintaining evaluation coverage and timeliness [81].

Despite these benefits, important challenges remain. LLM outputs can exhibit factual errors, calibration issues, and social biases; they are also sensitive to prompt design and may raise

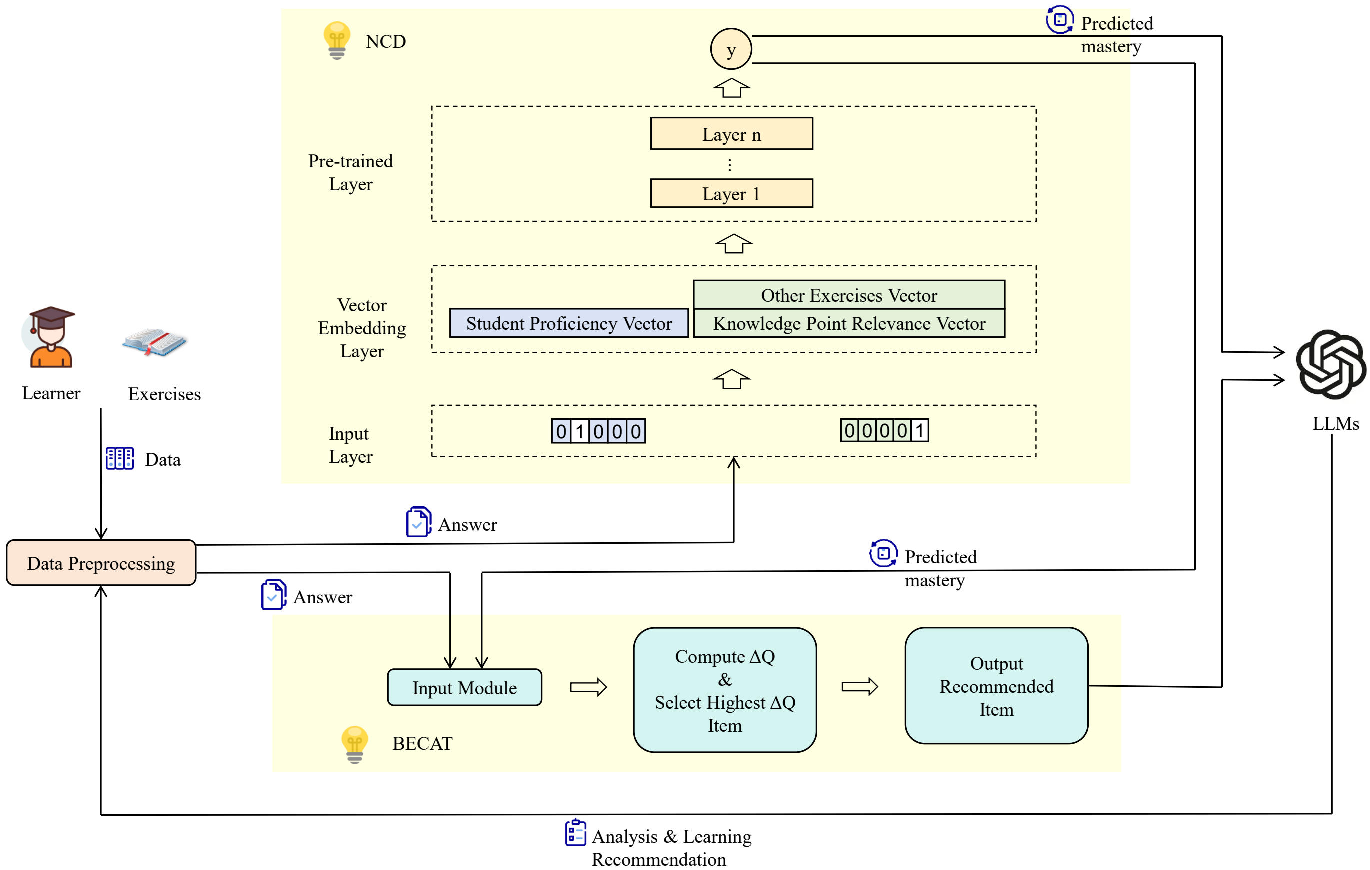

Fig. 1. The proposed EduLoop-Agent follows a four-stage workflow, forming a "Diagnosis–Recommendation–Feedback" closed loop.

privacy and safety concerns in classroom deployments [82]. Consequently, effective educational use typically pairs LLMs with grounding strategies (e.g., retrieval or tool integration), structured prompts aligned with instructional rubrics, human-in-the-loop review, and explicit safeguards. Building on these insights, our system employs LLMs not as standalone tutors but as feedback generators tethered to structured evidence: diagnostic posteriors from NCD and item selections from BECAT. This design aims to transform model decisions into interpretable, actionable guidance while mitigating generic or unsupported feedback.

III. Proposed Method

EduLoop-Agent is an end-to-end, closed-loop personalized learning framework that operationalizes a "Diagnosis–Recommendation–Feedback" cycle as shown in Fig. 1. We first outline the workflow and then detail how NCD produces fine-grained mastery estimates, BECAT selects informative items in real time, and a prompt-integration mechanism grounds LLM feedback in model evidence to yield interpretable, actionable guidance.

A. Workflow of Proposed EduLoop-Agent

The proposed EduLoop-Agent follows a four-stage workflow, forming a "Diagnosis–Recommendation–Feedback" closed loop:

1) Data Preprocessing: The ASSISTments dataset is standardized to generate input files suitable for the NCD model and the BECAT strategy. These include an item–knowledge mapping table, student response records, a knowledge graph, and auxiliary mapping files (e.g., mapping between knowledge point IDs and textual descriptions), ensuring that both item information and knowledge structures are effectively utilized.

2) Model Training: Preprocessed data are fed into training scripts to train the NCD model and optimize its parameters. The outputs include trained model files and a table estimating students' mastery levels across knowledge points, providing the basis for personalized recommendations.

3) Adaptive Recommendation: Leveraging the student mastery table and knowledge graph, data are input into the item recommendation scripts. The BECAT strategy dynamically selects items, enhancing the relevance and efficiency of recommendations while maintaining assessment accuracy.

4) Personalized Feedback: Student mastery data, recommended items, and relevant textual information are input into a Large Language Model (LLM), which generates interpretable feedback and actionable learning suggestions. This stage supports targeted interventions and the optimization of adaptive learning paths.

B. Neural Cognitive Diagnosis Model

The NCD model provides fine-grained and interpretable modeling of students' knowledge mastery:

1) Embedding and Feature Interaction: Students, items, and knowledge points are mapped into low-dimensional vector spaces to form the ability vector $\vartheta_s \in \mathbb{R}^d$ of a student s, the difficulty vector $\beta_q \in \mathbb{R}^d$ of an item q , and the discrimination scalar $\alpha_q \in \mathbb{R}^+$ of an item q. The interaction between a student s and an item q is represented as:

$$x_{s,q} = \alpha_q \cdot \quad (\vartheta_s - \beta_q) \tag{1}$$

This interaction feature captures the difference between student ability and item difficulty, modulated by the item discrimination, reflecting the relative impact of the item on the student's performance.

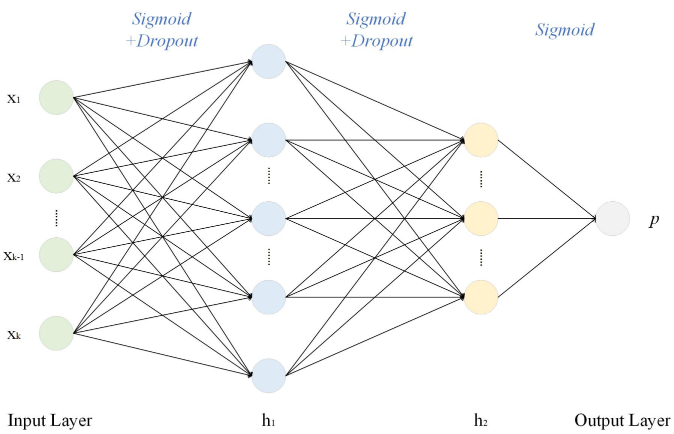

Fig. 2. The embedding layer of the Neural Cognitive Diagnosis model.

2) Prediction Network: As shown in Fig. 2, the interaction feature is input into a multi-layer perceptron (*MLP*), which maps it nonlinearly to the probability of a correct response:

$$\hat{r}_{s,q} = \sigma(MLP(x_{s,q})) \tag{2}$$

where $\sigma(\cdot)$ denotes the Sigmoid function, and the output $\hat{r}_{s,q} \in (0, 1)$ represents the probability that student s answers item q correctly.

3) Loss Function: The model adopts binary cross-entropy (equivalent to negative log-likelihood) as the objective function:

$$\mathscr{L} = -\frac{1}{N} \sum_{(s,q)} \left[r_{s,q} \log \hat{r}_{s,q} + (1 - r_{s,q}) \log(1 - \hat{r}_{s,q}) \right] \tag{3}$$

where $r_{s,q} \in (0, 1)$ represents the actual response. This loss quantifies the discrepancy between predicted and observed outcomes.

C. Bounded Ability Estimation CAT Strategy

The BECAT strategy dynamically selects items based on students' estimated mastery levels:

1) Expected Model Change (EMC): In an adaptive testing context, *EMC* quantifies the contribution of each candidate item to updating student ability. Let $\Delta\theta^{(T)}$ and $\Delta\theta^{(F)}$ denote parameter changes under correct and incorrect responses. The expected model change for item q is:

$$EMC(q) = \hat{r}_{s,q} \cdot \| \Delta\theta^{(T)} \|_2 + \left(1 - \hat{r}_{s,q}\right) \cdot \| \Delta\theta^{(F)} \|_2 \tag{4}$$

Higher *EMC* values indicate greater potential to refine student ability estimates.

2) Item Dependency Modeling: A Bayesian weight matrix W captures complementary information among items:

$$W_{i,j} = C - \mathbb{E}[\| \nabla_\theta \ell_i - \nabla_\theta \ell_j \|] \tag{5}$$

where ℓ_i and ℓ_j are the loss functions for items i and j , and C ensures non-negativity. This matrix reflects inter-item dependencies and informs item selection.

3) Information Score and Gain: The information score *F(S)* measures the coverage of unselected items relative to the selected set S:

$$F(S) = \sum_{i \notin S} \max_{j \in S} w_{ij} \tag{6}$$

The information gain Δ_q of candidate q is defined as the score increment:

$$\Delta_q = F(S \cup \{q\}) - F(S) \tag{7}$$

A greedy selection of items with maximal gain approximates an optimal combination.

D. Prompt Integration Mechanism

The LLM prompt template consists of fixed and dynamic components. The fixed component defines the LLM role (educational assessment expert and teaching assistant), specifies task requirements (analyzing mastery, evaluating recommendations, and providing learning suggestions), and sets output rules (structured bullet points with length limits). The dynamic component, populated automatically by the system, includes student mastery levels, recommended item texts, and corresponding knowledge points. Mastery levels, output by NCD, are presented as "Knowledge Point Name + Mastery Value," while recommended items, output by BECAT, include item text and associated knowledge descriptions.

Output is required to be structured and actionable, with three sections: mastery analysis, recommendation evaluation, and personalized learning suggestions. This ensures relevance, readability, and effective integration of diagnosis, recommendation, and feedback into a coherent adaptive learning workflow.

IV. Experimental Results and Analysis

This section outlines the experimental setup, dataset preprocessing, and evaluation protocol used to assess EduLoop-Agent. We report results on the ASSISTments dataset, quantifying the NCD model's predictive and diagnostic performance (AUC/ACC/RMSE/MSE/Loss) and visualizing training dynamics, then illustrate the downstream effects of BECAT-driven item selection and LLM-generated feedback through representative student cases.

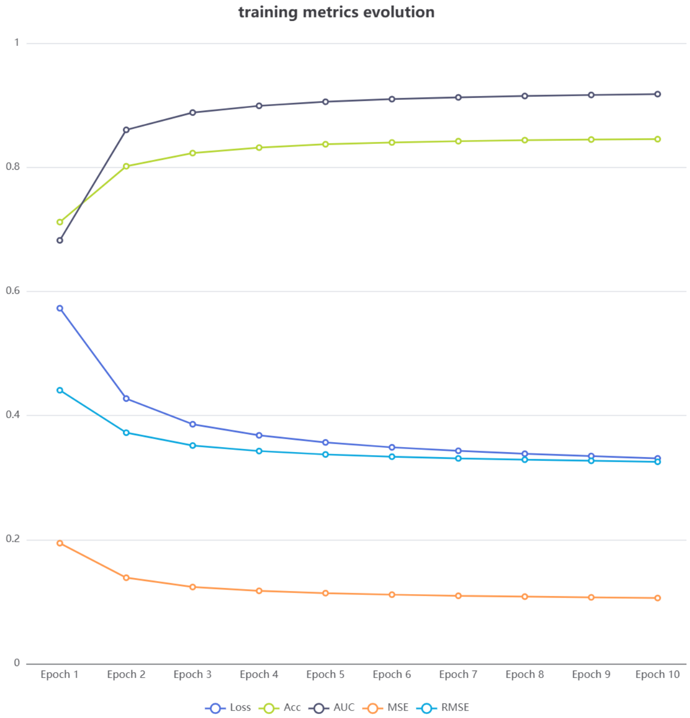

Fig. 3. The performance of the NCD model.

=== LLM Analysis of Student 3212 ===

Item Content: {'Scatter Plot'}

Knowledge Points: ['Box and Whisker', 'Counting Methods']

The student's mastery level in Box and Whisker is 0.48, which is slightly below the passing threshold. This indicates that the student's understanding of the construction, interpretation, and five-number summary of boxplots remains insufficient, with possible misconceptions in interpreting distributions or identifying outliers. In Counting Methods, the mastery level is 0.33, suggesting weak proficiency in fundamental counting techniques such as listing, permutations, and combinations. The student often makes errors or lacks familiarity when enumerating and systematically calculating possible outcomes.

Given the current mastery profile, this item—which involves both knowledge points—poses a certain challenge to the student. It can serve as a suitable practice problem for consolidating foundational knowledge and diagnosing weak points. However, for Counting Methods, the student may encounter confusion in reasoning or omit key steps, requiring additional scaffolding and prerequisite knowledge support.

Recommendations: The student should first systematically review the construction of boxplots, including the five-number summary, identification of outliers, drawing methods, and interpretation of data distributions. Practice with real datasets is recommended, alongside verbal explanations of each step's meaning. For Counting Methods, the student should begin with simple cases, focusing on classification listing and the basic principles of permutations and combinations, and then gradually progress to multi-step and more complex counting scenarios. Interactive exercises, video tutorials, and group discussions can be incorporated to enhance interest and deepen comprehension. When addressing integrated problems, the student should first attempt independently, followed by targeted review to address errors. For weakly mastered content, the emphasis should be placed on "consolidating fundamentals + progressive advancement" rather than tackling advanced tasks prematurely.

Fig. 4. EduLoop-Agent Generated Analysis Of Student 324.

A. Experimental Setups

The experimental environment of this study is based on Python 3.9, with the primary hardware being a computation server equipped with an NVIDIA GPU, running on Windows 11. For tool selection, PyTorch is adopted as the core deep learning framework, responsible for the training and inference of the Neural Cognitive Diagnosis (NCD) model. Meanwhile, the transformers library is used to load and invoke large language model (LLM) interfaces, enabling prompt-based feedback generation and natural language processing. To implement the adaptive recommendation strategy, the CAT library is employed to realize the Bounded Ability Estimation

=== LLM Analysis of Student 324 ===

Item Content: {'Rotations'}

Knowledge Points: ['Probability of a Single Event', 'Mean']

The student's mastery levels in Probability of a Single Event and Mean are 0.18 and 0.13, respectively, both of which indicate very weak proficiency. Specifically, Probability of a Single Event requires understanding the likelihood of an event occurring and accurately computing probabilities in simple contexts. Mean requires calculating the arithmetic average of a dataset. The student likely struggles with the basic probability formula, expressing probability as fractions or percentages, as well as understanding and correctly executing the steps for computing averages.

From a recommendation perspective, this item, which involves both knowledge points, presents both challenges and opportunities for learning. As the student's foundation is fragile, the exercise may serve as a valuable diagnostic tool for identifying deficiencies if the problem is moderately difficult. However, overly complex or multi-step items may impose excessive cognitive load and undermine confidence.

Recommendations: First, the student should review the basic concepts of probability, including "event," "sample space," and methods of expressing and computing probability. Reinforcement through everyday examples (e.g., coin tosses, drawing balls) is advised. Second, the student should repeatedly practice computing means with small datasets, beginning with step-by-step demonstrations, then oral calculations, and eventually full written solutions, to reinforce the conceptual understanding of averages. Visualization tools, abacuses, or small-group discussions may help maintain engagement. Once the fundamentals stabilize, the student can gradually transition to short and moderately difficult integrated problems combining probability and mean. Finally, after each exercise, the student should conduct error analysis, maintain an "error notebook," and revisit common mistakes regularly to strengthen retention.

Fig. 5. EduLoop-Agent Generated Analysis Of Student 1718.

=== LLM Analysis of Student 1718 ===

Item Content: {'Complementary and Supplementary Angles'}

Knowledge Points: ['Finding Percents']

The student's mastery level in Finding Percents is 0.27, which indicates a clear weakness. This knowledge point requires the ability to determine the percentage representation of a part relative to a whole, such as "What percent is 30 of 60?" or "What percent is 50/200?" Proficiency in this area demands both fluency in converting between fractions, decimals, and percentages, and a clear grasp of the computational process.At the current mastery level, the student may encounter several difficulties: (1) unfamiliarity with the basic formula part ÷ whole × 100%, leading to confusion; (2) frequent errors when converting between fractions, decimals, and percentages; and (3) challenges in connecting problem statements to abstract calculations in applied contexts.

From a recommendation perspective, this item is well-suited to the student's learning needs. Given the weak foundation, focusing on Finding Percents offers an opportunity to directly address deficiencies, which is more effective than attempting higher-level tasks at this stage. Emphasis should be placed on consolidating core skills and reinforcing basic concepts.

Recommendations: The student should first systematically review the fundamental concepts of percentages and the standard computational formula, with emphasis on the direct application of part ÷ whole × 100%. Next, practice with small exercises on converting fractions, decimals, and percentages should be incorporated to strengthen fluency. Additionally, contextualized problems (e.g., shopping discounts, statistical proportions) are recommended to enhance applied understanding. An error notebook should be maintained for reflection on frequent mistakes, supplemented with teacher guidance where appropriate. Gradual progression to multi-step application problems is advised once the foundation is solid. Ultimately, this approach will foster both greater problem-solving flexibility and improved confidence.

Fig. 6. EduLoop-Agent Generated Analysis Of Student 3212.

Computerized Adaptive Testing (BECAT) strategy, which is integrated with the NCD module.

In the feedback generation stage, the LLM API is called to integrate cognitive diagnostic results with recommended item information, which are then input into the model to produce personalized instructional feedback.

B. Dataset

The experiments are conducted on the publicly available ASSISTments dataset, which has been widely used in knowledge tracing and adaptive testing research, offering strong generalizability and comparability. The dataset contains student response records, item–knowledge point annotations, and correctness labels (correct/incorrect). During preprocessing, the raw response logs are cleaned and reformatted, ensuring consistency in the mapping of student IDs, item IDs, and knowledge point IDs.

C. Result Analysis

1) Performance of the NCD Model: Based on the NCD model, this study establishes a cognitive diagnosis framework to predict both the probabilities of students correctly answering items and their levels of mastery over knowledge points. The predictions are compared with the actual response data and the evaluation metrics, including AUC, ACC, RMSE, MSE, and LOSS are calculated. The experimental results are visualized using Echarts software.

As shown in Fig. 3, the results demonstrate that model training is stable: the loss metric decreases smoothly without significant oscillations, indicating a normal optimization process. Both ACC and AUC show notable improvement, with AUC exceeding 0.9, reflecting strong discriminative capability. The continuous decline of MSE and RMSE indicates that the deviation between predicted and actual values is steadily reduced, highlighting the increasing accuracy of the model. Overall, after 10 epochs of training, the NCD model demonstrates strong performance: it not only predicts student responses accurately but also provides reliable estimates of mastery levels, making it well-suited for personalized learning diagnosis and recommendation.

2) Item Recommendation Results and Feedback: For illustrative purposes, we randomly selected three students to demonstrate the analysis results as shown in Fig. 4, 5 and 6:

V. Conclusions

This paper presented EduLoop-Agent, an end-to-end, closed-loop personalized learning framework that unifies Neural Cognitive Diagnosis, a bounded-ability CAT strategy, and LLM-based feedback to align diagnosis, adaptive recommendation, and intervention within a single "Diagnosis–Recommendation–Feedback" cycle. Experiments on the ASSISTments dataset indicate stable training dynamics, strong response-prediction performance, improved recommendation relevance, and feedback that is interpretable and pedagogically actionable. Nonetheless, the study's external validity is limited by reliance on a single public dataset and offline simulations, and LLM-generated feedback can be sensitive to prompt design, exhibit variability, and inherit social biases, highlighting the need for better calibration and factual grounding. Future work will broaden evaluation across subjects, grade levels, and institutions; include ablations and user studies with teachers and students; enhance reliability via retrieval-grounded prompting, constrained decoding, rubric-aligned evaluation, and human-in-the-loop review; and conduct classroom deployments to measure learning gains, recommendation efficiency, and fairness under explicit privacy and safety safeguards.

References

[1] Z. Chen, D. Ren, Z. Wang, L. Gao, Y. Shi, H. Liu, and T. Yang, "Enhancing student academic performance in middle school classrooms by fostering engagement motivation through intelligent assessment of teacher praise emotional intensity," *Learning and Instruction*, vol. 99, p. 102185, Oct. 2025.

[2] Y. Wang, M. Zuo, X. He, and Z. Wang, "Exploring students online learning behavioral engagement in university: Factors, academic performance and their relationship," *Behavioral Sciences*, vol. 15, no. 1, p. 78, Jan. 2025.

[3] S. Dong, X. Niu, R. Zhong, Z. Wang, and M. Zuo, "Leveraging label semantics and meta-label refinement for multi-label question classification," *Knowledge-Based Systems*, vol. 318, p. 113412, Jun. 2025.

[4] X. Liao, X. Zhang, Z. Wang, and H. Luo, "Design and implementation of an ai-enabled visual report tool as formative assessment to promote learning achievement and self-regulated learning: An experimental study," *British Journal of Educational Technology*, vol. 55, no. 3, pp. 1253–1276, 2024.

[5] D. Chen, X. Xu, Z. Wang, and J. Shen, "Personalized art image generation model for smart art education," in *2024 International Conference on Intelligent Education and Intelligent Research (IEIR)*. Macau, China: IEEE, Nov. 2024, pp. 1–8.

[6] Z. Wang, M. Wang, C. Zeng, J. Yao, Y. Yang, and H. Xu, "Enhanced convolutional neural networks based learner authentication for personalized e-learning system," in *2023 International Conference on Intelligent Education and Intelligent Research (IEIR)*. Wuhan, China: IEEE, Nov. 2023, pp. 1–7.

[7] L. Ma, X. Zhang, Z. Wang, and H. Luo, "Designing effective instructional feedback using a diagnostic and visualization system: Evidence from a high school biology class," *Systems*, vol. 11, no. 7, p. 364, Jul. 2023.

[8] Z. Wang, W. Wu, C. Zeng, J. Yao, Y. Yang, and H. Xu, "Smart contract vulnerability detection for educational blockchain based on graph neural networks," in *2022 International Conference on Intelligent Education and Intelligent Research (IEIR)*. IEEE, Dec. 2022, pp. 8–14.

[9] L. Li, Z. Wang, J. M. Jose, and X. Ge, "Llm supporting knowledge tracing leveraging global subject and student specific knowledge graphs," *Information Fusion*, vol. 126, p. 103577, Feb. 2026.

[10] Z. Wang, M. Su, Y. Yang, C. Zeng, and L. Ye, "Cross-disciplinary cognitive diagnosis leveraging deep transfer learning for smart education," in *2024 International Conference on Intelligent Education and Intelligent Research (IEIR)*. Macau, China: IEEE, Nov. 2024, pp. 1–8.

[11] L. Li and Z. Wang, "Knowledge relation rank enhanced heterogeneous learning interaction modeling for neural graph forgetting knowledge tracing," *PLOS ONE*, vol. 18, no. 12, p. e0295808, Dec. 2023.

[12] Z. Wang, W. Yan, C. Zeng, Y. Tian, and S. Dong, "A unified interpretable intelligent learning diagnosis framework for learning performance prediction in intelligent tutoring systems," *International Journal of Intelligent Systems*, vol. 2023, p. e4468025, Feb. 2023.

[13] L. Li and Z. Wang, "Knowledge graph-enhanced intelligent tutoring system based on exercise representativeness and informativeness," *International Journal of Intelligent Systems*, vol. 2023, p. e2578286, Oct. 2023.

[14] Z. Wang, Z. Lu, C. Zeng, S. Dong, M. Zuo, and J. Sun, "Mmkt: Multimodal knowledge tracing in personalized e-learning systems for supporting lifelong learning," *IEEE Transactions on Computational Social Systems*, pp. 1–20, 2025.

[15] Y. Zhang, Y. Zhang, W. Xu, Z. Wang, and J. Sun, "Singpad: A knowledge tracing dataset based on music performance assessment," in *Proceedings of the 17th International Conference on Educational Data Mining*, 2024, pp. 332–340.

[16] Z. Wang, W. Wu, C. Zeng, H. Luo, and J. Sun, "Psychological factors enhanced heterogeneous learning interactive graph knowledge tracing for understanding the learning process," *Frontiers in Psychology*, vol. 15, May 2024.

[17] L. Li and Z. Wang, "Calibrated q-matrix-enhanced deep knowledge tracing with relational attention mechanism," *Applied Sciences*, vol. 13, no. 4, pp. 1–24, Jan. 2023.

[18] Z. Wang, Y. Hou, C. Zeng, S. Zhang, and R. Ye, "Multiple learning features–enhanced knowledge tracing based on learner–resource response channels," *Sustainability*, vol. 15, no. 12, p. 9427, Jan. 2023.

[19] Z. Wang, J. Wan, Y. Yang, C. Zeng, and J. Shen, "Unified knowledge tracing framework for subjective and objective assessments," in *2024 International Conference on Intelligent Education and Intelligent Research (IEIR)*. Macau, China: IEEE, Nov. 2024, pp. 1–8.

[20] S. Dong, X. Tao, R. Zhong, Z. Wang, M. Zuo, and J. Sun, "Advanced mathematics exercise recommendation based on automatic knowledge extraction and multi-layer knowledge graph," *IEEE Transactions on Learning Technologies*, pp. 1–17, 2023.

[21] P. Lou, "Learning path recommendation of intelligent education based on cognitive diagnosis," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 13, 2023.

[22] L. Lyu, Z. Wang, H. Yun, Z. Yang, and Y. Li, "Deep knowledge tracing based on spatial and temporal representation learning for learning performance prediction," *Applied Sciences*, vol. 12, no. 14, pp. 1–21, Jan. 2022.

[23] J. Meyer, T. Jansen, R. Schiller, L. W. Liebenow, M. Steinbach, A. Horbach, and J. Fleckenstein, "Using llms to bring evidence-based feedback into the classroom: Ai-generated feedback increases secondary students' text revision, motivation, and positive emotions," *Computers and Education: Artificial Intelligence*, vol. 6, p. 100199, 2024.

[24] A. Kinder, F. J. Briese, M. Jacobs, N. Dern, N. Glodny, S. Jacobs, and S. Leßmann, "Effects of adaptive feedback generated by a large language model: A case study in teacher education," *Computers and Education: Artificial Intelligence*, vol. 8, p. 100349, 2025.

[25] Z. Chen, L. Gao, Z. Wang, D. Ren, T. Yang, and Y. Shi, "Enhancing large language models for precise classification of teacher praise discourse: A fine-tuning approach," in *2024 International Conference on Intelligent Education and Intelligent Research (IEIR)*. Macau, China: IEEE, Nov. 2024, pp. 1–6.

[26] G. Rasch, *Probabilistic Models for Some Intelligence and Attainment Tests*, expanded ed. Chicago, IL: University of Chicago Press, 1993.

[27] M. D. Reckase, "Multidimensional item response theory models," in *Multidimensional Item Response Theory*. New York, NY: Springer, 2009, pp. 79–112.

[28] J. De La Torre, "The generalized dina model framework," *Psychometrika*, vol. 76, no. 2, pp. 179–199, 2011.

[29] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," *User Modeling and User-Adapted Interaction*, vol. 4, no. 4, pp. 253–278, 1995.

[30] Z. Wang, M. Wang, C. Zeng, and L. Li, "Scb-detr: Multiscale deformable transformers for occlusion-resilient student learning behavior detection in smart classroom," *IEEE Transactions on Computational Social Systems*, pp. 1–20, 2025.

[31] C. Zeng, F. Zou, S. Xia, and Z. Wang, "Sau-net: Saliency-based adaptive unfolding network for interpretable high-quality image compressed sensing in internet of things," *IEEE Internet of Things Journal*, vol. 12, no. 14, pp. 28 672–28 688, Jul. 2025.

[32] Z. Wang, L. Li, C. Zeng, S. Dong, and J. Sun, "Slbdetection-net: Towards closed-set and open-set student learning behavior detection in smart classroom of k-12 education," *Expert Systems with Applications*, vol. 260, p. 125392, Jan. 2025.

[33] C. Zeng, Y. Yu, Z. Wang, S. Xia, H. Cui, and X. Wan, "Gsista-net: Generalized structure ista networks for image compressed sensing based on optimized unrolling algorithm," *Multimedia Tools and Applications*, Mar. 2024.

[34] Z. Wang, L. Li, C. Zeng, S. Dong, and J. Sun, "Slb-mamba: A vision mamba for closed and open-set student learning behavior detection," *Applied Soft Computing*, vol. 180, p. 113369, Aug. 2025.

[35] C. Zeng, S. Xia, Z. Wang, and X. Wan, "Multi-channel representation learning enhanced unfolding multi-scale compressed sensing network for high quality image reconstruction," *Entropy*, vol. 25, no. 12, p. 1579, Dec. 2023.

[36] Z. Wang, M. Wang, C. Zeng, and L. Li, "Sbd-net: Incorporating multi-level features for an efficient detection network of student behavior in smart classrooms," *Applied Sciences*, vol. 14, no. 18, p. 8357, Jan. 2024.

[37] L. Li, Z. Wang, and T. Zhang, "Gbh-yolov5: Ghost convolution with bottleneckcsp and tiny target prediction head incorporating yolov5 for pv panel defect detection," *Electronics*, vol. 12, no. 3, pp. 1–15, Jan. 2023.

[38] Z. Wang, J. Yao, C. Zeng, L. Li, and C. Tan, "Students' classroom behavior detection system incorporating deformable detr with swin transformer and light-weight feature pyramid network," *Systems*, vol. 11, no. 7, p. 372, Jul. 2023.

[39] C. Zeng, J. Ye, Z. Wang, N. Zhao, and M. Wu, "Cascade neural network-based joint sampling and reconstruction for image compressed sensing," *Signal, Image and Video Processing*, vol. 16, no. 1, pp. 47–54, Feb. 2022.

[40] C. Zeng, K. Yan, Z. Wang, Y. Yu, S. Xia, and N. Zhao, "Abs-cam: A gradient optimization interpretable approach for explanation of convolutional neural networks," *Signal, Image and Video Processing*, pp. 1–8, Jul. 2022.

[41] Z. Wang, L. Li, C. Zeng, and J. Yao, "Student learning behavior recognition incorporating data augmentation with learning feature representation in smart classrooms," *Sensors*, vol. 23, no. 19, p. 8190, Jan. 2023.

[42] C. Zeng, Z. Wang, Z. Wang, K. Yan, and Y. Yu, "Image compressed sensing and reconstruction of multi-scale residual network combined with channel attention mechanism," *Journal of Physics: Conference Series*, vol. 2010, no. 1, p. 012134, Sep. 2021.

[43] Z. Wang, Z. Wang, C. Zeng, Y. Yu, and X. Wan, "High-quality image compressed sensing and reconstruction with multi-scale dilated convolutional neural network," *Circuits, Systems, and Signal Processing*, pp. 1–24, Sep. 2022.

[44] C. Zeng, Z. Wang, and Z. Wang, "Image reconstruction of iot based on parallel cnn," in *2020 International Conferences on Internet of Things (iThings)*. Rhodes, Greece: IEEE, Nov. 2020, pp. 258–263.

[45] Z. Wang, J. Yao, C. Zeng, W. Wu, H. Xu, and Y. Yang, "Yolov5 enhanced learning behavior recognition and analysis in smart classroom with multiple students," in *2022 International Conference on Intelligent Education and Intelligent Research (IEIR)*. IEEE, Dec. 2022, pp. 23–29.

[46] Y. Tian, X. Wang, H. Yao, J. Chen, Z. Wang, and L. Yi, "Occlusion handling using moving volume and ray casting techniques for augmented reality systems," *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 16 561–16 578, Jul. 2018.

[47] Z. Wang, C. Zuo, and C. Zeng, "Sae based unified double jpeg compression detection system for web image forensics," *International Journal of Web Information Systems*, vol. 17, no. 2, pp. 84–98, Apr. 2021.

[48] F. Wang, Q. Liu, E. Chen *et al.*, "Neural cognitive diagnosis for intelligent education systems," in *AAAI*, 2020.

[49] Q. Zheng, Z. Chen, Z. Wang, G. Hou, C. Liu, C. Zhao, and C. Zou, "Enhancing domain adaptation in speaker verification via partially shared adversarial network," in *Advanced Intelligent Computing Technology and Applications*, D.-S. Huang, Q. Zhang, C. Zhang, and W. Chen, Eds. Singapore: Springer Nature, 2025, pp. 158–168.

[50] Z. Wang, L. Li, and C. Zeng, "Dwmgrad: An innovative neural network optimization approach using dynamic window data for adaptive updating of momentum and learning rate," *Applied Intelligence*, vol. 55, no. 15, p. 1026, Oct. 2025.

[51] C. Zeng, Y. Zhao, and Z. Wang, "Smdrl: Self-supervised mobile device representation learning framework for recording source identification from unlabeled data," *Expert Systems with Applications*, vol. 282, p. 127635, Jul. 2025.

[52] Z. Chen, C. Zhao, Z. Wang, C. Liu, Q. Zheng, and C. Zou, "Ds-btian: A novel deep-shallow bidirectional transformer interactive attention network for multimodal emotion recognition," in *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2025, pp. 1–5.

[53] C.-Y. Zeng, C.-F. Ma, Z.-F. Wang, and J.-X. Ye, "Stacked autoencoder networks based speaker recognition," in *2018 International Conference on Machine Learning and Cybernetics (ICMLC)*. Chengdu: IEEE, Jul. 2018, pp. 294–299.

[54] Z. Chen, C. Liu, Z. Wang, C. Zhao, M. Lin, and Q. Zheng, "Mtlser: Multi-task learning enhanced speech emotion recognition with pre-trained acoustic model," *Expert Systems with Applications*, vol. 273, p. 126855, May 2025.

[55] Q. Zheng, Z. Chen, Z. Wang, H. Liu, and M. Lin, "Meconformer: Highly representative embedding extractor for speaker verification via incorporating selective convolution into deep speaker encoder," *Expert Systems with Applications*, vol. 244, p. 123004, Jun. 2024.

[56] C. Zeng, Y. Zhao, Z. Wang, K. Li, X. Wan, and M. Liu, "Squeeze-and-excitation self-attention mechanism enhanced digital audio source recognition based on transfer learning," *Circuits, Systems, and Signal Processing*, Sep. 2024.

[57] Z. Wang, S. Duan, C. Zeng, X. Yu, Y. Yang, and H. Wu, "Robust speaker identification of iot based on stacked sparse denoising auto-encoders," in *2020 International Conferences on Internet of Things (iThings)*. Rhodes, Greece: IEEE, Nov. 2020, pp. 252–257.

[58] C. Zeng, Y. Zhao, and Z. Wang, "Mobileformer: Cross-scale and multi-level representation learning based mobile recording device recognition for consumer electronics," *IEEE Transactions on Consumer Electronics*, pp. 1–1, 2024.

[59] Z. Wang, Q. Liu, J. Chen, and H. Yao, "Recording source identification using device universal background model," in *2015 International Conference of Educational Innovation through Technology (EITT)*. Wuhan, China: IEEE, Oct. 2015, pp. 19–23.

[60] C. Zeng, K. Li, and Z. Wang, "Enfformer: Long-short term representation of electric network frequency for digital audio tampering detection," *Knowledge-Based Systems*, vol. 297, p. 111938, Aug. 2024.

[61] Z.-Y. Zhu, Q.-H. He, X.-H. Feng, Y.-X. Li, and Z.-F. Wang, "Liveness detection using time drift between lip movement and voice," in *2013*

International Conference on Machine Learning and Cybernetics, vol. 02, Jul. 2013, pp. 973–978.
[62] C. Zeng, S. Kong, Z. Wang, K. Li, Y. Zhao, X. Wan, and Y. Chen, "Discriminative component analysis enhanced feature fusion of electrical network frequency for digital audio tampering detection," *Circuits, Systems, and Signal Processing*, Jul. 2024.
[63] Z.-F. Wang, G. Wei, and Q.-H. He, "Channel pattern noise based playback attack detection algorithm for speaker recognition," in *2011 International Conference on Machine Learning and Cybernetics*, vol. 4, Jul. 2011, pp. 1708–1713.
[64] C. Zeng, S. Kong, Z. Wang, K. Li, Y. Zhao, X. Wan, and Y. Chen, "Digital audio tampering detection based on spatio-temporal representation learning of electrical network frequency," *Multimedia Tools and Applications*, pp. 1–21, Mar. 2024.
[65] C. Zeng, S. Kong, Z. Wang, S. Feng, N. Zhao, and J. Wang, "Deletion and insertion tampering detection for speech authentication based on fluctuating super vector of electrical network frequency," *Speech Communication*, vol. 158, p. 103046, Mar. 2024.
[66] C. Zeng, S. Feng, Z. Wang, X. Wan, Y. Chen, and N. Zhao, "Spatio-temporal representation learning enhanced source cell-phone recognition from speech recordings," *Journal of Information Security and Applications*, vol. 80, p. 103672, Feb. 2024.
[67] C. Zeng, S. Feng, Z. Wang, Y. Zhao, K. Li, and X. Wan, "Audio source recording device recognition based on representation learning of sequential gaussian mean matrix," *Forensic Science International: Digital Investigation*, vol. 48, p. 301676, Mar. 2024.
[68] C. Zeng, S. Kong, Z. Wang, K. Li, and Y. Zhao, "Digital audio tampering detection based on deep temporal–spatial features of electrical network frequency," *Information*, vol. 14, no. 5, p. 253, May 2023.
[69] Z.-F. Wang, J. Wang, C.-Y. Zeng, Q.-S. Min, Y. Tian, and M.-Z. Zuo, "Digital audio tampering detection based on enf consistency," in *2018 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*. Chengdu: IEEE, Jul. 2018, pp. 209–214.
[70] C. Zeng, S. Feng, D. Zhu, and Z. Wang, "Source acquisition device identification from recorded audio based on spatiotemporal representation learning with multi-attention mechanisms," *Entropy*, vol. 25, no. 4, p. 626, Apr. 2023.
[71] Z. Wang, J. Zhan, G. Zhang, D. Ouyang, and H. Guo, "An end-to-end transfer learning framework of source recording device identification for audio sustainable security," *Sustainability*, vol. 15, no. 14, p. 11272, Jan. 2023.
[72] Z. Chen, M. Lin, Z. Wang, Q. Zheng, and C. Liu, "Spatio-temporal representation learning enhanced speech emotion recognition with multi-head attention mechanisms," *Knowledge-Based Systems*, vol. 281, p. 111077, Dec. 2023.
[73] C. Zeng, Y. Yang, Z. Wang, S. Kong, and S. Feng, "Audio tampering forensics based on representation learning of enf phase sequence," *International Journal of Digital Crime and Forensics*, vol. 14, no. 1, pp. 1–19, Jan. 2022.
[74] Z. Wang, Y. Yang, C. Zeng, S. Kong, S. Feng, and N. Zhao, "Shallow and deep feature fusion for digital audio tampering detection," *EURASIP Journal on Advances in Signal Processing*, vol. 2022, no. 69, pp. 1–20, Aug. 2022.
[75] C. Zeng, D. Zhu, Z. Wang, M. Wu, W. Xiong, and N. Zhao, "Spatial and temporal learning representation for end-to-end recording device identification," *EURASIP Journal on Advances in Signal Processing*, vol. 2021, no. 1, p. 41, 2021.
[76] C. Zeng, D. Zhu, Z. Wang, Z. Wang, N. Zhao, and L. He, "An end-to-end deep source recording device identification system for web media forensics," *International Journal of Web Information Systems*, vol. 16, no. 4, pp. 413–425, Aug. 2020.
[77] Z.-F. Wang, Q.-H. He, X.-Y. Zhang, H.-Y. Luo, and Z.-S. Su, "Playback attack detection based on channel pattern noise," *Journal of South China University of Technology*, vol. 39, no. 10, pp. 7–12, 2011.
[78] C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas, and J. Sohl-Dickstein, "Deep knowledge tracing," in *NIPS*, vol. 28, 2015.
[79] A. Ghosh, N. Heffernan, and A. S. Lan, "Context-aware attentive knowledge tracing," in *ACM SIGKDD*, 2020, pp. 2330–2339.
[80] H. Nakagawa, Y. Iwasawa, and Y. Matsuo, "Graph-based knowledge tracing: Modeling student proficiency using graph neural network," in *2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Thessaloniki, Greece, 2019, pp. 156–163.
[81] E. Kasneci *et al.*, "Chatgpt for good? on opportunities and challenges of large language models for education," *Computers and Education: Artificial Intelligence*, vol. 4, p. 100179, 2023.
[82] B. Nguyen *et al.*, "Towards generalized methods for automatic question generation in education," in *Proceedings of the European Conference on Technology Enhanced Learning (ECTEL)*, 2022.