Enhancing Cognitive Game Tracing via Diverse Information and Time-aware Modeling

Yang Yu, Bo Ma, Liangyu Chen, Mingsong Chen*

Shanghai Institute of Artificial Intelligence for Education, East China Normal University, Shanghai, China
MoE Engineering Research Center of SW/HW Co-Design Tech. and App., East China Normal University, Shanghai, China

Abstract

With the surge in cognitive gaming data, understanding players' learning patterns and cognitive growth has become increasingly important. These data offer valuable opportunities to study individual cognitive development during learning. However, the diversity of player profiles and the complexity of gaming tasks pose significant challenges for accurate skill prediction. Specifically, the heterogeneity of player profiles leads to diverse and complex learning trajectories; data sparsity and temporal dynamics further exacerbate these challenges. To address these challenges, we propose the MFCGT (Multi-feature Forget Cognitive Game Tracing) model. First, we perform multi-feature selection to extract key features from player behavior data to reduce noise and improve prediction accuracy. Second, we introduce a time-aware decay mechanism that simulates skill degradation using an exponential decay function, ensuring the model captures the impact of forgetting on learning trajectories. Finally, we incorporate an attention mechanism to dynamically identify the most relevant historical performance for the current task, thereby enhancing the model's predictive capability. Experimental results show that MFCGT significantly outperforms traditional models in skill prediction tasks. Additionally, MFCGT effectively captures players' learning dynamics and forgetting effects, providing more accurate learning predictions.

Keywords: Cognitive Game; Knowledge Tracing; Skill Acquisition and Learning.

Introduction

In recent years, cognitive games have gained significant attention as a powerful tool to understand human learning and cognitive development. Platforms offer a wide range of gamified cognitive tasks that assess various cognitive abilities, such as memory, attention, language, and reasoning. These games provide a unique opportunity to study how people learn and improve over time, especially in large-scale settings where millions of users interact with the platform. The ability to track and predict cognitive growth through these games has become a critical area of research, with applications in personalized learning, cognitive training, and educational technology (Maji, Acharya, Paul, Chakraborty, & Basu, 2025). Knowledge Tracing (KT), a well-established approach in educational data mining, aims to model and predict students' knowledge states based on their learning trajectories (Tsutsumi, Guo, Kinoshita, & Ueno, 2024). While traditional KT focuses on academic subjects such as mathematics or language, Cognitive Game Tracing (CGT) extends this concept to cognitive games, where the goal is to track and predict the development of cognitive skills of players over time (Maji et al., 2025). However, CGT faces unique challenges due to the complexity and diversity of cognitive tasks, as well as the heterogeneity of player backgrounds and learning behaviors (Guzmán, Rengifo, & Cena, 2024).

Despite the significant potential of cognitive games for studying learning dynamics, several challenges hinder the development of effective Cognitive Game Tracing (CGT) models (Yeratziotis et al., 2024). One major challenge is player heterogeneity, as players come from diverse backgrounds with varying ages, educational levels, and cognitive abilities. This diversity leads to highly varied learning trajectories, making it difficult to generalize models across different groups of players. Another challenge is task complexity, as cognitive games often involve multiple cognitive processes, such as memory, attention, and reasoning. The complexity of these tasks, combined with adaptive difficulty levels, adds difficulty to modeling learning trajectories (Guzmán et al., 2024). Additionally, data sparsity poses a significant issue, as players may not complete all tasks or may engage with certain tasks more frequently than others (Christie, Cook, & Rafferty, 2024). This results in sparse data, where many learning trajectories are incomplete, making it challenging to build robust models. Finally, temporal dynamics present a critical challenge, as cognitive skills evolve over time, and players may experience skill decay if they do not practice regularly. Capturing these temporal dynamics is crucial for accurate tracing, but it remains a significant challenge in existing models.

To address these challenges, several methods have been proposed to model learning trajectories, such as Bayesian Knowledge Tracing (BKT) and Deep Knowledge Tracing (DKT) (Saric-Grgic, Grubisic, & Gaspar, 2024b). These methods typically rely on historical performance data to predict future performance. However, they often fail to account for the multi-dimensional nature of cognitive tasks and the temporal dynamics of skill acquisition and decay. More advanced methods, such as those based on transformer architectures, have been developed to capture complex learning patterns and model sequential learning data. Others employ factor analysis to identify latent cognitive abilities. Despite these advancements, the existing methods still face several limitations. One key issue is limited feature utilization, as many models rely solely on task performance data, ignoring other important features such as player ages, task difficulty, and practice frequency. This narrow focus restricts their ability

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

to capture the full complexity of learning dynamics. Another limitation is the inadequate handling of forgetting effects, as most models do not explicitly account for skill decay over time, which is a critical aspect of cognitive learning. In addition, there is a lack of attention to task-specific variations. To overcome the limitations of existing methods, we propose the Multi-feature Forget Cognitive Game Tracing (MFCGT) model, which integrates multiple innovative components to effectively capture players' learning dynamics and cognitive growth. Our approach addresses the key challenges in Cognitive Game Tracing (CGT) by leveraging multi-feature selection, a time-aware forgetting mechanism, and an attention mechanism. First, we utilize a gradient boosting framework, to select the most relevant features from players' behavioral data. Second, recognizing that cognitive skills decay over time, we introduce a forgetting mechanism that simulates skill degradation. This mechanism uses an exponential decay function to adjust players' skill levels based on the time elapsed since their last practice, enabling the model to more accurately predict long-term skill retention. Finally, we employ an attention mechanism that dynamically weights the contributions of different tasks to the overall skill prediction. This allows the model to focus on the most relevant historical performance data for each task. Our work makes key contributions to the field of cognitive game tracing:

- Multi-feature integration incorporates a wide range of features, such as player demographics and task-specific data, to provide a more comprehensive understanding of learning dynamics in cognitive games.
- The skill decay and contextual attention mechanism introduces a time-aware approach to capture the temporal dynamics of skill acquisition and decay, addressing critical limitations of existing methods.
- Task-specific attention allows the model to adapt to the unique cognitive demands of different cognitive tasks, enhancing its ability to predict performance.
- Improved prediction accuracy is demonstrated through experiments, showing that our MFCGT model outperforms existing methods in predicting player cognitive performance, offering an accurate and robust approach to cognitive game tracing.

Related Work

Gamification in Learning

Cognitive games have emerged as a promising tool for studying human learning and cognitive development. Platforms provide a diverse range of gamified tasks designed to assess various cognitive abilities, including memory, attention, and reasoning.

The integration of gamification into e-learning has become a popular strategy to boost student engagement and motivation. By incorporating game-like elements such as points, badges, and leaderboards, educators aim to create a more immersive learning environment. Research has demonstrated that these elements can lead to improved learning performance (Gweon et al., 2015). However, not all gamified systems achieve their intended educational goals. Some systems are criticized for prioritizing entertainment over learning, resulting in designs that resemble animations or games rather than effective educational tools (Schodde, Bergmann, & Kopp, 2017).

The impact of gamification on learning performance has been a focal point of research. Studies suggest that well-implemented gamified systems can enhance learning outcomes by providing immediate feedback, fostering a sense of accomplishment, and maintaining student engagement over time (Yeratziotis et al., 2024). However, the success of gamification depends heavily on how well the game elements align with the learning objectives (Ciuffreda et al., 2024). For example, adaptive gamification, which tailors game elements to individual player types, has shown promise in improving learning performance (Kantharaju et al., 2019). This highlights the importance of personalization and adaptability in maximizing the educational benefits of gamification (Taghavi, Ghorbani, & Delrobaei, 2024).

User Modeling

User modeling is essential for understanding and predicting learner behavior, enabling the design of more effective educational interventions.

To assess students' knowledge proficiency, researchers have developed various psychometric models, such as Item Response Theory (IRT) and Deterministic Inputs, Noisy "And" Gate (DINA) models (Welling, Sheng, & Zhu, 2024). For example, IRT models represent students' knowledge ability as a continuous variable and use logistic functions to predict the likelihood of correctly answering a question. IRT models represent students' knowledge ability as a continuous variable, and DINA models use a binary latent vector to indicate whether a student has mastered the required skills. One of the most widely used methods for knowledge tracing is Bayesian Knowledge Tracing (BKT), which is based on Hidden Markov Models (Saric-Grgic, Grubisic, & Gaspar, 2024a). Traditional models like BKT and DKT have been adapted for use in educational games, but they often struggle with the dynamic and multi-skill nature of these environments. Recent approaches, such as the Hierarchical Memory Network (HMN), attempt to address these challenges by simulating human memory processes, but their effectiveness in cognitive games remains an open question (S. Liu et al., 2021). Similarly, (Hooshyar, Huang, & Yang, 2022) introduces a deep learning-based approach to model players' knowledge states in educational games. The study emphasizes the challenges of tracing cognitive skills in games, where players often deploy multiple overlapping skills simultaneously. In conclusion, while significant progress has been made in cognitive games and user modeling, there is still much to explore, particularly in educational games.

Problem formulation

Given a set of users $U = \{u_1, u_2, \ldots, u_N\}$, a set of games $G = \{g_1, g_2, \ldots, g_M\}$, our goal is to model the cognitive development of each user over time based on their gameplay interactions. For each user u_j , we have a sequence of interactions with games, denoted as $I_j = \{(i_1, p_{i1}), (i_2, p_{i2}), \ldots, (i_k, p_{ik})\}$, where i_m is the game ID, p_{im} is the performance of the game and k is the total number of interactions. Each user u_j is characterized by different attributes such as age, level of education, etc., and a history of gameplay, denoted as $C_j = \{c_j^1, c_j^2, c_j^3, \ldots, c_j^s, his_j\}$.

The objective of Cognitive Game Tracing (CGT) is to model the knowledge state of each user u_j over time, based on their interaction sequence I_j and characteristics C_j, and to predict their performance on future gameplay interactions, specifically at the next game i_{k+1} in the sequence. The prediction task is to develop a model that takes as input the interaction sequence I_j and characteristics C_j of a user u_j and predicts the performance in the subsequent gameplay interaction $(i_{k+1}, p_{i_{k+1}})$.

Methodology

The Cognitive Game Tracing (CGT) framework is composed of a series of interconnected modules designed to encapsulate the complexity of cognitive gaming data. As shown in Figure 1, the framework is systematically structured into distinct components, including data preprocessing, feature selection and fusion, skill decay and contextual attention mechanism, and model prediction. Each component plays a critical role in the pipeline, from refining raw data to generating predictions.

The initial phase of our framework involves cleansing and organizing the data set. This step is crucial for subsequent analyses. The dataset is preprocessed to ensure that it is consistent and representative of user interactions within the gaming environment. The preprocessing phase employs temporal alignment and missing value imputation to construct uniform interaction sequences. Feature selection and fusion combines decision trees for identifying predictive patterns with cross-encoding layers to dynamically integrate temporal interactions and static user attributes. Skill decay and contextual attention models knowledge retention via exponential decay, enhanced by attention mechanism that weight historical states based on task similarity and temporal relevance.

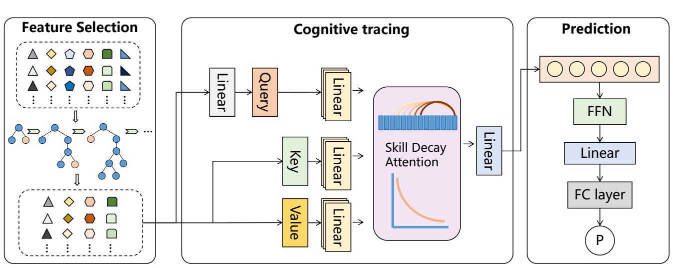

Figure 1: The framework of MFCGT model.

Algorithm 1 Implementation of MFCGT

Input: i) $U = \{u_1, u_2, \ldots, u_N\}$, set of users; ii) $G = \{g_1, g_2, \ldots, g_M\}$, set of games; iii) $I_j = \{(i_1, p_{i1}), (i_2, p_{i2}), \ldots, (i_k, p_{ik})\}$, interaction sequence for user u_j; iv) $C_j = \{(c_j^1, c_j^2, \ldots, c_j^s, his_j)\}$, user characteristics; v) T_1, T_2, number of training epochs;

Output: Predicted performance $p_{i_{k+1}}$;

$F \leftarrow \{f_1, f_2, \ldots, f_d\}$
$F_{\text{ranked}} \leftarrow [f_{(1)}, f_{(2)}, \ldots, f_{(d)}]$
$F_{\text{opt}} \leftarrow \{f_{(1)}, f_{(2)}, \ldots, f_{(k)}\}$
for $t \leftarrow 1$ to T_1 **do**
  for each user $u_j \in U$ **do**
    $x_t \leftarrow (I_j, F_{\text{opt}})$
  end for
end for
for $t \leftarrow 1$ to T_2 **do**
  for $u_j \in U$ **do**
    $s(i,j) \leftarrow \exp(-\lambda \cdot \Delta(i,j)) \cdot \sigma(q_i, k_j)$
    $\Delta(i,j) \leftarrow |i-j| \cdot \sum_{k=j+1}^{i} w_{k,j}$
    $\alpha_{i,j} \leftarrow \frac{\exp(s(i,j))}{\sum_{k=1}^{T} \exp(s(i,k))}$
    $h_i \leftarrow \sum_{j=1}^{T} \alpha_{i,j} \cdot h_j$
  end for
end for
return $p_{i_{k+1}} = \text{FC}(\text{FFN}(\text{Linear}(h_i)))$

Feature Selection

Our model extends the traditional knowledge tracing paradigm by incorporating a rich set of user features. The feature selection process in our Cognitive Game Tracing framework is a critical step that involves identifying and selecting the most informative features from the dataset. We employ a histogram-based algorithm to evaluate the gain of each feature. This algorithm quantifies the information gain for each feature, providing a preliminary ranking of feature importance. The features are then sorted based on their gain. Using the initial ranking, we construct a decision tree using a leaf-wise growth strategy. This strategy involves growing each leaf of the tree, which allows for the optimization of the model at each node. After the decision tree is constructed, we calculate the residual values, which represent the differences between the predicted and actual outcomes. The residual from the previous tree is used as the training sample for the next iteration. This iterative process continues, with each new tree being trained on the residuals of the previous one. At the end of the iterative process, the trees generated in each round are weighted and summed according to their performance to get the final model. This algorithm evaluates the significance of each feature within a dataset by counting the frequency at which they serve as decision-making thresholds. The features are then ranked in descending order according to their calculated significance. This ranking guides a systematic review of all features, where the least significant ones are considered for exclusion based on their impact on predictive accuracy. The process iterates through the entire feature set to determine the most effective subset, denoted as F_{opt}.

Algorithm 2 Feature Selection for Game Tracing

Input: i) $F = \{f_1, f_2, \ldots, f_d\}$, feature set; ii) y, target performance values; iii) k, number of folds for cross-validation; iv) τ, importance threshold;
Output: Optimal feature subset $F_{\text{opt}} \subseteq F$;

```
Initialize model M
{D_1, D_2, ..., D_k} ← split(D, k)
for i ← 1 to k do
    Train M on training data ⋃_{j≠i} D_j
    I_j^(i) ← compute_importance(f_j)
end for
I_j ← (1/k) ∑_{i=1}^{k} I_j^(i)
Ĩ_j ← I_j / max_k I_k
if max_j I_j > 2E[I_j] then
    Set τ ← 0.5 max_j I_j
else
    Set τ ← E[I_j] + σ(I_j)
end if
F_opt ← {f_j | Ĩ_j ≥ τ}
return F_opt
```

Feature Fusion

Following the feature selection stage in CGT, the subsequent step involves encoding the selected features to construct the input. This encoding process encompasses direct concatenation and cross-feature encoding. The direct concatenation method merges the game performance data with the optimal features, creating a comprehensive feature vector that serves as the model's input at each timestep. For cross-feature encoding, we employ a formula that integrates game ID, performance outcome, and selected features:

$$C(g_t, p_t) = g_t + [\max(g) + 1] \cdot p_t,$$

$$x_t = C(g_t, p_t) \oplus C(F_t, p_t) \oplus F_t.$$

Here, C denotes the cross-encoding operation, $\oplus$ signifies vector concatenation. g_t and p_t correspond to the game ID and performance, respectively, while F_t represents the identified optimal features.

Skill Decay and Contextual Attention Mechanism

To address the limitations of existing methods in cognitive game tracing, we propose a skill decay and contextual attention mechanism. This mechanism integrates a skill decay function and an attention mechanism to model the temporal dynamics of skill acquisition and decay, as well as task-specific variations in learning trajectories. The skill decay function is designed to simulate the gradual degradation of cognitive skills over time. The formula for the skill decay function is as follows:

$$s(i, j) = \exp\left(-\lambda \cdot \Delta(i, j)\right) \cdot \sigma(q_i, k_j),$$

where $s(t, \tau)$ represents the skill retention score at time t for a skill last practiced at time τ. It quantifies how much of the skill is retained over time. $\exp\left(-\lambda \cdot \Delta(i, j)\right)$ is the exponential decay term, where λ is a decay rate parameter that controls how quickly skills degrade over time. $\sigma(q_i, k_j)$ represents the similarity score between the current task q_i and the past task k_j. It is computed using a dot product or other similarity measures, scaled by a normalization factor.

The contextual time distance is a key component of the skill decay function. It not only considers the absolute time difference between i and j but also incorporates the contextual relevance of the tasks performed during that time. The formula for $\Delta(i, j)$ is defined as:

$$\Delta(i, j) = |i - j| \cdot \sum_{k=j+1}^{i} w_{k,j},$$

where $|i - j|$ is the absolute time difference between the current time i and previous practice time j. $w_{k,j}$ is a contextual weight that measures the relevance of the task performed at time k to the task performed at time j. It is computed using a softmax function over the similarity scores between tasks:

$$w_{k,j} = \frac{\exp(\sigma(q_k, k_j))}{\sum_{m=j+1}^{i} \exp(\sigma(q_m, k_j))}.$$

Model Prediction

Building upon the feature selection and skill decay mechanisms, the weighted feature representation is passed through a fully connected layer to generate the predicted performance score for the next cognitive task. To train the MFCGT model, we employ a loss function that measures the difference between the predicted performance $\hat{p}_{i_{k+1}}$ and the actual performance $p_{i_{k+1}}$. The loss function is defined as follows:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} \left(p_{i_{k+1}} - \hat{p}_{i_{k+1}}\right)^2,$$

where N is the total number of samples in the sequence, $p_{i_{k+1}}$ is the ground truth performance for the i-th sample, $\hat{p}_{i_{k+1}}$ is the predicted performance for the i-th sample. The goal of the training process is to minimize this loss function and improve the model's accuracy in predicting cognitive performance.

Experiment

Experimental Settings

Dataset. The dataset in this article comes from the project page of the Lumosity Cognitive Training Dataset (Steyvers & Schafer, 2020). Lumosity is an online platform that offers a variety of games to train and test users' cognitive abilities. The dataset has a very rich set of user data, with 36,296 users, including information such as user identifier, gender, age, and education level. Users were registered between August 1, 2013 and December 31, 2016 and are between the ages of 18 and 90. Users may be from the United States, Canada, or Australia, and their preferred language is English. At least 99% of the user's game history is done on the web product, not the mobile app. We also have a lot of game data in our dataset, with a total of more than 50 cognitive tasks, including the game ID, name, main domain, and main attributes. Games involve different cognitive domains such as attention, memory, math, etc. In our experiment, the overall dataset consists of six subsets: Attention, Flexibility, Language, Math, Memory, and Reasoning. There are a total of 36,736,284 sets of data in the game records and each user has a record of playing games.

Baseline Methods and Evaluation Metrics. Baseline models encompass classical architectures for processing sequential and non-sequential data. Recurrent Neural Networks (RNNs) (Grossberg, 2013) are designed to handle sequential data by propagating temporal information through hidden states. Long Short-Term Memory (LSTM) networks (Hochreiter & Schmidhuber, 1997) extend RNNs to address long-term dependency learning via specialized gating mechanisms. Gated Recurrent Units (GRUs) (Cho et al., 2014) further simplify LSTMs by merging gates into a unified update mechanism while retaining robust sequence modeling capabilities. For non-sequential tasks, the Multi-Layer Perceptron (MLP) (Kruse, Mostaghim, Borgelt, Braune, & Steinbrecher, 2022) provides a foundational feedforward structure with fully connected layers to learn input-output mappings. Deep Neural Networks (DNNs) (W. Liu et al., 2017) generalize this concept by stacking multiple layers to extract hierarchical representations from complex data. These models serve as fundamental baselines in domains such as time series forecasting and general machine learning tasks.

To assess model performance, we employ two widely used metrics: Mean Absolute Error (MAE) and Mean Squared Error (MSE). These metrics measure the discrepancy between predicted performances and ground truth values.

Experiment Result Analysis

During experiment, we observed high correlations among certain features, which led us to create a Kernel Density Estimation (KDE) plot to visually highlight these relationships. This plot is depicted in Figure 2, providing a view of the data distributions across different fields (Li, Zhang, & Journel, 2017). To systematically eliminate redundant features with strong collinearity, we employed a correlation heatmap, as shown in Figure 3. This heatmap effectively illustrated the interdependencies between factors.

The feature selection process prioritized features with high predictive importance while eliminating redundant variables with strong collinearity. As depicted in Figure 3, the correlation matrix revealed collinearity among specific factors. Notably, the pair nth_play_total and session exhibited a high correlation at a value of 0.95, suggesting near-identical measurements of engagement frequency. Guided by feature importance rankings, we retained nth_play_total with an importance score of 24 and excluded session due to its lower importance score of 19. Similarly, nth_play_master showed

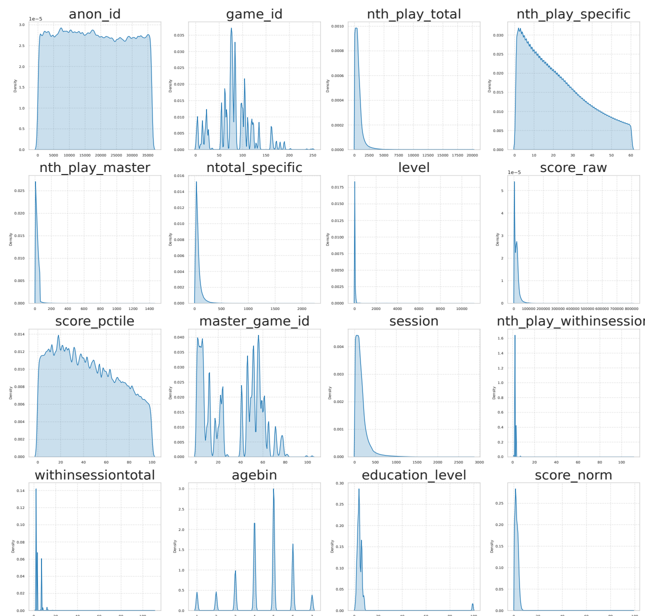

Figure 2: Kernel Density Graph.

a strong correlation with nth_play_specific at a value of 0.72, leading to the retention of the former, which held a higher importance score. The finalized feature set, including level, game_id, master_game_id, agebin, nth_play_master, and education_level, achieved minimal redundancy while preserving essential predictive signals. This optimized feature set effectively balances information richness and computational efficiency, mitigating overfitting risks inherent in high-dimensional behavioral datasets (Rumsey, 2011).

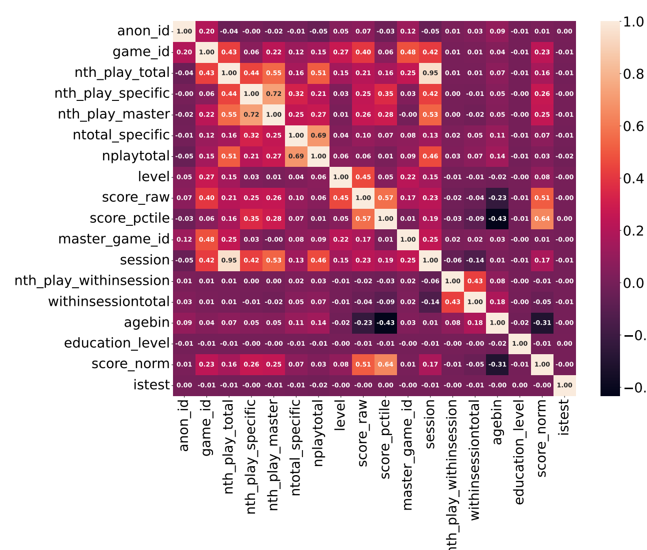

Figure 3: Correlation Heatmap Graph.

The change in the value of the loss over the course of our experimental results is shown in Figure 5. The results of our experiments are shown in the Table 1. As evidenced

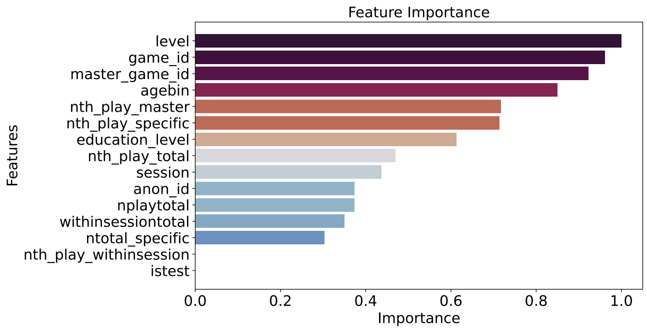

Figure 4: Normalized Feature Importance.

Table 1: Performance Comparison of Different Game Types.

Type	Model	MAE	MSE	Type	Model	MAE	MSE
Attention	MFCGT	**0.3298**	**0.3074**	Flexibility	MFCGT	**0.3881**	**0.3789**
	LSTM	0.3545	0.3339		LSTM	0.4390	0.4098
	MLP	3.6093	1.6687		MLP	0.9471	0.7485
	DNN	3.6092	1.6688		DNN	2.2588	1.2548
	GRU	0.3948	0.3594		GRU	0.4739	0.4392
	RNN	0.4245	0.3719		RNN	0.4725	0.4359
Language	MFCGT	**0.3688**	**0.3577**	Math	MFCGT	**0.2651**	**0.2443**
	LSTM	0.4152	0.3871		LSTM	0.3582	0.3218
	MLP	1.7401	1.1050		MLP	1.5956	1.0117
	DNN	1.7378	1.1056		DNN	1.5957	1.0114
	GRU	0.3831	0.3880		GRU	0.3740	0.3280
	RNN	0.4458	0.4080		RNN	0.3966	0.3539
Memory	MFCGT	**0.1177**	**0.1541**	Reasoning	MFCGT	**0.0963**	**0.1449**
	LSTM	0.2582	0.2459		LSTM	0.2943	0.2844
	MLP	2.7786	1.3066		MLP	0.8056	0.6501
	DNN	1.0114	1.3030		DNN	2.7789	1.3030
	GRU	0.3996	0.3353		GRU	0.3123	0.3058
	RNN	0.4059	0.3507		RNN	0.2327	0.2499

by the training trajectories in Figure 5, while all models exhibit decreasing loss values during training, MFCGT achieves the lowest final loss values with particularly notable advantages in complex tasks such as Memory and Reasoning. The model's MAE scores of 0.118 and 0.096 in these respective domains show absolute reductions of 0.141 and 0.198 over LSTM baselines, suggesting enhanced capability in capturing temporal dependencies. Sequential architectures generally outperform non-temporal models, with MFCGT showing 6.9% MAE reduction over LSTM and 16.5% over GRU in attention-related tasks. The model demonstrates task-specific strengths that align with its innovations. For mathematical tasks, MFCGT shows 0.093 MAE reduction over LSTM. In language tasks, it achieves 3.7% MAE reduction over GRU, while MFCGT yields 13.7% MSE reduction over GRU in flexibility tasks. These results validate the core design principles of our approach. The consistent outperformance spans diverse cognitive domains, including 0.085 average MAE reduction over LSTM and 0.106 over GRU across all tasks.

Ablation Study

To assess the individual contributions of key components within our MFCGT model, we conducted an ablation study with two variant models: MCGT, which omits the Skill Decay and Contextual Attention Mechanism, and FCGT lacks multi-feature selection. The results underscore the critical importance of both components. Without the attention mechanism, MCGT shows increased MAE and MSE across all domains, indicating its essential role in capturing temporal dynamics and task-specific variations. Similarly, FCGT's absence of comprehensive feature integration leads to degraded performance, emphasizing the value of diverse features. This study validates the synergistic effects of multi-feature selection, skill decay modeling and contextual attention in our MFCGT framework, each addressing unique challenges in cognitive game tracing.

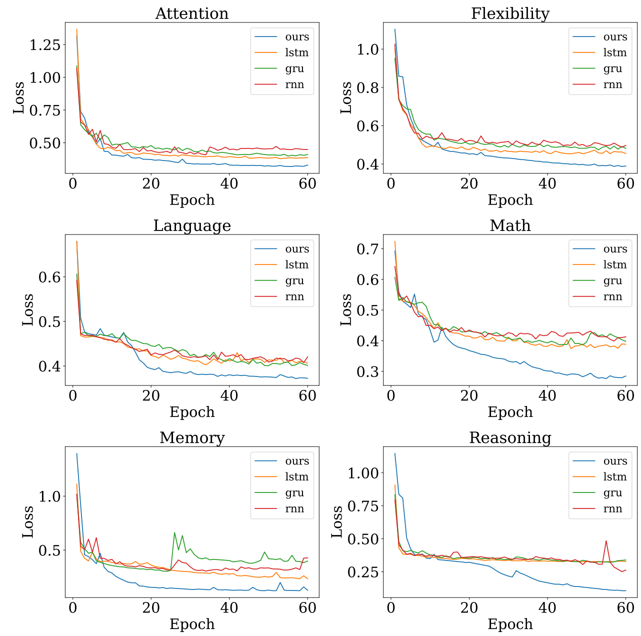

Figure 5: Visualization of Loss Convergence.

Table 2: Ablation Study Across Different Game Types.

Type	Model	MAE	MSE	Type	Model	MAE	MSE
Attention	FCGT	0.3421	0.3197	Flexibility	FCGT	0.4138	0.3995
	MCGT	0.3735	0.3452		MCGT	0.4487	0.4241
Language	FCGT	0.3973	0.3772	Math	FCGT	0.3167	0.2859
	MCGT	0.3784	0.3676		MCGT	0.3325	0.2997
Memory	FCGT	0.1882	0.2058	Reasoning	FCGT	0.2053	0.2245
	MCGT	0.3293	0.2965		MCGT	0.2233	0.2437

Conclusion

Our proposed model notably advances cognitive game tracing through multi-feature selection and skill decay and contextual attention mechanism. It shows superior performance in skill prediction across diverse cognitive domains, tackling challenges like user heterogeneity and temporal dynamics. Experimental results show MFCGT outperforms traditional models with significant improvements in prediction performance. This advancement also has the potential to enhance cognitive assessment and training through cognitive gaming.

References

Chalnick, A., & Billman, D. (1988). Unsupervised learning of correlational structure. In *Proceedings of the tenth annual conference of the cognitive science society* (pp. 510–516). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Christie, S. T., Cook, C., & Rafferty, A. N. (2024). Uncertainty-preserving deep knowledge tracing with state-space models. In D. A. Joyner, B. Paaßen, & C. D. Epp (Eds.), *Proceedings of the 17th international conference on educational data mining, EDM 2024.* International Educational Data Mining Society.

Ciuffreda, I., Morresi, N., Chieh, H., Revel, G. M., Hsu, Y., Nap, H. H., & Casaccia, S. (2024). Measure the cognitive decline of people with dementia using games for cognitive training. In *IEEE international symposium on medical measurements and applications* (pp. 1–5). IEEE.

Feigenbaum, E. A. (1963). The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought.* New York: McGraw-Hill.

Grossberg, S. (2013). Recurrent neural networks. *Scholarpedia*, *8*(2), 1888.

Guzmán, D. E., Rengifo, C. F., & Cena, C. E. G. (2024). Serious games for cognitive rehabilitation in older adults: A conceptual framework. *Multimodal Technol. Interact.*, *8*(8), 64.

Gweon, G., Lee, H., Dorsey, C., Tinker, R., Finzer, W., & Damelin, D. (2015). Tracking student progress in a game-like learning environment with a monte carlo bayesian knowledge tracing model. In *Proceedings of the fifth international conference on learning analytics and knowledge, LAK '15* (pp. 166–170).

Hill, J. A. C. (1983). A computational model of language acquisition in the two-year old. *Cognition and Brain Theory*, *6*, 287–317.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Hooshyar, D., Huang, Y., & Yang, Y. (2022). Gamedkt: Deep knowledge tracing in educational games. *Expert Syst. Appl.*, *196*, 116670.

Kantharaju, P., Alderfer, K. B., Zhu, J., Char, B. W., Smith, B. K., & Ontañón, S. (2019). Tracing player knowledge in a parallel programming educational game. *CoRR*, *abs/1908.05632*.

Kruse, R., Mostaghim, S., Borgelt, C., Braune, C., & Steinbrecher, M. (2022). Multi-layer perceptrons. In *Computational intelligence: a methodological introduction* (pp. 53–124). Springer.

Li, J., Zhang, C., & Journel, A. (2017). Visualising data distributions with kernel density estimation and reduced chi-squared statistic. *Journal of Geostatistics*, *5*(2).

Liu, S., Zou, R., Sun, J., Zhang, K., Jiang, L., Zhou, D., & Yang, J. (2021). A hierarchical memory network for knowledge tracing. *Expert Syst. Appl.*, *177*, 114935.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26.

Maji, P. K., Acharya, S., Paul, P., Chakraborty, S., & Basu, S. (2025). Deep learning inspired game-based cognitive assessment for early dementia detection. *Eng. Appl. Artif. Intell.*, *142*, 109901.

Matlock, T. (2001). *How real is fictive motion?* Doctoral dissertation, Psychology Department, University of California, Santa Cruz.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohlsson, S., & Langley, P. (1985). *Identifying solution paths in cognitive diagnosis* (Tech. Rep. No. CMU-RI-TR-85-2). Pittsburgh, PA: Carnegie Mellon University, The Robotics Institute.

Rumsey, D. J. (2011). Statistics for dummies. *Hoboken, NJ: Wiley Publishing*, 284.

Saric-Grgic, I., Grubisic, A., & Gaspar, A. (2024a). Correction: Twenty-five years of bayesian knowledge tracing: a systematic review. *User Model. User Adapt. Interact.*, *34*(5), 2115.

Saric-Grgic, I., Grubisic, A., & Gaspar, A. (2024b). Twenty-five years of bayesian knowledge tracing: a systematic review. *User Model. User Adapt. Interact.*, *34*(4), 1127–1173.

Schodde, T., Bergmann, K., & Kopp, S. (2017). Adaptive robot language tutoring based on bayesian knowledge tracing and predictive decision-making. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction, HRI 2017* (pp. 128–136). ACM.

Shrager, J., & Langley, P. (Eds.). (1990). *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.

Steyvers, M., & Schafer, R. J. (2020). Inferring latent learning factors in large-scale cognitive training data. *Nature Human Behaviour*, *4*(11), 1145–1155.

Taghavi, M. F., Ghorbani, F., & Delrobaei, M. (2024). Development of an augmented-reality-based serious game: A cognitive assessment study. *IEEE Trans. Cogn. Dev. Syst.*, *16*(3), 1087–1094.

Tsutsumi, E., Guo, Y., Kinoshita, R., & Ueno, M. (2024). Deep knowledge tracing incorporating a hypernetwork with independent student and item networks. *IEEE Trans. Learn. Technol.*, *17*, 951–965.

Welling, W. S., Sheng, Y., & Zhu, M. M. (2024). Cuda-aware MPI implementation of gibbs sampling for an IRT model. *Clust. Comput.*, *27*(2), 1821–1830.

Yeratziotis, A., Fotiadis, T., Achilleos, A., Savvides, S., Mettouris, C., Christoforou, C., … Mytides, A. (2024). A game-based cognitive intervention for young learners with reading difficulties. *SN Comput. Sci.*, *5*(6), 701.