AAAI-25
February 25 – March 4, 2025
Philadelphia, USA

USTC Campus Buildings

Agent4Edu: Generating Learner Response Data by Generative Agents for Intelligent Education Systems

Weibo Gao[1] Qi Liu[1,2] Linan Yue[1] Fangzhou Yao[1]
Rui Lv[1] Zheng Zhang[1] Hao Wang[1] Zhenya Huang[1,2]

Contact Us: : weibogao@mail.ustc.edu.cn : qiliuql@ustc.edu.cn

[1] University of Science and Technology of China (USTC)

Reporter: Weibo Gao[1]

[2] State Key Lab. of Cognitive Intelligence, Hefei, China

Outline

Background

- **Personalized Learning** represents a promising educational strategy within intelligent education systems, aiming to **enhance learners' practice efficiency**

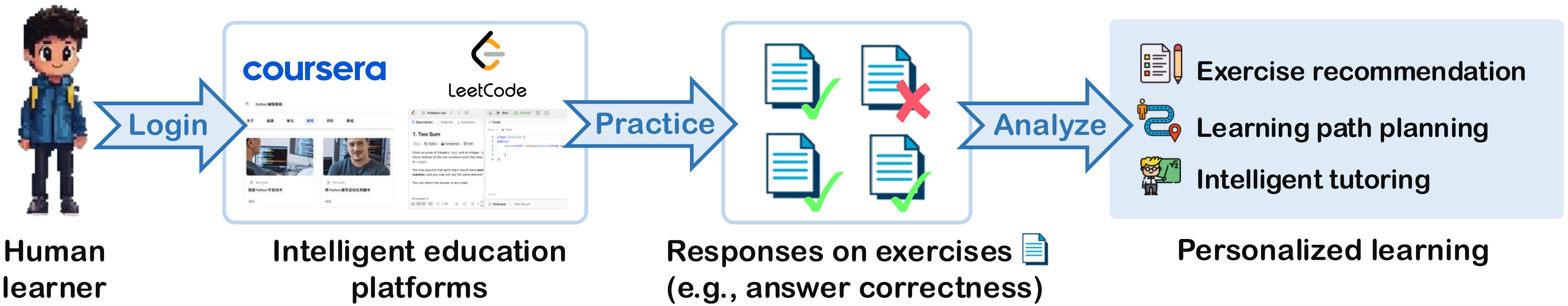
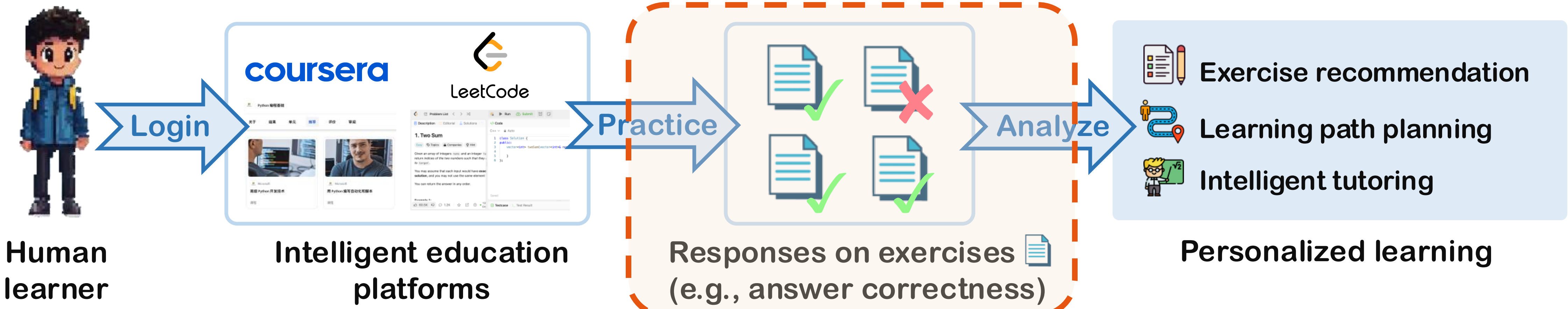


Fig. Personalized learning in intelligent education systems

Background

- **Personalized Learning** represents a promising educational strategy within intelligent education systems, aiming to **enhance learners' practice efficiency**



⚠ The **lack of offline response data** and potential biases with online practice create a **gap between offline and online performance** of personalized learning algorithms

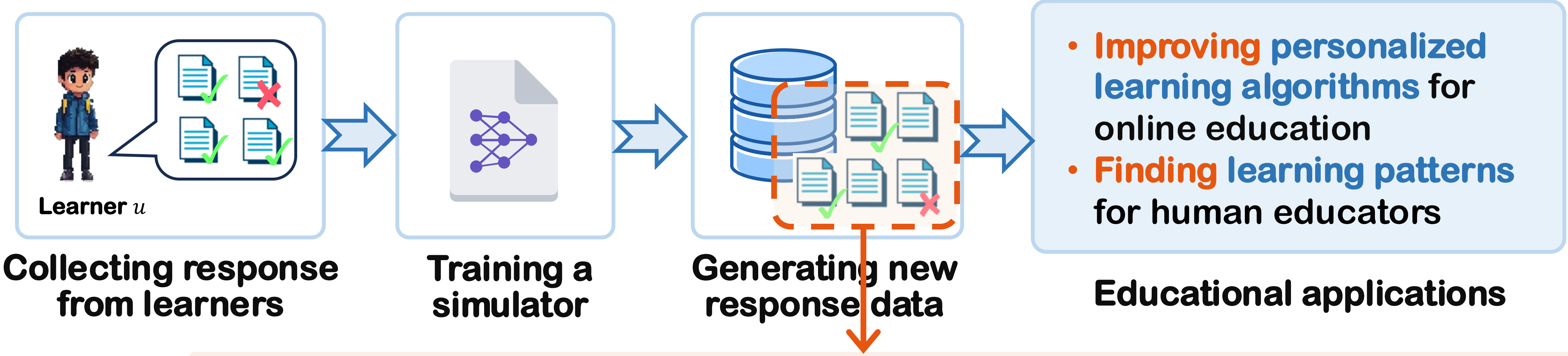
Learner Response Data Simulation

- **Simulating learners' response data** is a promising approach
  - Faithfully **captures** human learners' **response patterns**
  - Seamlessly **interacts** with personalized learning algorithms

Fig. The Current research pattern in simulating learners' response data

Limitations

- **Simulating learners' response data** is a promising approach
  - Faithfully **captures** human learners' **response patterns**
  - Seamlessly **interacts** with personalized learning algorithms

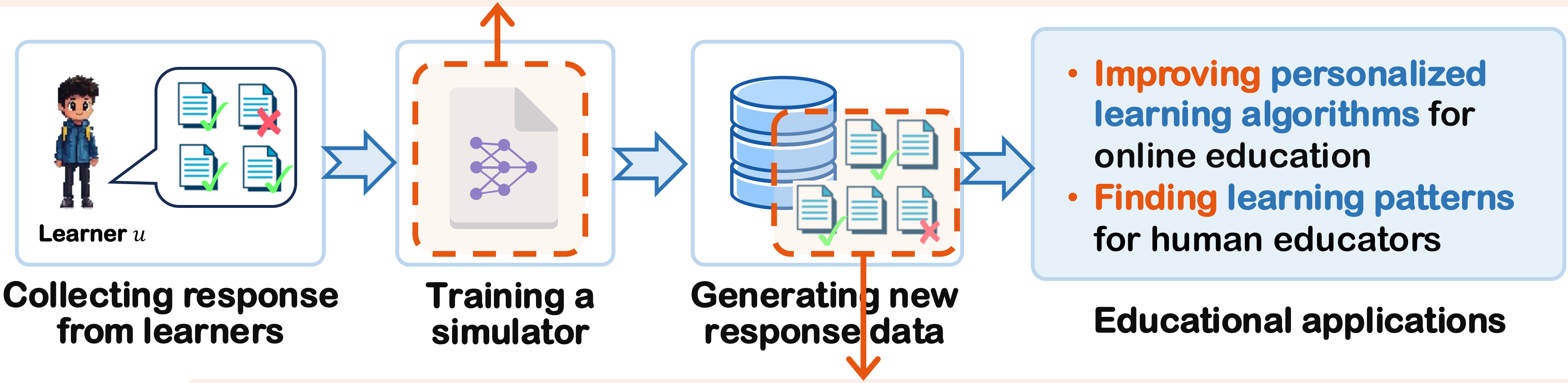


⚠ **Limitation 1:** Existing methods **lack explainability and reliability**, as they **only generate response results** (i.e., the answer correctness), not the problem-solving process

Limitations

- **Simulating learners' response data** is a promising approach
  - Faithfully **captures** human learners' **response patterns**
  - Seamlessly **interacts** with personalized learning algorithms

⚠ **Limitation 2:** Existing models **rely on real responses**, **limiting zero-shot simulation** in cold-start scenarios



⚠ **Limitation 1:** Existing methods **lack explainability and reliability**, as they **only generate response results** (i.e., the answer correctness), not the problem-solving process

Inspiration from Generative Agents

- **Generative (LLM-based) agents** can solve these limitations
  - Rich, **pre-trained knowledge and human-like intelligence** in LLMs can simulate intricate practice behaviors to improve **explainability and reliability**
  - **In-context learning** ability allows LLMs to **perform zero-shot simulations**

Stanford Town

AgentCourt

ChatDev

Fig. Representative LLM-based Agents

Inspiration from Generative Agents

- Generative (LLM-based) agents can solve these limitations
  - Rich, pre-trained knowledge and human-like intelligence in LLMs can simulate intricate practice behaviors to improve explainability and reliability
  - In-context learning ability allows LLMs to perform zero-shot simulations

We propose Agent4Edu, an LLM-based generative agent, to simulate the human practice process and responses reliably, explainably, and interactively

Outline

Task Formulation

- **Given**
  - $|U|$ learners and $|E|$ exercises in an intelligent educational system
  - The response data of learner $u \in U$ are denoted as a time-ordered set:

Exercise $e_i \in E$

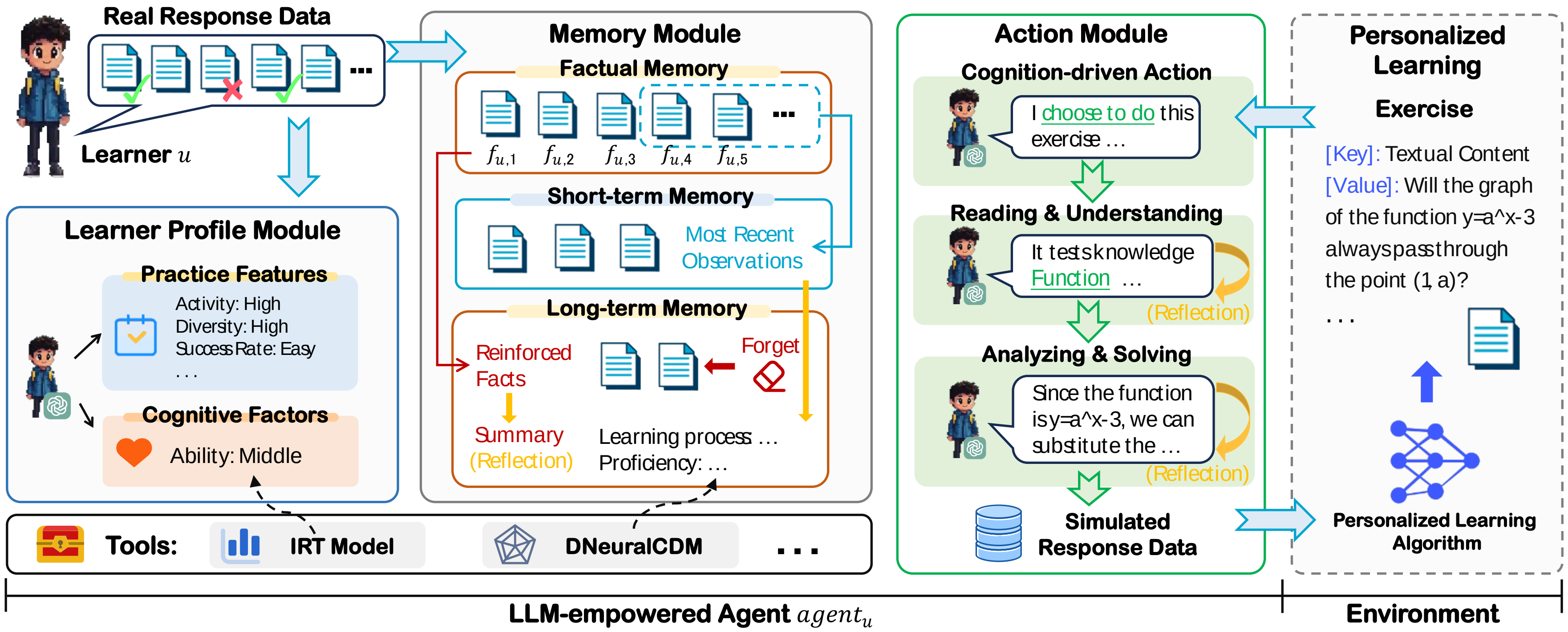
Textual content and knowledge concepts of exercise e_i

$$l_u = \{(\boxed{e_1}, c_{e_1}, y_{u,e_1}), (e_2, c_{e_2}, \boxed{y_{u,e_2}}), \ldots, (e_n, \boxed{c_{e_n}}, y_{u,e_n})\}$$

Learner u's response to exercise e_i,
i.e., if learner u answers e_i correctly, $y_i = 1$ otherwise $y_i = 0$

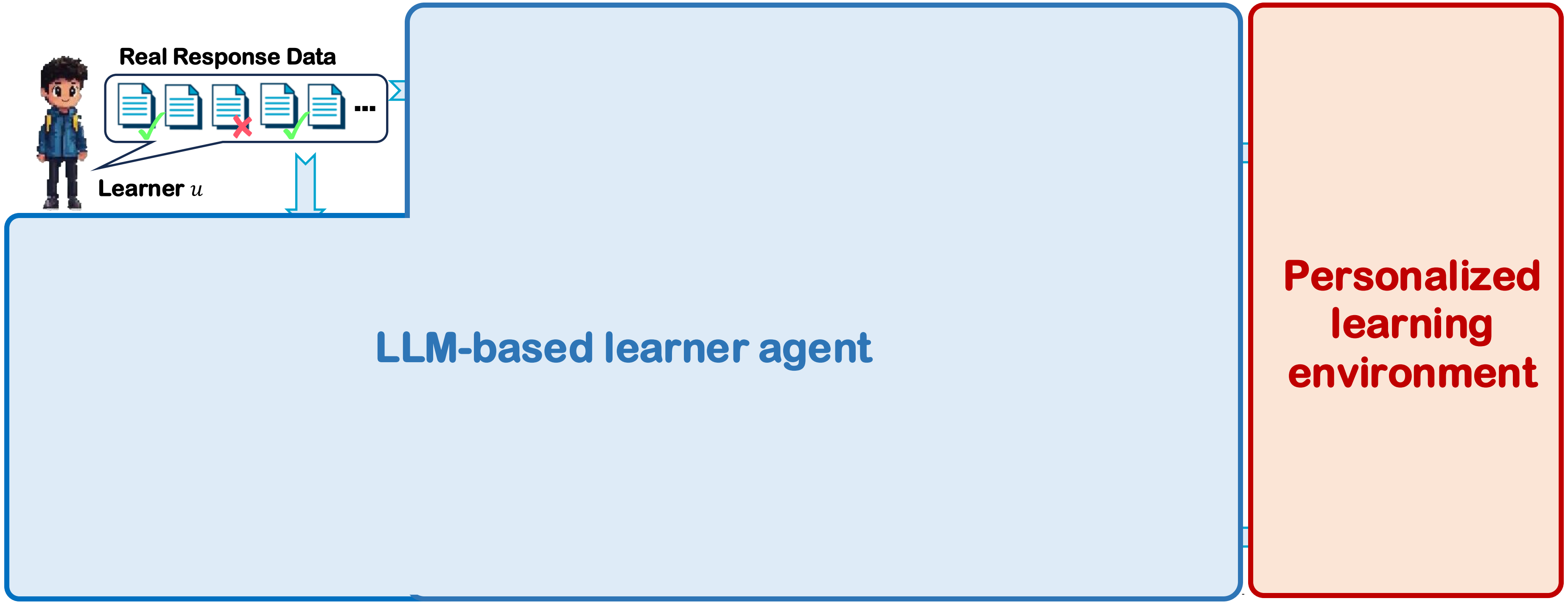
- **Goal**
  - **Accurately generating human learners' future response data** on unseen exercises by distilling their **learning patterns and cognitive preferences from historical response data**

Agent4Edu for Learning Simulation

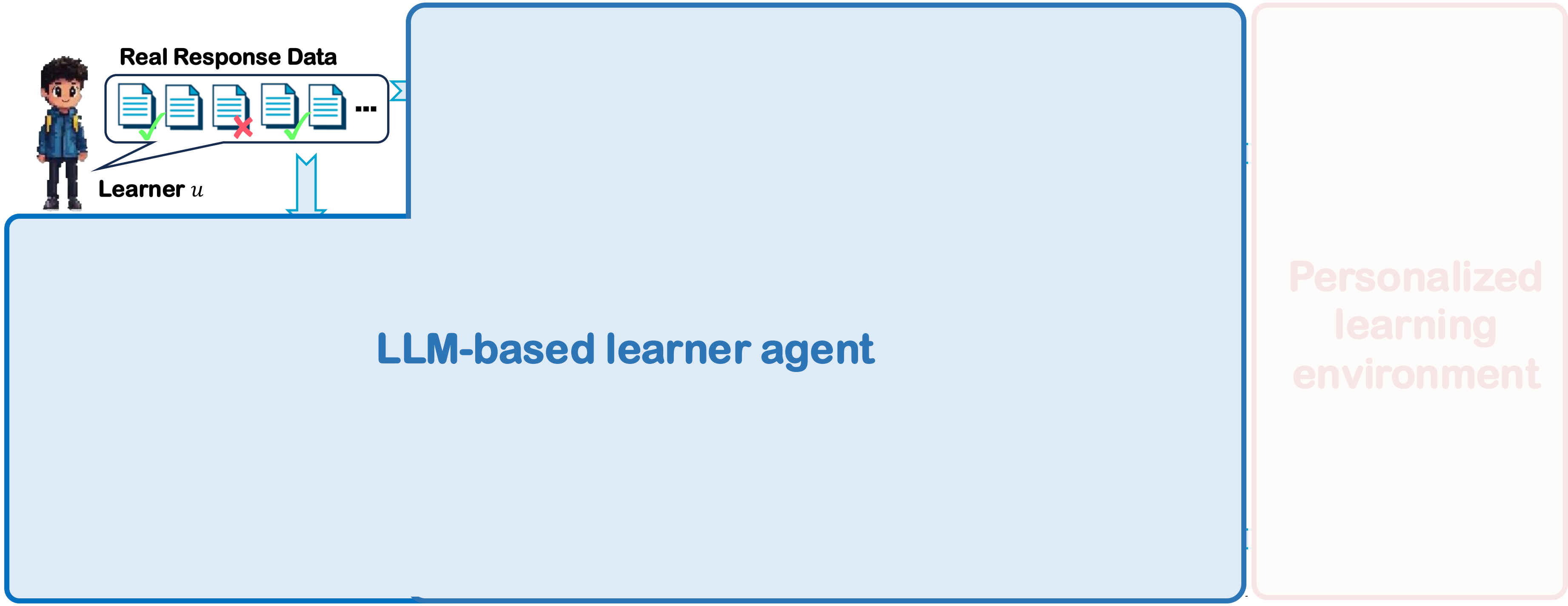


We propose Agent4Edu, an LLM-based agent to generate human learner response data by personalized learning simulation

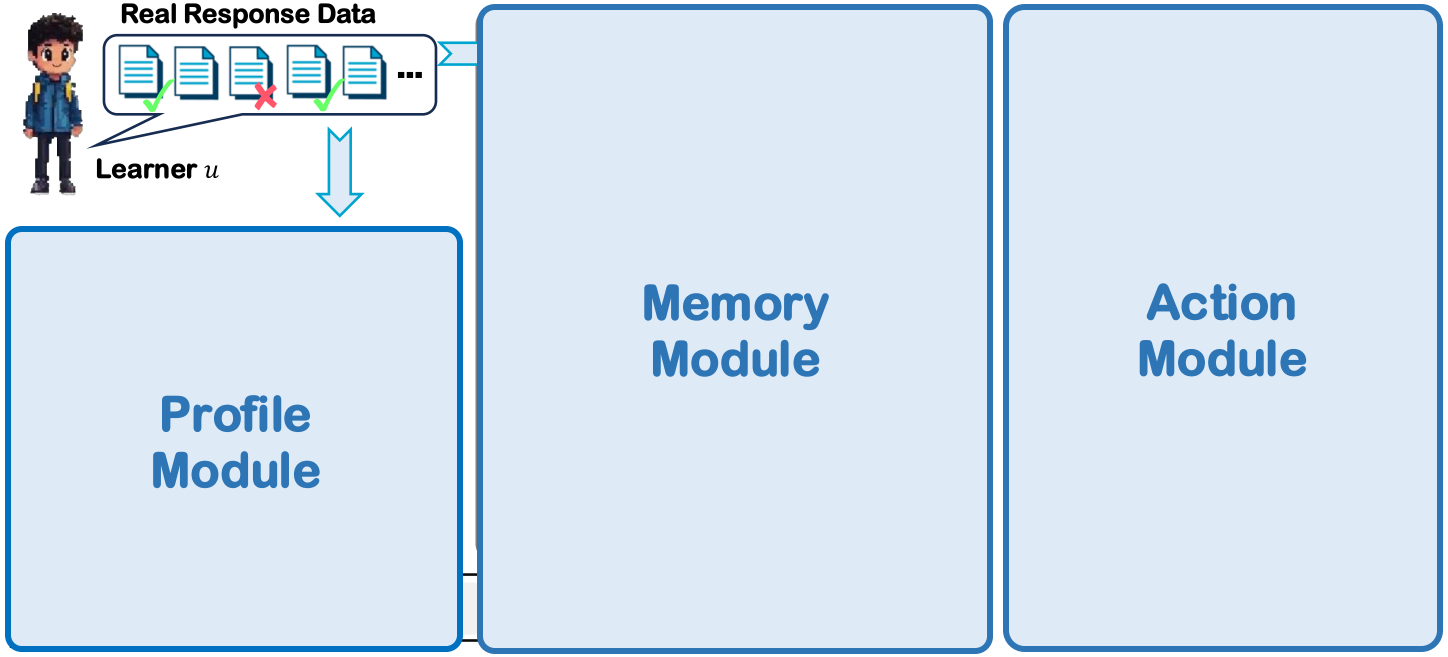
Agent4Edu for Learning Simulation



LLM-based Learner Agent

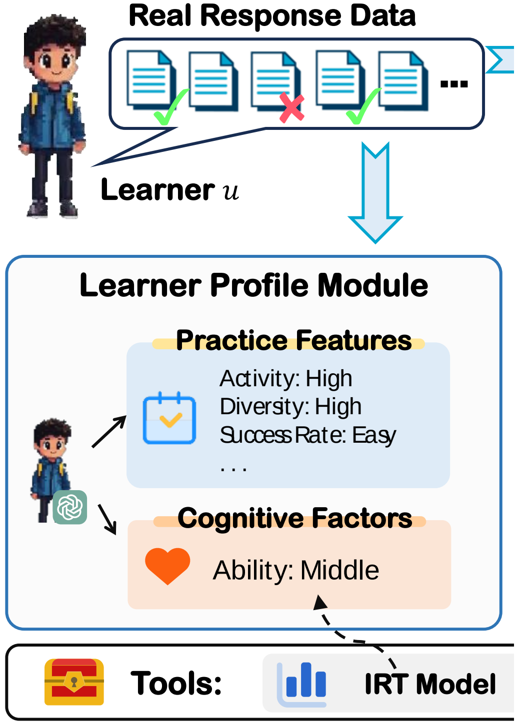


LLM-based Learner Agent



The **LLM-based learner agent** initialized with real human response data or manually pre-configured setup, aimed at **capturing learner learning patterns and cognitive preferences**. **Each real learner** is equipped with **an independent agent**.

LLM-based Learner Agent: Profile



Profile Module

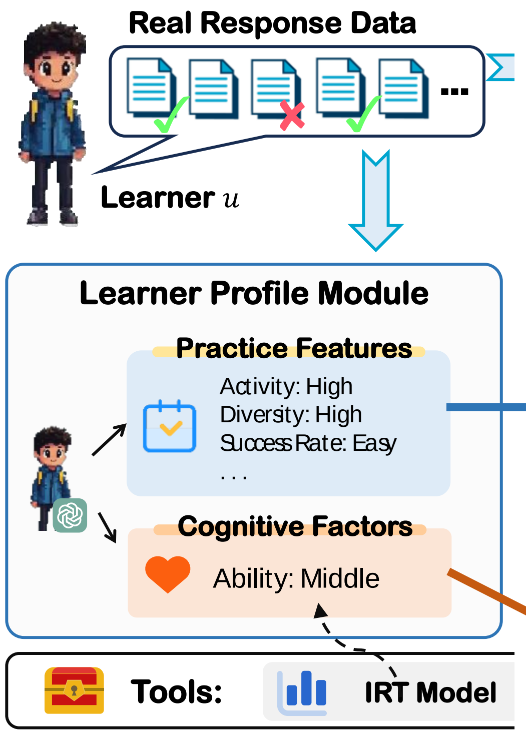
Practice Features

- Activity: Proportion of Practiced Exercises
- Diversity: Proportion of Practiced Knowledge
- Success rate: Response Accuracy
- Preference: Most frequently practiced knowledge

Cognitive Factors

- Ability: Problem-solving ability obtained by using human ability measure **tool** (a pretrained **IRT model**)

LLM-based Learner Agent: Profile



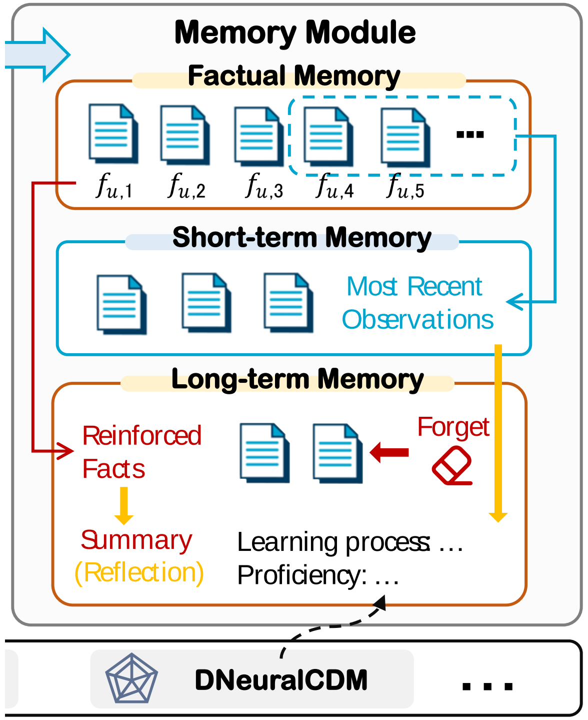
Profile Example

You are a high school student engaging in self-directed exercising on a online learning platform.

The exercises on this learning platform involve Math and Physics. During your online studies, you exhibit a high level of enthusiasm, which means you maintain a high level of online exercise activity and you tend to practice frequently. You have a low curiosity in the learning, indicating that you tend to choose problems that explore limited categories of knowledge. The knowledge concept you practice most often is: leng-chi's law. Your success rate is high.

Additionally, you possess common analytical and problem-solving ability.

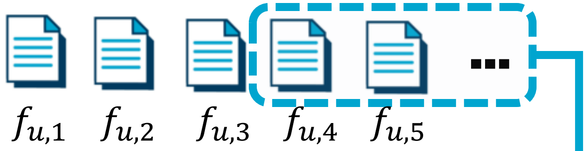
LLM-based Learner Agent: Memory



Memory Module

Factual Memory

- All the response logs of the learner



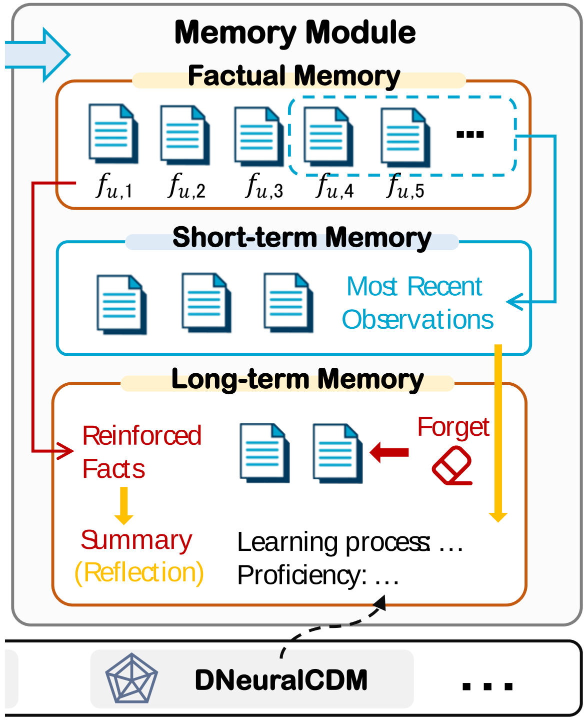
Short-term Memory

- Recent response logs

Long-term Memory

- High-level learning summary and ideas

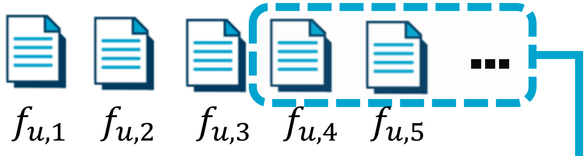
LLM-based Learner Agent: Memory



Memory Module

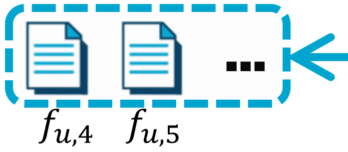
Factual Memory

- All the response logs of the learner



Short-term Memory

- Recent response logs

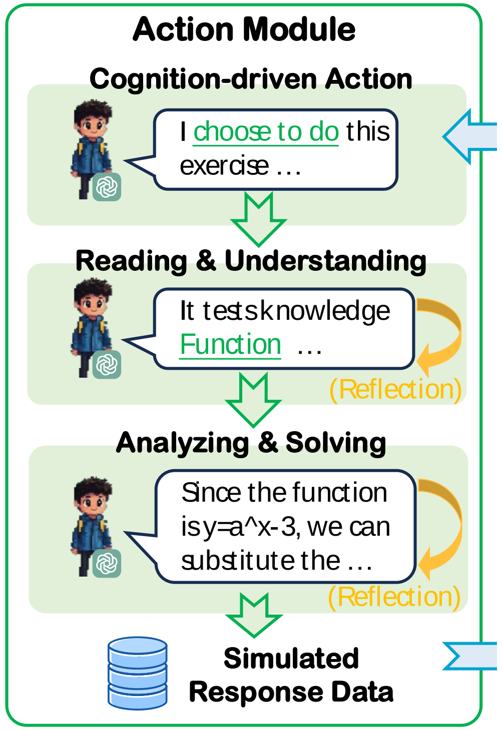


Long-term Memory

- High-level learning summary and ideas

- **Reinforced Facts: Repeated** identical or similar exercise responses
- **Learning Process Summary: Summarization** of the learner's learning process by the LLM (Reflection)
- **Knowledge Proficiency:** Dynamic inference of the learner's **knowledge mastery evolution** after each exercise, based on a human knowledge mastery measurement **tool** (a pre-trained **DNeuralCDM**)
- **Forgetting:** Knowledge not practiced for a long time will be **forgotten**

LLM-based Learner Agent: Action



Action Module

Given an exercise…

Will the graph of the function y=a^x-3 always pass through the point (1, a)?

…

Cognition-driven Action

- The agent reads the exercise's content and decides whether or not to practice it, based on current cognitive factors

> I choose to do this exercise ...

Reading and Understanding Exercise

- Simulating the process of reading and understanding exercises, similar to how humans approach them

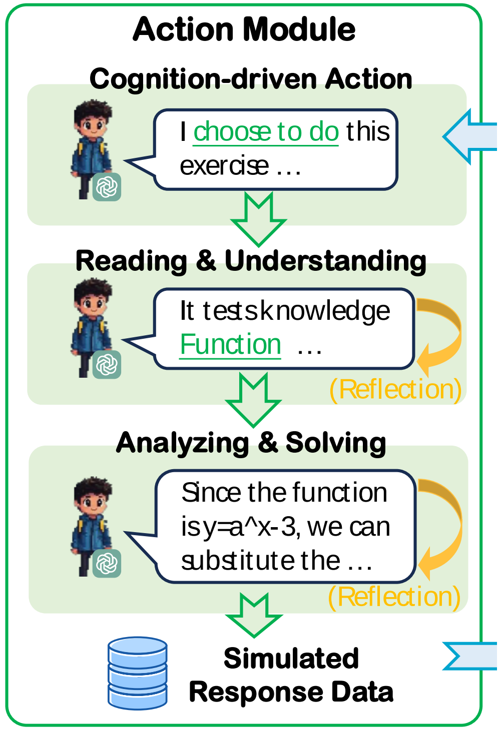
> It tests knowledge *Function* ...

Analyzing and Solving Exercise

- Writing the problem-solving idea and then output the answer and response

> Since the function is y=a^x-3, I can substitute the ...

LLM-based Learner Agent: Action



Action Module

Given an exercise…

Will the graph of the function y=a^x-3 always pass through the point (1, a)?

. . .

Cognition-driven Action

- The agent reads the exercise's content and decides whether or not to practice it, based on current cognitive factors

I choose to do this exercise ...

Reading and Understanding Exercise

- Simulating the process of reading and understanding exercises, similar to how humans approach them

It tests knowledge *Function* ...

(Reflection)

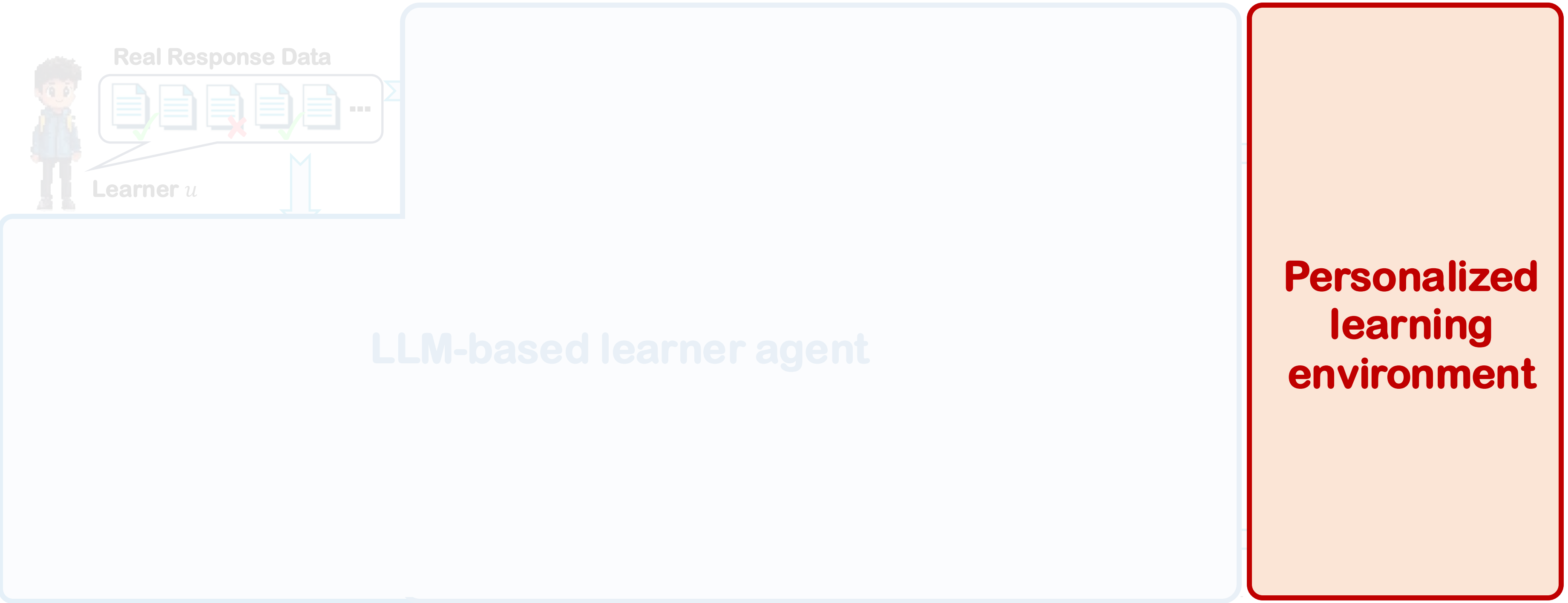
Analyzing and Solving Exercise

- Writing the problem-solving idea and then output the answer and response

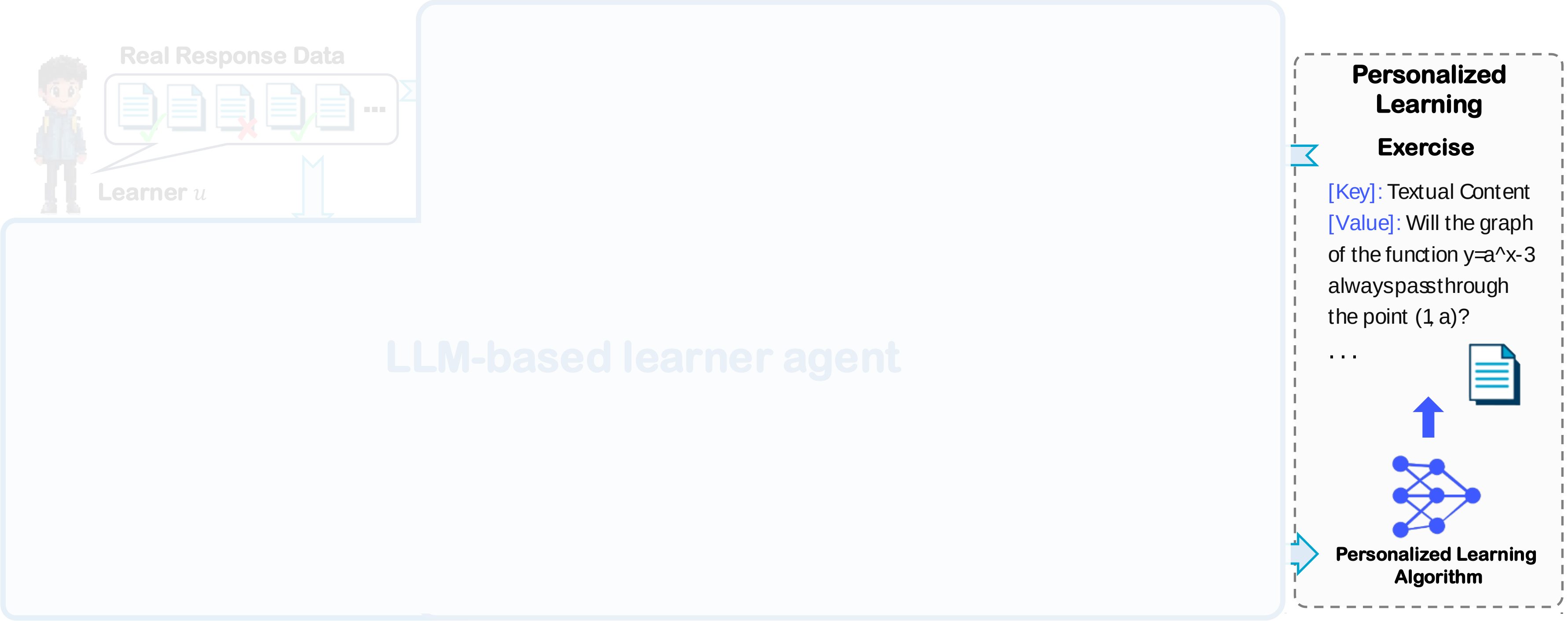
Since the function is y=a^x-3, I can substitute the ...

(Reflection)

Personalized Learning Environment



Personalized Learning Environment



The **personalized learning environment** to provide exercises for **LLM-based learner agents**' practice **interactively**, incorporating a series of **personalized learning algorithms**

Outline

Dataset and LLM Setup

EduData

- EduData dataset is provided by iFLYTEK Co., Ltd. It comprises **time-ordered response records** on **mathematics** and **physics**
- Each record includes **the exercise ID, correctness, and timestamp**
- Each exercise testing one knowledge concept

500 Learners

1,032 Exercises

18,045 Responses

LLM Setup

- We use **GPT-3.5-turbo-1106** and **GPT-4-turbo** through **OpenAI**'s API service to construct the **LLM-based agent** for experimentations
- The **temperature** parameter of the LLM is set to **0** to avoid randomness

LLM-based Agent Simulation Evaluation

Performance on learner response prediction

- Baseline models: 5 traditional models with supervised training

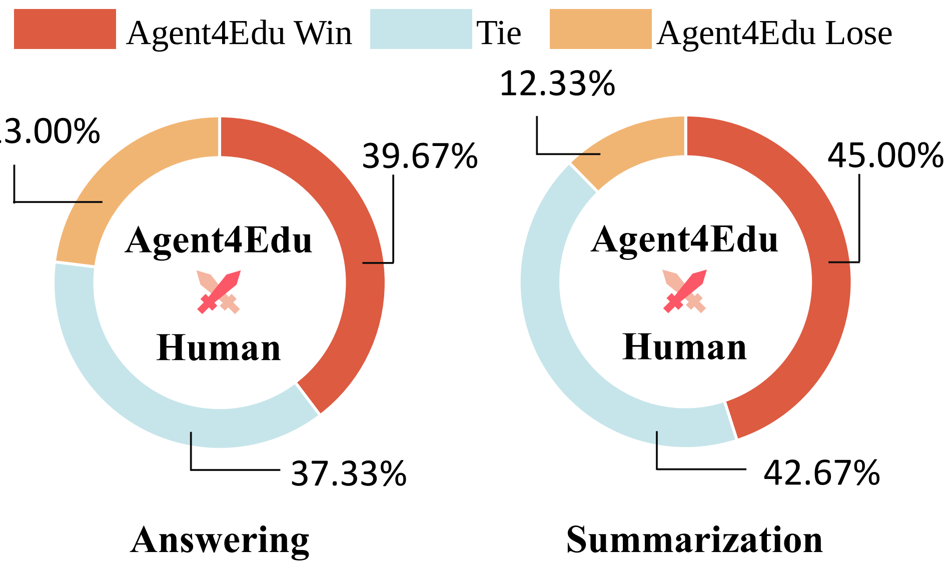
Model	ACC ↑	F1-score ↑	ROUGE-3 ↑
KES	50.11	58.32	25.77
DKVMN	64.39	76.70	37.24
EERNN	<u>65.72</u>	76.06	**43.55**
SAKT	65.52	<u>78.33</u>	31.09
DAISIM	65.63	78.25	31.72
Agent4Edu (GPT-3.5-turbo)	**66.70**	**79.84**	<u>37.97</u>
Agent4Edu (GPT-3.5-turbo)$_{100}$	65.40	78.72	35.14
Agent4Edu (GPT-4)$_{100}$	66.51	79.53	34.86

The proposed Agent4Edu model nearly outperforms all baseline models

LLM-based Agent Simulation Evaluation

Zero-shot simulation evaluation

- Agent4Edu **VS.** Human: Evaluating whether Agent4Edu's simulated problem-solving process mimics human behavior using GPT-3.5-turbo



The Agent4Edu's simulation is closely aligned with real humans

LLM-based Agent Simulation Evaluation

Comparison of simulated and real data distributions

- Visualizing the correct answer rates of 500 learners

The simulated data distribution on the correct answer rate by Agent4Edu is similar to the real data (Ground-Truth)

Personalized Learning Improvement

USTC Stone Statue

AAAI-25

- Exploring whether synthetic response data generated by Agent4Edu can improve the training of personalized learning algorithms
  - Learning Environment: Computerized Adaptive Testing (CAT), which aims to estimate learners' ability or knowledge proficiency with minor exercises, as the experimental environment

	Testing length is 5			Testing length is 10		
Model	F1-score	F1-score+	Imp.	F1-score	F1-score+	Imp.
FSI	80.11	82.39	+2.28	81.10	82.51	+1.41
KLI	79.45	81.84	+2.39	80.63	82.82	+2.19
MAAT	81.77	81.97	+0.20	81.71	81.88	+0.17

Agent4Edu can effectively enhance CAT strategies

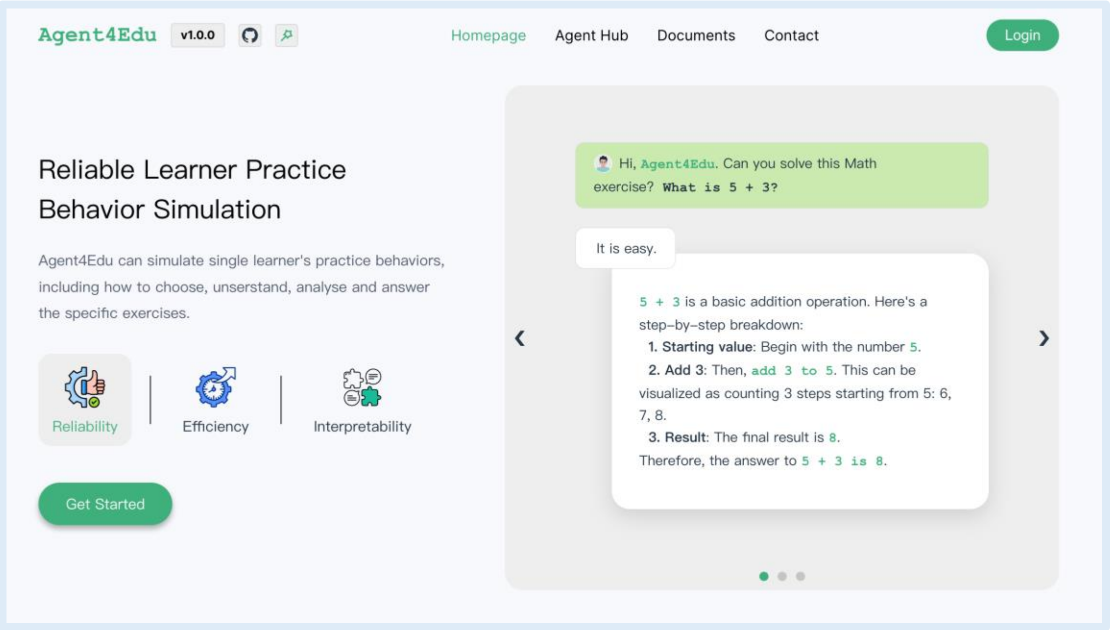
Outline

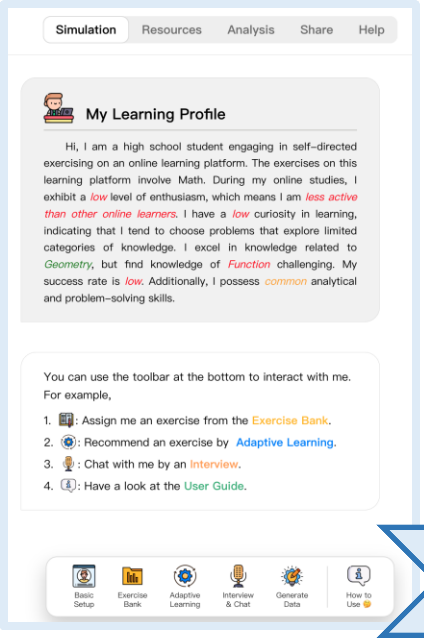
Conclusion

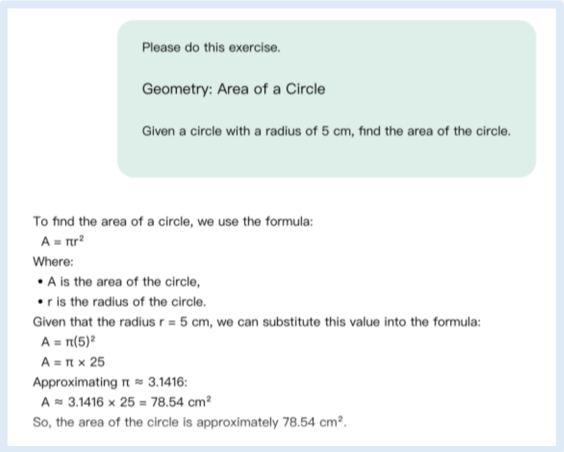
- **Agent4Edu**, an **LLM-based agent** to generate human learner **response data** by personalized learning simulation
  - **LLM-based learner agent**: Capturing learner learning patterns and cognitive preferences, and **generating practice behaviors and responses**
  - **Personalized learning environment**: Providing exercises for LLM-based learner agents' practice **interactively**, incorporating a series of **personalized learning algorithms**
- **Experimental Findings**
  - **Agent4Edu** can **simulate human learning** **reliably and explainably**
  - **Agent4Edu** can perform **effective** **zero-shot response simulations**
  - **Agent4Edu** can **improve** **personalized learning algorithms**

Application

- We have **developed a website** for **interactive learner practice simulation**, integrated with **Agent4Edu**







Project Code of Agent4Edu

The website is currently in **internal testing** and will be launched soon.
Please follow our GitHub project for updates~

AAAI-25
February 25 – March 4, 2025
Philadelphia, USA

University of Science and Technology of China (USTC)

State Key Lab. of Cognitive Intelligence, Hefei, China

Thank You
Q & A

Full Paper of Agent4Edu

Homepage of Weibo Gao (The first author)

Project Code of Agent4Edu

Contact Us: : weibogao@mail.ustc.edu.cn : qiliuql@ustc.edu.cn