AI Agents in Education: Four Trends and a Practical Workflow

Jiali Zhang

Department of Mathematics and Statistics, Missouri University of Science and Technology
Rolla, Missouri, 65401, USA
jztk7@umsystem.edu

Abstract

AI agents based on large language models (LLMs) are transforming education by offering personalized tutoring, faster grading, innovative content generation, and streamlined administrative tasks. This paper explores four major trends of AI agents in the education industry: tutoring and mentoring, automated grading, content creation, and administrative automation. This study also introduces a common AI agent prototype, focusing on customization through fine-tuning and retrieval-augmented generation (RAG), and highlights the role of external tools in enhancing versatility. By reviewing current applications and technical methods, this work demonstrates the growing role of AI in improving instructional quality and operational efficiency.

Current Trends

An AI agent refers to a framework designed to perform specific tasks. In this paper, the core component of the agent, referred to as the "brain," is a large language model (LLM), and the tasks are defined within the context of the education industry. The rise of chatbots such as ChatGPT in late 2022 showcased how LLMs can engage in human-like dialogue, create new content, and analyze data patterns. In schools and universities, these AI agents now serve as virtual tutors, content developers, and administrative aids. With the global AI-in-education market topping $5–6 billion in 2024 (Grand View Research 2024), the adoption of such technologies continues to accelerate.

The integration of AI agents in education is reshaping the industry by offering innovative solutions across multiple dimensions. The four primary areas of impact are outlined below.

AI-Powered Tutoring & Mentoring: AI-driven tutoring systems have emerged as a powerful tool for personalized learning. These systems assist students with homework, self-study, and planning tailored learning paths based on individual performance. For instance, since January 2024, Gauth (Gauth 2024) has become a popular platform in the US. Unlike traditional homework helpers such as Chegg, which rely on user-supplied solutions, Gauth employs a specialized large language model (LLM) trained on STEM subjects to solve complex math, physics, and chemistry problems—even handling professor-crafted questions that deviate from standard textbook examples. Additionally, Khan Academy's Khanmigo (Shetye 2024), built on GPT-4, uses Socratic questioning techniques to guide students, fostering critical thinking instead of merely providing answers.

Automated Grading & Feedback: AI agents are transforming the grading process by providing consistent, rapid, and objective evaluations of student work. Large language models can assess essays for grammar, structure, and clarity, often matching or exceeding human performance in evaluation tasks. For example, in November 2024, Xindongfang (NetEase News 2024) introduced an AI grading tool that reduced the average grading time per essay from 10 minutes to just 2 minutes. Despite these successes in language-based assignments, challenges persist in grading STEM subjects. Handwritten math and physics assignments, with their complex notations and unstructured formats, remain difficult to evaluate—even when digitized using optical character recognition (OCR) technologies.

Content Generation: AI agents are also serving as creative partners for both students and educators. For students, these tools help generate ideas, organize thoughts, and develop coherent outlines for essays and projects. For teachers, AI simplifies content creation by automatically generating slide decks, crafting alternative exam questions, and producing customized practice problems. Platforms like EduAide and Curipod enable educators to input a topic, and receive a curated set of teaching resources and interactive lesson materials, thereby streamlining lesson preparation and allowing teachers to focus on personalized instruction.

Administrative Automation: Educational institutions are increasingly adopting AI to improve operational efficiency and enhance oversight. On the administrative side, AI chatbots integrated with academic systems can handle routine tasks, such as answering common questions about financial aid and registration. For instance, Georgia State University has deployed an AI chatbot to manage student inquiries, thus reducing barriers to enrollment and boosting administrative efficiency (Georgia State University 2022). Furthermore, AI analytics are used to monitor educational quality by tracking student engagement. U.S. ClassDojo's AI module, for example, provides parents with visual charts of classroom

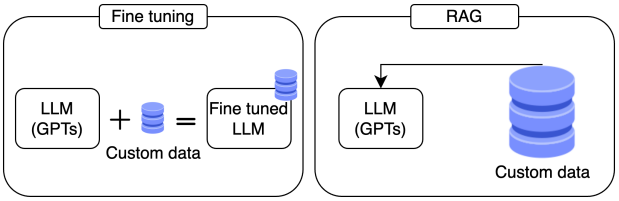

Figure 1: Customizing LLMs: Fine-Tuning vs. Retrieval-Augmented Generation (RAG)

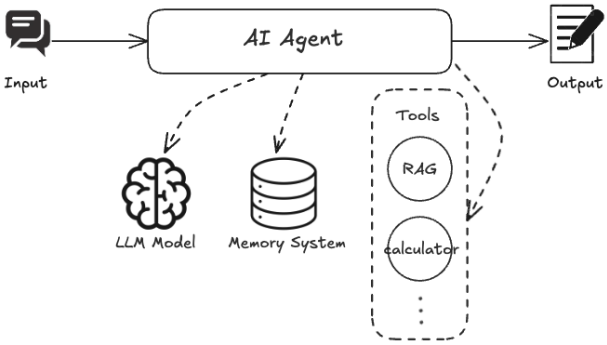

Figure 2: Prototype Architecture of an AI Agent

participation, enabling timely interventions and contributing to improved academic outcomes.

System Architecture Overview

The foundation of AI agents lies in their core language models, which can be customized through fine-tuning or retrieval-augmented generation (RAG) (Lewis et al. 2020), as illustrated in Figure 1. Fine-tuning adjusts a base model's parameters via domain-specific training, as illustrated by Gauth's specialized LLM for STEM problems. By contrast, RAG draws on external knowledge bases without modifying model parameters, enabling real-time adaptability and reducing computational overhead—an approach well-suited for administrative automation.

While fine-tuning offers deep customization, it is resource intensive and operationally complex. RAG, however, is more efficient and avoids the overhead associated with model retraining. In addition, a full-fledged AI agent often integrates external tools—such as OCR, calculators, and code interpreters—to extend its capabilities beyond text generation.

Figure 2 illustrates a prototype architecture for an AI agent. The system operates through the following workflow:

1. **Input:** The agent receives a user request.
2. **Processing:** The AI agent determines the appropriate action, which may include:
   - Interacting with a large language model (e.g., ChatGPT) for language-based reasoning or text generation.
   - Accessing a memory system to recall context from previous interactions.
   - Utilizing external tools (such as RAG, calculators, web search, email interfaces, or APIs) to perform specific functions or retrieve data.
3. **Output:** The agent compiles the final response and delivers it to the user.

This roadmap offers a clear framework for developing AI agents. It emphasizes not only core customization through fine-tuning or RAG, but also the integration of external tools to boost overall performance and versatility.

Conclusion

This paper explores four key ways AI is used in education: personalized tutoring, automated grading, content creation, and administrative tasks. Real-world examples such as Gauth, Khanmigo, and Xindongfang demonstrate how AI can be customized through fine-tuning or retrieval-augmented generation (RAG). These approaches allow AI systems to incorporate external knowledge sources for enhanced performance. A prototype system also shows how tools like OCR (for reading handwritten text) and calculators can make AI even more powerful.

Despite these advancements, the adoption of AI in education introduces several challenges that must be addressed by various stakeholders. Students should be encouraged to maintain their critical thinking skills and avoid over-reliance on AI-generated responses. Educators require ongoing professional development to integrate AI tools effectively and ethically into their teaching practices. Administrators are responsible for ensuring data privacy, managing system integration, and aligning AI adoption with institutional goals. EdTech providers must prioritize the accuracy, fairness, and transparency of their models. Policymakers play a vital role in establishing regulatory frameworks and fostering public trust in AI technologies within the education sector.

References

Gauth. 2024. Gauth - Your Ultimate AI Homework Helper. https://www.gauthmath.com.

Georgia State University. 2022. Classroom Chatbot Improves Student Performance, Study Says. https://news.gsu.edu/2022/03/21/classroom-chatbot-improves-student-performance-study-says/.

Grand View Research. 2024. AI In Education Market Size & Trends. https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report.

Lewis, P.; Perez, E.; Piktus, A.; Petroni, F.; Karpukhin, V.; Goyal, N.; Küttler, H.; Lewis, M.; Yih, W.-t.; Rocktäschel, T.; et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33: 9459–9474.

NetEase News. 2024. Xindongfang's AI Grading Tool Announcement (in Chinese). https://www.163.com/dy/article/JIU3TT6J0512D3VJ.html.

Shetye, S. 2024. An Evaluation of Khanmigo, a Generative AI Tool, as a Computer-Assisted Language Learning App. *Studies in Applied Linguistics and TESOL*, 24(1).