

THIẾT KẾ MÔ HÌNH AI TÁC NHÂN HỖ TRỢ ĐÁNH GIÁ, PHẢN HỒI TRONG DẠY HỌC TOÁN Ở TRƯỜNG PHỔ THÔNG

DESIGNING AN AI AGENT MODEL TO SUPPORT ASSESSMENT AND FEEDBACK
IN MATHEMATICS TEACHING AT SECONDARY SCHOOLS

Nguyễn Thu Phương^{1,+},
Trần Hải Anh²,
Trần Trung³

¹Trường Đại học Sư phạm Hà Nội;

²Trường Đại học Đại Nam;

³Học viện Dân tộc

+Tác giả liên hệ • Email: ngthuphuong2012@gmail.com

Article history

Received: 27/01/2026

Accepted: 18/3/2026

Published: 29/4/2026

Keywords

Multi-agent system, learning data analysis, human in the loop, secondary school mathematics, design-based research

ABSTRACT

Modern mathematics education requires the timely diagnosis of students' systematic errors and misconceptions. This study designs and evaluates the effectiveness of a Multi-Agent System operating within a human-in-the-loop framework, focusing on three core functions: (1) in-depth analysis of error data; (2) automated generation of analogous learning content to reinforce conceptual understanding; and (3) provision of adaptive feedback tailored to individual learning needs. The research employs a Design-Based Research (DBR) approach to develop an integrated framework that combines principles of mathematics didactics with Generative AI technologies. The findings contribute a feasible theoretical framework and a practical tool to support teachers' timely instructional interventions, thereby enhancing personalized learning outcomes and contributing to the digital transformation of mathematics education.

1. Mở đầu

Trí tuệ nhân tạo (Artificial Intelligence - AI) bắt nguồn từ mục tiêu mô phỏng các quá trình nhận thức và ra quyết định của con người thông qua thuật toán và máy tính (Garzón và cộng sự, 2025). Sự bùng nổ của cuộc Cách mạng công nghiệp 4.0 đã đưa AI thành mắt xích quan trọng thay đổi phương thức vận hành của nhiều lĩnh vực, từ y tế, kinh tế đến giáo dục. Đặc biệt, sự xuất hiện của AI tạo sinh (Generative AI) với các mô hình ngôn ngữ lớn (LLMs) như GPT-4 đã đánh dấu một bước ngoặt chưa từng có, cho phép máy móc không chỉ phân tích dữ liệu mà còn tạo ra nội dung, ý tưởng và giải pháp mới với tốc độ vượt trội (Kasneji và cộng sự, 2023). Trong giáo dục, AI không chỉ được sử dụng để tối ưu hóa các tác vụ hành chính mà còn hỗ trợ cá nhân hóa học tập, phân tích dữ liệu học tập và cung cấp phản hồi thích ứng cho người học. Tuy nhiên, phần lớn các hệ thống này vẫn hoạt động theo mô hình đơn tác nhân, trong đó một mô hình AI đảm nhiệm toàn bộ quá trình đọc hiểu, đánh giá và phản hồi. Chúng hoạt động thụ động dựa trên câu lệnh (prompt), thiếu khả năng tự chủ trong quy trình làm việc phức tạp và tiềm ẩn nguy cơ cung cấp thông tin sai lệch ("ảo giác AI") nếu không có cơ chế kiểm soát (Trần Trung và cộng sự, 2025). Để khắc phục điều này, xu hướng công nghệ đang chuyển dịch mạnh mẽ từ các mô hình đơn lẻ sang các hệ thống đa tác nhân (Multi-Agent Systems - MAS) nơi các AI có thể phối hợp, lập kế hoạch và thực thi chuỗi hành động khép kín (Xi và cộng sự, 2023).

Trong việc đánh giá học tập của HS, nhiều nghiên cứu trước đây đã tập trung vào các hệ thống chấm điểm tự động (Automated Essay Scoring - AES) hoặc các mô hình AI hỗ trợ phản hồi học tập (Zhai và cộng sự, 2021). Nhưng hoạt động đánh giá không chỉ dừng lại ở việc xác định đúng - sai mà còn đòi hỏi phân tích lỗi sai, nhận diện các nhận thức sai lầm và cung cấp phản hồi kịp thời để hỗ trợ điều chỉnh tư duy của HS. Theo Hattie và Timperley (2007), phản hồi chỉ phát huy hiệu quả khi được cung cấp đúng thời điểm và gắn với tiến trình học tập cụ thể của người học. Tuy nhiên, trong thực tiễn dạy học, GV thường gặp khó khăn khi phải xử lý số lượng lớn bài làm trong thời gian ngắn, đặc biệt ở các hoạt động kiểm tra thường xuyên hoặc luyện tập trên lớp. Điều này dẫn đến việc phản hồi học tập chưa thực sự kịp thời và khó đảm bảo tính cá nhân hóa cho từng HS (Maksimchuk và cộng sự, 2025). Nhằm khắc phục những hạn chế trên, xu hướng phát triển AI Agent và mô hình "Con người trong vòng lặp" (Human-in-the-loop) đang được xem là một hướng tiếp cận tiềm năng trong giáo dục hiện đại. Khác với các hệ thống AI hoạt động hoàn toàn tự động, mô hình này duy trì vai trò kiểm soát chuyên môn của GV trong quá trình đánh giá và phản

hồi, đồng thời tận dụng năng lực xử lý dữ liệu của AI để giảm tải các công việc mang tính lặp lại (Yingzhe, 2025). Sự kết hợp giữa AI Agent và cơ chế Human-in-the-loop không chỉ góp phần nâng cao hiệu quả xử lý dữ liệu học tập mà còn đảm bảo tính an toàn, độ tin cậy và tính sư phạm của hệ thống hỗ trợ học tập.

Hiện nay những nghiên cứu về AI trong giáo dục hiện nay chủ yếu tập trung vào các môi trường học tập đại học hoặc các hệ thống hội thoại thông minh, chatbot, tích hợp vào hệ thống quản lý (LMS) trong khi các nghiên cứu thực nghiệm về hệ thống đa tác nhân AI (multi-agent AI systems) trong dạy học Toán ở trường phổ thông còn tương đối hạn chế, đặc biệt trong ứng dụng hỗ trợ GV dạy học trực tiếp tại lớp. Trên cơ sở đó, nghiên cứu đặt ra hai mục tiêu chính: (1) Thiết kế khung mô hình hệ thống đa tác nhân AI hỗ trợ đánh giá, phản hồi trong dạy học Toán dưới sự kiểm soát của GV; (2) Tiến hành thử nghiệm bước đầu để đánh giá tính khả thi của mô hình.

2. Kết quả nghiên cứu

2.1. Phương pháp nghiên cứu

Thiết kế nghiên cứu. Áp dụng phương pháp nghiên cứu dựa trên thiết kế (Design-Based Research - DBR). Đây là phương pháp tối ưu để phát triển các giải pháp công nghệ giáo dục thông qua các chu trình lặp lại giữa thiết kế, triển khai, phân tích và cải tiến trong bối cảnh thực tiễn. Nghiên cứu được triển khai qua 3 giai đoạn: (1) Phân tích bối cảnh và thiết kế khung lý thuyết ban đầu cho hệ thống đa tác nhân; (2) Thử nghiệm hệ thống sử dụng mô hình ngôn ngữ cơ bản. Ghi nhận sai sót do hạn chế về khả năng suy luận của mô hình; (3) Cải tiến kỹ thuật bằng cách nâng cấp sang mô hình ngôn ngữ lớn mạnh hơn và tinh chỉnh cấu trúc điều phối agent. Thực nghiệm lại trên cùng tập dữ liệu để kiểm chứng tính ổn định.

Khách thể và bối cảnh nghiên cứu. Thực nghiệm được tiến hành tại một trường THCS ở Hà Nội với đối tượng là 46 HS lớp 7 (chia thành 08 nhóm) trong tiết học “Luyện tập: Số thực”. Nhóm tác giả sử dụng phương pháp chọn mẫu thuận tiện có chủ đích, tập trung vào nhóm HS có kỹ năng công nghệ cơ bản để đảm bảo dữ liệu đầu vào phản ánh đúng tư duy toán học thay vì rào cản thao tác máy tính.

Tiêu chí đo lường và công cụ thu thập dữ liệu. Để đánh giá hệ thống, nghiên cứu thiết lập bộ tiêu chí đo lường cụ thể: (1) Độ chính xác chấm điểm: Tỷ lệ khớp lệnh giữa máy và GV trên tổng số câu hỏi; (2) Chất lượng phản hồi sư phạm: Phân tích dựa trên khung lý thuyết của Hattie và Timperley (2007) về phản hồi hiệu quả; (3) Tính khả thi kỹ thuật: Khả năng vận hành trên hạ tầng máy tính phổ thông (Core i5, 16GB RAM). Dữ liệu được thu thập qua phiếu bài làm file word, quan sát của GV và email báo cáo do AI Agent tạo ra.

2.2. Tác nhân AI và khung lý thuyết ứng dụng trong thiết kế mô hình

“AI Agent” được định nghĩa là một hệ thống có khả năng quan sát môi trường, phân tích dữ liệu và tự đưa ra quyết định nhằm đạt được các mục tiêu nhất định thông qua chu trình nhận thức - ra quyết định - thực hiện hành động. Khác với các phần mềm tự động truyền thống, AI Agent hiện đại dựa trên nền tảng các mô hình ngôn ngữ lớn (LLMs) như Gemini hay GPT, cho phép hiểu và tương tác linh hoạt với ngữ cảnh giáo dục phức tạp. Tuy nhiên, các hệ thống đơn tác nhân thường gặp khó khăn trong việc phân tách nhiệm vụ, dễ dẫn đến hiện tượng ưu tiên đưa ra điểm số thay vì cung cấp phản hồi gợi mở. Do đó, nghiên cứu hiện đại đang chuyển dịch mạnh mẽ sang hệ thống đa tác nhân - nơi các tác nhân chuyên biệt (như tác nhân chấm điểm, tác nhân phản hồi) phối hợp để đảm bảo tính chuyên sâu và cá nhân hóa cao nhất cho người học.

Sự xuất hiện của các LLMs đã thay đổi hoàn toàn cục diện này. Để khắc phục nhược điểm “ảo giác” của Generative AI, khái niệm “AI Agent” hoạt động theo cơ chế “Con người trong vòng lặp” đang trở thành xu hướng nghiên cứu chủ đạo. Thay vì để AI hoạt động hoàn toàn tự động mô hình này đặt GV vào vị trí người kiểm duyệt cuối cùng. Cách tiếp cận này tận dụng được tốc độ xử lý dữ liệu của AI Agent để giảm tải công việc lặp lại, đồng thời giữ được sự chính xác và tính sư phạm thông qua sự giám sát của GV. Đây được xem là mô hình tối ưu để triển khai AI trong trường học giai đoạn hiện nay, đảm bảo sự cân bằng giữa hiệu suất công nghệ và an toàn giáo dục (Yingzhe, 2025).

“Agent framework” được hiểu là một hệ thống hoặc nền tảng kiến trúc dùng để hỗ trợ việc thiết kế, xây dựng, triển khai và vận hành các AI Agent (Li, 2026). Khái niệm này không chỉ đề cập đến một agent đơn lẻ, mà bao hàm tập hợp các công cụ, thư viện và cơ chế điều phối giúp các agent có thể hoạt động một cách có cấu trúc và hiệu quả. Trong bối cảnh các LLM-based agents, agent framework thường được xây dựng xoay quanh các mô hình ngôn ngữ lớn như GPT-4o hoặc các mô hình tương đương, cho phép agent thực hiện các nhiệm vụ tự động thông qua tương tác bằng ngôn ngữ tự nhiên, kết hợp với khả năng gọi công cụ, quản lý bộ nhớ và phối hợp hành động. Hiện nay, nhiều agent framework tiêu biểu đã được đề xuất và phát triển, chẳng hạn như LangChain, AutoGPT, AgentVerse và CrewAI, đóng vai trò là nền tảng quan trọng cho việc nghiên cứu và triển khai các hệ thống AI Agent trong thực tế.

Có thể phân loại AI Agent theo mức độ phức tạp trí tuệ và mô hình kiến trúc quyết định, từ phản xạ đơn giản đến các agent học tập thích ứng, mỗi loại agent có những đặc điểm và ứng dụng riêng biệt trong các hệ thống tự động hóa và tương tác phức tạp. Các loại AI Agent cơ bản gồm: simple reflex agents, model-based reflex agents, goal-based agents, utility-based agents và learning agents (Huang, 2024). Nhìn chung, các loại này thể hiện sự tiến hóa từ phản ứng đơn thuần tới ra quyết định phức tạp trong môi trường động và không xác định. Nghiên cứu này lựa chọn Goal-based agents được xây dựng trên nền tảng tự động hóa mã nguồn mở n8n, kết hợp với mô hình ngôn ngữ lớn. Lựa chọn này không chỉ dựa trên cơ sở lý thuyết, mà còn xuất phát từ khả năng triển khai thực tế và mở rộng trong bối cảnh giáo dục Toán học hướng tới đối tượng là GV THCS có thể tự triển khai. Khung n8n cho phép hiện thực hóa AI Agent dưới dạng một workflow có điều phối, trong đó tác nhân AI đóng vai trò trung tâm điều phối giữa LLMs, bộ nhớ, dữ liệu và các công cụ hỗ trợ giảng dạy.

2.3. Thiết kế mô hình AI Agent hỗ trợ đánh giá, phản hồi trong dạy học

Khung AI Agent được đề xuất trong nghiên cứu này tập trung hỗ trợ ở ba chức năng chính: chấm điểm bài làm của HS, đưa ra nhận xét và phản hồi học tập trực tiếp cho GV và phụ huynh, đề xuất bài tập phù hợp. Thông qua việc tự động hóa các tác vụ mang tính lặp lại và tốn thời gian, hệ thống góp phần giảm tải công việc cho GV, đồng thời cung cấp phản hồi nhanh chóng và nhất quán cho người học. Nghiên cứu đề xuất một kiến trúc hệ thống dựa trên mô hình “Human-in-the-loop” kết hợp sức mạnh xử lý ngôn ngữ tự nhiên của các LLMs đóng vai trò là “bộ não” với sự kiểm soát sự phạm của GV, được thiết kế dưới dạng hệ thống đa tác nhân (Multi-Agent System), trong đó các tác nhân phối hợp thực hiện quy trình đánh giá khép kín. Quy trình vận hành được chia thành 3 tầng xử lý: (1) Tầng dữ liệu (Input Layer): Tiếp nhận dữ liệu bài làm của HS; (2) Tầng xử lý AI (AI Processing Layer): Nơi 4 Agent chuyên biệt hoạt động song song và tuần tự; (3) Tầng kiểm soát và phân phối (Control & Distribution Layer): Nơi GV thực hiện quyền duyệt trước khi hệ thống gửi phản hồi đến phụ huynh/HS.

Để đảm bảo an toàn thông tin và tính chính xác sự phạm, hệ thống đề xuất GV như một chốt kiểm soát. GV truy cập vào giao diện quản lý để xem lại các email nháp. Tại đây, GV có thể: chỉnh sửa lời nhận xét nếu AI diễn đạt chưa phù hợp; xác nhận gửi hàng loạt hoặc gửi từng cá nhân. Quy trình này giúp giảm thiểu tối đa các lỗi còn lại của AI, đồng thời đảm bảo chuẩn tính sự phạm khi gửi điểm cũng như nhận xét của các em tới phụ huynh. Hệ thống bao gồm 4 Agent được thiết kế với các vai trò và nhiệm vụ chuyên biệt. Để làm rõ phạm vi và vai trò của hệ thống được đề xuất, bảng 1 dưới trình bày các chức năng chính của khung AI Agent hỗ trợ GV trong quá trình dạy học.

Bảng 1. Các chức năng tác nhân AI sử dụng trong mô hình (Nguồn: Tác giả)

Agent	Vai trò	Nhiệm vụ	Đầu ra
Agent 1: Tác nhân Chấm điểm	Giám khảo độc lập	Đối chiếu câu trả lời của HS với bộ đáp án chuẩn. Chấm điểm nhị phân (Đúng/Sai) cho câu hỏi trắc nghiệm và chấm theo từ khóa/logic cho câu hỏi điền khuyết.	Bảng điểm số lượng hóa và danh sách các câu sai cụ thể.
Agent 2: Tác nhân Phản hồi Sự phạm	Người hướng dẫn	Phân tích lỗi sai dựa trên kết quả từ Agent 1; xác định nguyên nhân sai lầm và sinh nhận xét sự phạm bằng ngôn ngữ tự nhiên, mang tính khích lệ và định hướng.	Nhận xét chi tiết, mang tính sự phạm cho từng lỗi sai.
Agent 3: Tác nhân Cá nhân hóa Luyện tập	Cố vấn học tập	Dựa trên dạng bài HS làm sai, truy xuất cơ sở dữ liệu để đề xuất 3-5 bài tập tương tự nhằm khắc phục lỗ hổng kiến thức.	Danh sách bài tập bổ trợ được cá nhân hóa.
Agent 4: Tác nhân gửi thông tin	Trợ lý hành chính	Tổng hợp điểm số, nhận xét và bài tập bổ trợ để soạn thảo email gửi phụ huynh; lưu email ở trạng thái chờ duyệt.	Email hoàn chỉnh gửi phụ huynh (có xác nhận của GV).

Dựa trên khung lý thuyết đề xuất trình bày bên trên, khung cấu trúc các hệ thống AI Agent được triển khai như một thực thể trung tâm, kết nối các thành phần sau (xem hình 1):

- User: là tác nhân khởi tạo tương tác với hệ thống, thông qua việc gửi các yêu cầu sự phạm như giao bài tập, yêu cầu đánh giá kết quả học tập, hoặc đề xuất nội dung ôn luyện cho HS.

- LLM: là thành phần suy luận cốt lõi của AI Agent. Mô hình này đảm nhiệm việc phân tích yêu cầu của GV, suy luận logic toán học, giải thích lời giải theo từng bước, cũng như sinh nội dung giảng dạy và phản hồi bằng ngôn ngữ tự nhiên. Trong kiến trúc agent, LLM không hoạt động độc lập mà được điều khiển bởi agent thông qua prompt, bộ nhớ và dữ liệu hỗ trợ.

- Prompt: đóng vai trò như một cơ chế điều khiển hành vi của agent, mã hóa tri thức sự phạm, mục tiêu giảng dạy môn Toán và các ràng buộc về phong cách trình bày. Prompt giúp định hướng cách suy luận của LLM, đảm bảo rằng các phản hồi không chỉ đúng về mặt toán học mà còn phù hợp với trình độ người học và mục tiêu giáo dục đặt ra.

- Bộ nhớ (Memory): là thành phần cho phép agent duy trì ngữ cảnh dài hạn và ngắn hạn trong quá trình tương tác. Bộ nhớ lưu trữ lịch sử trao đổi, kết quả làm bài của HS và các phản hồi trước đó, từ đó hỗ trợ cá nhân hóa quá trình giảng dạy. Trong ngữ cảnh LLM agents, bộ nhớ đóng vai trò biểu diễn trạng thái giúp agent tránh lặp lại thông tin và điều chỉnh chiến lược giảng dạy theo tiến trình học tập của từng HS.

- Nguồn dữ liệu (Data): bao gồm các tài nguyên tri thức được lưu trên Google Drive nhằm phục vụ giảng dạy Toán, như ngân hàng đề toán, tài liệu học tập, đáp án mẫu và các biểu mẫu đánh giá.

- Máy chủ SMTP: đại diện cho các công cụ bên ngoài mà agent có thể gọi để thực hiện hành động. SMTP cho phép AI Agent tự động gửi hàng loạt email tới HS hoặc GV.

- Output: là kết quả cuối cùng do AI Agent tạo ra và phân phối thông qua email được cấp phát. Nội dung đầu ra bao gồm đánh giá kết quả làm bài của HS, nhận xét chi tiết về ưu điểm và hạn chế, cũng như đề xuất các bài tập tự luyện phù hợp.

2.4. Thực nghiệm mô hình tác nhân AI

2.4.1. Mô tả thực nghiệm

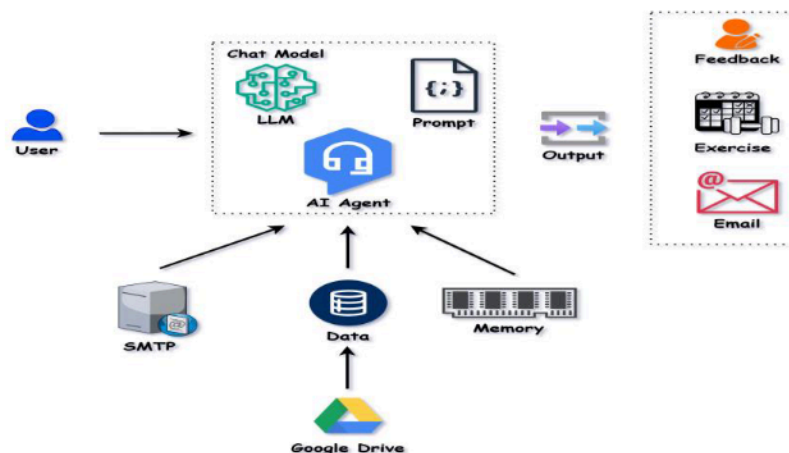
Thực nghiệm được thực hiện tại một lớp học ở khu vực trung tâm Hà Nội, mẫu thực nghiệm bao gồm 46 HS, đại diện cho nhóm đối tượng HS THCS tại khu vực đô thị với khả năng tiếp cận công nghệ ở mức độ khá. Thực nghiệm được tích hợp vào tiết học: “Luyện tập chung: Số thực” (Toán 7). Đây là nội dung kiến thức nền tảng lớn, đồng thời cũng là trọng tâm thi trong chương trình học kì 1 theo phân phối chương trình của các em, nên có nhu cầu luyện tập lớn. Để đảm bảo tính khách quan của dữ liệu và tối ưu hóa sự tương tác, tiết học được tổ chức kết hợp phương pháp dạy học theo Góc - Trạm với sự hỗ trợ của thiết bị số:

- Phân nhóm: Lớp học được chia thành 08 nhóm học tập ngẫu nhiên (5-6 HS/nhóm) nhằm đảm bảo sự đa dạng về trình độ trong mỗi nhóm.

- Thiết bị: Mỗi nhóm được trang bị 01 máy tính xách tay kết nối internet để nhập liệu đáp án.

- Kiểm soát dữ liệu: Chúng tôi vẫn giữ nguyên hoạt động truyền thống của phương pháp Góc - Trạm bằng việc tách các câu hỏi riêng ở từng trạm, mỗi câu hỏi ở 1 mảnh giấy được gấp lại để các em HS phải tự đưa ra chiến lược đọc, hiểu, tổng hợp, phân công cho nhóm mình, đồng thời tránh sao chép giữa các bài làm so với làm bài cá nhân thông thường. Bên cạnh đó kết hợp nhập liệu trên máy tính thay vì chữ viết tay ra giấy đảm bảo dữ liệu đầu vào cho AI phản ánh chính xác năng lực tư duy của từng nhóm. Tổng hợp 4 trạm có 19 câu hỏi, trong đó có 8 câu hỏi trắc nghiệm, 11 câu hỏi dạng điền đáp án tương ứng với các nội dung học (trắc nghiệm tổng quát, thực hiện phép tính, tìm x, toán lời văn).

Các thí nghiệm trong nghiên cứu này được thực hiện trên máy tính xách tay sử dụng Microsoft Windows 11 Home (64-bit). Hệ thống được trang bị bộ vi xử lý Intel Core i5-9300H (4 nhân, 8 luồng) và 16 GB RAM, đáp ứng yêu cầu triển khai các dịch vụ AI Agent ở quy mô vừa và nhỏ. Nền tảng Docker được sử dụng để đóng gói và vận



Hình 1. Khung cấu trúc của hệ thống AI Agent đề xuất (Nguồn: Tác giả)

hành hệ thống trong môi trường cục bộ. Cấu hình này phản ánh điều kiện phần cứng hạn chế của các cá nhân, đồng thời đảm bảo tính khả thi và khả năng tái lập của mô hình đề xuất. Quy trình thực nghiệm diễn ra theo 4 giai đoạn:

- Bước 1: Thực hiện nhiệm vụ học tập HS làm việc theo nhóm để giải quyết phiếu bài tập chủ đề “Số thực” lần lượt ở 4 trạm dưới dạng hoạt động nhóm, tại mỗi trạm sẽ có số câu hỏi tương ứng. Đại diện các nhóm thực hiện điền kết quả vào phiếu trả lời do GV thiết kế sẵn khung, đảm bảo dữ liệu đầu vào được chuẩn hóa để AI Agent xử lý chính xác nhất.

- Bước 2: Xử lý và phân tích bởi AI Agent ngay sau khi các nhóm nộp bài, hệ thống đa tác nhân AI (như đã mô tả ở mục 2.3) tự động kích hoạt: *Agent Chấm điểm*: So sánh kết quả với đáp án chuẩn để tính điểm số; *Agent Nhận xét*: Phân tích các lỗi sai đặc trưng của từng nhóm; *Agent Đề xuất*: Tự động chọn lọc các bài tập bổ trợ tương ứng với điểm yếu của nhóm.

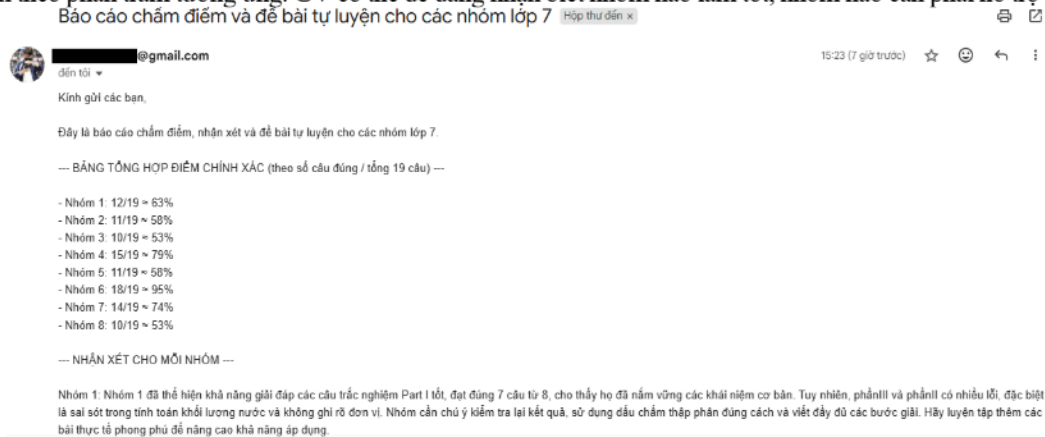
- Bước 3: Kiểm duyệt và phản hồi: GV truy cập trang quản trị để rà soát nhanh các nội dung do AI soạn thảo. Sau khi GV xác nhận “Duyệt”, hệ thống đồng loạt gửi email cá nhân hóa đến từng phụ huynh/HS. GV dùng chính kết quả Agent đã chấm và thấy ngay được các phần nội dung HS dễ sai cần chữa lại trong tiết học.

- Bước 4: Kiểm tra đối chiếu lại kết quả của model Gemini Agent thực hiện lần đầu, ghi lại các lỗi, hạn chế của mô hình, tiếp tục đưa ra phương án nâng cấp. Thực hiện phân tích kết quả HS với mô hình Agent mới (Nemotron Agent).

2.4.2. Phân tích kết quả AI phản hồi

Phân tích từ kết quả của 08 nhóm HS (46 em) thông qua dữ liệu được trả về qua Gmail của hệ thống AI Agent là Gemini Agent (lần phân tích 1) và Nemotron Agent (lần phân tích 2).

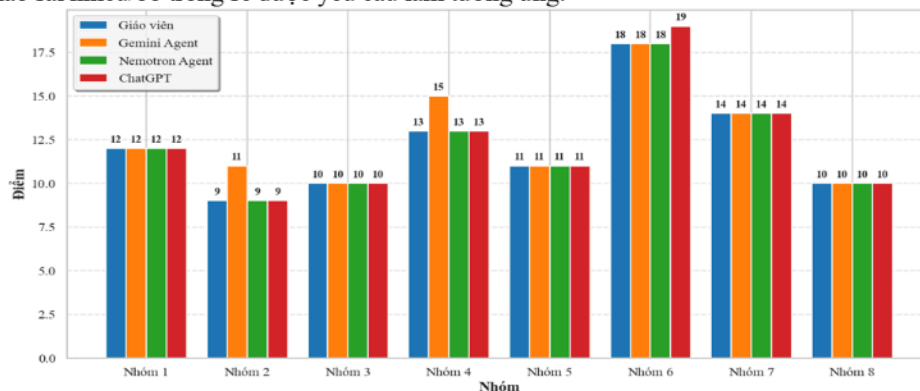
- Điểm số: Máy đã thực hiện chấm điểm tự động ngay sau khi nhận bài. Kết quả điểm (số câu đúng) được thông báo kèm theo phần trăm tương ứng. GV có thể dễ dàng nhận biết nhóm nào làm tốt, nhóm nào cần phải hỗ trợ (hình 2).



Hình 2. Kết quả phản hồi của Gemini Agent được gửi qua email của GV (Nguồn: Tác giả)

- Nhận xét sự phạm: Dựa trên dữ liệu đầu ra cho thấy AI đã cung cấp các phản hồi mang tính xây dựng thay vì chỉ thông báo điểm số. Cấu trúc sự phạm trong nhận xét rõ ràng: khen ngợi phần làm tốt, góp ý chỉ ra lỗi sai ở các câu tương ứng, động viên HS cố gắng (hình 2).

Phân luồng học tập dựa trên kết quả: AI gửi bài (GV chuẩn bị) để luyện tập cho các nhóm dựa trên kết quả chấm điểm, phần nào sai nhiều/bò trống sẽ được yêu cầu làm tương ứng.



Hình 3. So sánh điểm giữa GV và các mô hình AI

Để đánh giá hiệu quả bước đầu của mô hình AI tác nhân đề xuất, nghiên cứu đã thực hiện so sánh kết quả chấm điểm giữa hệ thống với kết quả chấm tay của GV và các LLMs đơn tác nhân phổ biến (hình 3 và bảng 2). Kết quả thực nghiệm cho thấy sự khác biệt đáng kể về độ chính xác: trong khi mô hình tác nhân dựa trên Nemotron-3-Nano đạt mức tương đồng tuyệt đối (100%) với GV trên cả 8 nhóm, thì các mô hình khác lại ghi nhận tỉ lệ sai số nhất định.

Bảng 2. Sai số dự báo điểm của các mô hình

Mô hình	MAE	RMSE
Gemini Agent	0.500	1.000
Nemotron Agent	0.000	0.000
ChatGPT	0.125	0.354

Về mặt kĩ thuật, sự vượt trội của Nemotron-3-Nano (30B-A3B) xuất phát từ kiến trúc lai (hybrid) với 30 tỉ tham số, được NVIDIA tối ưu hóa cho khả năng suy luận logic tập trung khi triển khai dưới dạng AI Agent trên nền tảng n8n. Điều này được minh chứng cụ thể qua bảng 2 với các chỉ số sai số thấp nhất trong các mô hình thử nghiệm với MAE = 0.000 và RMSE = 0.000. Ngược lại, Gemini-2.5-flash-lite dù vận hành trong hệ thống Agent trên n8n lại ghi nhận độ lệch điểm số tại 2/8 nhóm (Nhóm 2 và Nhóm 4) với các chỉ số sai số cao nhất (MAE = 0.500; RMSE = 1.000) do đặc tính của dòng Flash vốn ưu tiên tốc độ hơn là chiều sâu suy luận chuyên sâu. Trong khi đó, ChatGPT-4o dù sở hữu kiến trúc MoE với hơn 1 nghìn tỉ tham số, nhưng do thử nghiệm ở phương thức chấm trực tiếp đơn lẻ trên giao diện web, mô hình này vẫn gặp lỗi tại Nhóm 6, dẫn đến sai số ở mức trung bình (MAE = 0.125; RMSE = 0.354) khi thiếu đi lớp điều phối và kiểm soát.

Bên cạnh sai số về định lượng, phương thức chấm trực tiếp trên ChatGPT-4o hoàn toàn thiếu khả năng tự động hóa quy trình hậu xử lí. Trong khi hệ thống đa tác nhân dựa trên Nemotron và Gemini có thể thực hiện chuỗi hành động khép kín từ chấm điểm đến tự động soạn thảo và gửi phản hồi qua email cho HS, thì ChatGPT-4o chỉ dừng lại ở việc phản hồi văn bản trên giao diện hội thoại, đòi hỏi GV phải thao tác thủ công để chuyển tiếp thông tin. Việc chủ động tích hợp các mô hình mã nguồn mở trên nền tảng n8n không chỉ tối ưu hóa chi phí mà còn cho thấy tiềm năng của việc điều phối trong việc bù đắp hạn chế của các mô hình ngôn ngữ nhỏ, tạo ra quy trình hỗ trợ GV. Tuy nhiên do thiết kế mô hình hướng tới GV có thể tự tạo được mô hình chạy được trên máy tính xách tay, không tốn chi phí, các mô hình AI Agent vẫn còn những lỗi về ngôn ngữ trong nhận xét, GV cần tinh chỉnh trước khi gửi mail, hay chuẩn bị phiếu đáp án đối chiếu để có kết quả chấm chuẩn xác nhất. Đồng thời, kết quả cũng cho thấy vai trò không thể thay thế của GV trong việc giám sát và đảm bảo chất lượng đầu ra cuối cùng.

3. Kết luận

Nghiên cứu đã đề xuất và tiến hành thử nghiệm bước đầu khung mô hình hệ thống đa tác nhân AI hoạt động theo cơ chế “con người trong vòng lặp” nhằm hỗ trợ quá trình đánh giá và phản hồi trong dạy học toán phổ thông. Thay vì thay thế vai trò của GV, hệ thống được định hướng như một trợ lí sư phạm thông minh, hỗ trợ tự động hóa một số tác vụ như chấm điểm, phân tích lỗi sai và đề xuất định hướng hỗ trợ học tập phù hợp cho từng nhóm HS. Kết quả bước đầu cho thấy mô hình có tiềm năng góp phần giảm tải công việc thường xuyên cho GV, đồng thời tăng cường khả năng phản hồi kịp thời và cá nhân hóa trong môi trường học tập số.

Bên cạnh những kết quả ban đầu, nghiên cứu vẫn còn một số giới hạn cần tiếp tục được mở rộng trong các giai đoạn tiếp theo. Thử nghiệm hiện tại chủ yếu được triển khai trên quy mô nhỏ và mang tính thăm dò ban đầu, do đó chưa phản ánh đầy đủ sự đa dạng của bối cảnh dạy học phổ thông. Ngoài ra, việc đánh giá hiệu quả hệ thống hiện mới tập trung vào dữ liệu học tập và phản hồi từ quá trình sử dụng, trong khi các chỉ báo về trải nghiệm người học, mức độ chấp nhận công nghệ và sự thay đổi về năng lực tự học vẫn cần được khảo sát sâu hơn.

Từ đó, các nghiên cứu tiếp theo có thể tập trung theo hướng mở rộng quy mô thực nghiệm và thiết kế các nghiên cứu đối chứng nhằm đánh giá rõ hơn tác động của hệ thống đối với kết quả học tập, chất lượng phản hồi và mức độ tham gia học tập của HS. Đồng thời, việc tích hợp hệ thống đa tác nhân AI vào các nền tảng quản lí học tập (LMS) hiện có sẽ góp phần hình thành môi trường dữ liệu học tập đồng bộ, hỗ trợ các hoạt động và phản hồi thích ứng trong thời gian thực. Bên cạnh khía cạnh kĩ thuật, các nghiên cứu dài hạn về sự thay đổi trong nhận thức, động cơ học tập và mức độ tự chủ của HS khi tương tác thường xuyên với AI Agent cũng là hướng đi cần thiết nhằm làm rõ vai trò của AI như một thành phần hỗ trợ trong hệ sinh thái học tập số tương lai.

Tài liệu tham khảo

Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction*, 9(8), 84. <https://doi.org/10.3390/mti9080084>

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Huang, Y. (2024). Levels of AI agents: From rules to large language models [Preprint]. *arXiv*. <https://arxiv.org/abs/2405.06643>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Li, H. (2026). General framework of AI agents. *Journal of Computer Science and Technology*. <https://doi.org/10.1007/s11390-025-5951-5>
- Maksimchuk, M., Roeber, E., & Store, D. (2025). Generative AI in the K-12 Formative Assessment Process: Enhancing Feedback in the Classroom. In *Proceedings of the Artificial Intelligence in Measurement and Education Conference (AIME-Con): Full Papers* (pp. 107-110).
- Trần Trung, Đỗ Mạnh Hùng, Nguyễn Thị Kim Sơn (2025). Hệ thống AI tác nhân trong phân tích nội dung tự động của tập dữ liệu giáo dục. *Tạp chí Giáo dục*, 25(11), 1-4. <https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/3440>
- Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., ... & Gui, T. (2023). *The rise and potential of large language model based agents: A survey*. arXiv preprint arXiv:2309.07864.
- Yingzhe, L. I. (2025). Addressing “hallucinations” in AI-generated content: Strategies for developing student fact-checking and information evaluation skills. *Artificial Intelligence Education Studies*, 1(2), 48-62. <https://doi.org/10.6914/aiese.010204>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., & Spector, J. M. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021.