NGHIÊN CỨU XÂY DỰNG CHATBOT AI HỖ TRỢ GIẢNG DẠY

Hoàng Thị Mai, Trần Văn Hoàng, Hà Đặng Cao Tùng
Trường Đại học Thủ Đô Hà Nội

Tóm tắt: *Bài báo này trình bày một giải pháp xây dựng chatbot sử dụng trí tuệ nhân tạo (Artificial Intelligence) nhằm hỗ trợ quá trình giảng dạy, với mục tiêu tối ưu hóa công việc cho giáo viên và nâng cao chất lượng học tập. Chatbot này được phát triển dựa trên việc kết hợp các công nghệ tiên tiến như Xử lý ngôn ngữ tự nhiên (Natural Language Processing), Mô hình ngôn ngữ lớn (Large Language Model) và kỹ thuật Tạo tăng cường truy xuất (Retrieval-Augmented Generation). Những công nghệ này giúp chatbot có khả năng hiểu ngữ cảnh, xử lý các câu hỏi chuyên môn phức tạp và cung cấp các phản hồi nhanh chóng, chính xác. Kết quả thử nghiệm cho thấy chatbot đạt hiệu suất cao trong việc cung cấp các câu trả lời đúng và tự nhiên, phản ánh khả năng xử lý tốt nhiều loại câu hỏi khác nhau. Hệ thống đã chứng minh vai trò là một công cụ đáng tin cậy, hỗ trợ giáo viên trong các công việc như soạn thảo bài giảng, tìm kiếm thông tin và giải đáp thắc mắc của học sinh. Các phản hồi của chatbot được đánh giá là mạch lạc, dễ hiểu và phù hợp với ngữ cảnh giáo dục. Bài báo cũng chỉ ra một số hạn chế cần được cải thiện, bao gồm khả năng hiểu các ngữ cảnh phức tạp còn chưa hoàn thiện và phạm vi kiến thức vẫn còn giới hạn. Các hướng nghiên cứu trong tương lai sẽ tập trung vào việc nâng cấp mô hình Xử lý ngôn ngữ tự nhiên, mở rộng cơ sở dữ liệu và tích hợp các tính năng đa phương tiện như nhận diện giọng nói và hình ảnh. Mục tiêu cuối cùng là biến chatbot trở thành một công cụ hỗ trợ giáo dục thông minh và toàn diện hơn, góp phần thúc đẩy việc cá nhân hóa trải nghiệm học tập trong thời đại số.*

Từ khóa: *Chatbot giáo dục, Xử lý Ngôn ngữ Tự nhiên, Mô hình ngôn ngữ lớn, Tạo tăng cường truy xuất.*

Nhận bài ngày 18.04.2025; gửi phản biện, chỉnh sửa, duyệt đăng ngày 30.05.2025
Liên hệ tác giả: Hoàng Thị Mai; email: htmai@daihocthudo.edu.vn

1. ĐẶT VẤN ĐỀ

Hiện nay, trí tuệ nhân tạo (Artificial Intelligence - AI) đã trở thành một phần không thể thiếu trong giáo dục, giúp tối ưu hóa công việc của giáo viên và nâng cao chất lượng học tập. Trong số các ứng dụng AI, chatbot nổi lên như một công cụ hỗ trợ hiệu quả, cung cấp tài liệu học tập nhanh chóng, tự động hóa các công việc như chấm điểm và theo dõi tiến độ học tập, cung cấp phản hồi tức thời và hỗ trợ giáo viên cá nhân hóa bài học để đáp ứng nhu cầu đặc thù của từng học sinh [1] [2]. Ngoài ra, chatbot còn có thể đóng vai trò như một trợ lý ảo, giúp giáo viên tổ chức lớp học một cách hiệu quả mà không bị quá tải, giảm đáng kể khối lượng công việc hành chính như chấm điểm, theo dõi kết quả học tập và tổ chức các bài kiểm tra trực tuyến, quản lý tiến độ học tập của học sinh một cách hiệu quả hơn [3] [4] [5].

Nhiều quốc gia trên thế giới đang đẩy mạnh ứng dụng AI để hỗ trợ giáo viên giảm tải công việc và nâng cao chất lượng giảng dạy [6] [7]. Dù vậy, việc triển khai chatbot trong giáo dục vẫn gặp phải một số thách thức. Khả năng hiểu ngữ cảnh, xử lý các câu hỏi chuyên môn và tương tác với giáo viên trong những tình huống giảng dạy phức tạp là những vấn đề

cần được giải quyết [8]. Điều này đặt ra yêu cầu cần nghiên cứu và phát triển các mô hình chatbot thông minh hơn, phù hợp với nhu cầu thực tiễn của giáo dục hiện đại. Trong bài báo này, nhóm tác giả đề xuất một giải pháp xây dựng Chatbot trên nền tảng các công nghệ tiên tiến như Xử lý Ngôn ngữ Tự nhiên (Natural Language Processing -NLP) với Mô hình ngôn ngữ lớn (Large Language Model -LLM) và Tạo tăng cường truy xuất (Retrieval-Augmented Generation -RAG), giúp cung cấp phản hồi nhanh chóng, chính xác nhằm trợ giúp giáo viên trong việc giảng dạy.

2. NỘI DUNG

2.1 Cơ sở lý luận

2.1.1. Xử lý ngôn ngữ tự nhiên

Học máy (Machine Learning - ML) là "một lĩnh vực nghiên cứu cho phép máy tính có khả năng tự học mà không cần hướng dẫn cụ thể" [9]. Học máy tập trung vào việc cho phép hệ thống "học" tự động từ dữ liệu, cung cấp cho chatbot khả năng trở nên thông minh hơn và phản hồi nhanh hơn. Các thuật toán học máy cho phép chatbot học hỏi từ các tương tác trước đó, xác định các mẫu và đưa ra dự đoán với sự can thiệp tối thiểu của con người [10]. Xử lý ngôn ngữ tự nhiên (Natural Language Processing - NLP) là một lĩnh vực chuyên sâu trong Trí tuệ nhân tạo (AI), đóng vai trò như cầu nối giữa giao tiếp của con người và khả năng hiểu biết của máy móc, kết hợp các nguyên lý ngôn ngữ học tính toán, học máy và các kỹ thuật học sâu. Mục tiêu của NLP là xử lý và phân tích một lượng lớn dữ liệu ngôn ngữ tự nhiên, như văn bản và giọng nói, sao cho có thể mang lại ý nghĩa và giá trị cho máy móc [11]. NLP đóng vai trò quan trọng trong việc phát triển chatbot, giúp hệ thống hiểu và phản hồi ngôn ngữ tự nhiên của con người một cách chính xác. Thông qua NLP, chatbot có thể nhận diện ý định người dùng, trích xuất thông tin từ câu hỏi và tạo ra phản hồi phù hợp với ngữ cảnh [12] [13].

2.1.2. Mô hình ngôn ngữ lớn (LLM)

Việc ứng dụng mô hình ngôn ngữ lớn (LLM) mang lại nhiều lợi ích vượt trội trong phát triển chatbot: LLM có khả năng phân tích và xử lý ngữ cảnh một cách toàn diện, giúp chatbot hiểu chính xác ý nghĩa của câu hỏi hoặc yêu cầu từ người học. Khác với các phương pháp truyền thống chỉ dựa vào từ khóa hoặc cụm từ cố định, LLM sử dụng các cơ chế như Attention để nhận diện mối quan hệ giữa các từ và cụm từ trong câu [14]. Điều này giúp chatbot đưa ra các phản hồi chính xác, ngay cả khi câu hỏi có tính phức tạp hoặc được đặt dưới dạng tự nhiên, không chuẩn hóa.

Một trong những ưu điểm nổi bật của LLM là khả năng xử lý đa dạng các yêu cầu từ người dùng, bất kể lĩnh vực hay độ khó của câu hỏi. LLM còn có thể thích nghi với nhiều phong cách giao tiếp khác nhau của người dùng, như cách đặt câu hỏi trực tiếp, yêu cầu chi tiết hay mong muốn nhận thêm thông tin liên quan. Điều này làm cho chatbot trở thành một công cụ hỗ trợ linh hoạt và hiệu quả, đáp ứng nhu cầu cá nhân hóa của từng người dùng [13].

Khả năng sinh ngôn ngữ tự nhiên là một điểm mạnh khác của LLM. Nhờ được huấn luyện trên một lượng dữ liệu khổng lồ, LLM có thể tạo ra các câu trả lời mạch lạc, tự nhiên và dễ hiểu như cách con người giao tiếp [15]. Điều này giúp người dùng cảm thấy thoải mái và dễ dàng tương tác mà không gặp phải cảm giác gượng gạo khi sử dụng chatbot. Một số mô hình LLM hiện đại còn được thiết kế để học thêm từ các tương tác với người dùng, cho phép chatbot không ngừng cải thiện chất lượng phản hồi và ngày càng đáp ứng tốt hơn nhu cầu của người dùng [16]. Theo thời gian, chatbot có thể điều chỉnh phong cách giao tiếp hoặc tập trung vào các lĩnh vực mà người học gặp khó khăn, từ đó nâng cao hiệu quả giảng dạy. Với việc sử dụng các mô hình LLM đã được huấn luyện trước (pre-trained) như GPT-4,

BERT, và các biến thể chuyên biệt, chatbot có thể tiết kiệm thời gian và tài nguyên, đồng thời mang lại khả năng tổng quát hóa cao.

2.1.3. Tạo tăng cường truy xuất (RAG)

Để tối ưu hóa khả năng trả lời các câu hỏi liên quan đến thông tin doanh nghiệp, một phương pháp phổ biến là tinh chỉnh các mô hình ngôn ngữ lớn (LLMs) sao cho phù hợp với cấu trúc câu hỏi và câu trả lời mong muốn. Cách làm này giúp nâng cao độ chính xác của các phản hồi [17]. Tuy nhiên, đối với các doanh nghiệp vừa và nhỏ, chi phí phần cứng và thời gian để huấn luyện lại mô hình lại là một thử thách lớn. Hơn nữa, khi dữ liệu của doanh nghiệp thay đổi, việc cần phải thực hiện lại quá trình huấn luyện chỉ làm tăng thêm sự phức tạp và chi phí [18].

Để khắc phục vấn đề này, kỹ thuật Retrieval-Augmented Generation (RAG) đã được ứng dụng. Thay vì tinh chỉnh toàn bộ mô hình, RAG cung cấp thông tin ngữ cảnh từ bên ngoài cho các LLMs, giúp giảm thiểu việc huấn luyện lại khi dữ liệu thay đổi [19]. Bằng cách truy xuất thông tin từ nguồn dữ liệu ngoài và sử dụng thông tin này trong quá trình tạo ra câu trả lời, RAG giúp giảm thiểu nhu cầu huấn luyện lại khi dữ liệu thay đổi. Điều này đặc biệt hữu ích trong việc duy trì tính hiệu quả và cập nhật thông tin kịp thời mà không tiêu tốn quá nhiều tài nguyên.

Retrieval-Augmented Generation (RAG) là một kỹ thuật mạnh mẽ kết hợp khả năng sinh ngôn ngữ tự nhiên của các mô hình ngôn ngữ lớn với tri thức từ các nguồn bên ngoài. Khi người dùng đưa ra câu hỏi, hệ thống sử dụng các công cụ truy xuất để tìm kiếm thông tin từ cơ sở dữ liệu và tích hợp vào quá trình tạo câu trả lời của LLM [20]. Quá trình truy xuất này sử dụng các kỹ thuật embedding để biểu diễn và tìm kiếm thông tin một cách hiệu quả, với sự hỗ trợ của các mô hình như Sentence-Transformers. Để tìm kiếm và quản lý các véc-tơ nhúng, hệ thống tính toán sự tương quan ngữ nghĩa giữa câu hỏi và các đoạn văn bản thông qua các phương pháp đo khoảng cách như Cosine, Euclid hoặc tích chập [21]. Ngoài ra, để lưu trữ và truy vấn các véc-tơ nhúng này, cần có các hệ thống quản lý cơ sở dữ liệu đặc biệt, hiện nay đã có nhiều hệ thống cơ sở dữ liệu hỗ trợ kỹ thuật RAG và véc-tơ embedding.

Mô hình Retrieval-Augmented Generation (RAG) là sự kết hợp giữa truy xuất thông tin và sinh ngôn ngữ tự nhiên, giúp tăng cường chất lượng của các câu trả lời bằng cách mở rộng bối cảnh thông tin. Kiến trúc của RAG gồm ba phần chính: Bộ Mã hóa (Encoder), Bộ Truy xuất Thông tin (Retriever) và Bộ Sinh Ngôn ngữ (Generator). Các phần này phối hợp nhịp nhàng để đảm bảo tính chính xác và mạch lạc của thông tin được tạo ra [22].

2.2 Phương thức xây dựng chatbot

Trong nghiên cứu này, ba công nghệ AI chính được áp dụng để phát triển chatbot giáo dục:

Thứ nhất, Xử lý Ngôn ngữ Tự nhiên (NLP) đóng vai trò quan trọng trong việc giúp chatbot hiểu và phản hồi các câu hỏi của người dùng một cách chính xác. Các kỹ thuật NLP như Tokenization, Embeddings được sử dụng để trích xuất thông tin từ câu hỏi của người dùng và tạo phản hồi phù hợp.

Thứ hai, việc ứng dụng và tích hợp Large Language Model (LLM) vào chatbot giúp cải thiện đáng kể khả năng hiểu và tạo phản hồi tự nhiên. Trong hệ thống này, LLMs được sử dụng để tạo phản hồi dựa trên dữ liệu truy xuất từ cơ sở tri thức, thay vì chỉ dựa vào dữ liệu huấn luyện ban đầu. Điều này giúp giảm thiểu tình trạng "ảo tưởng" (hallucination) – khi mô hình tạo ra thông tin không chính xác hoặc không có căn cứ. Ngoài ra, quá trình fine-tuning hoặc instruction tuning có thể được thực hiện để điều chỉnh mô hình theo nhu cầu giáo dục cụ thể, giúp chatbot hiểu sâu hơn về nội dung học tập và phong cách giảng dạy.