|  |  |
| --- | --- |
| **HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**  **---------------------------------------** | |
|  |  |
|  | |
|  | |
|  | |
|  | |
| **Họ và tên tác giả đề án tốt nghiệp** | |
| **Trần Tiến Vượng** | |
|  | |
|  | |
| TÊN ĐỀ TÀI ĐỀ ÁN TỐT NGHIỆP  NGHIÊN CỨU ỨNG DỤNG CHATBOT AI THIẾT KẾ MÔ HÌNH CHAT NHẬP VAI TRONG GIẢNG DẠY | |
|  | |
|  | |
| **Chuyên ngành: Hệ Thống Thông Tin** | |
| **Mã số: B23CHIS031** | |
| **TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ** | |
|  | |
|  | |
|  | |
| HÀ NỘI - NĂM 2025 | |

Đề án tốt nghiệp được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS. Nguyễn Trung Kiên

Phản biện 1: …………………………………………………………………………

Phản biện 2: …………………………………………………………………………

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: ....... giờ ....... ngày ....... tháng ....... .. năm ...............

Có thể tìm hiểu đề án tốt nghiệp tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

# **MỞ ĐẦU**

## **Lý do chọn đề tài**

Trong bối cảnh giáo dục trực tuyến ngày càng phát triển, nhu cầu nâng cao tính tương tác giữa học viên và giảng viên trở nên cấp thiết. Chatbot nhập vai, với khả năng hóa thân thành nhân vật cụ thể trong các tình huống giảng dạy, mang lại trải nghiệm học tập sinh động, trực quan và hiệu quả hơn so với các chatbot truyền thống. Sự phát triển của AI tạo sinh và NLP đã tạo điều kiện thuận lợi cho việc ứng dụng các mô hình ngôn ngữ lớn vào giáo dục. Chatbot nhập vai có thể tùy chỉnh theo ngữ cảnh, học hỏi từ tương tác người dùng và sử dụng nguồn tài liệu chuyên biệt để đưa ra phản hồi chính xác, phù hợp với nội dung đào tạo. Đây là giải pháp tiềm năng giúp cải thiện chất lượng học tập và giao tiếp trong môi trường giáo dục số.

## **Mục đích nghiên cứu**

**Chatbot AI nhập vai (AI RP Bot)** đang thay đổi cách học viên tương tác với công nghệ bằng cách mang lại trải nghiệm học tập cá nhân hóa, hấp dẫn và bám sát chương trình đào tạo. Nhờ được tích hợp các bộ kiến thức chuyên ngành, chatbot có thể nhập vai vào các nhân vật khác nhau để tạo ra các tình huống học tập phong phú, giúp giảm sự nhàm chán.

Các chatbot này sử dụng công nghệ xử lý ngôn ngữ tự nhiên (NLP) kết hợp với mô hình ngôn ngữ lớn (LLM) để tạo phản hồi phù hợp cho từng người dùng. Đặc biệt, chatbot nhập vai được thiết kế theo kịch bản cụ thể với phong cách, tính cách và hành vi riêng, đồng thời có khả năng khai thác tri thức từ nguồn tài liệu chuyên biệt và bên ngoài để tạo ra phản hồi mang tính cá nhân hóa cao.

## **Đối tượng và phạm vi nghiên cứu**

### ***Đối tượng nghiên cứu***

Trong đề án này, mô hình chatbot nhập vai được phát triển nhằm hỗ trợ và cá nhân hóa quá trình học tập, giúp học viên ôn luyện kiến thức một cách có hệ thống, bám sát nội dung bài học và tài liệu giảng dạy.

Để đạt được mục tiêu trên, cần ứng dụng các kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) để nâng cao khả năng hiểu ngữ cảnh và sinh phản hồi tự nhiên. Tuy nhiên, NLP hiện vẫn gặp nhiều thách thức như khó hiểu ngữ nghĩa chính xác, biểu cảm, cấu trúc câu phức tạp, và ngôn ngữ không chính thức.

Bên cạnh đó, việc xây dựng kịch bản nhân vật mang tính cá nhân hóa cao là yếu tố quan trọng để tăng tính ứng dụng của chatbot nhập vai. Ngoài ra, hệ thống **RAG (Retrieval-Augmented Generation)** được xem là giải pháp tối ưu, giúp chatbot kết hợp truy xuất thông tin và sinh câu trả lời phù hợp với ngữ cảnh và nội dung đào tạo mới nhất.

### ***Phạm vi nghiên cứu***

Trong phạm vi đề án, em xây dựng hệ thống chatbot nhập vai hỗ trợ quá trình học tập và giảng dạy, kết hợp giữa công nghệ xử lý ngôn ngữ tự nhiên (NLP) và kỹ thuật truy xuất tri thức nhằm nâng cao hiệu quả giao tiếp và chất lượng phản hồi.

Chatbot được thiết kế theo kịch bản định sẵn, đóng vai nhân vật cụ thể để tương tác với học viên, giải đáp thắc mắc, kiểm tra kiến thức và hỗ trợ học tập. Hệ thống không chỉ dựa vào tri thức huấn luyện sẵn có mà còn tích hợp dữ liệu từ nguồn ngoài, giúp phản hồi chính xác và sát với thực tế giảng dạy.

## **Những luận điểm cơ bản và đóng góp của tác giả**

Mục tiêu của đề án là thiết kế và xây dựng chatbot nhập vai nhằm hỗ trợ hiệu quả cho quá trình học tập. Các đóng góp chính bao gồm:

* Phát triển chatbot có khả năng đối thoại theo ngữ cảnh trong kịch bản nhập vai.
* Ứng dụng mô hình ngôn ngữ lớn để tạo nội dung tự nhiên, phù hợp với nhân vật.
* Khai thác dữ liệu hội thoại để cải thiện khả năng hiểu ngữ cảnh.
* Tích hợp tri thức ngoài để nâng cao độ chính xác, tin cậy và cá nhân hóa phản hồi.

# **CHƯƠNG 1**

# **CƠ SỞ LÝ LUẬN**

## **Mô hình ngôn ngữ lớn**

### ***Mô hình ngôn ngữ lớn là gì ?***

Mô hình ngôn ngữ lớn (LLM) là một thuật toán học sâu tiên tiến dựa trên kiến trúc Transformer, được huấn luyện trên lượng dữ liệu khổng lồ để thực hiện các tác vụ xử lý ngôn ngữ tự nhiên như tạo văn bản, dịch thuật, tóm tắt, phân loại và trả lời câu hỏi. LLM hoạt động như một mạng nơ-ron sâu lấy cảm hứng từ bộ não con người, với hàng tỷ tham số được điều chỉnh trong quá trình huấn luyện nhằm mã hóa kiến thức về ngữ pháp, ngữ nghĩa và mối liên hệ giữa các yếu tố ngôn ngữ. Nhờ đó, LLM có thể ứng dụng hiệu quả trong nhiều lĩnh vực như y tế, tài chính, giáo dục, lập trình và chatbot.

### ***Các thành phần cơ bản của LLM***

Mô hình ngôn ngữ lớn (LLM) là hệ thống mạng nơ-ron sâu với kiến trúc Transformer,

gồm bốn thành phần chính:

* **Embedding layer**: Chuyển từ ngữ thành vector, giúp nắm bắt ngữ nghĩa và cú pháp.
* **Feedforward layer**: Trích xuất đặc trưng trừu tượng, tăng khả năng hiểu ngữ cảnh.
* **Recurrent layer**: Xử lý văn bản theo trình tự thời gian, lưu giữ thông tin từ trước.
* **Attention layer**: Giúp mô hình tập trung vào các từ quan trọng trong câu.

LLM có ba loại mô hình chính:

* **Tổng quát:** Dự đoán từ tiếp theo, dùng cho sinh văn bản đa dụng.
* **Điều chỉnh theo hướng dẫn**: Phản hồi theo yêu cầu cụ thể, dùng cho tạo nội dung, mã, phân tích.
* **Điều chỉnh đối thoại:** Tối ưu hóa cho hội thoại, phù hợp chatbot và trợ lý ảo.

**Quy trình phát triển LLM gồm 3 giai đoạn:**

1. Mô hình ngôn ngữ lớn: Được xây dựng dựa trên kiến trúc Transformer, một kỹ thuật học sâu tiên tiến cho phép xử lý dữ liệu tuần tự.
2. Huấn luyện trên dữ liệu lớn (Pre-Training): Trên dữ liệu lớn, học các mẫu ngôn ngữ cơ bản.
3. **Tinh chỉnh nâng cao**: Dữ liệu chuyên biệt cho các tác vụ cụ thể.

### ***Lịch sử phát triển và sự ra đời của LLM – NLP và Neural Networks***

Việc phát triển mô hình ngôn ngữ lớn (LLM) là kết quả của quá trình tiến hóa lâu dài trong lĩnh vực xử lý ngôn ngữ tự nhiên. Ban đầu, các hệ thống ngôn ngữ dựa trên quy tắc (rule-based) sử dụng câu lệnh if-else để xử lý từ khóa và tạo phản hồi cố định. Bước tiến lớn tiếp theo là sự ra đời của mạng nơ-ron, được giới thiệu từ năm 1943 bởi McCulloch và Pitts, lấy cảm hứng từ hoạt động của nơ-ron sinh học. Mạng nơ-ron trở thành nền tảng cốt lõi cho sự phát triển của các mô hình học sâu hiện đại, đặt nền móng cho sự hình thành của LLM ngày nay.

* Sự phát triển của LLM – Embeddings, LSTM, Attention & Transformers:

Sự phát triển của mô hình ngôn ngữ lớn (LLM) trải qua nhiều giai đoạn với các bước tiến công nghệ quan trọng. Kỹ thuật **embedding** là nền tảng giúp mô hình hiểu mối quan hệ ngữ nghĩa giữa các từ bằng cách

chuyển đổi văn bản thành vector, phục vụ hiệu quả cho truy xuất thông tin.

Sau đó, **LSTM**và**GRU** ra đời nhằm khắc phục hạn chế của RNN trong xử lý chuỗi dài, giúp mô hình ghi nhớ thông tin tốt hơn, cải thiện hiệu quả trong các tác vụ như dịch máy và nhận dạng giọng nói.

* LLM hiện đại – Attention, Transformers:

Việc đưa ra **cơ chế chú ý (attention mechanism)**là bước đột phá trong xử lý ngôn ngữ tự nhiên, giúp mô hình tập trung vào các phần quan trọng của đầu vào. Kiến trúc **Transformer** đã thay thế các mô hình tuần tự như RNN, LSTM nhờ khả năng xử lý

song song hiệu quả.

Sự ra đời của các mô hình như **BERT** (Google, 2018) và **GPT-1 → GPT-3** (OpenAI, 2018–2020) đã mở đường cho các LLM hiện đại, với khả năng hiểu ngữ cảnh sâu và sinh văn bản mượt mà.

Các xu hướng và công nghệ thúc đẩy LLM hiện nay:

1. **Mixture of Experts (MoE):** Sử dụng cơ chế định tuyến động để chỉ kích hoạt một phần mô hình phù hợp với từng đầu vào, giúp **tối ưu chi phí tính toán** mà vẫn duy trì hiệu suất cao.
2. **Retrieval Augmented Generation (RAG):** Kết hợp LLM với khả năng **truy xuất dữ liệu từ cơ sở tri thức ngoài**, giúp phản hồi chính xác hơn, có căn cứ và cập nhật theo thời gian thực.
3. **Meta-learning (Học siêu nhận thức)**: Cho phép mô hình học cách học, linh hoạt thích nghi với nhiệm vụ mới với ít dữ liệu. Bao gồm:

* **Few-shot learning**: Học với vài ví dụ mẫu.
* **Self-supervised learning**: Học từ dữ liệu không gán nhãn.
* **Reinforcement learning**: Học qua tương tác và phản hồi từ môi trường.

## **Tình hình nghiên cứu và ứng dụng mô hình LLM**

### ***Giới thiệu***

### Các mô hình ngôn ngữ lớn (LLM) đóng vai trò quan trọng trong chuyển đổi số, đặc biệt trong xử lý ngôn ngữ tự nhiên và ứng dụng vào chatbot, dịch thuật, tạo nội dung. Nhờ khả năng sinh văn bản mạch lạc và phù hợp ngữ cảnh, LLM được triển khai rộng rãi trong giáo dục, tài chính, y tế. Phần này tập trung khảo sát các mô hình tiêu biểu trong và ngoài nước để làm cơ sở lựa chọn công nghệ phù hợp cho việc xây dựng chatbot nhập vai.

### ***Tình hình nghiên cứu và ứng dụng trên thế giới***

Hiện nay, các hệ thống mô hình ngôn ngữ lớn (LLM) đã đạt được những bước tiến đáng kể, trở thành công cụ nền tảng trong lĩnh vực trí tuệ nhân tạo và được ứng dụng rộng rãi trong nhiều lĩnh vực như xử lý ngôn ngữ tự nhiên, tự động hóa, phân tích dữ liệu, hỗ trợ ra quyết định, giáo dục và tăng cường trải nghiệm người dùng. Nhiều ứng dụng thực tiễn như Google Translate, Notion AI, IBM Watson Health, Replika hay Khanmigo đã cho thấy tiềm năng và sự đa dạng của các LLM trong việc giải quyết các bài toán phức tạp, cải thiện hiệu quả và cá nhân hóa dịch vụ.

Về mặt kỹ thuật, các mô hình LLM tiêu biểu như GPT-3/4 (OpenAI), BERT (Google), LLaMA (Meta), và Grok (xAI) đều sở hữu những ưu điểm riêng về khả năng sinh ngôn ngữ tự nhiên, hiểu ngữ cảnh, hiệu suất xử lý và khả năng mở rộng. GPT nổi bật với năng lực sinh văn bản linh hoạt và đa ngôn ngữ, BERT với khả năng hiểu ngữ cảnh hai chiều và dễ dàng tinh chỉnh cho nhiều tác vụ NLP, trong khi LLaMA là một mô hình mã nguồn mở hiệu quả, tối ưu về tài nguyên và dễ tích hợp. Grok của xAI lại chú trọng vào dữ liệu thời gian thực và tương tác tự nhiên theo phong cách gần gũi, tuy vẫn còn hạn chế về tính học thuật và khả năng đa ngôn ngữ.

Việc khảo sát các LLM tiêu biểu trên thế giới không chỉ giúp đánh giá toàn diện tiềm năng công nghệ, mà còn là cơ sở quan trọng để lựa chọn mô hình phù hợp cho mục tiêu ứng dụng cụ thể, trong đó có thiết kế chatbot nhập vai phục vụ giáo dục cá nhân hóa – trọng tâm của đề tài này.

### ***Tình hình nghiên cứu và ứng dụng trong nước***

Dù còn đối mặt với nhiều thách thức về hạ tầng và tài nguyên ngôn ngữ, nhưng chúng ta đã có những bước tiến đáng ghi nhận trong việc bắt kịp xu hướng phát triển mô hình ngôn ngữ lớn (LLM) và công nghệ AI tạo sinh. Nhiều doanh nghiệp công nghệ trong nước đã chủ động nghiên cứu và xây dựng các mô hình ngôn ngữ phù hợp với tiếng Việt, tiêu biểu như ViGPT của VinBigdata, KiLM của Zalo AI và nền tảng GenAI của FPT Smart Cloud. Các mô hình này đánh dấu giai đoạn đầu trong việc nội địa hóa công nghệ LLM và hướng đến mục tiêu ứng dụng vào các lĩnh vực thực tiễn.

Đặc biệt, chatbot thông minh – một ứng dụng điển hình của LLM – đang được triển khai rộng rãi trong nhiều ngành trong nước. Trong lĩnh vực dịch vụ khách hàng, các doanh nghiệp như Viettel, VNPT, Shopee, FPT Telecom sử dụng chatbot để hỗ trợ người dùng 24/7, xử lý yêu cầu và tối ưu trải nghiệm. Trong ngành tài chính – ngân hàng, chatbot được tích hợp vào các ứng dụng của TPBank và VIB để tư vấn sản phẩm, quản lý tài khoản và giao dịch trực tuyến. Lĩnh vực y tế chứng kiến sự phát triển của các chatbot như Vinmec Chatbot và eDoctor, hỗ trợ tư vấn sức khỏe, đặt lịch khám và cung cấp thông tin khám chữa bệnh. Trong giáo dục, các trường đại học và nền tảng học tập như Zalo (Kiki Chatbot) và Đại học FPT ứng dụng chatbot nhằm hỗ trợ luyện thi, giải bài tập và cung cấp thông tin đào tạo.

Tổng thể, các mô hình và ứng dụng này phản ánh sự năng động và tiềm năng phát triển của hệ sinh thái LLM trong nước, đồng thời tạo tiền đề quan trọng cho việc triển khai các hệ thống chatbot nhập vai có tính cá nhân hóa và phù hợp ngữ cảnh trong giáo dục và các lĩnh vực chuyên biệt khác.

## **Mô hình LLM điển hình – OpenAI và một số kỹ thuật liên quan**

### ***OpenAI (GPT)***

Các mô hình GPT (Generative Pre-trained Transformer) của OpenAI đánh dấu bước tiến lớn trong lĩnh vực xử lý ngôn ngữ tự nhiên (NLP), với khả năng thực hiện đa tác vụ mà không cần huấn luyện giám sát. Dưới đây là quá trình phát triển qua các thế hệ chính:

* **GPT-1 (2018)**: Mô hình đầu tiên sử dụng pre-training trên BookCorpus (7000 sách) với 117 triệu tham số, cho thấy khả năng **zero-shot learning** mà không cần huấn luyện lại cho từng nhiệm vụ.
* **GPT-2 (2019)**: Mở rộng quy mô lên 1.5 tỷ tham số, huấn luyện trên **WebText** (40GB), sử dụng cơ chế **task conditioning**, cho phép thực hiện nhiều tác vụ với cùng một mô hình không giám sát.
* **GPT-3 (2020)**: Nhảy vọt lên **175 tỷ tham số**, được huấn luyện trên tập dữ liệu hỗn hợp (Common Crawl, Books, Wikipedia…), thể hiện vượt trội trong cả zero-shot và few-shot learning, nhưng cũng đặt ra nguy cơ **lạm dụng và thiên kiến xã hội**.
* **GPT-4 (2023)**: Nâng cao độ chính xác, giảm “ảo giác thông tin”, hỗ trợ **32.768 token**, và cải thiện hiệu suất rõ rệt trên các bài kiểm tra tiêu chuẩn đa ngôn ngữ – bao gồm cả tiếng Việt. GPT-4 phù hợp với các ứng dụng giáo dục, pháp lý và khoa học nhờ khả năng hiểu ngữ cảnh và tạo phản hồi chính xác hơn.

### ***Fine tuning***

**Fine-tuning**là quá trình điều chỉnh các mô hình ngôn ngữ lớn (LLM) đã được huấn luyện trước bằng cách tiếp tục huấn luyện trên tập dữ liệu nhỏ hơn và chuyên biệt, nhằm nâng cao hiệu suất cho các tác vụ cụ thể. Quá trình này giúp cá nhân hóa mô hình, tiết kiệm tài nguyên và cải thiện độ chính xác mà không cần huấn luyện lại từ đầu.

#### **Lợi ích của Fine-tuning:**

* Điều chỉnh mô hình cho tác vụ cụ thể.
* Giảm chi phí huấn luyện và dữ liệu.
* Tăng độ chính xác trong ứng dụng thực tế.

#### **Quy trình Fine-tuning gồm các bước:**

1. **Khởi tạo từ mô hình pre-trained.**
2. **Thay đổi lớp đầu ra** cho phù hợp với tác vụ.
3. **Huấn luyện lại** với dữ liệu chuyên biệt.
4. **Tùy chỉnh siêu tham số** để tránh overfitting/underfitting.

### **Các phương pháp Fine-tuning tiêu biểu:**

* **Instruction Fine-tuning**: Tinh chỉnh dựa trên các chỉ dẫn (prompts) nhằm cải thiện khả năng thực hiện tác vụ theo yêu cầu.
* **Full Fine-tuning**: Cập nhật toàn bộ tham số của mô hình, cho hiệu quả cao nhưng đòi hỏi nhiều tài nguyên và dữ liệu.
* **PEFT (Parameter Efficient Fine-tuning)**: Chỉ tinh chỉnh một phần nhỏ tham số quan trọng, giúp giảm thiểu chi phí bộ nhớ và tránh quên kiến thức cũ. Các tham số không quan trọng được "đóng băng", duy trì cấu trúc mô hình ban đầu.

### ***RAG***

RAG (Retrieval-Augmented Generation) là một framework AI kết hợp truy xuất dữ liệu từ cơ sở kiến thức bên ngoài với mô hình ngôn ngữ lớn, giúp cung cấp thông tin bổ sung, cập nhật và giảm ảo giác. RAG khắc phục hạn chế của các mô hình tạo sinh truyền thống bằng cách tích hợp dữ liệu từ cơ sở dữ liệu nội bộ, tệp, hoặc nội dung công khai như bài báo, trang web. Điều này cho phép tạo ra nội dung chính xác, phù hợp ngữ cảnh và linh hoạt hơn cho các nhiệm vụ yêu cầu thông tin cụ thể.

#### Kiến trúc của RAG

Kiến trúc RAG gồm ba phần chính:

1. **Nguồn kiến thức bên ngoài**: Là dữ liệu phi tham số, thường chuyên biệt cho nhiệm vụ và nằm ngoài dữ liệu huấn luyện ban đầu. Dữ liệu này thường được lưu trong cơ sở dữ liệu vector với đa dạng chủ đề và định dạng.
2. **Mẫu lời nhắc**: Cấu trúc chuẩn hóa lời nhắc truyền yêu cầu đến mô hình, bao gồm truy vấn, hướng dẫn và ngữ cảnh. Trong RAG, dữ liệu truy xuất từ nguồn bên ngoài được chèn vào mẫu lời nhắc để cung cấp thông tin theo ngữ cảnh giúp mô hình tạo phản hồi chính xác hơn.
3. **Mô hình ngôn ngữ lớn (LLM)**: Nhận lời nhắc đã được tăng cường thông tin từ nguồn kiến thức bên ngoài, kết hợp kiến thức nội tại của mô hình với dữ liệu mới để tạo ra phản hồi cuối cùng chính xác và phù hợp.

#### Các công cụ triển khai RAG

**LangChain**, **LlamaIndex** và **DSPy** là ba thư viện phổ biến mã nguồn mở hỗ trợ xây dựng và tối ưu hóa các đường dẫn RAG và ứng dụng LLM:

* **LangChain**: Cung cấp các khối xây dựng và tích hợp bên thứ ba để phát triển ứng dụng LLM, hỗ trợ kết hợp với **LangGraph** (xây dựng RAG pipelines) và **LangSmith** (đánh giá hiệu suất RAG).
* **LlamaIndex**: Framework tập trung vào kết nối LLM với dữ liệu ngoài, hỗ trợ bởi **LlamaHub** – kho tài nguyên gồm trình tải dữ liệu, công cụ và bộ dữ liệu phục vụ tạo RAG pipelines.
* **DSPy**: Framework mô-đun cho phép cấu hình cả mô hình ngôn ngữ (LLM) và mô hình truy xuất (RM), hỗ trợ tối ưu hóa toàn diện quy trình RAG.

## **Tổng kết chương 1**

Chương 1 trình bày toàn diện cơ sở lý thuyết về mô hình ngôn ngữ lớn (LLM), bao gồm khái niệm, kiến trúc, cơ chế hoạt động, lịch sử phát triển và các kỹ thuật liên quan. Nội dung cũng phân tích ưu nhược điểm của các mô hình tiêu biểu như GPT, BERT, LLaMA, đồng thời tổng quan tình hình nghiên cứu và ứng dụng LLM trên thế giới và trong nước.

Trong chương 2, đề tài sẽ đi sâu phân tích mô hình chatbot nhập vai trong bối cảnh giáo dục, làm rõ cơ chế hoạt động, vai trò và giá trị mà công nghệ này mang lại, đồng thời tổng hợp các phương pháp tiếp cận và cách thức xây dựng hệ thống phù hợp với đặc thù môi trường học tập hiện đại.

# **CHƯƠNG 2**

# **MÔ HÌNH CHATBOT NHẬP VAI TRONG GIÁO DỤC**

## **2.1. Khái niệm chatbot nhập vai**

### ***Giới thiệu***

Phần này tập trung làm rõ khái niệm chatbot nhập vai và các công nghệ nền tảng hỗ trợ như AI tạo sinh, mô hình khuếch tán, GANs, Transformer, LLM, RLHF và học dựa trên lời nhắc. Nội dung nhấn mạnh vai trò của LLM không chỉ trong việc sinh văn bản, mà còn thúc đẩy ứng dụng sáng tạo như chat nhập vai để tăng tính tương tác, cá nhân hóa và mô phỏng thực tế trong giáo dục.

### ***Chat nhập vai (role-playing chat)***

Chat nhập vai (Role-Playing Chat) là một ứng dụng nổi bật của mô hình ngôn ngữ lớn (LLM), cho phép chatbot đảm nhận vai trò nhân vật cụ thể nhằm tạo ra cuộc đối thoại sinh động, tương tác và giàu ngữ cảnh. Phương pháp này được áp dụng rộng rãi trong giáo dục, huấn luyện kỹ năng, chăm sóc sức khỏe tâm lý và các lĩnh vực khác. Các đặc trưng chính bao gồm:

* Khả năng giả lập vai trò như giáo viên hoặc chuyên gia
* Tương tác linh hoạt theo thời gian thực
* Tùy chỉnh kịch bản theo bối cảnh
* Mang lại trải nghiệm giao tiếp tự nhiên, cá nhân hóa cao.

### ***Chat nhập vai trong giáo dục***

Chat nhập vai (Role-Playing Chat) là một ứng dụng giàu tiềm năng trong giáo dục, nhờ khả năng tạo ra trải nghiệm học tập sinh động, tương tác và cá nhân hóa dựa trên các mô hình ngôn ngữ lớn (LLM). Công nghệ này giúp:

* **Tăng cường tương tác và hứng thú học tập:** Học viên giao tiếp với các nhân vật ảo trong các bối cảnh cụ thể, từ đó nâng cao sự chú ý và khả năng tiếp thu kiến thức.
* **Cá nhân hóa nội dung học**: Bài giảng, tài liệu và nhiệm vụ học tập được điều chỉnh phù hợp với trình độ, mục tiêu và phong cách học của từng cá nhân.
* **Thúc đẩy học tập qua thực hành**: Chatbot nhập vai cung cấp môi trường an toàn để luyện tập kỹ năng và xử lý tình huống thực tế, cho phép học viên lặp lại, thử sai và cải thiện hiệu suất học tập.

### ***Mô hình tạo sinh***

Mô hình tạo sinh (Generative Models) là một thành phần cốt lõi trong trí tuệ nhân tạo hiện đại, đặc biệt trở nên nổi bật với sự phát triển của các mô hình ngôn ngữ lớn (LLM) như GPT, PaLM, Claude và LLaMA. Khác với các mô hình phân loại hay hồi quy vốn chỉ đưa ra dự đoán trong một không gian đầu ra giới hạn, mô hình tạo sinh có khả năng học và mô phỏng lại phân phối xác suất của dữ liệu huấn luyện. Điều này cho phép chúng sinh ra nội dung mới, giàu tính sáng tạo và tương tự về mặt ngữ nghĩa với dữ liệu gốc. Một số công nghệ cốt lõi trong mô hình tạo sinh gồm:

* **LLM (Mô hình ngôn ngữ lớn)**: Dựa trên kiến trúc Transformer, có khả năng sinh văn bản, hội thoại và hỗ trợ nhiều tác vụ xử lý ngôn ngữ tự nhiên.
* **Diffusion Models**: Mô hình thống kê học lại quá trình loại bỏ nhiễu để tái tạo dữ liệu, ứng dụng nhiều trong sinh ảnh.
* **GANs (Generative Adversarial Networks)**: Gồm hai thành phần Generator và Discriminator, nổi bật trong tạo ảnh, video, âm thanh.
* **Transformer**: Kiến trúc mạng nơ-ron sử dụng Self-Attention, nền tảng cho nhiều mô hình sinh hiện đại.
* **RLHF (Reinforcement Learning from Human Feedback)**: Kết hợp học tăng cường với phản hồi con người để cải thiện tính tự nhiên và an toàn.
* **Prompt-based Learning**: Cho phép điều khiển đầu ra của LLM bằng lời nhắc, hỗ trợ zero-shot và few-shot learning.
* **Seq2Seq và VAE (Variational Autoencoder)**: Các mô hình biến đổi chuỗi và biểu diễn tiềm ẩn để sinh nội dung mới.
* **Zero-shot / Few-shot Learning**: Giúp mô hình giải quyết nhiệm vụ mới dù không có hoặc chỉ có rất ít dữ liệu huấn luyện.

### ***Mô hình truy xuất thông tin***

Mô hình truy xuất thông tin là một thành phần quan trọng trong các hệ thống xử lý ngôn ngữ tự nhiên, đóng vai trò then chốt trong việc truy tìm và cung cấp dữ liệu phù hợp từ kho tri thức phi cấu trúc dựa trên truy vấn đầu vào. Trong bối cảnh xây dựng chatbot nhập vai, mô hình này hỗ trợ chatbot khai thác tri thức bên ngoài, từ đó tăng cường chất lượng phản hồi và tính nhất quán trong hội thoại.

Quy trình hoạt động của mô hình truy xuất thông tin được chia thành hai giai đoạn chính: **tiếp thu dữ liệu (ingestion)**và **suy luận (inference).**

* **Tiếp thu dữ liệu**: Giai đoạn này xử lý và chuẩn hóa dữ liệu đầu vào từ nhiều nguồn, chuyển dữ liệu sang định dạng vector thông qua quá trình nhúng (embedding). Các vector này được lưu trữ trong cơ sở dữ liệu vectơ (vector database), cho phép truy xuất ngữ nghĩa nhanh chóng và hiệu quả.
* **Suy luận**: Gồm ba bước chính:
  1. **Truy xuất (retrieval)**: Câu truy vấn được mã hóa thành vector và so sánh với các vector trong cơ sở dữ liệu để tìm ra các điểm dữ liệu liên quan nhất dựa trên độ tương đồng.
  2. **Tăng cường (augmentation)**: Các thông tin truy xuất được tích hợp vào cấu trúc lời nhắc, cung cấp ngữ cảnh và tri thức nền cho mô hình tạo sinh.
  3. **Tạo sinh (generation)**: Mô hình ngôn ngữ kết hợp dữ liệu nội bộ và dữ liệu đã tăng cường để tạo ra phản hồi ngôn ngữ tự nhiên, phù hợp với truy vấn và bối cảnh hội thoại.

### ***AI tạo sinh***

AI tạo sinh (Generative AI) là lĩnh vực AI chuyên tạo ra nội dung mới dựa trên dữ liệu đầu vào, như văn bản, hình ảnh, âm thanh, video và mã lập trình. Khác với AI truyền thống chỉ phân tích hoặc dự đoán, Generative AI có khả năng sáng tạo nội dung ngôn ngữ tự nhiên thông qua các mô hình ngôn ngữ lớn (LLM).

Một xu hướng quan trọng hiện nay là tích hợp Generative AI với hệ thống truy xuất thông tin – còn gọi là **RAG (Retrieval-Augmented Generation)**. Kiến trúc này giúp khắc phục giới hạn về kiến thức tĩnh của mô hình bằng cách truy xuất nội dung thời gian thực từ kho dữ liệu chuyên ngành và kết hợp với đầu vào ngữ cảnh (prompt) để tạo ra phản hồi chính xác và phù hợp hơn.

RAG gồm ba thành phần cốt lõi:

* **Prompt**: Hướng dẫn hành vi sinh ngôn ngữ của mô hình dựa trên bối cảnh học tập.
* **Lịch sử hội thoại**: Giữ ngữ cảnh dài hạn giúp cá nhân hóa và duy trì dòng hội thoại.
* **Kết quả truy xuất**: Cung cấp thông tin nền từ cơ sở dữ liệu làm cơ sở cho phản hồi.

## **Ứng dụng mô hình ngôn ngữ lớn trong giáo dục**

Sự phát triển của mô hình ngôn ngữ lớn (LLM) đang mở ra nhiều cơ hội đổi mới trong lĩnh vực giáo dục thông qua khả năng xử lý ngôn ngữ tự nhiên, sinh văn bản và phản hồi ngữ cảnh. Các ứng dụng tiêu biểu của LLM trong giáo dục hiện đại bao gồm:

* **Cá nhân hóa học tập**: LLM cho phép điều chỉnh nội dung giảng dạy theo trình độ và phong cách học của từng cá nhân, hỗ trợ học tập linh hoạt và chủ động.

**Tại** Phần Lan đã ứng dụng LLM để xây dựng chương trình học tùy chỉnh cho học sinh.

* **Sinh tài liệu học tập tự động:** LLM giúp giáo viên tạo đề kiểm tra, bài tập, giáo trình một cách nhanh chóng và hiệu quả.

Tại Vương quốc Anh, Đại học Cambridge tích hợp AI để sinh tài liệu giảng dạy đa ngôn ngữ.

* **Mô phỏng và nhập vai giáo dục**: Mô hình có thể tái hiện nhân vật lịch sử hoặc tình huống học tập nhằm phát triển tư duy phản biện.

Tại Hoa Kỳ, LLM được sử dụng trong lớp học lịch sử để nhập vai các nhân vật nổi tiếng như Abraham Lincoln.

* **Hỗ trợ giảng dạy trực tuyến và quản lý lớp học**: LLM đóng vai trò trợ giảng ảo, hỗ trợ phản hồi cá nhân hóa và theo dõi tiến trình học tập.

**Tại** Hàn Quốc triển khai mô hình trả lời theo thời gian thực và đề xuất lộ trình học phù hợp cho từng học viên.

* **Phân tích dữ liệu học tập và dự đoán**:LLM có thể phân tích bài làm và lịch sử học tập để phát hiện điểm yếu, dự đoán xu hướng và hỗ trợ cải tiến chương trình.

Tại New Zealand sử dụng công cụ AI để xây dựng chính sách giáo dục dựa trên dữ liệu phân tích.

## **Ứng dụng chatbot nhập vai trong giáo dục**

Trong bối cảnh giáo dục hiện đại, việc tích hợp AI – đặc biệt là chatbot nhập vai – đang mở ra hướng đi mới nhằm khắc phục sự khô khan, thiếu tương tác của phương pháp giảng dạy truyền thống. Chatbot nhập vai là ứng dụng của công nghệ AI tạo sinh, cho phép mô phỏng các cuộc hội thoại hoặc tình huống giáo dục mang tính tương tác và thực hành cao.

Công nghệ này giúp học viên tham gia vào các kịch bản giả định như trò chuyện với giáo viên ảo, nhân vật lịch sử hay nhân vật hư cấu, từ đó phát triển kỹ năng giao tiếp, tư duy phản biện và giải quyết vấn đề trong môi trường an toàn và sinh động. Chatbot không chỉ cung cấp phản hồi thời gian thực mà còn giúp học viên luyện tập kiến thức lớp học một cách linh hoạt, giảm áp lực và tăng hứng thú học tập.

Các vai trò nổi bật của chatbot trong giáo dục bao gồm:

* **Giáo viên/Trợ giảng ảo**: hỗ trợ giải thích kiến thức, cung cấp tài liệu, thiết kế bài tập và kiểm tra.
* **Nhân vật mô phỏng**: tạo tình huống học tập phong phú qua trò chơi hóa và tương tác trực quan.
* **Hỗ trợ cá nhân hóa**: theo dõi tiến trình học, đưa ra phản hồi, gợi ý tài liệu phù hợp với từng học viên.
* **Tự động hóa thông tin học vụ**: phản hồi các câu hỏi phổ biến như lịch học, lịch thi, deadline,…

## **Một số chatbot phổ biến hiện nay trong giáo dục**

Các hệ thống chatbot trong giáo dục hiện nay đang ngày càng phát huy vai trò quan trọng trong hỗ trợ giảng dạy, học tập và quản lý nhờ vào khả năng tương tác linh hoạt, phản hồi thời gian thực và cá nhân hóa dựa trên công nghệ trí tuệ nhân tạo.

**Các nhóm ứng dụng tiêu biểu gồm**:

1. **Hỗ trợ học tập cá nhân hóa**: Chatbot tích hợp LLM giúp phân tích hành vi học tập, nhận diện điểm yếu và gợi ý tài liệu phù hợp.

* ChatGPT: giải thích khái niệm, luyện thi, phát triển tư duy phản biện.
* Khanmigo: trợ lý AI mô phỏng giáo viên, đưa ra phản hồi theo ngữ cảnh cụ thể.

1. **Học ngôn ngữ**: Ứng dụng NLP và nhận diện giọng nói giúp phát triển kỹ năng nghe, nói, đọc, viết.

* ELSA Speak: đánh giá phát âm bằng nhận diện âm thanh.
* Duolingo Max: tích hợp GPT để mở rộng hội thoại và cá nhân hóa học tập.
* Grammarly AI: sửa lỗi ngữ pháp, cải thiện văn phong.
* The Coach: phản hồi tức thì kỹ năng nói cho người học trung cấp trở lên.

1. **Hỗ trợ quản lý giáo dục**: Tự động hóa tư vấn tuyển sinh, hỗ trợ học vụ, trả lời thông tin hành chính.

* Ivy.ai: hỗ trợ tài chính, tuyển sinh, học vụ tại nhiều đại học Mỹ.
* AdmitHub (Mainstay): theo dõi hành vi học viên, đưa cảnh báo sớm và gợi ý cải thiện kết quả học tập.

## **Cách thức xây dựng chatbot nhập vai trong giáo dục**

### ***Xây dựng bối cảnh***

Việc thiết kế cốt truyện và bối cảnh hấp dẫn trong chatbot nhập vai đóng vai trò quan trọng trong việc tăng tính tương tác và hứng thú học tập. Tuy nhiên, các tình huống cần đảm bảo tính thực tiễn, bám sát nội dung học liệu và câu hỏi thực tế từ giảng viên và học viên để tránh gây hiểu sai. Đồng thời, việc xây dựng tình huống cần đa dạng, gần gũi với môi trường học tập nhằm tăng tính ứng dụng và hiệu quả tiếp thu kiến thức.

### ***Thiết kế nhân vật***

Thiết kế nhân vật trong chatbot nhập vai (role-playing chatbot) là bước then chốt để đảm bảo tính chân thực, hiệu quả và hấp dẫn trong tương tác với người học. Nhân vật có thể được xây dựng dựa trên hình mẫu thực tế, nhân vật lịch sử hoặc hư cấu, nhưng cần đảm bảo sự nhất quán và phù hợp với mục đích sử dụng.

Việc thiết kế bao gồm các yếu tố chính như:

* **Hồ sơ nhân vật (profile):** xác định tên, vai trò, độ tuổi, trình độ, sở thích để tạo cảm giác rõ ràng và gần gũi.
* **Câu chuyện nền (backstory):** giúp định hình cá tính, động lực và tạo chiều sâu cho nhân vật.
* **Tính cách và kỹ năng:** lựa chọn các đặc điểm như hài hước, nghiêm túc, thân thiện,… phù hợp với vai trò; kèm theo đó là các năng lực như giải thích kiến thức, phản hồi linh hoạt, tư vấn học tập,...
* **Phong cách giao tiếp (tone & style):** cần xác định rõ mức độ trang trọng, học thuật hay thân mật tùy thuộc vào đối tượng người học; ngôn ngữ đặc trưng (tiếng lóng, học thuật, biểu cảm,...) cũng được cân nhắc để tăng tính thuyết phục và tự nhiên.

Những yếu tố này giúp chatbot tạo ra trải nghiệm nhập vai mạch lạc, tăng sự gắn kết và hỗ trợ hiệu quả cho hoạt động học tập tương tác.

### ***Tích hợp dữ liệu học tập***

Việc tích hợp dữ liệu học tập và học liệu chuyên sâu vào chatbot nhập vai là yếu tố then chốt để nâng cao chất lượng tương tác và đảm bảo độ chính xác học thuật. Thay vì chỉ phản hồi dựa trên khả năng ngôn ngữ tổng quát, chatbot cần được thiết kế để truy xuất các nguồn học liệu chính thống, phù hợp với nội dung đào tạo và mục tiêu giảng dạy.

Sự tích hợp này giúp chatbot phản hồi dựa trên tri thức đáng tin cậy từ giáo trình, tài liệu giảng viên, và các nguồn học thuật được kiểm duyệt. Đồng thời, hệ thống có thể phân tích dữ liệu học tập cá nhân – như lịch sử tương tác, kết quả học tập, thói quen học – nhằm đưa ra các đề xuất học liệu cá nhân hóa, từ đó hỗ trợ lộ trình học tập hiệu quả hơn cho từng học viên.

### ***Cá nhân hóa trải nghiệm học tập***

Cá nhân hóa trải nghiệm học tập là yếu tố cốt lõi trong giáo dục hiện đại, đặc biệt khi ứng dụng chatbot nhập vai. Khác với điều chỉnh nội dung đơn thuần, cá nhân hóa ở đây bao gồm việc thiết kế toàn bộ hành trình học dựa trên đặc điểm, nhu cầu, tiến độ và phong cách học tập của từng cá nhân. Chatbot nhập vai có khả năng nhận diện và phân tích dữ liệu học tập cá nhân như mức độ hiểu, thói quen, phản hồi và khó khăn gặp phải, từ đó đưa ra các phản hồi và gợi ý phù hợp theo thời gian thực.

Ngoài ra, chatbot giúp duy trì sự tương tác liên tục bằng cách theo dõi tiến trình học và tự động điều chỉnh độ khó, nhịp độ giảng dạy để phù hợp với từng người học. Điều này không chỉ nâng cao khả năng tiếp thu và áp dụng kiến thức, mà còn thúc đẩy sự chủ động, linh hoạt trong học tập, góp phần xây dựng một hệ thống giáo dục cá nhân hóa toàn diện, phát triển cả kiến thức, kỹ năng và thái độ học tập của học viên.

### ***Đánh giá và phản hồi hiệu suất học tập***

Đánh giá và phản hồi hiệu suất học tập là yếu tố thiết yếu trong hệ thống giáo dục ứng dụng chatbot nhập vai, không chỉ nhằm đo lường kết quả học tập mà còn theo dõi tiến trình và sự phát triển của người học theo thời gian thực. Chatbot có khả năng ghi nhận dữ liệu như mức độ hoàn thành, thời gian phản hồi, tương tác và hành vi học tập để đưa ra phản hồi cá nhân hóa, linh hoạt và kịp thời.

Khác với các phương pháp đánh giá truyền thống chủ yếu mang tính định lượng, chatbot nhập vai cho phép đánh giá định tính dựa trên tiến trình, kỹ năng và hành vi học tập. Hệ thống có thể nhận diện lỗi, gợi ý sửa sai, tái tạo tình huống để học viên luyện tập, từ đó giúp cải thiện kỹ năng mà không gây áp lực như các bài kiểm tra thông thường.

Ngoài ra, chatbot còn xây dựng các chỉ số hiệu suất động như độ chính xác, tốc độ phản hồi, mức độ chủ động tương tác và khả năng ứng dụng kiến thức vào thực tiễn. Những dữ liệu này hỗ trợ việc điều chỉnh kế hoạch học tập cá nhân, góp phần nâng cao động lực, phát triển kỹ năng tự học và tối ưu hóa quá trình học trong môi trường giáo dục số.

## **Tổng kết chương 2**

Trong chương 2, em đã phân tích mô hình chatbot nhập vai trong giáo dục, nhấn mạnh khả năng mô phỏng tương tác, cá nhân hóa học tập, đánh giá hiệu suất và phản hồi linh hoạt. Chương cũng trình bày các bước xây dựng chatbot từ thiết kế bối cảnh, nhân vật đến tích hợp tri thức và học liệu.

Trong chương 3, em sẽ tập trung vào triển khai kỹ thuật, bao gồm thiết kế nhân vật, xây dựng hội thoại, lựa chọn kiến trúc hệ thống, tích hợp mô hình ngôn ngữ lớn và cá nhân hóa phản hồi dựa trên dữ liệu học tập, đảm bảo tính nhập vai và phù hợp mục tiêu giáo dục.

# **CHƯƠNG 3**

# **XÂY DỰNG HỆ THỐNG THỬ NGHIỆM CHATBOT NHẬP VAI TRONG ĐÀO TẠO**

## **Ý tưởng về mô hình chatbot nhập vai trong đào tạo**

Trong chương này, em sẽ trình bày ý tưởng thiết kế mô hình chatbot nhập vai trong vai trò giáo viên ảo, tập trung vào khả năng tương tác ngôn ngữ mạch lạc, duy trì hành vi và phong cách giao tiếp đặc trưng trong môi trường giáo dục. Mô hình hướng đến bốn mục tiêu chính:

* Phát triển chatbot có khả năng nhập vai ổn định nhờ công nghệ xử lý ngôn ngữ tự nhiên và mô hình ngôn ngữ lớn (LLM)
* Thiết kế kịch bản nhân vật với thông tin đầy đủ về tính cách, ngôn ngữ, bối cảnh và kiến thức nền
* Tích hợp tri thức chuyên ngành nhằm đảm bảo tính chính xác và cập nhật của phản hồi
* Kết hợp kịch bản định hướng với khả năng sinh ngôn ngữ để tạo ra giao tiếp tự nhiên, linh hoạt theo ngữ cảnh.

## **Thiết kế nhân vật**

Trong quá trình phát triển hệ thống chatbot, việc thiết kế nhân vật là yếu tố then chốt trong hệ thống chatbot nhập vai, ảnh hưởng trực tiếp đến tính cách, giọng điệu và nội dung giao tiếp của AI. Một hồ sơ nhân vật rõ ràng và nhất quán giúp đảm bảo sự chân thực, tự nhiên và hấp dẫn trong tương tác, từ đó nâng cao hiệu quả nhập vai. Do đó, quá trình xây dựng nhân vật cần được thực hiện nghiêm túc, bao gồm việc xác định các đặc điểm nhân trắc học, hành vi ngôn ngữ và nền tảng tri thức phù hợp với mục tiêu sử dụng.

Các bước thiết kế được thực hiện để đảm bảo tính chân thực và hiệu quả giao tiếp:

1. Thiết lập hồ sơ nhân vật: Xác định thông tin cơ bản (họ tên, nghề nghiệp,…), phong cách giao tiếp (thân thiện, trang trọng,…), năng lực chuyên môn (giảng dạy Toán, tư vấn,…) và ranh giới hành vi (giới hạn chủ đề, ngôn ngữ,…) nhằm duy trì phản hồi nhất quán.
2. Tích hợp tri thức chuyên ngành: Chuyển đổi tài liệu (giáo trình, bài báo,…) thành vector ngữ nghĩa, lưu trữ trong cơ sở dữ liệu vector và truy xuất theo thời gian thực để đảm bảo phản hồi chính xác.
3. Cấu hình phương thức học tập: Tùy chỉnh chatbot theo các chế độ như đối thoại định hướng, đặt câu hỏi kích thích tư duy hoặc sinh bài kiểm tra, nhằm nâng cao hiệu quả học tập và khả năng tự đánh giá của học viên.

## **Xây dựng bối cảnh**

Việc xây dựng bối cảnh là yếu tố then chốt trong thiết kế chatbot nhập vai nhằm đảm bảo tương tác có định hướng, phù hợp với môi trường giáo dục, góp phần nâng cao tính chân thực, mức độ nhập vai và giá trị sư phạm trong quá trình triển khai chatbot giáo dục. Ba thành phần chính được xác định gồm:

* **Môi trường tương tác**: Chatbot được đặt vào các bối cảnh giả lập như lớp học trực tuyến, phòng tư vấn hay buổi hội thảo, có tính đến thời điểm trong ngày và tình huống cụ thể để điều chỉnh ngữ điệu và phản hồi phù hợp.
* **Vai trò và mối quan hệ tương tác**: Chatbot đảm nhận các vai trò như giáo viên, cố vấn học tập hoặc giảng viên phản biện, từ đó lựa chọn phong cách giao tiếp và nội dung tương tác nhất quán với kỳ vọng của người học.
* **Mục tiêu tương tác**: Từng lượt hội thoại cần gắn với mục tiêu rõ ràng như cung cấp kiến thức, phản biện, đánh giá năng lực hoặc tư vấn học tập, giúp chatbot định hướng chiến lược phản hồi một cách linh hoạt và hiệu quả.

## **Kiến trúc hệ thống**

## ***Giao diện***

Giao diện người dùng (UI) là yếu tố quan trọng quyết định chất lượng trải nghiệm khi học viên tương tác với chatbot nhập vai. Trong nghiên cứu này, UI được thiết kế theo hướng tối giản và trực quan bằng nền tảng Streamlit, giúp tối ưu hóa khả năng tích hợp với backend Python và hỗ trợ phản hồi thời gian thực. Giao diện tạo không gian trò chuyện tự nhiên, cung cấp đầy đủ thông tin và lịch sử tương tác, đồng thời tích hợp các chức năng học tập như đặt câu hỏi, làm bài kiểm tra hoặc gợi ý kiến thức. Các thành phần chính bao gồm khung hội thoại, thanh nhập văn bản, bảng thông tin nhân vật, chức năng đính kèm tài liệu và bảng điều hướng tiện ích, từ đó nâng cao tính tương tác và hiệu quả học tập.

### ***Kiến trúc Backend***

Thành phần backend đóng vai trò trung tâm trong hệ thống chatbot nhập vai, đảm nhận việc xử lý truy vấn, điều phối mô hình ngôn ngữ lớn (LLM), truy xuất tri thức và duy trì trạng thái hội thoại. Được xây dựng bằng Python, backend tiếp nhận yêu cầu từ giao diện Streamlit, thực hiện tiền xử lý ngôn ngữ như chuẩn hóa văn bản và phân tích mục đích, sau đó chuyển truy vấn thành vector để tìm kiếm ngữ nghĩa trong cơ sở dữ liệu tri thức. Các thông tin liên quan được tổng hợp thành ngữ cảnh và gửi đến LLM để sinh phản hồi. Kết quả sau cùng được trả về giao diện và lưu trữ phục vụ cá nhân hóa và cải tiến hệ thống.

### ***Dữ liệu***

Dữ liệu là yếu tố trung tâm quyết định chất lượng phản hồi của chatbot, đồng thời hỗ trợ lưu trữ và tối ưu hóa khả năng vận hành trong tương lai. Hệ thống sử dụng hai loại cơ sở dữ liệu chính: cơ sở dữ liệu có cấu trúc và cơ sở dữ liệu vector, được xử lý và tổ chức để đáp ứng nhu cầu huấn luyện, truy xuất và sinh phản hồi.

1. Cơ sở dữ liệu có cấu trúc: Hệ thống áp dụng MySQL, một hệ quản trị cơ sở dữ liệu quan hệ mạnh mẽ, có khả năng mở rộng, phù hợp cho các hệ thống học tập với lượng lớn học viên và yêu cầu truy vấn phức tạp. Các chức năng chính bao gồm:

* Lưu trữ thông tin định danh và hồ sơ học viên.
* Quản lý cấu hình nhân vật nhập vai.
* Phân loại và lưu trữ phiên trò chuyện.
* Ghi lại lịch sử hội thoại.
* Quản lý các tệp tài liệu được tích hợp trong chatbot.

1. Cơ sở dữ liệu vector: Cơ sở dữ liệu vector, sử dụng Milvus DB, đóng vai trò quan trọng trong việc truy xuất tri thức ngữ nghĩa hiệu quả. Khác với cơ sở dữ liệu quan hệ, vector database lưu trữ các biểu diễn ngữ nghĩa (embedding) của văn bản, cho phép tìm kiếm dựa trên mức độ tương đồng ý nghĩa thay vì từ khóa.

* Tài liệu học tập được chuyển đổi thành vector thông qua mô hình nhúng, sau đó được lập chỉ mục trong Milvus DB.
* Milvus DB hỗ trợ lưu trữ, quản lý và tìm kiếm vector, phù hợp với các ứng dụng trí tuệ nhân tạo như Retrieval-Augmented Generation (RAG) và chatbot ngữ nghĩa.

### ***Mô hình truy xuất thông tin***

#### Quy trình chuyển đổi tài liệu sang vector

Quy trình chuyển đổi tài liệu văn bản thành cơ sở dữ liệu vector là một bước cốt lõi trong triển khai phương pháp Retrieval-Augmented Generation (RAG), cho phép truy xuất thông tin ngữ nghĩa và sinh phản hồi ngôn ngữ tự nhiên với độ chính xác cao. Các bước thực hiện bao gồm:

1. **Tài liệu đầu vào**: Sử dụng các nguồn văn bản thô như sách, giáo trình, bài báo hoặc văn bản chuyên ngành làm cơ sở tri thức cho hệ thống.
2. **Phân đoạn**: Tài liệu được chia thành các đoạn nhỏ (kích thước 200-500 từ) để tăng tính linh hoạt và hiệu quả trong xử lý và truy xuất.
3. **Chuyển đổi thành vector**: Mỗi đoạn văn được đưa vào mô hình nhúng để chuyển thành vector số, biểu diễn ngữ nghĩa dưới dạng mảng số thực có độ dài cố định (e.g., [-4.9, 3.6, 0.9, 7.8, 3.6]).
4. **Biểu diễn vector**: Các vector phản ánh mối quan hệ ngữ nghĩa, với các đoạn văn có ý nghĩa tương đồng được biểu diễn gần nhau trong không gian vector đa chiều.
5. **Lưu trữ trong cơ sở dữ liệu vector**: Các vector ngữ nghĩa được lưu trữ trong cơ sở dữ liệu vector (Vector DB), hỗ trợ tìm kiếm dựa trên mức độ tương đồng ý nghĩa, thay vì chỉ dựa trên từ khóa.

#### Quy trình truy xuất thông tin

Quy trình sinh tích hợp Retrieval-Augmented Generation (RAG) kết hợp truy xuất thông tin và mô hình ngôn ngữ lớn (LLM) nhằm cung cấp phản hồi chính xác, dựa trên tri thức thực tế. Quy trình này bao gồm các bước sau:

1. Truy vấn người dùng: Học viên gửi truy vấn (câu hỏi hoặc yêu cầu thông tin), được xử lý để phục vụ quá trình truy xuất thông tin ngữ nghĩa.
2. Truy xuất từ cơ sở dữ liệu vector:
   * Truy vấn được chuyển đổi thành vector ngữ nghĩa thông qua mô hình nhúng.
   * Vector truy vấn được so sánh với các vector lưu trữ trong cơ sở dữ liệu vector (Milvus DB) để xác định các đoạn văn bản có độ tương đồng ngữ nghĩa cao nhất.
   * Các đoạn văn bản gốc (chunks) tương ứng với vector được chọn, cùng thông tin liên quan, được trả về.
3. Xây dựng ngữ cảnh: Các đoạn văn bản truy xuất được tổng hợp thành một khối ngữ cảnh thống nhất, đóng vai trò dữ liệu bổ sung cho mô hình ngôn ngữ lớn (LLM) để sinh phản hồi phù hợp và chính xác.

### ***Mô hình tạo phản hồi***

Mô hình tạo phản hồi đóng vai trò cốt lõi trong kiến trúc chatbot, chịu trách nhiệm sinh các câu trả lời phù hợp với ngữ cảnh hội thoại, đặc điểm nhân vật nhập vai và mục tiêu hỗ trợ học tập. Trong nghiên cứu này, quá trình tạo phản hồi được triển khai dựa trên mô hình ngôn ngữ lớn (LLM), kết hợp với các thành phần bổ trợ bao gồm lời nhắc (prompt), lịch sử hội thoại và tri thức ngữ nghĩa được truy xuất, nhằm đảm bảo độ chính xác, tính tự nhiên và sự phù hợp trong giao tiếp.

#### Kiến trúc tổng quan

Quy trình hoạt động của mô hình ngôn ngữ lớn (LLM) trong bối cảnh truy xuất tri thức ngữ nghĩa, tích hợp và xử lý các thành phần đầu vào để tạo phản hồi. Các yếu tố chính bao gồm:

* Tri thức truy xuất: Các tài liệu liên quan (Documents) được truy xuất từ cơ sở dữ liệu để cung cấp thông tin bổ sung.
* Ngữ cảnh hội thoại: Lịch sử tin nhắn (Message history), đảm bảo tính liên tục và phù hợp trong giao tiếp.
* Lời nhắc: Prompt, định hướng dẫn mô hình tạo phản hồi theo mục tiêu cụ thể.

#### Mô hình sử dụng

Trong đề án này, mô hình ngôn ngữ lớn GPT-4 của OpenAI được lựa chọn nhờ các đặc tính vượt trội, đáp ứng yêu cầu của hệ thống chatbot nhập vai. Các lý do chính bao gồm:

1. Hiểu ngữ cảnh sâu: GPT-4 có khả năng duy trì mạch hội thoại, đảm bảo phản hồi nhất quán và liên kết chặt chẽ với ngữ cảnh.
2. Tính linh hoạt trong phản hồi: Mô hình hỗ trợ tùy chỉnh phong cách giao tiếp thông qua lời nhắc (prompt) được thiết kế riêng, phù hợp với các đặc điểm nhân vật nhập vai.
3. Tích hợp hiệu quả: GPT-4 dễ dàng tích hợp vào hệ thống thông qua API OpenAI, tương thích với môi trường phát triển Python và giao diện Streamlit.
4. Hỗ trợ đa ngôn ngữ: Mô hình xử lý tốt tiếng Việt, đáp ứng nhu cầu triển khai trong lĩnh vực giáo dục tại Việt Nam.

#### Các thành phần bổ trợ

Trong nghiên cứu này, việc tích hợp agent trong LangChain với mô hình ngôn ngữ lớn (LLM) và các thành phần bổ trợ nhằm nâng cao khả năng tạo sinh văn bản thông minh và phù hợp ngữ cảnh. Các yếu tố chính bao gồm:

1. Lời nhắc (Prompt): Được thiết kế để định hướng vai trò và nhiệm vụ cụ thể của mô hình, đảm bảo phản hồi phù hợp với mục tiêu.
2. Lịch sử hội thoại (Chat history): Lưu trữ và cung cấp ngữ cảnh từ các lượt tương tác trước, hỗ trợ duy trì tính liên tục và bổ sung thông tin ngữ cảnh.
3. Kiến thức bên ngoài (External knowledge): Cung cấp dữ liệu bổ sung từ các nguồn tri thức bên ngoài, nâng cao độ chính xác và tính thực tiễn của phản hồi.

Sự kết hợp này cho phép mô hình sinh văn bản một cách thông minh, chính xác và phù hợp với bối cảnh ứng dụng.

### ***Các bước triển khai thiết kế***

Mô hình chatbot nhập vai được thiết kế dựa trên năm thành phần chính: LLM, tri thức ngoài, cơ sở dữ liệu vector, hệ thống truy xuất thông tin, và cơ sở dữ liệu lưu trữ kịch bản nhập vai. Quá trình triển khai gồm bốn bước:

1. Lựa chọn và tùy chỉnh mô hình ngôn ngữ lớn (LLM):

* Khả năng hiểu ngữ cảnh trong các đoạn hội thoại dài, đảm bảo tính liên kết.
* Tốc độ phản hồi ổn định theo thời gian thực.
* Khả năng nhập vai hiệu quả, duy trì vai trò nhất quán suốt quá trình giao tiếp.

1. Lưu trữ tri thức bên ngoài trong cơ sở dữ liệu vector:

* Tài liệu tri thức được phân đoạn thành các đoạn nhỏ (chunks).
* Áp dụng mô hình nhúng (embedding) để chuyển đổi các đoạn văn bản thành vector ngữ nghĩa.
* Các vector được lưu trữ trong cơ sở dữ liệu vector để hỗ trợ truy xuất ngữ nghĩa nhanh chóng và hiệu quả.

1. Xây dựng hệ thống truy xuất thông tin:

* Chuyển đổi truy vấn người dùng thành vector ngữ nghĩa.
* Tìm kiếm các vector có độ tương đồng cao.
* Tối ưu hóa số lượng kết quả truy xuất.

1. Tích hợp mô hình ngôn ngữ lớn với agent tạo sinh văn bản:

* Lời nhắc (Prompt): Định hướng vai trò và nhiệm vụ của mô hình.
* Lịch sử hội thoại (Chat history): Lưu trữ và cung cấp ngữ cảnh từ các lượt tương tác trước.
* Tri thức bên ngoài (External knowledge): Bổ sung dữ liệu từ cơ sở tri thức để tăng độ chính xác và tính thực tiễn của phản hồi.

### ***Kết luận***

Lựa chọn mô hình ngôn ngữ lớn (LLM) là yếu tố then chốt đảm bảo chất lượng hội thoại và duy trì khả năng nhập vai ổn định theo ngữ cảnh. Hệ thống tận dụng kỹ thuật nhúng và lưu trữ tri thức trong cơ sở dữ liệu vector để truy xuất thông tin hiệu quả. Việc ghi nhớ kịch bản nhập vai và lịch sử tương tác giúp duy trì mạch hội thoại và cá nhân hóa phản hồi. Nhìn chung, hệ thống đáp ứng tốt yêu cầu về giao tiếp tự nhiên trong môi trường giáo dục. Các hướng phát triển tương lai bao gồm mở rộng dữ liệu tri thức, tích hợp đa phương thức và tối ưu hóa hiệu năng thời gian thực.

## **Tổng kết chương 3**

Chương 3 trình bày quá trình thiết kế và phát triển mô hình chatbot nhập vai (RP Chatbot), kết hợp công nghệ xử lý ngôn ngữ tự nhiên (NLP) và mô hình ngôn ngữ lớn (LLM) GPT đã tinh chỉnh để tạo phản hồi linh hoạt và ngữ cảnh phù hợp. Nội dung tập trung vào:

* Thiết kế nhân vật với hành vi và phong cách nhất quán
* Xây dựng kiến trúc hệ thống gồm giao diện người dùng, backend, truy xuất tri thức và mô hình RAG
* Ứng dụng các kỹ thuật tối ưu như vector hóa tri thức, quản lý hội thoại và thiết kế prompt.

Chương tiếp theo sẽ đánh giá hiệu quả mô hình, so sánh với phương pháp khác và phân tích tính khả thi trong môi trường giáo dục.

# **CHƯƠNG 4**

# **ĐÁNH GIÁ, SO SÁNH CÁC KẾT QUẢ DỰ BÁO VÀ BÀN LUẬN**

## **Mô hình triển khai thử nghiệm**

Trong chương 4, em sẽ trình bày quá trình đánh giá hiệu quả và khả năng ứng dụng của mô hình chatbot nhập vai bằng thực nghiệm, sử dụng GPT-4 kết hợp cơ chế RAG nhằm hỗ trợ học tập cá nhân hóa. Nghiên cứu áp dụng phương pháp đánh giá định tính và định lượng, tập trung vào độ chính xác tri thức, khả năng giữ vai, mức độ phù hợp ngữ cảnh và tính thích nghi với người học. Mô hình Giskard được sử dụng để kiểm thử lỗi ngữ nghĩa, định kiến và sai sót, đảm bảo độ tin cậy. Kết quả cho thấy mô hình vượt trội so với chatbot truyền thống, đặc biệt trong xử lý câu hỏi phức tạp và hội thoại phong phú.

## **Phương pháp đánh giá hệ thống thử nghiệm**

### ***Mục tiêu***

Nghiên cứu đề xuất mô hình chatbot nhập vai tích hợp RAG nhằm khắc phục hạn chế phản hồi chung của LLM truyền thống, hướng đến phản hồi chuyên biệt theo từng nhân vật. Hệ thống gồm giao diện Streamlit, backend Python, cơ sở dữ liệu MySQL và MilvusDB, cùng mô hình GPT-4 triển khai theo kiến trúc RAG. Dữ liệu được cá nhân hóa theo từng giáo viên ảo nhằm tăng tính chuyên biệt. Mô hình được đánh giá qua tốc độ phản hồi, độ chính xác chủ đề và mức độ lặp lại, cho thấy khả năng cung cấp phản hồi linh hoạt, sát ngữ cảnh học tập.

### ***Đánh giá mô hình sinh trên các tập kiểm tra***

Trong nghiên cứu này, tác giả sử dụng công cụ Giskard để thiết lập một quy trình kiểm thử tự động và khách quan nhằm đánh giá hiệu quả của mô hình chatbot nhập vai trong việc phản hồi các câu hỏi. Quy trình kiểm thử bao gồm sáu bước chính: nạp dữ liệu kiểm thử định dạng .jsonl bằng lớp QATestset, chuyển đổi dữ liệu thành Test Suite, định nghĩa hàm sinh phản hồi batch\_prediction\_fn, xây dựng mô hình giskard.Model, chạy kiểm thử và hiển thị kết quả.

Kết quả đánh giá được phân tích theo từng loại câu hỏi, cho thấy mô hình đạt độ chính xác cao với câu hỏi tình huống (100%) và câu hỏi phức tạp (90%), phản ánh khả năng hiểu ngữ cảnh và suy luận tốt. Tuy nhiên, độ chính xác thấp ở các câu hỏi hội thoại và có yếu tố gây nhiễu (đều ở mức 50%) cho thấy mô hình còn hạn chế trong việc xử lý các tình huống có tính linh hoạt hoặc đánh lạc hướng cao.

## **Phân tích và nhận xét**

### ***Phân tích mô hình***

Trong bối cảnh hiện nay, mặc dù có nhiều mô hình ngôn ngữ lớn (LLMs) mã nguồn mở và thương mại, nhưng khả năng ứng dụng trực tiếp vào các bài toán chuyên biệt và cá nhân hóa vẫn còn hạn chế. Do đó, để đạt hiệu quả thực tiễn cao, việc tinh chỉnh các mô hình ngôn ngữ theo nhu cầu cụ thể của cá nhân, doanh nghiệp hoặc tổ chức là cần thiết.

Trong đề án này, em đã sử dụng mô hình GPT và thực hiện tinh chỉnh để xây dựng một **mô hình giáo viên ảo** phù hợp với yêu cầu giảng dạy và học tập. Mô hình này được so sánh với chatbot truyền thống thông qua các tiêu chí cụ thể. Kết quả cho thấy chatbot AI nhập vai (AI RP Chatbot) vượt trội hơn về mức độ cá nhân hóa, khả năng cung cấp kiến thức, phương pháp giảng dạy đa dạng, cũng như tính linh hoạt trong cập nhật và điều chỉnh chương trình học. Trong khi đó, chatbot truyền thống chủ yếu phản hồi ở mức cơ bản, thiếu khả năng tương tác sâu và thích ứng theo tiến trình học của người dùng.

### ***Đóng góp và ý nghĩa của đề án***

Nghiên cứu đề xuất một phương pháp mới phát triển chatbot nhập vai dựa trên mô hình ngôn ngữ lớn (LLM), đảm bảo vai trò nhân vật ổn định và phản hồi phù hợp theo bối cảnh. Hệ thống cung cấp phản hồi chính xác về tri thức, duy trì phong cách giao tiếp nhất quán và thích nghi với đặc điểm riêng của học viên.

Điểm nổi bật là khả năng mô phỏng hành vi nhập vai thông qua hồ sơ nhân vật chi tiết, bao gồm tính cách, phong cách ngôn ngữ và giới hạn hành vi, tạo trải nghiệm học tập chân thực và hỗ trợ giáo dục cá nhân hóa. Đề tài đóng góp về mặt kỹ thuật, mô hình hóa hành vi, và mở ra tiềm năng ứng dụng trong đào tạo mô phỏng, học tập tương tác và tự học chủ động, vượt qua hạn chế của các chatbot giáo dục hiện nay.

## **Tổng kết chương 4**

Trong chương 4, em đánh giá hiệu quả mô hình chatbot nhập vai dựa trên các tiêu chí kỹ thuật và so sánh với chatbot truyền thống. Mô hình thể hiện ưu thế về phản hồi mạch lạc, phù hợp ngữ cảnh, cá nhân hóa cao và duy trì ngôn ngữ đặc trưng nhân vật. Việc tích hợp RAG giúp tăng tốc và độ chính xác trong truy xuất tri thức. Tuy nhiên, mô hình còn hạn chế ở hiệu suất thời gian thực, độ nhất quán trong hội thoại dài, và xử lý câu hỏi gây nhiễu. Kết quả cho thấy mô hình tiềm năng nhưng cần cải thiện để ứng dụng rộng hơn.

# **KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TIẾP THEO**

Đề án nghiên cứu ứng dụng mô hình ngôn ngữ lớn (LLM) trong bối cảnh cách mạng công nghiệp 4.0, nhấn mạnh vai trò then chốt của AI trong giáo dục. Đề án trình bày tổng quan về LLM, nguyên lý hoạt động, kiến trúc và xu hướng phát triển, đồng thời đề xuất giải pháp tối ưu hóa cho giáo dục. Một mô hình chatbot nhập vai được xây dựng, có khả năng mô phỏng phong cách giao tiếp, tính cách và kiến thức chuyên môn theo kịch bản, cung cấp phản hồi tự nhiên, mạch lạc, ghi nhớ tương tác trước để tạo trải nghiệm học tập cá nhân hóa. Hệ thống tích hợp tri thức bên ngoài, đảm bảo phản hồi chính xác, độ tin cậy cao và khả năng mở rộng cho nhiều vai trò giáo dục.

Kết quả chính:

* Xây dựng nguyên mẫu chatbot nhập vai hỗ trợ học tập theo kịch bản.
* Tích hợp tri thức bên ngoài nâng cao tính chính xác và độ tin cậy.
* Mô hình có tính mở rộng, dễ dàng thích ứng với các nhiệm vụ giáo dục.

Hạn chế:

* Chỉ xử lý văn bản, chưa hỗ trợ âm thanh hoặc hình ảnh.
* Tùy biến theo người học còn hạn chế, thiếu kỹ thuật tinh chỉnh tiên tiến.
* Hiệu năng giảm trên thiết bị hạn chế tài nguyên.

Hướng phát triển:

* Tối ưu hóa mô hình cho trợ lý học tập cá nhân, trợ giảng ảo và huấn luyện kỹ năng mềm.
* Ứng dụng kỹ thuật như Prompt Engineering, Adapter Tuning, RAFT.
* Nâng cấp hệ thống đa kênh, hỗ trợ giọng nói, hình ảnh, video và PDF để tạo môi trường học tập toàn diện.