I.Tóm tắt nội dung tuần 2 của khóa học về LLM

Ngày 1: Sử dụng nhiều API AI - OpenAI, Claude, và Gemini

A.Nội dung chính:

- Giới thiệu về các API tiên tiến cho LLM: OpenAI (GPT-4), Claude (Anthropic), và Gemini (Google).

- Cách sử dụng các API này để gọi mô hình ngôn ngữ lớn (LLM) và xử lý đầu ra dưới nhiều định dạng (Markdown, JSON, Streaming).

- So sánh ưu và nhược điểm của các API, cách chọn mô hình phù hợp với nhu cầu cụ thể.

- Cách thiết lập môi trường lập trình với JupyterLab, quản lý khóa API một cách an toàn bằng tệp `.env`.

B.Kỹ năng đạt được:

- Thành thạo trong việc gọi API của OpenAI, Anthropic, Google.

- Biết cách kết hợp nhiều API để tối ưu hóa kết quả đầu ra.

- Hiểu về prompt engineering, bao gồm các kỹ thuật như one-shot prompting và cách phối hợp giữa hệ thống prompt và user prompt.

Ngày 2: Xây dựng giao diện AI với Gradio

A.Nội dung chính:

- Giới thiệu Gradio – một framework nhanh chóng để tạo giao diện AI.

- Cách sử dụng Gradio để tạo UI tương tác với OpenAI GPT, Claude, và Gemini.

- Cách tích hợp các mô hình LLM vào ứng dụng thực tế.

B.Kỹ năng đạt được:

- Thành thạo trong việc tạo giao diện AI đơn giản với Gradio.

- Biết cách kết nối giao diện với API của LLM để tạo chatbot hoặc ứng dụng AI.

- Cách sử dụng Markdown và Streaming Output trong giao diện AI.

Ngày 3: Xây dựng Chatbot AI - Mastering Gradio cho hỗ trợ khách hàng
A.Nội dung chính:

- Tạo giao diện chatbot AI với Gradio.

- Quản lý lịch sử hội thoại và cách LLM lưu giữ ngữ cảnh qua các lượt tương tác.

- Kỹ thuật multi-shot prompting để cải thiện phản hồi của mô hình AI.

B.Kỹ năng đạt được:

- Biết cách xây dựng một chatbot hỗ trợ khách hàng sử dụng OpenAI GPT.

- Cách duy trì ngữ cảnh trò chuyện bằng cách truyền lịch sử hội thoại vào API.

Ngày 4: Sử dụng công cụ AI với LLM
A.Nội dung chính:

- Khái niệm AI tools và cách chúng mở rộng khả năng của LLM.

- Xây dựng trợ lý AI cho hãng hàng không – một chatbot có khả năng cung cấp giá vé máy bay theo thời gian thực.

B.Kỹ năng đạt được:

- Hiểu về tool-augmented LLM, cách bổ sung công cụ để giúp mô hình AI có thể thực hiện nhiều nhiệm vụ hơn.

- Biết cách tạo chatbot có thể tra cứu dữ liệu và thực hiện hành động dựa trên yêu cầu của người dùng.

Ngày 5: AI đa phương thức – Tích hợp tạo hình ảnh và âm thanh
A.Nội dung chính:

- Khám phá LLM Agents – các thực thể AI có thể tự động hóa quy trình phức tạp.

- Cách tích hợp DALL-E 3 để tạo hình ảnh bằng AI.

- Tạo trợ lý AI đa phương thức có thể vẽ hình và tạo âm thanh dựa trên yêu cầu của người dùng.

B.Kỹ năng đạt được:

- Biết cách tích hợp DALL-E 3 để tạo hình ảnh bằng AI.

- Xây dựng mô hình AI có thể tạo ra nhiều dạng nội dung khác nhau (văn bản, hình ảnh, âm thanh).

II.Từ khóa quan trọng cho nghiên cứu và phát triển LLM

API Integration for LLM:

Tích hợp API của các mô hình LLM hàng đầu (GPT-4, Claude, Gemini) vào ứng dụng AI, bao gồm xác thực, gọi API, và xử lý đầu ra.

Gradio UI Development:

Sử dụng Gradio để tạo giao diện AI giúp người dùng dễ dàng tương tác với mô hình LLM mà không cần lập trình phức tạp.

Prompt Engineering Techniques:

Kỹ thuật tối ưu hóa lời nhắc (prompt) để mô hình AI phản hồi chính xác và hiệu quả, bao gồm zero-shot, one-shot, và few-shot prompting.

Chatbot Development:

Xây dựng chatbot AI có khả năng hiểu ngữ cảnh, duy trì hội thoại tự nhiên và phản hồi theo thời gian thực.

Context Handling in LLM:

Cách quản lý ngữ cảnh trong API của LLM để đảm bảo phản hồi mạch lạc và liên kết với thông tin trước đó.

Function Calling in LLM:

Sử dụng OpenAI Function Calling để kết nối LLM với công cụ bên ngoài như cơ sở dữ liệu, hệ thống tìm kiếm hoặc API bổ sung.

Tool-Augmented LLM:

Mở rộng khả năng của LLM bằng cách tích hợp các công cụ bên ngoài, giúp AI thực hiện các tác vụ như phân tích dữ liệu hoặc lập lịch.

AI Agents & Multi-Step Reasoning:

Ứng dụng AI Agents để thực hiện nhiệm vụ phức tạp theo nhiều bước, giúp AI tự động hóa các quy trình.

Multimodal AI (Text, Image, Audio):

Xây dựng hệ thống AI có thể xử lý dữ liệu đa phương thức như văn bản, hình ảnh và âm thanh để cung cấp trải nghiệm toàn diện.

DALL-E Image Generation:

Sử dụng DALL-E để tạo hình ảnh từ mô tả văn bản, áp dụng trong sáng tạo nội dung, thiết kế sản phẩm, và minh họa.

III.Các công nghệ được đề cập trong tuần 2 khóa học LLM

OpenAI GPT-4:

Mô hình ngôn ngữ lớn hàng đầu hiện nay, cung cấp khả năng xử lý ngôn ngữ tự nhiên mạnh mẽ.

Claude AI (Anthropic):

Mô hình AI tiên tiến từ Anthropic, tối ưu hóa cho các ứng dụng chatbot và hỗ trợ khách hàng.

Gemini AI (Google):

Mô hình AI của Google tập trung vào tính đa dụng và hiệu suất cao.

Gradio:

Framework giúp tạo giao diện người dùng nhanh chóng để thử nghiệm và triển khai AI.

JupyterLab:

Môi trường lập trình phổ biến cho việc phát triển AI và khoa học dữ liệu.

OpenAI Function Calling:

Công nghệ cho phép mô hình AI gọi các hàm bên ngoài để thực hiện các tác vụ cụ thể.

DALL-E 3:

Mô hình tạo hình ảnh từ văn bản mô tả do OpenAI phát triển.

LLM Agents:

Các thực thể AI có khả năng tự động thực hiện nhiệm vụ theo từng bước.

Multimodal AI:

Công nghệ AI có thể xử lý nhiều loại dữ liệu như văn bản, hình ảnh, và âm thanh.

API Key Management:

Phương pháp quản lý khóa API an toàn để bảo vệ dữ liệu và hạn chế truy cập trái phép.

III.Các công nghệ được đề cập trong tuần 2 khóa học LLM

OpenAI GPT-4:

- Cung cấp mô hình ngôn ngữ mạnh mẽ để xử lý và tạo nội dung văn bản.
- Được sử dụng làm nền tảng cho các chatbot AI, trợ lý ảo và sinh mã code.
- Hỗ trợ Function Calling để tích hợp với công cụ bên ngoài, mở rộng khả năng xử lý.

Claude AI (Anthropic):

- Chuyên về hỗ trợ hội thoại và tạo phản hồi tự nhiên.
- Tối ưu hóa để đảm bảo sự an toàn và đạo đức khi sử dụng AI.
- Được sử dụng như một lựa chọn thay thế cho OpenAI GPT-4 trong một số ứng dụng.

Gemini AI (Google):

- Hỗ trợ xử lý đa phương thức (văn bản, hình ảnh, âm thanh) trong một mô hình duy nhất.
- Tích hợp mạnh mẽ với hệ sinh thái Google để phân tích và tìm kiếm thông tin.
- Có khả năng học hỏi từ tài liệu mới và áp dụng vào phản hồi AI.

Gradio:

- Cung cấp giao diện đơn giản để thử nghiệm và triển khai mô hình LLM nhanh chóng.

- Giúp các nhà phát triển tạo UI tương tác mà không cần xây dựng ứng dụng web phức tạp.
- Cho phép kết nối trực tiếp API LLM vào các ứng dụng thử nghiệm.

JupyterLab:

- Môi trường lập trình phổ biến để thử nghiệm và phát triển AI.
- Hỗ trợ ghi chép, chạy mã Python, và hiển thị kết quả trực tiếp.
- Dùng để triển khai API LLM và tạo mô hình thử nghiệm.

OpenAI Function Calling:

- Cho phép mô hình LLM gọi hàm bên ngoài để thực hiện các tác vụ đặc biệt.
- Giúp tích hợp LLM vào các hệ thống thực tế như truy vấn cơ sở dữ liệu, đặt lịch, và xử lý tác vụ kinh doanh.
- Mở rộng khả năng của AI vượt ra ngoài việc chỉ xử lý ngôn ngữ tự nhiên.

DALL-E 3:

- Tạo hình ảnh từ mô tả văn bản để hỗ trợ các ứng dụng sáng tạo.
- Mở rộng khả năng của LLM sang lĩnh vực thiết kế, nghệ thuật và mô phỏng thị giác.
- Được sử dụng để tạo nội dung trực quan trong các ứng dụng AI đa phương thức.

LLM Agents:

- Giúp AI thực hiện các nhiệm vụ phức tạp theo từng bước thay vì chỉ phản hồi một lần.
- Có khả năng lập kế hoạch, phân tích, và ra quyết định dựa trên mục tiêu cụ thể.
- Được sử dụng trong các hệ thống tự động hóa, trợ lý ảo, và phân tích dữ liệu.

Multimodal AI:

- Cho phép LLM xử lý nhiều loại dữ liệu khác nhau như văn bản, hình ảnh, và âm thanh.
- Mở rộng ứng dụng của AI sang các lĩnh vực như nhận diện giọng nói, phân tích hình ảnh, và chatbot đa phương thức.
- Kết hợp nhiều dạng đầu vào để cải thiện trải nghiệm người dùng.

API Key Management:

- Cung cấp phương pháp bảo mật khi sử dụng API của LLM để ngăn chặn rò rỉ dữ liệu.
- Hỗ trợ quản lý quyền truy cập để kiểm soát việc sử dụng API trong các hệ thống doanh nghiệp.
- Đảm bảo các mô hình LLM có thể chạy trong môi trường an toàn và đáng tin cậy.