ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO SEMINAR

Cách cài đặt chat bot trên ứng dụng web từ API key của ChatGPT

Môn: Mạng Máy Tính

Giảng viên: Th.S Đỗ Hoàng Cường

Họ và tên	MSSV
Nguyễn Quốc Trung	21120350
Trần Trọng Nghĩa	21120507
Trần Kỳ Thanh	21120556

Ngày 13 tháng 3 năm 2023

1 Giới thiệu

1.1 ChatGPT

Chat bot là một chương trình phần mềm được thiết kế để mô phỏng cuộc trò chuyện với con người, thường là qua Internet, cung cấp hỗ trợ trong nhiều công việc khác nhau như bán hàng, quảng bá sản phẩm, hỗ trợ khách hàng, giới thiệu nội dung...

ChatGPT là một mô hình máy học (machine learning) được phát triển bởi OpenAI, một tổ chức nghiên cứu trí tuệ nhân tạo được thành lập vào năm 2015. ChatGPT gây chú ý bởi khả năng sinh ra văn bản tự nhiên với độ chính xác cao, dưới nhiều hình thức nội dung khác nhau (giải thích kiến thức, soạn e-mail, viết CV...) theo yêu cầu của người dùng cuối. GPT là từ viết tắt tiếng Anh của Generative Pre-trained Transformer - một mạng lưới thần kinh AI được huấn luyện trước trên một lượng dữ liệu văn bản khổng lồ được cung cấp từ nhiều nguồn đa dạng về nội dung, chủ đề, ngôn ngữ và cấu trúc nhằm sinh ra các câu trả lời tự nhiên, giống với ngôn ngữ của con người nhất có thể.

Nền tảng cơ bản của ChatGPT là mạng thần kinh (neural network) - một mô hình học sâu bao gồm các neuron nhân tạo (được gọi là nút) tổ chức thành nhiều lớp (mỗi lớp gồm nhiều nút), được liên kết với nhau. Trong mạng thần kinh, thông tin được truyền đi thông qua mạng lưới kết nối phức tạp giữa các lớp và quá trình tính toán cụ thể được thực thi tại từng nút. Khi quá trình nhận dữ liệu và phản hồi được lặp đi lặp lại nhiều lần qua các lớp, mạng thần kinh dần xuất hiện khả năng nhận dạng các mẫu và dấu hiệu trong dữ liệu đầu vào, từ đó "học"được kiến thức mà người phát triển mong muốn, giúp đưa ra kết quả chính xác và đáp ứng được yêu cầu công việc.

Với 175 tỷ tham số trong mô hình, khối lượng dữ liệu training khổng lồ, ChatGPT trở thành một công cu tiên phong trong công việc xử lý ngôn ngữ tự nhiên và có ảnh hưởng to lớn đến khắp các lĩnh vực, từ thị trường công nghệ, việc làm, đến giáo dục và đào tạo...

1.2 Những tính năng nổi bật của ChatGPT

Một số tính năng nổi bật của ChatGPT bao gồm:

- Giao tiếp cơ bản với người dùng: chào hỏi, hội thoại cơ bản, phản hồi với nhận xét của người dùng.
- Trả lời những câu hỏi trực tiếp từ người dùng về hầu hết các chủ đề.
- Thông báo với người dùng nếu không có hiểu biết về chủ đề được yêu cầu.
- Soạn e-mail theo yêu cầu của người dùng về nội dung, văn phong, hình thức, mục đích.
- Đề xuất ý tưởng trong nhiều lĩnh vực như công việc, học tập, giải trí, sáng tạo nội dung.
- Viết code theo yêu cầu người dùng và giải thích code.
- Có thể dịch văn bản với nhiều ngôn ngữ khác nhau.
- Đọc hiểu văn bản: phân tích nội dung, bố cục, văn phong, ngữ nghĩa...
- Tạo ra các tác phẩm truyện, thơ, bài viết ngắn, tóm tắt, báo cáo và nội dung khác dựa trên thông tin được cung cấp.

- Phân tích dữ liệu và trích xuất thông tin từ các nguồn dữ liệu khác nhau như báo cáo, trang web, tài liệu học tập, tin tức và phương tiện truyền thông xã hội.
- Ghi nhớ lịch sử hồi thoại để duy trì đối thoại một cách tự nhiên nhất có thể và đính chính thông tin sai lệch nếu có.
- Từ chối tạo ra các nội dung nhạy cảm, không lành mạnh.
- Điều khiển và điều chỉnh các thiết bị thông minh: hỗ trợ người dùng điều khiển các thiết bị thông minh như đèn, máy lạnh, thiết bị giải trí qua các lệnh đơn giản.

Ngoài ra, ChatGPT vẫn còn tồn tại những vấn đề như:

- Có thể đưa ra thông tin sai lệch, không chính xác.
- Có thể hiểu sai yêu cầu hoặc ý định của người dùng và cung cấp phản hồi không chính xác hoặc không phù hợp.
- Thiếu khả năng tương tác thực tế: chỉ có thể tương tác với người dùng thông qua văn bản, không có khả năng tương tác trực tiếp với môi trường thực tế.
- Thiếu khả năng đánh giá và lựa chọn: ChatGPT hiện tại chưa có khả năng đánh giá và lựa chọn các phương án tối ưu nhất trong các tình huống phức tạp.
- Tính bảo mật: nguy cơ rò rỉ thông tin và xâm nhập tài khoản của người dùng.
- Khả năng bị lạm dụng: ChatGPT có thể bị lạm dụng để tạo ra các phản hồi giả mạo hoặc các thông tin sai lệch, gây ra những hậu quả tiêu cực cho người dùng và xã hội.
- Bị ảnh hưởng bởi thiên kiến của dữ liệu training.
- Không có hiểu biết đầy đủ về những sự kiện, hiện tượng xảy ra sau 2021.

1.3 Triển vọng

- Đẩy mạnh quá trình tương tác giữa người và máy.
- Trợ lý ảo và hệ thống tự động.
- Tra cứu thông tin, tóm tắt kiến thức, trích dẫn nguồn, kiểm định thông tin.
- Tư vấn người dùng về các lĩnh vực cụ thể (y tế, giáo dục, tài chính...).
- Hỗ trợ học sinh, sinh viên trong công tác học tập, chuẩn bị bài, nghiên cứu, hoàn thành bài tập.
- Hỗ trợ người dùng có các vấn đề sức khỏe, thương tật.
- Phát triển các trò chơi thông minh có tính tương tác cao.
- Hỗ trợ ngôn ngữ: dịch thuật, giải thích định nghĩa của từ.

2 Cài đặt chat bot

2.1 Tạo tài khoản OpenAI

Hiện tại, Việt Nam chưa có trong danh sách các quốc gia mà OpenAI trực tiếp hỗ trợ cung cấp dịch vụ. Khi truy cập vào trang web chính thức từ mạng Việt Nam, người dùng sẽ nhận được thông báo "ChatGPT is not available in your country", cho thấy ứng dụng này chưa được hỗ trợ tại quốc gia của chúng ta.

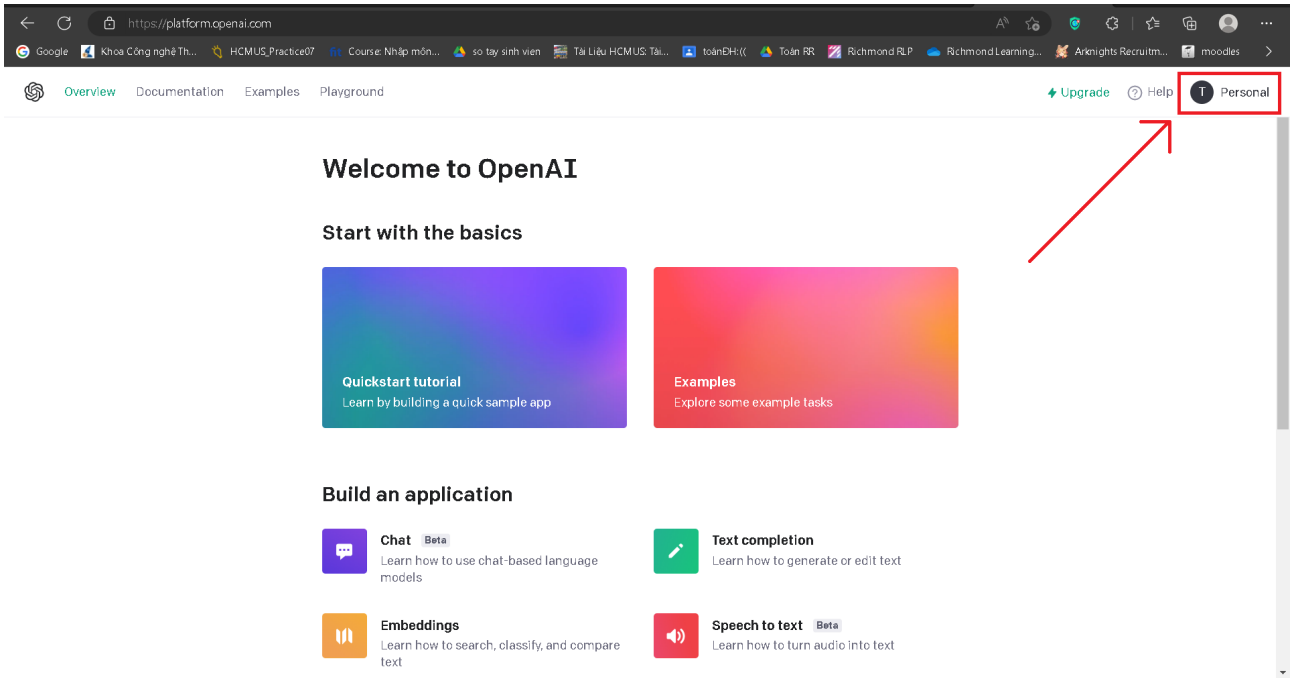
Do đó, giải pháp dành cho người dùng tại Việt Nam là thông qua các kênh thứ ba để đăng ký truy cập như sử dụng VPN để thay đổi sang địa chỉ IP của các nước được hỗ trợ dịch vụ, thuê số điện thoại nước ngoài để đăng ký tài khoản, hay mua các tài khoản có sẵn.
Tuy nhiên, việc sử dụng các kênh thứ ba mang lại một số rủi ro tiềm ẩn nhất định cho người dùng. Ví dụ, việc quá nhiều người cùng sử dụng một tài khoản có thể làm chậm quá trình nhận phản hồi và giảm hiệu quả công việc, thường xuyên gây tắc nghẽn. Ngoài ra, việc thuê số điện thoại nước ngoài hoặc mua tài khoản không uy tín chứa đựng nguy cơ bị lừa đảo. Hơn nữa, người dùng không sở hữu tài khoản thực sự vì số điện thoại đăng ký thuộc về bên thứ ba.

2.2 Nhận API key

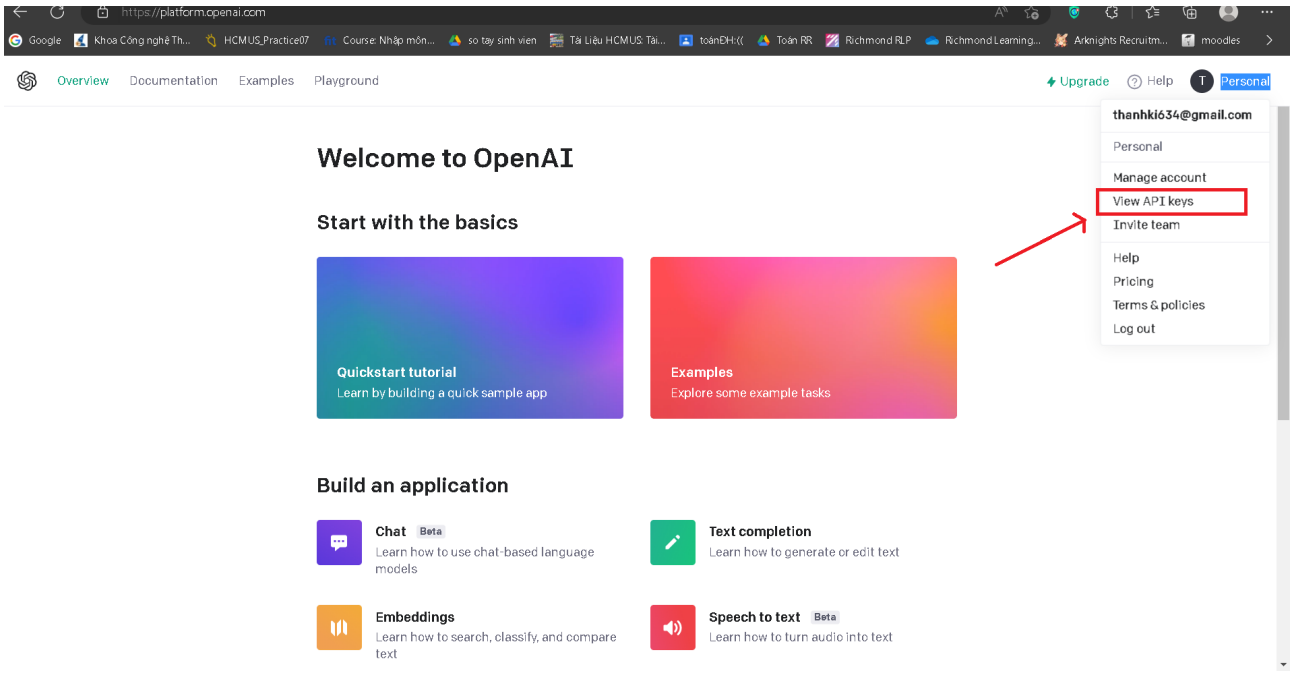
Bước 1: Truy cập vào trang web platform.openai.com bằng đường link Overview - OpenAI API hoặc beta.openai.com.

Bước 2: Sử dụng công cụ VPN (nếu cần thiết) và đăng nhập vào tài khoản OpenAI đã được đăng ký trước đó.

Bước 3: Sau khi đăng nhập thành công, mở trang chủ platform.openai.com và di chuyển chuột đến mục Personal ở góc trên bên phải màn hình, chọn mục "View API keys".

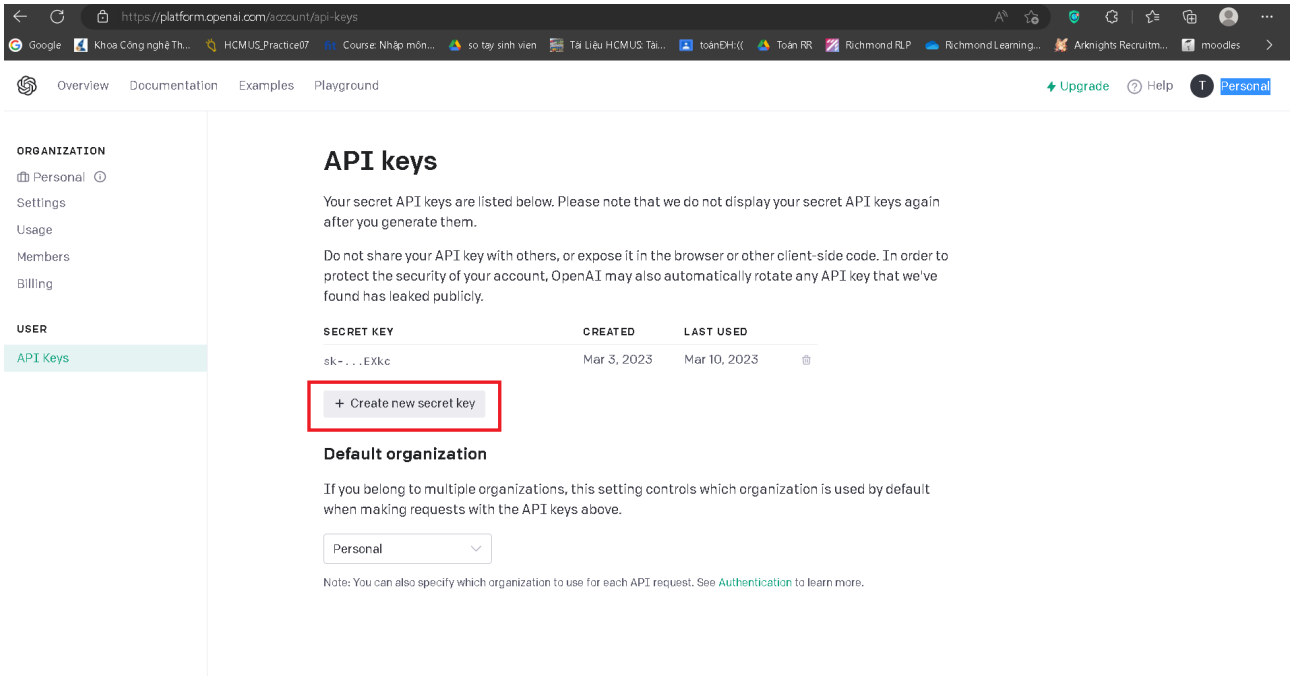


Hình 1: Di chuột vào mục personal để hiện các tùy chỉnh và thông tin cá nhân

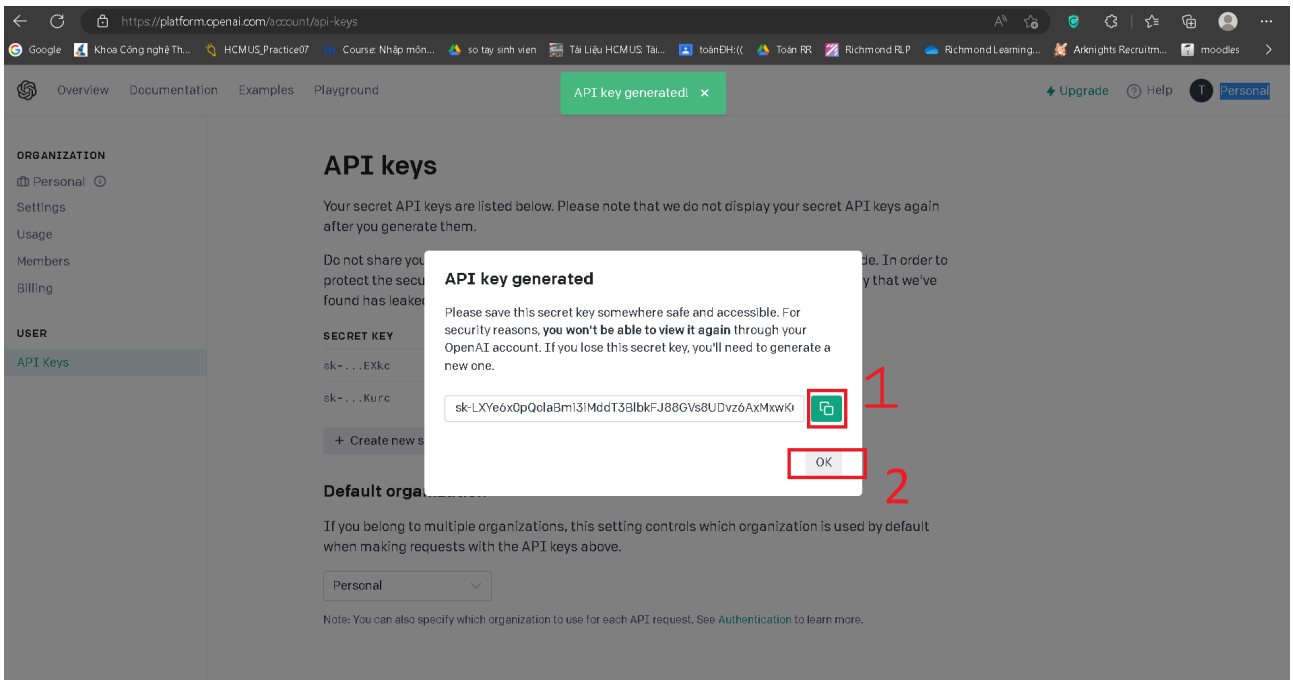


Hình 2: Chọn mục "View API keys"để được đưa đến trang kiểm soát API key

Bước 4: Bấm vào mục “Create new secret key” để khởi tạo một API key mới. Một hộp thoại sẽ hiện lên cùng với API key được khởi tạo. Lưu lại API key này ở một nơi an toàn vì sau khi tắt hộp thoại, sẽ không thể xem đầy đủ API key này được nữa.

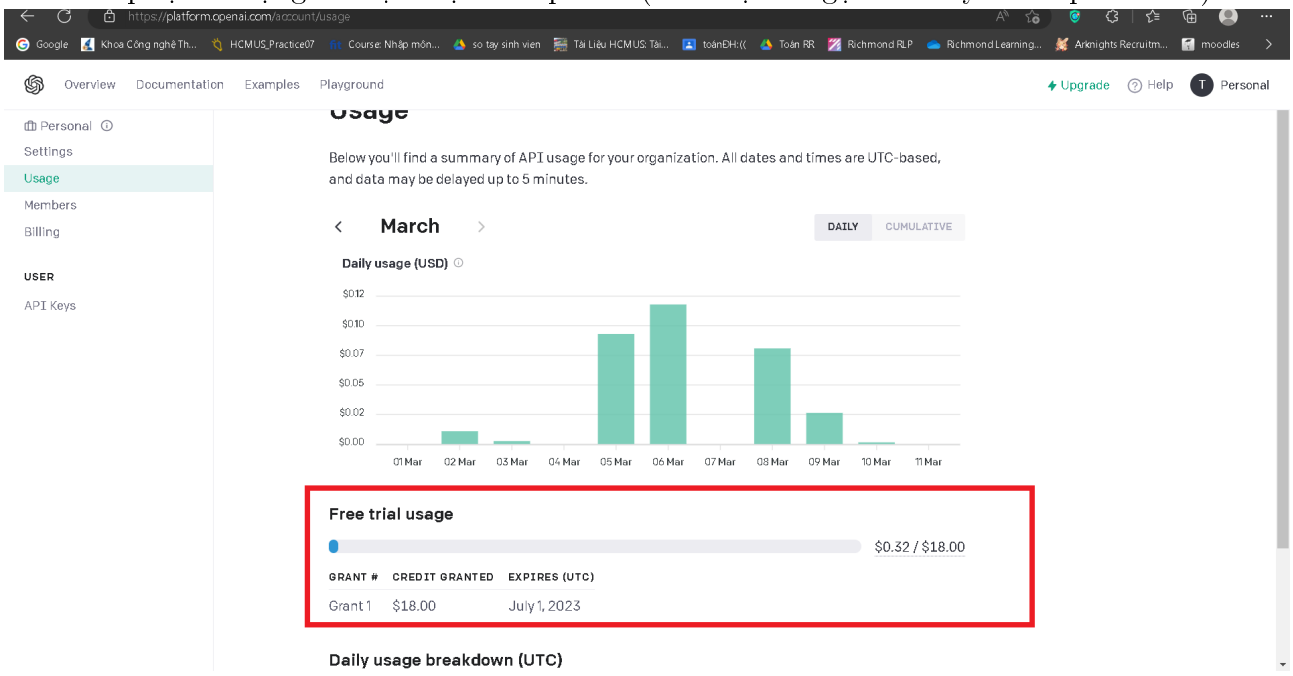


Hình 3: Chọn "Create new secret key"để khởi tạo API key mới



Hình 4: Sao chép key vừa được khởi tạo, lưu nó ở nơi an toàn. Sau đó tắt hộp thoại để hoàn tất quá trình khởi tạo

Lưu ý: Một tài khoản miễn phí được cung cấp cho người dùng bởi OpenAI có hạn mức sử dụng là $18, sau khi sử dụng hết hạn mức và quá thời hạn được quy định trước, người dùng cần nạp thêm tiền vào tài khoản để
có thể tiếp tục sử dụng các dịch vụ của OpenAI (Mỗi một lần gọi API key có chi phí $0.02.)



Hình 5: Hạn mức sử dụng và thời hạn sử dụng được hiển thị trong platform.openai.com/account/usage

3 Cài đặt chat bot

Để quá trình cài đặt được đơn giản và nhanh chóng, chúng tôi sử dụng ngôn ngữ lập trình Python để xây dựng chat bot trên nền tảng ứng dụng web local.

3.1 Các thư viện được sử dụng

3.1.1 openai

Openai là một gói thư viện (package) được viết cho ngôn ngữ lập trình Python, chỉ hỗ trợ các phiên bản Python từ 3.7.1 trở đi. Thư viện này gồm có các hàm, phương thức và lớp đối tượng được xây dựng sẵn để tương tác với các API của OpenAI như GPT-3, DALL-E và Codex, cũng như các công cụ để xác thực truy cập, định dạng đầu ra/vào và xử lý lỗi.

Trong thư viện này, phương thức gọi được sử dụng là `openai.ChatCompletion.create()`.

3.1.2 gradio

Thư viện gradio hỗ trợ xây dựng các ứng dụng tương tác cơ bản trên nền tảng web để vận dụng và chia sẻ các mô hình máy học, API và luồng công việc khoa học dữ liệu một cách tiện lợi, nhanh chóng và dễ dàng.

3.2 Cấu trúc mã nguồn

Mã nguồn của ứng dụng demo này được đặt trong một môi trường Python ảo với các thư viện cần thiết được cài đặt sẵn và script để kích hoạt.

Khi chạy mã nguồn, cửa sổ terminal sẽ cung cấp đường dẫn đến trang web demo tương ứng trên server .

Tổng quan, mã nguồn cài đặt các công việc chính:

1. Khai báo các thư viện sử dụng
2. Xác thực API key bằng Key đã được tạo trong đề mục 2.
3. Thêm dữ liệu đầu vào: Hàm `add_request(history, request)` nhận yêu cầu mới của người dùng dưới dạng chuỗi văn bản và chèn yêu cầu này vào cuối biến danh sách `history` với định dạng mà hàm sinh kết quả hiểu được.
4. Thêm dữ liệu đầu ra: Hàm `add_response(history, request)` nhận câu trả lời của AI dưới dạng chuỗi văn bản và chèn chúng vào cuối biến danh sách `history` với định dạng mà hàm sinh kết quả hiểu được.
5. Hàm sinh kết quả `generate_response` nhận vào danh sách lịch sử hội thoại và nội dung người dùng nhập vào, cập nhật lịch sử hội thoại rồi sau đó gọi hàm `openai.ChatCompletion.create()` để tạo câu trả lời. Hàm này duyệt qua toàn bộ lịch sử hội thoại được lưu trong `history` rồi tạo ra câu trả lời cho yêu cầu cuối cùng trong lịch sử hội thoại được sử dụng, đồng thời có thể được tùy chỉnh theo các tham số như model (mô hình), temperature (chỉ số ngẫu nhiên của kết quả)... Kết quả sinh ra được cập nhật trong lịch sử hộp thoại để sử dụng cho lần gọi tiếp theo.
6. Hàm `chat` nhận dữ liệu được nhập vào UI và dữ liệu phát sinh trong quá trình chạy ứng dụng (chính là lịch sử hội thoại), gọi các hàm đã giải thích ở trên, rồi lưu vào biến lưu trữ để ứng dụng xử lý tiếp.
7. Hàm tạo UI với khung nhập liệu, khung hiển thị lịch sử hội thoại. Khi kích hoạt, chúng sẽ gọi hàm `chat` với tham số là các biến lưu dữ liệu .

3.3 Nhận xét

Chức năng

- Có thể vận hành như 1 chat bot cơ bản, trả lời được đa số các câu hỏi đưa ra một cách rõ ràng, mạch lạc với độ tin cậy cao.
- Kết quả được trình bày thành trên đoạn, nhiều dòng.
- Định dạng riêng cho code.
- Có thể nhớ lịch sử hội thoại để tiếp tục nhận phản hồi, tiếp tục chủ đề được đề cập đến mà không bị ngắt quãng.

Khuyết điểm

- Thỉnh thoáng sẽ xuất hiện lỗi ngữ pháp, câu cú.
- Chỉ có duy nhất một khung chat.
- Không thể hủy quá trình sinh câu trả lời giữa chừng.

4 Tham khảo

[1] *Introducing ChatGPT*. OpenAI. Truy cập ngày 12 tháng 3 năm 2023. https://openai.com/blog/chatgpt

[2] *Definition: GPT-3*. TechTarget. Truy cập ngày 12 tháng 3 năm 2023. https://www.techtarget.com/searchenterpriseai/definition/GPT-3

[3] *What is deep learning?*. IBM. Truy cập ngày 12 tháng 3 năm 2023. https://www.ibm.com/topics/deep-learning

[4] *What are neural networks?*. IBM. Truy cập ngày 12 tháng 3 năm 2023. https://www.ibm.com/topics/neural-networks

[5] *openai 0.27.2*. The Python Package Index. Truy cập ngày 12 tháng 3 năm 2023. https://pypi.org/project/openai/

[6] *gradio 3.20.1*. The Python Package Index. Truy cập ngày 12 tháng 3 năm 2023. https://pypi.org/project/gradio/

[7] *API reference*. OpenAI. Truy cập ngày 12 tháng 3 năm 2023. https://platform.openai.com/docs/api-reference/

[8] Hamilton, A.C. (2022) *How to Create an Advanced Chatbot: A Comprehensive Guide to Using Open AI's Chat GPT*. Digital Age Media.