

21CNTThuc - Xử lý ngôn ngữ tự nhiên ứng dụng

Nhóm 5 - Teach me

19127620 - Nguyễn Văn Vũ Tường

20127349 - Trần Quốc Thuận

Đề tài “H.11” - AI English Tutor

Các ưu điểm của đề tài:

Sử dụng được mô hình ngôn ngữ lớn GPT-3.5 Turbo của OpenAI để đưa ra các chỉnh sửa cho các câu văn tiếng Anh, giúp cải thiện trình độ tiếng Anh của người dùng. Mô hình GPT-3.5 Turbo thỏa mãn phạm vi của đề tài và có chi phí sử dụng thấp.

Sử dụng được công nghệ chuyển đổi văn bản thành giọng nói của ElevenLabs để củng cố trải nghiệm học tập bằng âm thanh. ElevenLabs có âm thanh tự nhiên.

Sử dụng được thư viện Streamlit để tạo ra giao diện người dùng đẹp và dễ sử dụng.

Các khuyết điểm của đề tài:

Mô hình ngôn ngữ phi trạng thái và không có ký ức về các tương tác trong quá khứ. Mọi tương tác đều hoàn toàn độc lập với các tương tác trước đó.

Có thể cung cấp cho mô hình ngữ cảnh của các tương tác trước đó bằng cách gửi toàn bộ cuộc hội thoại hoặc một bản tóm tắt trở lại mô hình. Tuy nhiên, việc này sẽ sử dụng nhiều token hơn và bất kỳ tham chiếu nào đến dữ liệu bị mất sẽ gây ra ảo giác và thông tin không chính xác.

ElevenLabs có chi phí sử dụng cao và dễ dàng hết dung lượng.

Phần cuối có thêm một số khuyết điểm. Tuy nhiên, những khuyết điểm này nằm ngoài phạm vi của đề tài gốc nên không được thêm vào phần này.

Tóm tắt phương pháp của đề tài:

Đề tài sử dụng phương pháp “few-shot” prompting, cụ thể là “one-shot” prompting.

Ngoài hướng dẫn và dữ liệu đầu vào, prompt còn chứa thêm một minh họa. Minh họa này bao gồm một cặp dữ liệu đầu vào và đáp án.

Dựa vào minh họa đó, mô hình học rồi đưa ra câu trả lời thỏa mãn yêu cầu của đề tài: đưa ra các chỉnh sửa cho các câu văn tiếng Anh, giúp cải thiện trình độ tiếng Anh của người dùng.

Các vấn đề đã làm được (những gì đã làm khác so với code mẫu ban đầu):

Các API của OpenAI và ElevenLabs đã được cập nhật sau khi code mẫu ban đầu được tạo ra. Nhóm đã cập nhật mã nguồn để chạy được các API này.

ElevenLabs dễ dàng hết dung lượng. Vì vậy, nhóm đã thêm vào mã nguồn Google Text-to-Speech để có thể dễ dàng minh họa tính năng âm thanh.

Nhóm có thêm vào một số tính năng. Những tính năng này đơn giản và không mạnh bằng tính năng chính của code mẫu ban đầu. Tuy nhiên, chúng giúp người dùng có một trải nghiệm học tiếng Anh phong phú hơn. Những tính năng này là:

- Đưa ra một lời khuyên cho việc học tiếng Anh.
- Giải thích một lỗi mà nhiều người thường mắc phải khi sử dụng tiếng Anh.
- Giải thích một từ hoặc cụm từ, sau đó sử dụng nó trong một câu văn.
- Diễn đạt một ý tưởng theo một cách khác.
- Phát âm một câu văn.

Nói thêm về khuyết điểm:

Nhóm muốn thêm vào một số tính năng với mục đích là giúp người dùng chuẩn bị cho kỳ thi viết của IELTS hoặc TOEIC. Tuy nhiên, nhóm không có đủ kinh nghiệm và thời gian để thực hiện điều này (tạo cấu trúc prompt phù hợp, thử nghiệm các prompt khác nhau, kiểm tra kết quả).

Hơn nữa, những kỳ thi này quan trọng và tốn nhiều tiền. Vì thế, người dùng không nên dựa vào mô hình ngôn ngữ mà nên tham khảo các giáo viên có trình độ chuyên môn.

Mô hình GPT-3.5 Turbo có những hạn chế nhất định. Mô hình có khả năng tạo ra các phản hồi dựa trên các tập dữ liệu khổng lồ. Tuy nhiên, nó không hiểu văn bản ở cấp độ con người. Nó có thể cung cấp thông tin nghe hợp lý nhưng không chính xác hoặc không nắm bắt được các vấn đề phức tạp.

Khi đánh giá các câu văn và các bài luận trả lời bài kiểm tra viết của IELTS hoặc TOEIC, mô hình dường như không thể tuân theo các đề mục của IELTS hoặc TOEIC. Nó cũng không thể đánh giá chính xác nội dung và tổ chức của các bài luận.

Vì vậy, những nhiệm vụ quan trọng như vậy nên được thực hiện bởi một giáo viên IELTS hoặc TOEIC có trình độ chuyên môn. Họ có thể cung cấp phản hồi thích hợp và tính điểm chính xác.

Tương lai và tiềm năng:

Việc sử dụng mô hình ngôn ngữ lớn để hỗ trợ việc học tiếng Anh đã trở thành một hiện thực.

Tính đến tháng 2 năm 2023, Duolingo là ứng dụng học ngôn ngữ phổ biến nhất thế giới dựa trên số lượt tải xuống hàng tháng, với gần 13.4 triệu người dùng tải xuống trong tháng đó. Dịch vụ Duolingo Max sử dụng mô hình GPT-4 để giải thích câu trả lời và tập luyện giao tiếp.

Các mô hình ngôn ngữ lớn sẽ ngày càng được cải thiện và dễ tiếp cận hơn. Trong tương lai, có thể sẽ xuất hiện các mô hình ngôn ngữ lớn hoặc các mô hình chuyên môn có khả năng chấm điểm các bài kiểm tra viết của IELTS hoặc TOEIC. Để làm được điều đó, sẽ cần sự quan tâm của các tổ chức liên quan để tạo các bộ dữ liệu, huấn luyện và kiểm tra các mô hình.

Ngày 23 tháng 4:

- “one-shot” prompting và “few-shot” prompting khác nhau. “one-shot” prompting mới đúng. Xem slide tuần 5 trang 45.
- Thư viện Python chọn version được.
- Tính năng IELTS thì cho vào khoảng 4 band làm “zero-shot” prompting. Không đủ token để làm “few-shot” prompting.