

Nội dung bài giảng

<https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/>

Tactics để đưa ra câu prompt hiệu quả:

- Sử dụng delimiter và cho ChatGPT biết delimiter ở đâu để phân tách câu
- Provide thêm role của mình là gì
- Provide thêm format đầu ra
- Provide thêm điều kiện đầu ra, ví dụ:
 - o Yêu cầu đầu ra không được quá 50 chữ
 - o Tập trung vào khía cạnh nào đó trong message thôi
- Few shot prompting: Đưa ra ví dụ và yêu cầu bot viết theo
- Give the model time to think, đưa ra instruction: Vì mô hình là thuộc dạng predict word by word, nên ta chỉ ra được instruction cho model làm theo sẽ cho nó đi đúng hướng hơn và có thêm thời gian để predict word.

Ứng dụng: Summarizing:

- Summarize tập trung vào khía cạnh gì đó
- Paraphrase câu review để gửi cho một bên liên quan khác.

Ứng dụng: Inferring:

- Suy luận và đánh giá một câu comment của KH
- Cho biết được emotion của câu trả lời ntn

Ứng dụng: Transforming:

- Language translation
- Grammar, spelling checking
- Tone adjustment
- Format conversion

Ứng dụng: Automated Reply Service

Ứng dụng: Chatbot:

- Đối với LLM, thì để tạo ra một Chatbot ta cần phải lưu lại message trước để bot mỗi lần chat có thể keep track được câu chuyện hiện tại

Now, that should give you an idea. Before we start, here are some more things you should know.

 - OpenAI's ChatGPT has a token limit of 4096 (1 token = ~ 4 characters)
 - This limit includes both your prompt and the response (completion tokens) that is returned from ChatGPT
 - Anything that has to be generated post this token limit will be ignored abruptly without even a single warning
 - So, ideally we should be keeping our prompt token limit with 2048 (half the limit) and leave the rest for ChatGPT's completion. This buffer changes based on your use case.

ChatGPT OpenAI chỉ có token limit 4096 nên cuộc đối thoại với chatGPT chỉ giới hạn ở 4096*4 ký tự, bao gồm những câu hỏi và trả lời ngắn của cả người dùng và bot.

Note:

- *ChatGPT không giống với Google Search là ở chỗ, tất cả những gì người dùng nhập đều là dữ liệu để ChatGPT search kiếm kết quả, đôi khi trong câu message thì người dùng chỉ yêu cầu ChatGPT tuân theo những nguyên tắc đầu ra, thái độ trả lời, và cách trả lời. ChatGPT sẽ đọc và hiểu message người dùng hơn là Google Search.*
- *API openAI GPT: Mỗi lần call API thì memory sẽ được tạo mới, nên khi call phải nhập vào luôn message history.*

Resource bài giảng:

Câu message 1:

Bạn đang cần tìm một đồng hồ thông minh nhằm theo dõi sức khỏe, tập luyện thể thao với mức giá dễ tiếp cận cùng thiết kế đẹp mắt? Vậy Redmi Watch 3 Active sẽ là một sự lựa chọn cực kỳ phù hợp cho bạn với màn hình 1.83 inch hiển thị nhiều nội dung, khả năng theo dõi hơn 100 chế độ tập luyện, đa dạng tính năng sức khỏe và nhiều công cụ thông minh hỗ trợ người dùng. Redmi Watch 3 Active được đóng gói trong một hộp đựng nhỏ gọn, đi kèm với cáp sạc nam châm, sách hướng dẫn sử dụng. Chiếc đồng hồ thông minh đến từ Redmi có thiết kế hiện đại, thanh lịch với kiểu dáng mặt đồng hồ vuông hiện cũng được Apple sử dụng cho các sản phẩm Apple Watch của hãng. Mặt đồng hồ vuông sẽ giúp thiết bị hiển thị được nhiều nội dung hơn so với mặt tròn truyền thống và mang lại cảm giác sắc sảo, thanh thoát. Thiết bị có kích thước 46.94 x 38.88 x 10.94 mm và nặng 41.7 gram (có dây đeo). Sản phẩm sẽ phù hợp với kích thước cổ tay từ 13.5 - 20 cm và cũng cho cảm giác đeo thoải mái, dễ chịu, không bị mỏi hay quá nặng tay khi dùng liên tục. Redmi Watch 3 Active Vỏ đồng hồ hình chữ nhật cổ điển và đơn giản có thân được xử lý NCVM, với lớp hoàn thiện bằng kim loại làm tăng thêm tính thẩm mỹ. Nhìn thoáng qua, chiếc đồng hồ này có ngoại hình khá giống với Apple Watch. Thiết bị cũng hỗ trợ kháng nước 5ATM, đủ để người dùng sử dụng máy khi tắm, bơi lội ở vùng nước nông, đi mưa. Các cạnh của đồng hồ được hoàn thiện bóng tạo điểm nhấn nổi bật. Nhờ các cạnh cong nhẹ, thiết bị không chỉ duy trì tổng thể vuông vức mà còn tránh làm máy trở nên quá góc cạnh, tạo cảm giác thoải mái và không gây khó chịu khi đeo trên tay. Ở cạnh phải có một nút bấm vật lý, giúp bạn theo dõi nhanh một số chức năng nhằm hỗ trợ người dùng trong các trường hợp không tiện chạm vào màn hình như khi đang đeo găng tay, đang bơi hay khi tay bị ướt,... Mặt lưng của thiết bị là nơi đặt các cảm biến, đây chính là khu vực quan trọng nhất giúp ghi lại tình trạng sức khỏe và trạng thái tập luyện của người dùng. Phần mặt dưới được hoàn thiện nhám để giúp đồng hồ bám chắc hơn khi đeo, các cạnh cũng được bo cong mềm mại nhằm tạo sự thoải mái, tránh bị cấn, gây khó chịu. Redmi Watch 3 Active Một nâng cấp lớn của Redmi Watch 3 Active so với các thế hệ tiền nhiệm là màn hình kích thước lên đến 1.83 inch. Màn hình này sử dụng tấm nền LCD độ phân giải 240 x 280 Pixels và được thiết kế cong 2.5D cho cảm giác vuốt từ các cạnh rất dễ tay và được bảo vệ bởi kính cường lực giúp người dùng an tâm hơn. Màn hình lớn cho phép người dùng xem nhiều nội dung hơn, dễ dàng thao tác và hạn chế bị bấm nhầm. Tuy nhiên, do các viền màn hình hơi dày nên sẽ làm cho tổng thể máy khá lớn, đồng thời, nếu bạn để hình nền là các màu

sáng sẽ bị lộ nhược điểm này ra, trường chưa được thẩm mỹ cho lắm. Redmi Watch 3 Active Về trải nghiệm thực tế thì màn hình này có khả năng hiển thị sắc nét khi so với các đối thủ khác, độ sáng màn hình tối đa 450 nits ở mức vừa phải, không cao như trên các đồng hồ dùng tấm nền AMOLED nhưng vẫn đủ để xem nội dung khi đang chạy xe ngoài trời nắng. Redmi Watch 3 Active Phần dây đeo đi kèm Redmi Watch 3 Active sử dụng chất liệu silicon mềm mại. Đây là loại chất liệu được mình đánh giá là rất phù hợp cho các sản phẩm đồng hồ thông minh và vòng đeo tay thông minh giá rẻ vì đặc trưng bền bỉ, thoải mái và ít gây kích ứng da khi đeo. Redmi Watch 3 Active Phần chốt cài dây đeo vào đồng hồ cũng rất cứng cáp nhưng vẫn đảm bảo rằng người dùng có thể dễ dàng tháo, lắp và thay thế dây đeo. Khóa cài trên dây đồng hồ cũng rất chắc chắn, đảm bảo thiết bị không bị rơi ra khi vận động mạnh và được thiết kế khéo léo để ẩn đi phần dây thừa. Redmi Watch 3 Active Để sử dụng Redmi Watch 3 Active, người dùng chỉ cần kết nối với điện thoại qua Bluetooth và ứng dụng Mi Fitness. Sau đó, chiếc đồng hồ này có thể cập nhật các dữ liệu sang điện thoại, đồng thời, còn hiển thị thông báo, nhận hoặc từ chối cuộc gọi, cho phép điều khiển một số tính năng trên điện thoại,... Redmi Watch 3 Active Chiếc đồng hồ từ Redmi có khả năng theo dõi hơn 100 chế độ tập luyện thể thao bao gồm 10 chế độ thể thao chuyên nghiệp. Nhờ vào các cảm biến tích hợp và thuật toán thể thao chuyên nghiệp mà thiết bị có thể theo dõi và phân tích cơ thể bạn khi tập luyện, từ đó, bạn có thể xây dựng lộ trình tập phù hợp, xem xét tính hiệu quả của việc tập luyện,... Redmi Watch 3 Active Bên cạnh đó, thiết bị cũng hỗ trợ theo dõi nhiều chỉ số sức khỏe phổ biến như: Độ oxy trong máu, nhịp tim, giấc ngủ,... Từ đó giúp bạn nắm được thể trạng hiện tại của bản thân, phát hiện sớm một số nguy cơ sức khỏe. Redmi Watch 3 Active Thiết bị cũng được trang bị mic giúp bạn nghe, gọi trực tiếp trên đồng hồ. Đây là một tính năng rất hữu ích khi bạn đang tập luyện hoặc khi không tiện lấy điện thoại từ trong túi ra để nghe, gọi. Ở chế độ sử dụng thông thường, đồng hồ có thể hoạt động được khoảng 12 ngày và với các tác vụ nặng thì thiết bị vẫn trụ được 8 ngày. Đây là thời lượng rất tốt, giúp những người hay quên sạc đồng hồ như mình yên tâm hơn khi sử dụng và có thể mang sản phẩm theo trong những chuyến du lịch dài ngày. Redmi Watch 3 Active Tóm kết lại thì mình nhận xét Redmi Watch 3 Active là một sản phẩm đồng hồ thông minh có thiết kế đẹp mắt, hiện đại, màn hình lớn cùng đa dạng các tính năng tập luyện, sức khỏe hữu ích. Thiết bị hiện đang được bán tại CellphoneS với mức giá 1.29 triệu đồng và có hai tùy chọn màu sắc là đen và xám. Các bạn có thể ghé ngay hệ thống ``

- Đánh giá khen hoặc chê
- Suy luận những thông tin khác
- Dịch ngôn ngữ
- Automated Reply mail service dựa vào feedback người dùng

Message 2:

``mulan 2020 là do phụ nữ đạo diễn. nhấn mạnh sức mạnh, vị trí của ng phụ nữ. nhưng kết ra thì là phụ nữ muốn có sức mạnh, muốn có vị trí đứng chỗ đáng có của mình thì phải sinh ra ở trước vạch đích, phải là người được chọn. ko thì bạn chỉ là con nữ phụ vô dụng, ko ai quan tâm như nv con chị mulan trong phim. so với bản 98 thì các giá trị dc lồng ghép vào bay màu hết. đây là trường ca sử thi về 1 anh hùng trong truyền thuyết go brrrr chứ chả phải câu chuyện về 1 ng phụ nữ từ chân yếu tay mềm đến trở người hùng trung hoa nữa. Cũng ko trách dc. 4 thg viết kịch bản thì tây lông cả 4 thg. mà tây lông 2020 mà. nữ quyền geishit ko nào thì dễ chứ giá trị văn hoá nhân văn thì chịu.``

Thêm một lí do nữa: Mk nghĩ Disney đã đưa ra thêm 1 sai lầm nữa khi quyết định ngôn ngữ của phim là tiếng anh trong khi diễn viên toàn là người Trung. Việc không được thành thạo trong tiếng anh cộng thêm câu thoại hơi hợt đã thực sự ảnh hưởng đến diễn xuất của các diễn viên, đặc biệt là Cung Lợi (Thôi nào Disney, bh là 2020 rồi mọi người đọc phụ đề suốt và Parasite cũng vừa thắng Oscar cho phim nói tiếng ngoài xuất sắc nhất đấy thôi)

Ẩn bớt

Ai đã coi Mulan bản gốc sẽ biết Mulan là một nhân vật có chiều sâu về mặt cảm xúc cũng như suy nghĩ, cô cũng thể hiện rất nhiều cảm xúc đủ loại xuyên suốt câu chuyện! Nghe Mulan live action do LDP đóng thì mình biết ngay Mulan này chả khác gì 1 bức tượng di động cả! Mới thấy diễn viên thôi đã không muốn coi rồi:((((

Bài Tập và Project

Bài tập 1:

Message: Vấn đề biến đổi khí hậu đang ngày càng trở nên nghiêm trọng và cần phải có biện pháp giảm thiểu tác động của con người lên môi trường. Một trong những cách quan trọng để đối phó với vấn đề này là tăng cường sử dụng năng lượng tái tạo, như năng lượng mặt trời và gió.

Câu trả lời mẫu: Để giảm tác động của con người lên môi trường do biến đổi khí hậu, cần tăng cường sử dụng năng lượng tái tạo như năng lượng mặt trời và gió.

Nhiệm vụ của người học: Hãy tạo một câu prompt để tạo ra một câu trả lời tương tự với câu trả lời mẫu trên.

Bài tập 2:

Message: Tiệm bánh Fr*unch P*stry - Mình đặt chiếc bánh như này hôm 19/8 và đến đây mình đưa mẫu rồi hỏi tiệm là "dạ tiệm có vẽ được mẫu như này không ạ, có gì trang trí thêm hoa quả cho em nhé ạ". Mình hỏi xong tiệm đáp là "dạ bên mình có vẽ được bạn nhé, bao giờ bạn lấy bánh nhỉ" thì mình kêu 20/8 (tức là hôm nay) mình lấy thì tiệm cũng ok.

Đến chiều hôm nay mình nhờ người đi lấy bánh hộ thì bánh mang về lại như này mà hết tổng cộng 250k. Mình đã gọi cho tiệm hỏi thì tiệm nói với giọng rất là thái độ như kiểu chửi ngược lại mình

vậy. Tiệm kêu tiệm vẽ được mà mang về lại có mỗi cái bánh trắng tinh rồi mảnh giấy in hình chứ không có 1 cái gì thêm.

Cách làm ăn của tiệm rất là chán, tiệm còn sai cả tuổi chứ không mỗi gì bánh. Mình đã gọi để hỏi thì lại không xin lỗi mà còn thái độ ngược với khách hàng.

Hãy áp dụng các ứng dụng: Inferring, Summarizing, Transforming và Automated Reply Mail. Đây là câu hỏi mở, nhiệm vụ của Châu là liệt kê câu prompt ứng với từng ứng dụng để hỏi với message review ở trên.

Project 1:

Build một con Bot về Order hàng, sau khi người dùng chốt đơn thì sẽ xuất ra một mẫu json chứa thông tin đơn hàng bao gồm:

- Tên người mua
- Địa chỉ giao tới
- Thông tin loại và số lượng sản phẩm
- Tổng giá cả sản phẩm
- Ngày giờ đặt hàng
- Hình thức thanh toán

Ví dụ: Dùng ChatGPT tạo ra một con bot bán Pizza, với menu như sau:

- pizza loại 1: hải sản, bánh dày 2mm, bán kính 15cm, giá 3.5\$
- pizza loại 2: xúc xích, thịt viên, dày 2mm, bán kính 15cm, giá 3\$
- pizza loại 3: thập cẩm, dày 2mm, bán kính 17cm, giá 4.5\$

Trong quá trình chat, thì Bot cần phải chủ động đưa ra câu hỏi về những thông tin KH và menu để cho KH trả lời.

Discussion Answer:

- **Role “system” có thực sự cần thiết?**
 - Theo bài, <https://community.openai.com/t/separate-chatcompletion-api-calls-for-system-and-user/297732/14>
người dùng có nhắc đến vấn đề dùng message system ở từng đoạn API call. Nhưng vấn đề sẽ nằm ở việc take cost của OpenAI trên token
 - Họ prevent việc ChatGPT sẽ trả lời out of scope bằng cách tạo ra một cuộc đối thoại giữa user và assistant: <https://community.openai.com/t/how-to-prevent-chatgpt-from-answering-questions-that-are-outside-the-scope-of-the-provided-context-in-the-system-role-message/112027/25>
 -
 - Ngoài ra, một số người bảo việc đặt system message ở cuối câu cho những dạng “Act as a ...” sẽ tránh người dùng cố ý thay đổi cách nói chuyện của Bot trong khi sử dụng.

<https://community.openai.com/t/the-system-role-how-it-influences-the-chat-behavior/87353/2>

Nhưng chung quy lại thì role system mà OpenAI build là optional. Nhưng đối với những người dùng engineering, họ tận dụng cái tên là “system” khác biệt và tự đặt ra rule cho system để điều phối cuộc nói chuyện.

- **Việc sử dụng delimiter khác thay vì triple backtick ``` có thay đổi gì không?**

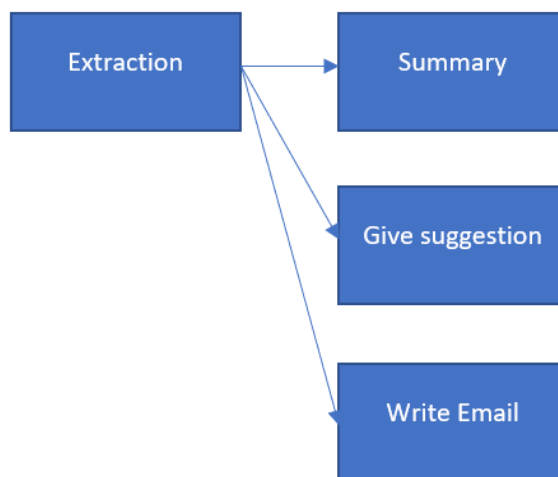
Theo ChatGPT trả lời thì: triple backtick là phân vùng code ở Markdown (.md file) cho nên việc GPT-3 đã học nhiều và nhận biết được triple backtick là vùng độc lập về codeblock hoặc information gì đó là đúng. Tuy nhiên không có nhiều dẫn chứng cho thấy việc dùng triple backtick sẽ tốt hơn dùng delimiter khác, chỉ cần đảm bảo là message viết ra dễ hiểu, rõ ràng.

OpenAI đề xuất các delimiter như là: ``` , """" , < > , <tag> </tag> Có thể các dấu delimiter trên được học nhiều trên tập dữ liệu training của mô hình GPT hiện tại của OpenAI.

- **Summary vs Extraction trong ứng dụng Prompt Engineering:**

- Theo như tên gọi, summary là bước tổng hợp và cho ra output ngắn gọn về thông tin, trong khi extraction chỉ là giải nén vấn đề, như trang DeepLearning.ai đã ghi:

“you can also ask it to extract information rather than summarize it. So here's a prompt that says you're tasked to extract relevant information to give feedback to the shipping department. And now it just says, product arrived a day earlier than expected without all of the other information, which was also helpful in a general summary, but less specific to the shipping department if all it wants to know is what happened with the shipping.”



Thì nếu Châu bảo model extraction, thì sẽ hướng dẫn mô hình đi tới các application khác chính xác hơn, do mô hình đã được học trên data workflow như trên.

Còn nếu Summary thì đã bao gồm luôn bước extraction và tới summary.

1

2

3