**TRƯỜNG ĐẠI HỌC KINH TẾ QUỐC DÂN**

**VIỆN CÔNG NGHỆ THÔNG TIN VÀ KINH TẾ SỐ**

**----o0o----**



**BÁO CÁO THỰC TẬP CHUYÊN ĐỀ**

**TÊN ĐỀ TÀI**

***Cá nhân hóa ChatGPT với dữ liệu đề án tuyển sinh ĐH kinh tế quốc dân***

***Giảng viên hướng dẫn*  : ThS. Phạm Thảo**

***Họ và tên* : Lù Chín Trình**

***Mã sinh viên* : 11208209**

***Lớp* : Khoa học máy tính 62**

**HÀ NỘI, 2023**

**MỤC LỤC**

[LỜI NÓI ĐẦU 2](#_Toc151968445)

[I. TỔNG QUAN VỀ ĐỀ TÀI. 3](#_Toc151968446)

[1. Giới thiệu chung. 3](#_Toc151968447)

[1. 1. Khái niệm LLMs 3](#_Toc151968448)

[1. 2. Giới thiệu ChatGPT 3](#_Toc151968449)

[2. So sánh giữa chatbot thông thường và chatbot dựa trên LLMs 3](#_Toc151968450)

[2. 1. Cách mà ChatGPT thay đổi một số nghiệp vụ 3](#_Toc151968451)

[3. Yêu cầu đặt ra. 3](#_Toc151968452)

[4. mục tiêu đề tài. 3](#_Toc151968453)

[5. phạm vi đề tài. 3](#_Toc151968454)

[II. GIỚI THIỆU VỀ CÁC CÔNG NGHỆ SỬ DỤNG. 4](#_Toc151968455)

[1. Langchain 4](#_Toc151968456)

[2. Django 4](#_Toc151968457)

[2. 1. Django là gì? 4](#_Toc151968458)

[3. Xampp 12](#_Toc151968459)

[3. 1. Xampp là gì 12](#_Toc151968460)

[3. 2. Đặc điểm của xampp 13](#_Toc151968461)

[3. 3. Xampp dùng làm gì 14](#_Toc151968462)

[3. 4. Cách cài đặt xampp trên windows 15](#_Toc151968463)

[3. 5. Hướng dẫn sử dụng xampp 18](#_Toc151968464)

[4. ChatGPT API 19](#_Toc151968465)

[5. Llama2 20](#_Toc151968466)

[5. 1. Khái niệm 20](#_Toc151968467)

[5. 2. Ưu điểm, nhược điểm 21](#_Toc151968468)

[III. PHÂN TÍCH THIẾT KẾ THỐNG. 24](#_Toc151968469)

[1. Mô tả hệ thống 24](#_Toc151968470)

[1. 1. Sơ đồ kiến trúc hệ thống 24](#_Toc151968471)

[1. 2. Sơ đồ chức năng hệ thống 25](#_Toc151968472)

[1. 3. Chức năng đối với khách hàng 26](#_Toc151968473)

[1. 4. Chức năng nối với với admin 30](#_Toc151968474)

[2. Phân tích hệ thống về mặt chức năng 33](#_Toc151968475)

[2. 1. Biểu đồ luồng dữ liệu (BLD) mức khung cảnh 33](#_Toc151968476)

[2. 2. Biểu đồ luồng dữ liệu (BLD) mức cảnh 34](#_Toc151968477)

[2. 3. Biểu đồ luồng dữ liệu (BLD) mức dưới đỉnh 35](#_Toc151968478)

[3. Phân tích hệ thống về mặt cơ sở dữ liệu 36](#_Toc151968479)

[3. 1. Xác định các thực thể 36](#_Toc151968480)

[3. 2. Sơ đồ thực thể liên kết 39](#_Toc151968481)

[3. 3. Mô hình quan hệ giữa các thực thể 41](#_Toc151968482)

[3. 4. Danh sách các bảng dữ liệu 42](#_Toc151968483)

[3. 5. Mô tả chi tiết các bảng 42](#_Toc151968484)

[4. Một số thuật toán trong phần mềm 46](#_Toc151968485)

[4. 1. Thuật toán tiền xử lí dữ liệu huấn luyện. 46](#_Toc151968486)

[4. 2. Thuật toán tải file lên CSDL. 47](#_Toc151968487)

[IV. SẢN PHẦM DEMO VÀ KẾT QUẢ. 48](#_Toc151968488)

[1. Giao diện sản phẩm 48](#_Toc151968489)

[1. 1. Giao diện đăng nhập, đăng ký. 48](#_Toc151968490)

[1. 2. Giao điện chính của Chat 49](#_Toc151968491)

[2. Báo cáo kết quả 51](#_Toc151968492)

[KẾT LUẬN 55](#_Toc151968493)

[TÀI LIỆU THAM KHẢO 56](#_Toc151968494)

LỜI NÓI ĐẦU

Trong thời gian gần đây, trong ngành trí tuệ nhân tạo đã nổi lên một chatbot tên là ChatGPT. Với sự xuất hiện của công nghiệp này đã kéo theo sự thay đổi trong cách vận hành của rất ngành nghề. Vậy liệu chatGPT là gì? ChatGPT được hiểu là hệ thống chatbot trí tuệ nhân tạo được tạo ra dựa trên khoa học công nghệ GPT-3. Đây là một mô hình ngôn ngữ lớn có khả năng tạo văn bản giống như đang trò chuyện với người thật chỉ với những từ khóa cơ bản. Nó sẽ có phần mô phỏng hơi giống với những cuộc trò chuyện tự động với khách hàng nhưng được ứng dụng công nghệ AI hiện đại. Điều này giúp mang đến cho người dùng những trải nghiệm trò chuyện, tìm kiếm thông tin trở nên thú vị hơn.

Hệ thống chương trình này có thể giúp bạn trả lời câu hỏi ở nhiều khía cạnh và lĩnh vực khác nhau, hỗ trợ bạn viết code, viết email, học tập và làm việc. ChatGPT hỗ trợ bạn tương tác với điện thoại di động và máy tính tốt hơn và thậm chí là có thể thay thế con người làm việc trong một số yêu cầu nhất định.

Kể từ khi được tung ra thị trường, công cụ này đã khiến nhiều người không khỏi ngạc nhiên về [những lợi ích và độ thông minh của ChatGPT](https://dienmaycholon.vn/kinh-nghiem-mua-sam/bat-mi-nhung-loi-ich-cua-viec-su-dung-cong-cu-chatgpt-ma-ban-nen-biet) đem lại. Những tác vụ trước kia khiến bạn mất nhiều thời gian tìm kiếm trên Google giờ đây ChatGPT sẽ giúp bạn đưa ra câu trả lời trong vòng vài giây. Đây được xem là mối nguy lớn có thể thay thế Google trong tương lai. Tuy nhiên nó vẫn còn nhược điểm khá lớn đó là độ chính xác trên thực tế không đồng đều.

Từ những ưu điểm của chatGPT đã kể trên thì em đã chọn được một ý tưởng là sẽ huấn luyện chatGPT dựa trên dữ liệu cá nhân cung cấp và có thể tận dụng sự sáng tạo của chat kết hợp với thông tin được cấp để chatGPT có thể có thể phục vụ cho người dùng, khách hàng… Và dưới đây là phần nghiên cứu của em về đề tài này.

1. TỔNG QUAN VỀ ĐỀ TÀI.
   1. Giới thiệu chung.
      1. Khái niệm LLMs

Large language model là một loại mô hình ngôn ngữ được đào tạo bằng cách sử dụng các kỹ thuật học sâu trên tập dữ liệu văn bản khổng lồ. Các mô hình này có khả năng tạo văn bản tương tự như con người và thực hiện các tác vụ xử lý ngôn ngữ tự nhiên khác nhau.

Một mô hình ngôn ngữ có thể có độ phức tạp khác nhau, từ các mô hình ***n-gram*** đơn giản đến các mô hình mạng mô phỏng hệ thần kinh của con người vô cùng phức tạp. Tuy nhiên, thuật ngữ Large language model” thường dùng để chỉ các mô hình sử dụng kỹ thuật học sâu và có số lượng tham số lớn, có thể từ hàng tỷ đến hàng nghìn tỷ. Những mô hình này có thể phát hiện các quy luật phức tạp trong ngôn ngữ và tạo ra các văn bản y hệt con người.

Dưới đây là một số ví dụ về LLM trong thực tế

GPT-3 (Generative Pre-training Transformer 3) – Đây là một trong những Mô hình Ngôn ngữ Lớn lớn nhất được phát triển bởi OpenAI. Nó có 175 tỷ tham số và có thể thực hiện nhiều tác vụ, bao gồm tạo văn bản, dịch thuật và tóm tắt.

* BERT (Bidirectional Encoder Representations from Transformers) – Được phát triển bởi Google, BERT là một LLM phổ biến khác đã được đào tạo trên một kho dữ liệu văn bản khổng lồ. Nó có thể hiểu ngữ cảnh của một câu và tạo ra các câu trả lời có ý nghĩa cho các câu hỏi.
* XLNet – LLM này được phát triển bởi Đại học Carnegie Mellon và Google sử dụng một cách tiếp cận mới để lập mô hình ngôn ngữ được gọi là “permutation language modeling”. Nó đạt được hiệu suất cao nhất trong các tác vụ ngôn ngữ, bao gồm tạo ngôn ngữ và trả lời câu hỏi.
* T5 (Text-to-Text Transfer Transformer) – T5, do Google phát triển, được đào tạo về nhiều tác vụ ngôn ngữ và có thể thực hiện chuyển đổi văn bản, như dịch văn bản sang ngôn ngữ khác, tạo bản tóm tắt và trả lời câu hỏi.
* RoBERTa (Robustly Optimized BERT Pretraining Approach) – Được phát triển bởi Facebook AI Research, RoBERTa là phiên bản BERT cải tiến, hoạt động tốt hơn trên một số tác vụ ngôn ngữ.

Tuy nhiên trong đề tài nghiên cứu lần này em chỉ tập trung vào nghiên cứu mô hình chatGPT đang được biết đến rộng rãi hiện nay.

* + 1. Giới thiệu ChatGPT

ChatGPT (Chat Generative Pre-training Transformer) là một chatbot do OpenAI phát triển dựa trên mô hình Transformer của Google. Đây là một AI (trí tuệ nhân tạo) giúp bạn tạo các cuộc trò chuyện tự động và trả lời các câu hỏi về nhiều chủ đề và lĩnh vực khác nhau.

ChatGPT cho phép bạn có thể trò chuyện như con người, và điều đặc biệt hơn là nó có khả năng tương tác ở dạng các cuộc hội thoại, đàm thoại tương tự như cách hai con người với nhau. Điều này làm cho hầu hết người dùng vô cùng ngạc nhiên

Hình 1. OpenAI


Hình 1. OpenAI

Để đào tạo GPT, các nhà khoa học OpenAI đã thu thập một lượng lớn văn bản do con người viết từ Wikipedia, bách khoa toàn thư, các tờ báo lớn và các nguồn thông tin công khai khác, lên tới hàng trăm triệu tài liệu.

Sau đó, họ làm sạch và chọn nội dung trước khi giao nó cho mô hình AI để đọc và đào tạo nhiều lần. Bằng cách đọc khối dữ liệu này, mô hình AI sẽ học được các lớp nghĩa đằng sau các từ và cụm từ, và càng đọc nhiều, nó càng xây dựng được nhiều lớp nghĩa hơn.

Chat GPT được coi là một trong những mô hình ngôn ngữ tiên tiến nhất hiện nay với khả năng tự động học và làm việc với các loại dữ liệu lớn. Nó được trang bị các tính năng nổi bật như tự động hóa hội thoại, trả lời câu hỏi, tạo câu trả lời tự động và cải thiện khả năng tự học của mô hình.

Mọi thắc mắc, băn khoăn của bạn về bất kỳ lĩnh vực nào sẽ được Chat GPT giải đáp trôi chảy, đầy đủ chỉ trong tích tắc. Chat GPT đảm nhận tất cả các nhiệm vụ sáng tạo và nghệ thuật như viết thơ, soạn nhạc, viết thư, thiết kế và thậm chí tạo hoặc gỡ lỗi trong lập trình.

* 1. So sánh giữa chatbot thông thường và chatbot dựa trên LLMs

+ Chatbot thông thường

Chatbot truyền thống được xây dựng dựa trên các nguyên tắc và luồng hội thoại được xác định từ trước. Trong hầu hết các trường hợp, người dùng được yêu cầu nhấn vào các nút bấm, tương ứng với các tùy chọn có sẵn

chatbot cũng có thể nhận diện được ý định của người dùng, thông qua những từ khóa đã được lập trình sẵn (keyword-based). Tuy nhiên, trong trường hợp này, câu hỏi mà người dùng đưa ra phải trùng khớp với câu hỏi trong kịch bản. Nếu có bất kỳ yêu cầu nào nằm ngoài kịch bản, chatbot hoàn toàn không thể đáp ứng được.

Việc phụ thuộc vào kịch bản hay các luồng hội thoại được xây dựng từ trước khiến chatbot tương tác rập khuôn, kém tự nhiên với người dùng. Thực tế, điều này còn dẫn đến tình trạng chatbot không hiểu hoặc hiểu sai, từ đó đưa ra các phản hồi thiếu chính xác đối với yêu cầu của người dùng.

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, thiết kế

Mô tả được tạo tự động

Hình 2. Sự khác nhau giữa chat dựa trên AI và chat thông thường

+ ChatGPT

chatGPT là chương trình phần mềm được xây dựng và triển khai với sự trợ giúp của công nghệ trí tuệ nhân tạo để hỗ trợ người dùng và tạo ra các tương tác tự nhiên giữa bot và người. Trí tuệ nhân tạo (AI) là công nghệ được quan tâm nhiều nhất hiện nay. Công nghệ này có khả năng xử lý dữ liệu lớn, với tốc độ cao, để từ đó đưa ra các phán đoán mà không cần có sự can thiệp của con người.

ChatGPT bắt đầu hoạt động ngay khi người dùng nhập thông tin đầu vào, truy vấn hoặc câu hỏi mà họ muốn được giải đáp. Sau đó, công nghệ hiểu ngôn ngữ tự nhiên (Natural Language Understanding – NLU) được áp dụng để phân tách văn bản và phát hiện ý định, thái độ, quan điểm của câu lệnh, từ đó rút ra ý nghĩa của thông tin đầu vào.

Tiếp đến, công cụ học máy (Machine Learning – ML) khớp ý nghĩa này với cơ sở dữ liệu để tìm nạp thông tin liên quan. Cuối cùng, chatbot xác định câu trả lời tốt nhất, phù hợp với truy vấn của người dùng và tạo ra phản hồi bằng văn bản.

Ảnh có chứa văn bản, Phông chữ, phần mềm, ảnh chụp màn hình

Mô tả được tạo tự động

Hình 3. Mô hình tổng quan chatbot

Đáng nói, theo thời gian, AI chatbot sẽ học thêm các ý định hay cách nói mới, từ chính những lần tương tác thành công hay thất bại. Dựa trên trải nghiệm thực tế của người dùng, nó có khả năng tinh chỉnh phản hồi để đáp ứng tốt hơn truy vấn ở những lần tiếp theo.

+ So sánh giữa chatbot thông thường và chatGPT

Dựa vào những phân tích bước đầu, có thể thấy AI chatbot vượt trội hơn so với chatbot truyền thống bởi khả năng nhận diện ý định chính xác, phản hồi linh hoạt và năng lực tự học theo thời gian.

Bảng so sánh dưới đây sẽ giúp bạn có được cái nhìn chi tiết hơn về các tính năng của AI chatbot trong tương quan với chatbot truyền thống.

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, số

Mô tả được tạo tự động

Hình 4. Sự khác nhau giữa chatbot truyền thống và chatbot dựa trên AI

* Khả năng thấu hiểu

Chatbot truyền thống chủ yếu hiểu ý định của người dùng thông qua các từ khóa có sẵn trong kịch bản. Tuy nhiên, không thể phủ nhận rằng một ý định/yêu cầu có thể được diễn đạt theo rất nhiều cách khác nhau, chưa kể đến việc sử dụng từ ngữ địa phương.

Ngoài ra, các yếu tố khác như thành ngữ, tiếng lóng hoặc lỗi chính tả cũng có thể cản trở quá trình phân tích và nhận diện của chatbot. Do đó, chatbot truyền thống gặp trở ngại lớn để hiểu và đáp ứng các yêu cầu ngoài kịch bản.

Trong khi đó, chatbot AI có thể hiểu ngôn ngữ tự nhiên của con người, không chỉ nhận diện ý định, yêu cầu của người dùng, mà còn phân tích được cảm xúc, quan điểm (tích cực/tiêu cực/trung lập) mà họ ẩn chứa trong câu lệnh. Đây là điểm ưu việt giúp AI chatbot tương tác tự nhiên và mang đến cho khách hàng những trải nghiệm khác biệt

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, thiết kế

Mô tả được tạo tự động

Hình 5. Ví dụ về chatbot thông thường khi hỏi những câu chatbot không biết

* Đáp ứng đúng yêu cầu

Chatbot truyền thống thường đưa ra các tùy chọn (dạng nút ấn) để dẫn dắt người dùng tìm hiểu thêm thông tin hoặc đến trang web khác/gọi điện cho tư vấn viên/điền biểu mẫu…. Trong trường hợp phản hồi dưới dạng văn bản, câu trả lời của các chatbot truyền thống cũng thường rập khuôn theo kịch bản, ít có sự đa dạng, linh hoạt theo từng bối cảnh cuộc trò chuyện.

Một điểm hạn chế khác dễ thấy, do không thể nhận diện chính xác ý định của người dùng, nên việc chatbot truyền thống đưa ra phản hồi sai yêu cầu là hoàn toàn có thể xảy ra.

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, phần mềm

Mô tả được tạo tự động

Hình 6. Sự khác nhau khi trả lời của chatbot AI và chatbot thông thường

Ngược lại, AI chatbot phản hồi chính xác hơn, thậm chí tự nhiên và mang tính cá nhân hóa cao hơn. Nhờ khả năng thấu hiểu ý định và quan điểm của người dùng, AI chatbot online có thể đưa ra câu trả lời phù hợp, tối ưu cho từng yêu cầu.

* Tự động nâng cấp

Đối với chatbot truyền thống, việc nâng cấp phải được quản trị viên hoặc lập trình viên thực hiện thủ công. Một chatbot điển hình có thể gặp khó khăn trong việc mở rộng quy mô để đáp ứng nhu cầu của doanh nghiệp đang phát triển. Nó thiếu các công cụ cần thiết để thích ứng với bất kỳ kịch bản tiềm năng nào hoặc sự gia tăng đột biến của số lượng các yêu cầu.

Đối với AI chatbot, việc nâng cấp có thể được tự động tiến hành thông qua các tương tác hàng ngày với người dùng. Trong quá trình tương tác, chatbot sẽ tự động học từ chính những lần tương tác đúng hoặc sai. Qua đó, AI chatbot ngày một làm giàu thêm vốn từ, củng cố khả năng hiểu ngôn ngữ tự nhiên, dần dần, thích ứng linh hoạt với những kịch bản, bối cảnh trò chuyện mới.

* Thấu hiểu đa ngôn ngữ

Khả năng tương tác của chatbot truyền thống thường chỉ hạn chế trong một ngôn ngữ được thiết lập từ đầu. Việc chatbot không thể trò chuyện bằng các ngôn ngữ khác đòi hỏi doanh nghiệp phải xây dựng một chatbot riêng cho người dùng nước ngoài hoặc chuyển đổi tùy chọn ngôn ngữ nhiều lần. Tất cả điều này yêu cầu rất nhiều nguồn lực, thời gian và chi phí.

Mặt khác, AI chatbot có khả năng trò chuyện bằng nhiều ngôn ngữ – điều này không chỉ đơn thuần là dịch thuật, mà nâng cao hơn còn là năng lực nhận thức được các yếu tố văn hóa vùng miền của người dùng cuối.

Đứng trước bài toán bản địa hóa một thương hiệu, doanh nghiệp phải đảm bảo rằng khách hàng có thể tương tác bằng ngôn ngữ mà họ cảm thấy thoải mái nhất. Do đó, chatbot đa ngôn ngữ là câu trả lời hiệu quả và tiết kiệm chi phí nhất để đáp ứng nhu cầu này. AI chatbot đa ngôn ngữ cho phép doanh nghiệp trò chuyện với người dùng bằng nhiều ngôn ngữ khác nhau, nâng cao mức độ tương tác và chuyển đổi.

Ảnh có chứa văn bản, ảnh chụp màn hình, Trang web, Website

Mô tả được tạo tự động

Hình 7. Có khả năng tự động update chuyển đổi ngôn ngữ

* Tiết kiệm chi phí

Để triển khai chatbot truyền thống, doanh nghiệp có thể tiết kiệm chi phí đầu tư ban đầu, song tốn nhiều chi phí để duy trì đội ngũ chăm sóc khách hàng xử lý các yêu cầu phức tạp, và nâng cấp hệ thống một cách thủ công. Ngoài ra, trong trường hợp doanh nghiệp muốn tương tác với khách hàng ở nhiều thị trường khác nhau, thì việc xây dựng chatbot ở nhiều ngôn ngữ cũng khiến chi phí tăng lên theo cấp số nhân.

Thay vào đó, giải pháp tối ưu về cả chất lượng và chi phí là AI chatbot. Tuy chi phí đầu tư ban đầu cao hơn, song loại chatbot này hiệu quả và tiết kiệm hơn về lâu dài, chưa kể đến khả năng sinh lời từ việc đem đến cho người dùng trải nghiệm thuận tiện, tự nhiên, đột phá.

Ảnh có chứa văn bản, ảnh chụp màn hình, danh thiếp, Phông chữ

Mô tả được tạo tự động

Hình 8. ChatGPT tiết kiệm chi phí xây dựng chatbot

* Khả năng tạo ra nội dung

ChatGPT có khả năng tự động tạo ra nội dung dựa trên thông tin đầu vào. Điều này có thể đem lại lợi ích cho các bài viết, các nội dung marketing hoặc quảng cáo trên các trang web. Với khả năng này, ChatGPT có thể tạo ra những câu chữ, đoạn văn, thậm chí cả bài báo hoàn chỉnh, giúp tiết kiệm thời gian và công sức của những người viết nội dung. Tuy nhiên, vẫn còn nhiều thách thức cần phải vượt qua để ChatGPT có thể tạo ra nội dung chất lượng cao và đáp ứng nhu cầu của người đọc.

Một trong những thách thức lớn nhất là việc ChatGPT không có khả năng hiểu rõ được ý nghĩa và mục đích của nội dung. Do đó, việc kiểm soát chất lượng của nội dung được tạo ra bởi ChatGPT là rất cần thiết. Ngoài ra, các thuật toán của ChatGPT cũng cần được cập nhật và đào tạo thường xuyên để đảm bảo chất lượng của nội dung được tạo ra.

Nếu được sử dụng đúng cách, khả năng tự động tạo ra nội dung của ChatGPT có thể giúp cho các doanh nghiệp, tổ chức hay cá nhân có thể tăng cường năng suất và giảm thiểu chi phí trong việc sản xuất nội dung. Tuy nhiên, cần lưu ý rằng ChatGPT chỉ là một công cụ hỗ trợ, và người sử dụng vẫn phải có kiến thức và kinh nghiệm để tạo ra nội dung chất lượng.

* Khả năng học tập và cải thiện kết quả

Trong lĩnh vực chatbot, khả năng học tập và cải thiện kết quả là một yếu tố quan trọng đối với hiệu suất và trải nghiệm của người dùng. Và ChatGPT đã chứng tỏ được sự vượt trội của mình trong việc này.

Các chatbot truyền thống thường phải được lập trình sẵn để có thể xử lý các nội dung cụ thể và cung cấp các câu trả lời thích hợp. Tuy nhiên, ChatGPT sử dụng mô hình học sâu để học tập thông tin từ hàng triệu tài liệu trên Internet, từ đó cải thiện khả năng xử lý ngôn ngữ tự nhiên và đưa ra các câu trả lời thích hợp cho từng câu hỏi.

Thêm vào đó, ChatGPT có thể cập nhật thông tin mới và tiếp tục học tập để nâng cao khả năng trả lời câu hỏi và cải thiện kết quả ngay cả khi hệ thống đã được triển khai và đang hoạt động. Điều này giúp cho ChatGPT có khả năng cung cấp trải nghiệm tốt hơn cho người dùng và giúp họ giải quyết các vấn đề một cách nhanh chóng và chính xác hơn.

Tóm lại, ChatGPT đã chứng tỏ được khả năng học tập và cải thiện kết quả của mình so với các chatbot truyền thống. Khả năng học tập liên tục của nó giúp cho ChatGPT trở nên ngày càng thông minh và có khả năng cung cấp trải nghiệm tốt hơn cho người dùng. Các công ty và tổ chức nên cân nhắc sử dụng ChatGPT để tăng cường hiệu quả và trải nghiệm của chatbot của họ.

* + 1. Cách mà ChatGPT thay đổi một số nghiệp vụ

ChatGPT - một ứng dụng đột phá trong lĩnh vực trí tuệ nhân tạo, đang mang lại sự thay đổi quan trọng và tích cực đối với nhiều ngành nghề hiện nay. Với khả năng linh hoạt và đa dạng, ChatGPT đã được áp dụng một cách sáng tạo và hiệu quả. Trong lĩnh vực dịch thuật, ChatGPT giúp kết nối các ngôn ngữ và văn hóa khác nhau, làm cho việc giao tiếp trở nên mượt mà và hiệu quả hơn. Đồng thời, nó còn là một trợ lý đắc lực trong hỗ trợ khách hàng trực tuyến, giải đáp nhanh chóng các thắc mắc và yêu cầu của khách hàng. Ở lĩnh vực y tế, ChatGPT có khả năng phân tích và tư vấn về thông tin y tế, đem lại lợi ích cho bác sĩ và bệnh nhân. Nó cũng là một trợ lý quan trọng trong giáo dục, hỗ trợ giáo viên và học sinh với kiến thức bổ ích và đa dạng. Với khả năng tư vấn về tài chính cá nhân và quản lý tài sản, ChatGPT hỗ trợ người dùng trong việc đưa ra quyết định thông minh và có lợi cho tương lai. Không chỉ dừng lại ở đó, ChatGPT còn có ứng dụng trong lĩnh vực viết bài viết chất lượng và quản lý sự kiện quan trọng.

Ứng dụng của ChatGPT trong lĩnh vực giáo dục đã mang lại một sự đột phá đáng kể trong việc tối ưu hóa quá trình học tập và giao tiếp trong môi trường học đương đại. Thành công của nó nằm ở khả năng tư vấn, hỗ trợ, và cái nhìn đa dạng mà nó mang đến cho học sinh, sinh viên, giáo viên và những người tham gia vào quá trình giảng dạy và học tập. Giúp cho sinh sinh trải nghiệm học tập được cá nhân hóa phù hợp với nhu cầu cụ thể của họ.

ChatGPT không chỉ là một công cụ giải quyết các vấn đề học thuật; nó còn là người bạn đồng hành thông minh, sẵn sàng trả lời mọi câu hỏi và hỗ trợ trong việc tìm kiếm kiến thức. Nó giúp học sinh và sinh viên dễ dàng tiếp cận tài liệu học tập, tìm hiểu kiến thức mới, và giải quyết những khó khăn mà họ gặp phải trong quá trình học tập.

Đối với giáo viên, ChatGPT trở thành trợ thủ đắc lực, đưa ra những lời khuyên về cách giảng dạy hiệu quả và tạo nội dung bài giảng hấp dẫn. Nó cung cấp sự tư vấn và hướng dẫn trong việc phát triển mô hình giảng dạy tiến bộ và phù hợp với nhu cầu của học sinh. Cung cấp cho giáo viên và giáo sư cơ hội chấm điểm công việc của sinh viên nhanh hơn bằng cách tự động phân loại nó khi nó được gửi qua giao diện chatbot thay vì phải xem qua tất cả bằng tay, Tiết kiệm cho họ thời gian quý báu trong một ngày vốn đã bận rộn.

Không chỉ có tính hữu ích trong việc học tập, ChatGPT còn thúc đẩy sự tiến bộ trong quản lý và tổ chức công việc của giáo viên và học sinh. Từ việc tạo tài liệu giảng dạy đến quản lý lịch trình học tập, nó đóng vai trò quan trọng trong việc tối ưu hóa quá trình học tập và giảng dạy.

Với sự ứng dụng đa dạng và tiềm năng lớn, ChatGPT đang tạo ra sự thay đổi tích cực trong lĩnh vực giáo dục, tạo điều kiện tốt hơn cho sự phát triển và học tập hiệu quả của mọi người.

* 1. Yêu cầu đặt ra.

Qua phần giới thiệu trên thì ta thấy chatGPT là môt công cụ hỗ trợ trong công việc rất mạnh mẽ, tuy nhiên ChatGPT vẫn tồn tại những vấn đề như sẽ trả lời sai câu hỏi nếu là câu hỏi liên quan đến một doanh nghiệp nào đó ít được đề cập đến trên internet. Vậy liệu có cách nào để giải quyết vấn đề này? Trong phần nghiên cứu này thì sẽ em nghiên cứu cách mà có thể huấn luyện ChatGPT hiểu được dữ liệu cá nhân nhân được cung cấp bởi chính các tổ chức.

* 1. mục tiêu đề tài.

***Xử lí dữ liệu đã thu thập:*** Thực hiện xử lí dữ liệu từ dữ đề án tuyển sinh năm 2019-2022 của trường đại hcoj kinh tế quốc dân, chuyển dữ liệu bảng sang dữ liệu văn bản để phục vụ công việc huấn luyện chatGPT.

***Tạo một mô hình ChatGPT tùy chỉnh***: Mục tiêu chính là phát triển một mô hình Chatbot dựa trên ChatGPT APIs, phục vụ trong công tác tuyển sinh của nhà trường, phục vụ cho việc hỏi đáp của học sinh, sinh viên.

* 1. phạm vi đề tài.

***Thu thập và xử lý dữ liệu***: Phạm vi đề tài bao gồm việc thu thập và xử lý dữ liệu đề án tuyển sinh của ĐH Kinh tế quốc dân năm 2019-2022 để sử dụng trong việc đào tạo mô hình ChatGPT.

***Tích hợp mô hình vào ứng dụng thực tế***: Sử dụng dữ liệu đã qua quá trình đào tạo để huấn luyện cho Chatbot đã được xây dựng dựa trên ChatGPT APIs, có thể giải đáp những thắc của hs, sv.

1. GIỚI THIỆU VỀ CÁC CÔNG NGHỆ SỬ DỤNG.
   1. Langchain

LangChain là một frameword cho việc phát triển ứng dụng được thúc đẩy bởi các mô hình ngôn ngữ. Nó cho phép các ứng dụng có những đặc điểm sau:

* **Hiểu bối cảnh**: kết nối một mô hình ngôn ngữ với các nguồn thông tin bối cảnh (hướng dẫn, ví dụ few-shot, nội dung để cơ sở hóa phản hồi, vv.)
* **Tư duy**: dựa vào một mô hình ngôn ngữ để tư duy (về cách trả lời dựa trên bối cảnh đã cung cấp, các hành động cần thực hiện, vv.)

Các giá trị chính của LangChain bao gồm:

* **Các thành phần (Components)**: các khái niệm để làm việc với các mô hình ngôn ngữ, cùng với một bộ sưu tập các triển khai cho mỗi khái niệm. Các thành phần có tính modular và dễ sử dụng, cho dù bạn đang sử dụng phần còn lại của khung việc phát triển LangChain hay không.
* **Chuỗi sẵn sàng sử dụng (Off-the-shelf chains)**: một sự sắp xếp cấu trúc của các thành phần để hoàn thành các nhiệm vụ cấp cao cụ thể

Chuỗi sẵn sàng sử dụng giúp bạn dễ dàng bắt đầu. Đối với các ứng dụng phức tạp, các thành phần giúp bạn dễ dàng tùy chỉnh các chuỗi hiện có và xây dựng các chuỗi mới.

* 1. Django
     1. Django là gì?

Django là phần mềm bạn có thể sử dụng để phát triển các ứng dụng web một cách nhanh chóng và hiệu quả. Hầu hết các ứng dụng web có một số chức năng phổ biến, như xác thực, truy xuất thông tin từ cơ sở dữ liệu và quản lý cookie. Các nhà phát triển phải viết mã cho chức năng tương tự vào mọi ứng dụng web mà họ viết. Django giúp họ làm việc dễ dàng hơn bằng cách nhóm các chức năng khác nhau thành một tập hợp lớn các mô-đun có thể tái sử dụng, được gọi là một khung ứng dụng web. Các nhà phát triển sử dụng khung web Django để sắp xếp và viết mã của họ hiệu quả hơn và giảm đáng kể thời gian phát triển web.

* + - 1. Tại sao chọn django?
* Tốc độ phát triển

Khung Django có cấu trúc hợp lý và dễ cài đặt cũng như học hỏi, vì vậy, bạn có thể bắt đầu sử dụng trong vòng vài giờ. Những nhà thiết kế Django đã tạo ra khung để nhanh chóng triển khai bất kỳ kiến trúc web nào bằng mã. Nhờ có khung hỗ trợ, quá trình phát triển diễn ra nhanh chóng và thiết kế nhìn gọn ghẽ, thực tiễn. Bạn có thể viết mã chỉ bằng vài dòng vì Django cung cấp cấu trúc sẵn sàng sử dụng cho một số tác vụ phát triển web phổ biến, chẳng hạn như:

* Xác thực người dùng
* Quản trị nội dung
* Sơ đồ trang web
* Nguồn cấp dữ liệu RSS
* Tiết kiệm chi phí

Django là một dự án Python miễn phí và nguồn mở với một cộng đồng hoạt động tích cực trong việc đánh giá và bảo trì phần mềm. Django Software Foundation là tổ chức phi lợi nhuận thúc đẩy và hỗ trợ việc sử dụng cũng như duy trì Django. Tổ chức này tổ chức các buổi gặp mặt, hội họp và sự kiện cộng đồng thường xuyên nhằm khuyến khích những nhà phát triển khác đánh giá và đóng góp cho dự án Django. Qua đó, khung web thu được có chất lượng cao với tính năng phong phú, miễn phí.

* Phổ biến

Hàng nghìn dự án nguồn mở và trang web tầm cỡ sử dụng Django như:

* Instagram
* Mozilla Firefox
* Pinterest
* National Geographic

Vì được sử dụng phổ biến, khung không ngừng phát triển và có cơ sở hạ tầng hỗ trợ vững chắc. Một lượng lớn các cá nhân và công ty cung cấp gói hỗ trợ miễn phí và có trả phí cho bất kỳ thách thức nào trong quá trình phát triển mà bạn có thể đối mặt khi đang sử dụng Django.

* + - 1. Django hoạt động như thế nào?

Bất kỳ ứng dụng web nào cũng đều được cấu tạo từ hai phần: mã máy chủ và mã máy khách. Khách thể hoặc đối tượng truy cập trang web có trình duyệt. Khi nhập URL vào trình duyệt của mình, họ sẽ gửi một yêu cầu tới máy chủ web chứa ứng dụng web đang chạy. Các máy chủ xử lý yêu cầu bằng cách sử dụng cơ sở dữ liệu và gửi thông tin ngược về máy khách dưới dạng phản hồi. Mã máy khách hiển thị thông tin cho đối tượng truy cập dưới dạng trang web.

Django quản lý mã cho hệ thống yêu cầu và phản hồi này bằng cách sử dụng kiến trúc Mô hình-Khung nhìn-Mẫu (MVT).

* Model

Các mô hình Django đóng vai trò là giao diện giữa cơ sở dữ liệu và mã máy chủ. Chúng là nguồn thông tin chính xác duy nhất về dữ liệu của bạn. Những mô hình dữ liệu này chứa các trường và thao tác thiết yếu mà bạn cần có để tương tác với cơ sở dữ liệu của mình. Do đó, các mô hình Django chuyển đổi bảng cơ sở dữ liệu của bạn thành lớp hoặc đối tượng trong mã Python. Quá trình này được gọi là ánh xạ quan hệ đối tượng.

Nói chung, mỗi mô hình ánh xạ tới một bảng cơ sở dữ liệu duy nhất và có các thuộc tính đại diện cho các trường cơ sở dữ liệu. Ví dụ: nếu trang web của bạn bao gồm các chi tiết về nhân viên thì có thể được biểu diễn dưới dạng:

* Một bảng nhân viên có tên các nhân viên và trường địa chỉ.
* Một mô hình nhân viên gọi là Class Employee (Lớp nhân viên), trong đó có hai thuộc tính hay trường mô hình gọi là Name (Tên) và Address (Địa chỉ).
* Chế độ xem

Khung nhìn của Django sử dụng các mô hình để xử lý yêu cầu. Bạn có thể viết hàm khung nhìn cho từng loại yêu cầu mà đối tượng truy cập trang web có thể thực hiện đối với trang web của bạn. Hàm khung nhìn có thể lấy yêu cầu làm dữ liệu đầu vào và trả về phản hồi. Phản hồi có thể là mã lỗi, hình ảnh, tệp hoặc bất kỳ loại dữ liệu nào.

Django có tính năng trình ánh xạ URL hoặc trình điều phối URL, đóng vai trò ánh xạ hàm khung nhìn với URL của bạn. Bạn phải tạo tệp trình ánh xạ URL để viết các mẫu URL tại đó như nội dung bên dưới.

urlpatterns = [

path('employee/name', views.employee\_name),

path('employee/<int:year>/', views.year\_archive),

]

Ví dụ: nếu bạn muốn đối tượng truy cập trang web của mình xem danh sách gồm tất cả các nhân viên của bạn trong một năm cụ thể, vậy thì, bạn cần thiết lập đường dẫn URL employee/year number và viết hàm khung nhìn tương ứng của Django year\_archive. Khi đối tượng truy cập trang web của bạn nhập "yourwebsitename.com/employee/2020" vào trình duyệt của họ, sau đây là những bước sẽ diễn ra:

* Yêu cầu đến ứng dụng web của bạn.
* Khung web Django lấy số chỉ năm và tên hàm khung nhìn từ trình ánh xạ URL.
* Khung này chạy hàm khung nhìn year\_archive cho năm 2020.
* Year\_archive sử dụng mô hình nhân viên để lấy tất cả dữ liệu nhân viên từ cơ sở dữ liệu trong năm 2020.
* Khung web Django gửi trả dữ liệu dưới dạng phản hồi.
* Mẫu

Các mẫu Django quản lý cách trình bày trang web trong trình duyệt. Vì hầu hết các trang web đều sử dụng Ngôn ngữ đánh dấu siêu văn bản (HTML), bạn có thể viết mã mẫu Django theo cách tương tự như HTML. Tệp mẫu chứa một số thành phần nhất định:

* Các phần tĩnh của đầu ra HTML cuối cùng, chẳng hạn như hình ảnh, nút và tiêu đề.
* Cú pháp đặc biệt mô tả cách chèn nội dung hoặc dữ liệu động, thay đổi theo mọi yêu cầu.

Sau đây là những thành phần cấu tạo nên hệ thống mẫu Django.

* Ngôn ngữ mẫu

Ngôn ngữ mẫu là ngôn ngữ lập trình mà bạn sử dụng để viết mã mẫu HTML. Django hỗ trợ Ngôn ngữ mẫu Django của riêng mình và một lựa chọn thay thế phổ biến gọi là Jinja2.

* Công cụ mẫu

Công cụ mẫu xử lý các tệp mẫu và tạo đầu ra HTML cuối cùng. Công cụ này bao gồm dữ liệu từ phản hồi nằm trong đầu ra này.

Ví dụ: khi đối tượng truy cập trang web của bạn yêu cầu thông tin nhân viên, mẫu Django của bạn sẽ điền các mục sau vào trang web hiển thị với đối tượng truy cập: tiêu đề trang web của bạn, một bảng chứa tên và địa chỉ của tất cả nhân viên và một nút chọn Next (Tiếp theo).

* + - 1. Django có cứng nhắc không?

Chúng tôi gọi các khung web một cách không chính thức là “cứng nhắc” khi chúng buộc các nhà phát triển web tuân thủ một quy trình. Các khung web sở hữu một quan điểm hoặc cách hợp lý mà người phát triển phải hoàn thành những nhiệm vụ nhất định. Ví dụ: khung cứng nhắc thường hỗ trợ các ngành cụ thể phát triển hiệu quả bằng cách lưu giữ tài liệu chi tiết cho các tác vụ ứng dụng liên quan đến ngành đó.

Mặt khác, các khung không cứng nhắc ít bị hạn chế hơn về cách bạn có thể tích hợp các thành phần khác nhau của kiến trúc Mô hình-Khung nhìn-Mẫu (MVT) với nhau. Mặc dù linh hoạt hơn, những khung không cứng nhắc khiến tổ chức mã trở nên phức tạp vì các nhà phát triển khác nhau có thể sử dụng nhiều cách tiếp cận đa dạng cho cùng một tác vụ.

Django có phần hơi cứng nhắc. Django cung cấp nhiều thành phần đa dạng và bao gồm tài liệu về cách xử lý nhiều loại tác vụ phát triển web khác nhau. Các nhà phát triển có thể sử dụng kiến trúc tách biệt của Django để chọn trong một loạt các phương án, đồng thời, kiến trúc này thậm chí còn hỗ trợ thêm cho những phương án mới mà họ yêu cầu.

* + - 1. Bảo mật django

Tội phạm công nghệ cao thường chọn ứng dụng web làm mục tiêu để truy cập thông tin đăng nhập của người dùng, dữ liệu tài chính và những thông tin nhạy cảm khác. Khung web Django cung cấp một số tính năng bảo vệ các ứng dụng web và người dùng của bạn. Bạn sẽ tránh được nhiều lỗi bảo mật phổ biến bằng cách làm theo những biện pháp thực hành tốt nhất sau đây của Django. Sau đây là một vài ví dụ chúng tôi đưa ra.

* Bảo vệ chống chèn tập lệnh liên trang web

Các cuộc tấn công chèn tập lệnh liên trang web (XSS) xảy ra khi tội phạm công nghệ cao chèn mã độc vào trình duyệt của người dùng trang web của bạn. Chúng có thể tấn công người dùng của bạn bằng cách đánh lừa ứng dụng web của bạn theo một số cách như:

Lưu trữ tập lệnh độc hại trong cơ sở dữ liệu của bạn để máy chủ vô tình gửi đi mã độc trong lần phản hồi tiếp theo.

Đánh lừa người dùng của bạn nhấp vào liên kết, khiến mã độc của máy khách chạy trong trình duyệt của người dùng thay vì mã của bạn.

Ví dụ: mẫu Django có thể có {{name}}, hiển thị văn bản mà người dùng nhập vào trường tên trong hồ sơ của họ. Tội phạm công nghệ cao chiếm quyền kiểm soát phiên và thay đổi tên thành mã như <script>alert('hello')</script>. Mẫu thay đổi thành {{<script>alert('hello')</script>}}.

Mã này hiện đang chạy trong trình duyệt của người dùng của bạn nên văn bản thông báo "Hello" (Xin chào) bật lên trên màn hình của họ. Mặc dù đây là ví dụ giản lược, bạn có thể thấy cách tội phạm công nghệ cao chèn tập lệnh phức tạp để kiểm soát máy tính của người dùng của bạn.

Django bảo vệ bạn khỏi các cuộc tấn công như vậy bằng cách tự động thoát hoặc bỏ qua các ký tự cụ thể biểu thị mã độc như < và >, loại chúng khỏi đầu vào của người dùng.

* Bảo vệ chống truy cập trái phép

Các cuộc tấn công giả mạo yêu cầu liên trang web (CSRF) xảy ra khi tin tặc đánh cắp thông tin chứng thực của người dùng của bạn và gửi yêu cầu trái phép tới ứng dụng web của bạn. Django đã tích hợp biện pháp bảo vệ chống lại hầu hết các loại tấn công CSRF trong mô-đun Django CSRF. Cơ chế hoạt động là gửi một giá trị bí mật cho mọi người dùng trong lần đầu tiên họ đăng nhập.

Các yêu cầu từ máy khách mới bao gồm giá trị bí mật làm bằng chứng rằng máy khách đúng là đối tượng đang khai nhận. Vì chỉ trình duyệt của người dùng được cấp quyền mới nắm được giá trị bí mật, Django có thể tự động từ chối yêu cầu nếu những yêu cầu này đến từ một số máy khác đang giả mạo là người dùng. Bạn phải kích hoạt cài đặt mô-đun Django CSRF để biện pháp bảo vệ hoạt động.

* Bảo vệ chống đưa SQL vào

Tội phạm công nghệ cao sử dụng kỹ thuật đưa SQL vào để chèn mã SQL vào cơ sở dữ liệu thông qua các yêu cầu HTTP như POST. Mã độc có thể đánh cắp hoặc xóa dữ liệu chân thực của bạn. Django giải quyết vấn đề này bằng chính thiết kế của nó. Dữ liệu do người dùng gửi, gọi là thông số, được giữ riêng biệt, không chung với truy vấn cơ sở dữ liệu cho đến khi dữ liệu đi tới lớp mô hình. Sau đó, mô hình Django có thể tránh thoát các ký tự nguy hiểm khi đang tạo mã truy vấn.

* Bảo vệ nhờ hỗ trợ từ cộng đồng

Biện pháp bảo vệ trong Django vượt xa các tính năng bảo mật tích hợp của nó. Vì Django là khung nguồn mở, nhiều nhà phát triển dày dạn kinh nghiệm sử dụng và đánh giá các mô-đun của Django. Nhờ kiểm thử kỹ lưỡng, độ tin cậy của mã Django tăng lên cũng như phòng ngừa bỏ sót các lỗ hổng bảo mật ngẫu nhiên.

* + - 1. Khả năng điều chỉnh quy mô của django

Trong phát triển trang web, khả năng điều chỉnh quy mô đề cập tới khả năng trang web đồng thời xử lý nhiều yêu cầu từ máy khách. Các dự án Django có quy mô rất linh hoạt và có thể xử lý hàng nghìn yêu cầu. Bạn có thể điều chỉnh quy mô ứng dụng Django của mình theo những cách sau.

* Phần cứng

Đội ngũ Django đã thiết kế khung web để sử dụng hiệu quả phần cứng trong hệ thống của bạn. Với một kiến trúc không chia sẻ, Django tách biệt các thành phần như lớp cơ sở dữ liệu (các mô hình) và lớp ứng dụng (các khung nhìn). Bạn có thể thêm phần cứng ở bất kỳ cấp độ nào mà không ảnh hưởng tới phần còn lại của hệ thống. Bạn có thể thêm nhiều máy chủ cơ sở dữ liệu hoặc máy chủ ứng dụng vào hệ thống của mình và Django sẽ sử dụng những tài nguyên này một cách hiệu quả để quản lý nhiều đối tượng truy cập.

* Lưu bộ nhớ đệm

Lưu bộ nhớ đệm là quá trình lưu một số dữ liệu của trang web trên máy chủ phía khách hoặc trên các máy chủ trung gian để ứng dụng Django của bạn có thể xử lý yêu cầu nhanh hơn, qua đó tăng quy mô. Django cung cấp một hệ thống lưu bộ nhớ đệm mạnh mẽ với các cấp độ lưu bộ nhớ đệm khác nhau:

* Bạn có thể lưu toàn bộ trang web của mình vào bộ nhớ đệm.
* Bạn có thể lưu đầu ra của hàm khung nhìn cụ thể vào bộ nhớ đệm.
* Bạn có thể lưu nội dung cụ thể mà tốn nhiều thời gian tạo vào bộ nhớ đệm.

Các dự án Django cũng kết hợp hài hòa với các bộ nhớ đệm của bên thứ ba. Bạn có thể viết mã đưa ra gợi ý về những bộ nhớ đệm này và cho bộ nhớ đệm biết phần ứng dụng mà bạn muốn lưu vào bộ nhớ đệm.

* 1. Xampp
     1. Xampp là gì

**XAMPP** là một phần mềm cho phép giả lập môi trường server hosting ngay trên máy tính của bạn, cho phép bạn chạy demo website mà không cần phải mua hosting hay **VPS**. Chính vì vậy, XAMPP hay được phục vụ cho hoạt động học tập giảng dạy thực hành và phát triển web.

Ảnh có chứa văn bản, Phông chữ, ảnh chụp màn hình, biểu tượng

Mô tả được tạo tự động

Hình 9. Xampp

XAMPP được viết tắt của X + Apache + MySQL + PHP + Perl vì nó được tích hợp sẵn Apache, MySQL, PHP, FTP server, Mail Server. Còn X thể hiện cho sự đa nền tảng của XAMPP vì nó có thể dùng được cho 4 hệ điều hành khác nhau: Windows, MacOS, Linux và Solaris.

* + 1. Đặc điểm của xampp
* Ưu điểm

- Có thể chạy được linh hoạt được trên mọi hệ điều hành: Cross-platform, Windows, MacOS, Linux.

- Có cấu hình đơn giản nhưng mang đến nhiều chức năng hữu ích như: Lập Server giả định, lập Mail Server giả định và hỗ trợ SSL trên localhost.

- Tích hợp được nhiều tính năng với các thành phần quan trọng như: Apache, PHP, MySql. Vì thế, người dùng không cần cài đặt từng phần trên riêng lẻ mà chỉ cần cài XAMPP là có 1 web server hoàn chỉnh.

- Tạo mã nguồn mở: Giao diện quản lý dễ dàng và tiện lợi giúp người dùng luôn chủ động được trong chế độ khởi động lại hay bật/ tắt đối với máy chủ theo thời gian phù hợp nhất.

* Nhược điểm

- Do cấu hình khá đơn giản cho nên Xampp hoàn toàn không nhận được sự hỗ trợ về cấu hình Module, đồng thời cũng không có cả MySQL.

- Dung lượng tương đối nặng (141Mb).

- Không có nhiều phiên bản cho từng thành phần của server như PHP, Apache mà phải cài đặt riêng.

Ảnh có chứa đồ điện tử, lái xe, Thiết bị lưu trữ dữ liệu, đĩa cứng

Mô tả được tạo tự động

Hình 10. Hình ảnh mẫu về bộ nhớ

* + 1. Xampp dùng làm gì

XAMPP được dùng để xây dựng và phát triển website theo ngôn ngữ PHP. Ngoài ra, XAMPP còn được sử dụng để phát triển, nghiên cứu website thông qua localhost của máy tính cá nhân, biến máy tính cá nhân thành máy chủ, dùng chính ổ cứng của máy tính để làm nơi lưu trữ cho máy chủ trang web.

Tuy nhiên, đối với các trang web kinh doanh, các trang web cần vận hành liên tục thì việc sử dụng XAMPP để tạo lập không khả thi. Do localhost sử dụng máy tính để làm máy chủ, nên việc duy trì máy chủ sẽ rất khó khăn và không đảm bảo tốc độ lâu dài, thay vào đó các chủ trang web sẽ thuê server hosting. Chung quy lại XAMPP dùng chủ yếu để học tập, nâng cấp và thử nghiệm web.

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu tượng, Thương hiệu

Mô tả được tạo tự động

Hình 11. Xampp và các ứng dụng liên quan

* + 1. Cách cài đặt xampp trên windows

Một số lưu ý trước khi cài đặt XAMPP

- Gỡ hết phần mềm liên quan đến thiết lập localhost như: PHP, MySQL.

- Đổi port cho 1 số phần mềm đang chiếm dụng cổng 80 hoặc 443.

- Nếu máy chủ đã được cài đặt sẵn IIS thì không cần thiết phải cài thêm Xampp.

- Tắt cài đặt tường lửa trên Windows, cũng như tất cả các phần mềm Antivirus khác.

- Tắt User Account Control trên Windows để tránh bị giới hạn quyền truy cập.

Bước 1: Download phần mềm: [TẠI ĐÂY](https://www.apachefriends.org/download.html)

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, Trang web

Mô tả được tạo tự động

Hình 12. Tải xampp

Bước 2: Trong file vừa tải xuống, click chuột chọn tệp tin có đuôi .exe.

Ảnh có chứa văn bản, phần mềm, Trang web, Biểu tượng máy tính

Mô tả được tạo tự động

Hình 13. Xampp trong folder tải xuống

Bước 3: Cửa sổ Setup xuất hiện, nhấn Next.

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, biểu tượng

Mô tả được tạo tự động

Hình 14. Ấn next để tiếp tục

Bước 4: Chọn thư mục để cài đặt (hoặc để mặc định là C:\xampp) > Nhấn Next.

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, Biểu tượng máy tính

Mô tả được tạo tự động

Hình 15. Chọn nơi lưu trữ và tiếp tục

Bước 5: Bỏ chọn phần Learn more about Bitnami for XAMPP > Nhấn Next 2 lần nữa để bắt đầu quá trình cài đặt XAMPP.

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, Biểu tượng máy tính

Mô tả được tạo tự động

Hình 16. Tiếp tục cài đặt

Bước 6: Nhấn Finish để kết thúc cài đặt và mở bảng điều khiển của XAMPP.

* + 1. Hướng dẫn sử dụng xampp

- Khởi động localhost

Bước 1: Truy cập vào thư mục C:\xampp, mở file xampp-panel.exe để bảng điều khiển hiển thị.

Bước 2: Nhấn vào nút Start của 2 ứng dụng Apache và MySQL để bắt đầu khởi động web server. Khi 2 ứng dụng chuyển sang màu xanh thì localhost đã được khởi động.

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, Biểu tượng máy tính

Mô tả được tạo tự động

Hình 17. Khởi động mySQL

* 1. ChatGPT API

ChatGPT hay với tên gọi đầy đủ là Chat Generative Pre-training Transformer - một chatbot do công ty khởi nghiệp OpenAI phát triển. ChatGPT có thể được hiểu đơn giản là một AI (trí thông minh nhân tạo). Điểm đặc biệt của AI này nằm ở "kho" kiến thức mà ChatGPT đã học được.

ChatGPT có thể trả lời lưu loát đầy đủ các câu hỏi mà bạn đưa ra, bất kể là thắc mắc về lĩnh vực gì. Bên cạnh đó, ChatGPT còn có thể làm thơ, soạn nhạc, viết thư, thiết kế và thậm chí là cả sửa lỗi trong lập trình. Nhiều người đã sử dụng ChatGPT để làm những việc trên, điều này đã khiến cho AI ngày càng thông minh hơn.

Khi ChatGPT ra mắt, người dùng muốn tra cứu đã bắt đầu trò chuyện với AI này thay vì tìm kiếm thông tin trên Google. Với sự hiểu biết sâu rộng đa lĩnh vực, các thắc mắc của họ đã được giải quyết chỉ sau vài giây. Công cụ hoạt động giống như một cuộc hỏi đáp giữa người với người.

Điều này cũng đã xảy ra mặt tiêu cực, các câu trả lời của ChatGPT đôi khi vẫn chưa được chính xác. Nhưng công cụ này vẫn thể hiện rất tốt tiềm năng trong tương lai. Chính vì sự hiểu biết rộng cùng tính vận dụng cao, có vẻ như người dùng đã bắt đầu chuyển sang sử dụng công cụ này nhiều hơn, đây sẽ là 1 đối thủ đáng gờm sắp tới của Google.

Ảnh có chứa văn bản, Đồ họa, thiết kế đồ họa, ảnh chụp màn hình

Mô tả được tạo tự động

Hình 18. ChatGPT của OpenAI

* 1. Llama2
     1. Khái niệm

LLaMa là mô hình ngôn ngữ được Meta (trước đây là Facebook) giới thiệu và ra mắt lần đầu vào tháng 2 năm 2023. Phiên bản thứ hai của nó, LLaMa 2, được ra mắt vào tháng 7 cùng năm, với các bản phát hành mới dự kiến ​​​​định kỳ.

Sự khác biệt chính giữa dòng mô hình LLaMa và ChatGPT là mục đích chính của LLaMa không phải là một chatbot . Mặc dù nó chắc chắn có thể được sử dụng cho mục đích này nhưng nó chủ yếu có sẵn để người dùng tải xuống và tùy chỉnh cho trường hợp sử dụng cụ thể của họ. Về mặt này, LLaMa thực sự tương tự như dòng GPT gồm các mô hình ngôn ngữ lớn làm nền tảng cho chatbot ChatGPT.

Một điểm khác biệt so với GPT là LLaMa áp dụng cách tiếp cận gọn gàng hơn về mặt tham số. Hiện tại nó có sẵn ba kích thước tham số: 7 tỷ, 13 tỷ và 70 tỷ (so với 175 tỷ tham số của GPT 3.5).

Mặc dù điều này có nghĩa là chất lượng đầu ra có thể không cao bằng ChatGPT, nhưng nó cũng dẫn đến ít tài nguyên tính toán hơn. Điều này có nghĩa là các công ty khởi nghiệp và tổ chức nhỏ hơn có thể truy cập dễ dàng hơn nhiều, những người có thể tải xuống mô hình LLaMa miễn phí và tùy chỉnh mô hình cục bộ.ào tháng 7 cùng năm, với các bản phát hành mới dự kiến ​​​​định kỳ.

* + 1. Ưu điểm, nhược điểm

Dưới đây là một số ưu và nhược điểm cụ thể của mô hình ngôn ngữ lớn LLaMa:

* + - 1. Ưu điểm
* LLaMa là nguồn mở. Nó miễn phí cho nghiên cứu và sử dụng thương mại, khiến nó minh bạch hơn so với các mô hình AI đóng hoặc “hộp đen” (như ChatGPT).
* Với ít tham số hơn, LLaMa nhanh hơn nhiều mô hình khác, bao gồm cả dòng LLM GPT.
* LLaMa có thể được tải xuống cục bộ, tùy chỉnh và tinh chỉnh cho các tác vụ khác nhau mà không tiêu tốn quá nhiều tài nguyên tính toán.
* Nó được tối ưu hóa cho các trường hợp sử dụng đối thoại , lý tưởng cho việc đào tạo các chatbot dịch vụ khách hàng hoặc các công cụ tiếp thị kỹ thuật số tương tự.

Vì LLaMa 2 được đào tạo bằng cách sử dụng nhiều dữ liệu cập nhật hơn ChatGPT nên sẽ tốt hơn nếu bạn muốn tạo đầu ra liên quan đến các sự kiện hiện tại. Nó cũng có thể được tinh chỉnh bằng cách sử dụng dữ liệu mới hơn.

* + - 1. Nhược điểm
* Đó là nguồn mở… Đợi đã! Đây chẳng phải là một lợi thế sao? Đúng! Tuy nhiên, một số người lập luận rằng với công nghệ mạnh mẽ này, việc tạo ra mã nguồn mở có thể phản trực giác, khiến nó dễ bị kẻ xấu tấn công.
* LLaMa cần được tải xuống máy tính để bàn của bạn—nó không thể truy cập được trong trình duyệt của bạn hoặc dưới dạng ứng dụng. Đây có thể là một bất lợi nếu bạn đang tìm kiếm một chatbot sẵn sàng hoạt động.
* Khả năng của nó bị hạn chế ở một số khu vực—đầu ra có thể không sáng tạo bằng kết quả của các LLM lớn hơn và được đào tạo tốt hơn như những LLM hỗ trợ ChatGPT.
* Hiện tại, LLaMa không thể hỗ trợ các tác vụ toán học, mã hóa hoặc suy luận logic.
* Do kích thước nhỏ hơn nên không phải lúc nào nó cũng có thể nắm bắt được sắc thái ngôn ngữ hoặc thuật ngữ dành riêng cho từng miền thường gặp trong nhiều khía cạnh của tiếp thị kỹ thuật số (mặc dù vì dù sao nó cũng được thiết kế để tinh chỉnh nên điều này không phải lúc nào cũng là một bất lợi).
* Tinh chỉnh nó đòi hỏi chuyên môn kỹ thuật.
  + - 1. So sánh với chatGPT

LLaMa và ChatGPT cho tiếp thị kỹ thuật số: cái nào tốt hơn? Câu trả lời không phải là một câu trả lời phù hợp cho tất cả. Sự lựa chọn tốt hơn phụ thuộc vào mục tiêu, yêu cầu và nguồn lực cụ thể.

Dưới đây là một số cân nhắc bổ sung có thể giúp đưa ra quyết định sáng suốt:

* **Trường hợp sử dụng**: Xem xét các trường hợp sử dụng chính cho nỗ lực tiếp thị kỹ thuật số của bạn. Nếu bạn cần một công cụ chủ yếu dành cho chatbot dịch vụ khách hàng và tương tác nhanh, LLaMa có thể phù hợp hơn vì bạn có thể tinh chỉnh nó theo nhu cầu cụ thể của mình. Mặt khác, nếu nhu cầu của bạn rộng hơn—chẳng hạn như tạo nội dung, viết sáng tạo hoặc hỗ trợ viết mã—ChatGPT có lẽ là lựa chọn tốt nhất cho bạn.
* **Tài nguyên**: LLaMa có thể được lưu trữ riêng tư và các tùy chọn tham số nhỏ hơn của nó giúp các công ty khởi nghiệp và tổ chức nhỏ dễ tiếp cận hơn. ChatGPT chậm hơn rất nhiều, đòi hỏi sức mạnh tính toán đáng kể.
* **Cân nhắc về mặt đạo đức**: Nhìn chung, đạo đức AI là một chủ đề tranh luận sôi nổi hiện nay. Một số người cho rằng cách tiếp cận “hộp đen” của ChatGPT là phi đạo đức và các mô hình nguồn mở như LLaMa sẽ tốt hơn. Mặt khác, một số người cảm thấy rằng việc tạo ra nguồn mở cho một phần mềm mạnh mẽ như vậy sẽ khiến nó dễ bị lạm dụng. Không có câu trả lời rõ ràng ở đây, nhưng cả hai bên đều đáng được xem xét.
* **Hỗ trợ**: Cả LLaMa và ChatGPT đều có cộng đồng hỗ trợ trực tuyến đáng kể, nhưng với cộng đồng các nhà phát triển tích cực có thể mày mò LLaMa, nó có thể cung cấp nhiều hỗ trợ hơn. Điều này lý tưởng nếu bạn muốn điều chỉnh nó cho phù hợp với một trường hợp sử dụng cụ thể. Để sử dụng tổng quát hơn, cũng có một cộng đồng hỗ trợ tốt cho ChatGPT, nhưng cộng đồng này thường tập trung nhiều hơn vào cách tận dụng tối đa hệ thống nhắc nhở của nó .
* **Chi phí**: LLaMa được cung cấp miễn phí, cũng như ChatGPT 3.5. Tuy nhiên, đối với Chat GPT 4, bạn sẽ phải trả phí đăng ký. Tuy nhiên, mặc dù LLaMa miễn phí nhưng cũng đáng xem xét chi phí phát triển hoặc tinh chỉnh có thể cần thiết để điều chỉnh nó cho phù hợp với các trường hợp sử dụng chuyên biệt.

Lựa chọn giữa LLaMa và ChatGPT cuối cùng là vấn đề cân bằng các yếu tố được mô tả ở đây và xem xét nhu cầu sử dụng cũng như năng lực kỹ thuật của bạn. Mô hình “tốt nhất” cuối cùng sẽ là mô hình phù hợp nhất với các mục tiêu và ràng buộc tiếp thị kỹ thuật số cụ thể của bạn.

1. PHÂN TÍCH THIẾT KẾ THỐNG.
   1. Mô tả hệ thống

Hệ thống được thiết kế để cho hai đối tượng sử dụng:

* Đối với Người dùng: người dùng sau khi đã đăng ký tài khoản thành công và cập nhật thông tin cá nhân được sử dụng đầy đủ chức năng của hệ thống bao gồm chat với chatGPT, xem lại đoạn chat, xóa và đổi tên chat.
* Đối với admin: Sử dụng tài khoản trên trang quản trị để thực hiện các chức năng như quản lý người dùng, quản lý đánh giá kết quả của chatGPT, huấn luyện chatGPT,quản lí dữ liệu trên hệ thống, thống kê báo cáo.
  + 1. Sơ đồ kiến trúc hệ thống

Ảnh có chứa văn bản, biểu đồ, hàng, ảnh chụp màn hình

Mô tả được tạo tự động

Hình 19. Sơ đồ kiến trúc hệ thống

* + 1. Sơ đồ chức năng hệ thống

Sơ đồ mô tả chức năng hệ thống

Ảnh có chứa văn bản, biểu đồ, hàng, ảnh chụp màn hình

Mô tả được tạo tự động

Hình 20. Sơ đồ mô tả chức năng

* + 1. Chức năng đối với khách hàng
* Quản lí đăng nhập

+ Biểu đồ hoạt động:

Ảnh có chứa văn bản, hàng, ảnh chụp màn hình, biểu đồ

Mô tả được tạo tự động

Hình 21. Biểu đồ hoạt động Đăng nhập

* Chat với ChatGPT

+ biểu đồ hoạt động:

Ảnh có chứa văn bản, hàng, ảnh chụp màn hình, biểu đồ

Mô tả được tạo tự động

Hình 22. Biểu đồ hoạt động khi thực hiện chat

* Sửa tên đoạn chat đã dùng

+ Biểu đồ hoạt động:

Ảnh có chứa văn bản, hàng, ảnh chụp màn hình, Song song

Mô tả được tạo tự động

Hình 23. Biểu đồ hoạt động khi sửa đoạn chat

**Sơ đồ chức năng**

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, hàng

Mô tả được tạo tự động

Hình 24. Sơ đồ chức năng hệ thống

* + 1. Chức năng nối với với admin
* Quản lí tài khoản user

+ Biểu đồ hoạt động:

Ảnh có chứa văn bản, hàng, ảnh chụp màn hình, biểu đồ

Mô tả được tạo tự động

Hình 25. Biểu đồ hoạt động khi thực hiện quản lí tài khoản của admin

* Quản lí huấn luyện chatGPT

+ Biểu đồ hoạt động:

Ảnh có chứa văn bản, hàng, ảnh chụp màn hình, biểu đồ

Mô tả được tạo tự động

Hình 26. Biểu đồ hoạt động huấn luyện chatGPT

* Quản lí đánh giá đúng sai câu trả lời của chat
* Thống kê báo cáo
* Quản lí dữ liệu trên hệ thống

**Sơ đồ hệ thống**

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Phông chữ

Mô tả được tạo tự động

Hình 27. Sơ đồ hệ thống

* 1. Phân tích hệ thống về mặt chức năng
     1. Biểu đồ luồng dữ liệu (BLD) mức khung cảnh

Ảnh có chứa văn bản, biểu đồ, Kế hoạch, Bản vẽ kỹ thuật

Mô tả được tạo tự động

Hình 28. Biểu đồ BLD mức khung cảnh

* + 1. Biểu đồ luồng dữ liệu (BLD) mức cảnh

Ảnh có chứa văn bản, biểu đồ, bản phác thảo, Bản vẽ kỹ thuật

Mô tả được tạo tự động

Hình 29. Biểu đồ BLD mức cảnh

* + 1. Biểu đồ luồng dữ liệu (BLD) mức dưới đỉnh
* Đối với chức năng quản lí người dùng

Ảnh có chứa biểu đồ, văn bản, hàng, Bản vẽ kỹ thuật

Mô tả được tạo tự động

Hình 30. Biểu đồ BLD mức đỉnh với chức năng quản lí người dùng

* Đối với chức năng quản trị web

Ảnh có chứa văn bản, biểu đồ, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động

Hình 31. Biểu đồ BLD mức đỉnh với chức năng quản trị web

* 1. Phân tích hệ thống về mặt cơ sở dữ liệu
     1. Xác định các thực thể
* Thực thể categoryAccount

Thực thể cartegoryAccount gồm 2 thuộc tính ID, nameCategory

Ảnh có chứa vòng tròn, biểu đồ, bản phác thảo, thiết kế

Mô tả được tạo tự động

Hình 32. Thực thể categoryAccount

* Thực thể Account

Thực thể account gồm 9 thuộc tính: id, imageSource, categoryAccountID, name, userName, phoneNumber, age, password, startTime

Ảnh có chứa vòng tròn, biểu đồ, hàng, văn bản

Mô tả được tạo tự động

Hình 33. Thực thể Account

* Thực thể LoginHistory

Thực thể loginHistory gồm 2 thuộc tính: time, accountID

Ảnh có chứa vòng tròn, biểu đồ, thiết kế

Mô tả được tạo tự động

Hình 34. thực thể loginHistory

* Thực thể BlockChat

Thực thể blockChat gồm 4 thuộc tính: IDchat, TimeStart, accountID, titleChat

Ảnh có chứa biểu đồ, vòng tròn, văn bản, ảnh chụp màn hình

Mô tả được tạo tự động

Hình 35. Thực thể BlockChat

* Thực thể Questions

Thực thể question chứa câu hỏi trong blockChat gồm 3 thuộc tính: ID, blockChatID, timeQuestion

Ảnh có chứa biểu đồ, vòng tròn, hàng, thiết kế

Mô tả được tạo tự động

Hình 36. Thực thể Questions

* Thực thể Response

Thực thể response chưa thông tin phản hồi từ chatbot dựa trên câu hỏi đã cung cấp gồm 3 thuộc tính: responseMessage, TimeResponse, QuestionID

Ảnh có chứa văn bản, biểu đồ, vòng tròn, hàng

Mô tả được tạo tự động

Hình 37. Thực thể Response

* + 1. Sơ đồ thực thể liên kết
* Mỗi một tài khoản chỉ được tạo ra từ một loại tài khoản (phân cấp tài khoản như là tài khoản vip và tài khoản thường) nên liên kết giữa account và categoryAccount là nhiều-một



* Mỗi một tài khoản có thể đăng nhập nhiều lần nên liên kết giữa account và loginHistory là liên kết một-nhiều



* Mỗi một tài khoản có thể có nhiều khối chat nên là liên kết giữa account và blockChat là liên kết một-nhiều



* Mỗi một khối chat có thể có nhiều câu hỏi nên là liên kết giữa blockChat và question là liên kết một-nhiều



* Mỗi một câu hỏi có thể có nhiều câu trả lời vậy nên liên kết giữa thực thể question và response là liên kết một-nhiều



* + 1. Mô hình quan hệ giữa các thực thể

Ảnh có chứa văn bản, ảnh chụp màn hình, hàng, Phông chữ

Mô tả được tạo tự động

Hình 38. Mô hình quan hệ giữa các thực thể, trong đó nền xanh thể hiện khóa chính, khóa ngoại có nền màu xám

* + 1. Danh sách các bảng dữ liệu

|  |  |  |
| --- | --- | --- |
| **STT** | **Tên bảng** | **Diễn giải** |
| 1 | CatergoryAccount | chứa thông tin loại tài khoản |
| 2 | Account | Chứa thông tin tài khoản người dùng |
| 3 | LoginHistory | Chứa thông tin đăng nhập của từng người dùng |
| 4 | BlockChat | Chứa khối tin nhắn của người dùng |
| 5 | Question | Chứa câu hỏi của người dùng |
| 6 | Response | Chứa phản hồi của chatGPT |

Bảng 1. Danh sách các bảng dữ liệu

* + 1. Mô tả chi tiết các bảng

**Mục đích**: Có cái nhìn chi tiết về cơ sở dữ liệu của đề tài

**Khóa chính (Primary Key)**: là một giá trị dùng để phân biệt bản ghi này với bản ghi khác. Giá trị của khóa chính trong mỗi bản ghi là duy nhất trong mỗi bảng.

**Khóa ngoại (Foreign Key)**: là một trường hay một nhóm trường trong một bản ghi của một bảng, trỏ đến khóa của một bản ghi khác của một bảng. Thông thường, khóa ngoài trong một bảng trỏ đến khóa chính của một bảng khác. Kiểu tham chiếu này có thể liên các thông tin lại với nhau và nó là một phần quan trọng của quá trình chuẩn hóa.

**Bảng CATEGORYACCOUNT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên trường** | **Kiểu dữ liệu** | **kích thước** | **tính chất** | **mô tả** |
| ID | int | 10 | primary key | mã tài khoản |
| namecatergory | varchar | 50 |  | loại tài khoản |

Bảng 2. Bảng CategoryAccount

**Bảng ACCOUNT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên trường** | **Kiểu dữ liệu** | **kích thước** | **tính chất** | **mô tả** |
| ID | integer | 10 | primary key | mã tài khoản |
| categoryAccount | integer | 10 | foreign key | mã loại tk |
| username | varchar | 100 | not null | tên đăng nhập |
| password | varchar | 100 | not null | mật khẩu |
| starttime | date |  |  | ngày tạo tk |
| imagesource | varchar | 100 |  | ảnh đại diện |
| name | varchar | 100 | not null | tên người dùng |
| age | integer | 10 |  | tuổi |
| phonenumber | char | 10 |  | sdt |

Bảng 3. Bảng Account

**Bảng LOGINHISTORY**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên trường** | **Kiểu dữ liệu** | **kích thước** | **tính chất** | **mô tả** |
|  |  |  |  |  |
| time | date |  | not null | thời gian đăng nhập |
| accountid | integer | 10 | foreign key | mã tài khoản |

Bảng 4. Bảng LoginHistory

**Bảng BLOCKCHAT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên trường** | **Kiểu dữ liệu** | **kích thước** | **tính chất** | **mô tả** |
| idChat | integer | 10 | primary key | mã khối chat |
| timestart | date |  | not null | thời gian đăng nhập |
| accountid | integer | 10 | foreign key | mã tài khoản |
| titlechat | varchar | 50 |  | tiêu đề chat |

Bảng 5. Bảng BlockChat

**Bảng QUESTION**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên trường** | **Kiểu dữ liệu** | **kích thước** | **tính chất** | **mô tả** |
| id | integer | 10 | primary key | mã câu hỏi |
| blockchatidchat | integer | 10 | foreign key | mã khối chat |

Bảng 6. Bảng Question

**Bảng RESPONSE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tên trường** | **Kiểu dữ liệu** | **kích thước** | **tính chất** | **mô tả** |
| responsemessage | varchar | 1000 | not null | câu trả lời |
| timeresponse | date |  | not null | thời gian phản hồi |
| questionid | integer | 10 | foreign key | mã câu hỏi |

Bảng 7. Bảng Response

* 1. Một số thuật toán trong phần mềm
     1. Thuật toán tiền xử lí dữ liệu huấn luyện.

Thuật toán mô tả các bước xử lí dữ liệu khi có người upload file lên hệ thống

Ảnh có chứa biểu đồ, Bản vẽ kỹ thuật, Kế hoạch, hàng

Mô tả được tạo tự động

Bảng 8. Thuật toán tiền xử lí dữ liệu huấn luyện

* + 1. Thuật toán tải file lên CSDL.

Thuật toán dưới đây mô tả quy trình tải file lên của hệ thống khi có người upload

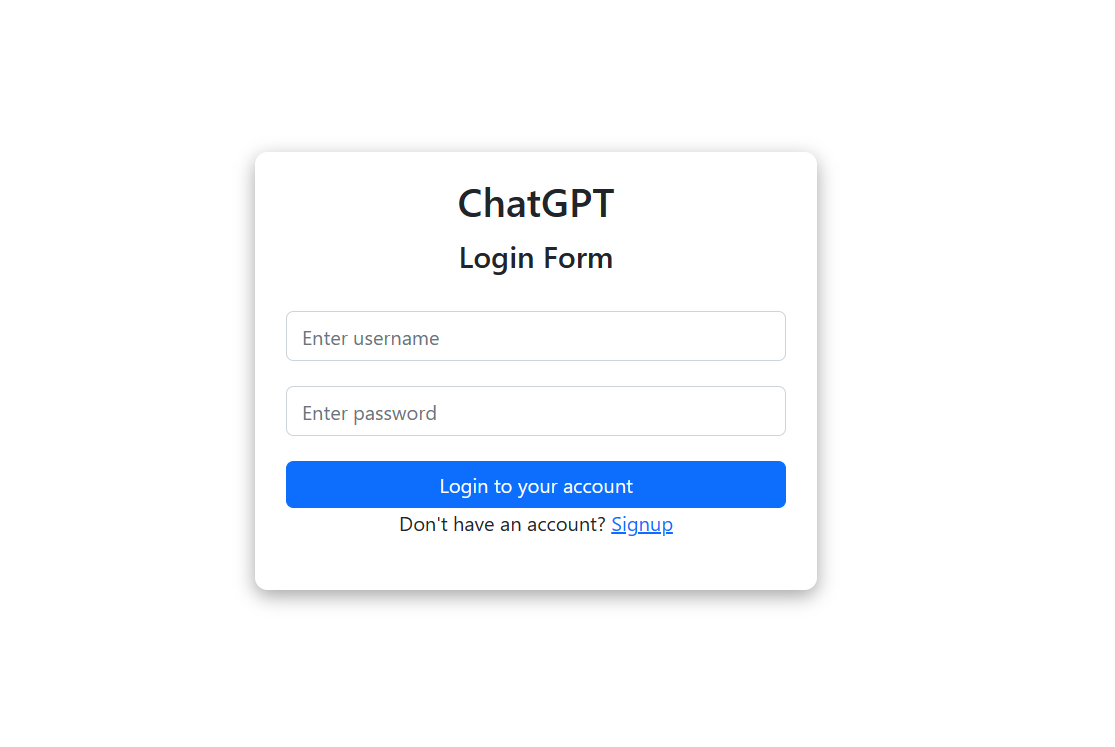
Ảnh có chứa biểu đồ, văn bản, Bản vẽ kỹ thuật, Kế hoạch

Mô tả được tạo tự động

Bảng 9. Thuật toán tải file lên csdl

1. SẢN PHẦM DEMO VÀ KẾT QUẢ.
   1. Giao diện sản phẩm
      1. Giao diện đăng nhập, đăng ký.

Để đăng nhập vào chat và sử dụng chat, đầu cần phải đăng nhập.



Hình 39. Giao diện đăng nhập của chat

Nếu chưa có tài khoản đăng nhập thì sẽ phải tiến hành đăng ký trước khi đăng nhập.

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, số

Mô tả được tạo tự động

Hình 40. Giao diện đăng ký của chat

* + 1. Giao điện chính của Chat

Sau khi đăng nhập thì trang web sẽ mở ra giao diện chat như dưới đây.

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, Phần mềm đa phương tiện

Mô tả được tạo tự động

Hình 41. Giao diện chat của sản phẩm

Khi sử dụng chat, giao diện sẽ hiển thị như sau:

Ảnh có chứa văn bản, ảnh chụp màn hình, phần mềm, Phần mềm đa phương tiện

Mô tả được tạo tự động

Hình 42. Giao diện của sản phẩm khi thực hiện chat.

* 1. Báo cáo kết quả

bảng tổng hợp những câu hỏi đánh giá trước và sau khi huấn luyện chatGPT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STT** | **Câu hỏi** | **Trước khi train** | **Sau khi train** | |
| 1 | đề án tuyển sinh của đại học kinh tế quốc dân | x | x | x |
| 2 | Ngành khoa học máy tính của đại học kinh tế quốc dân hà nội tuyển bao nhiêu chỉ tiêu năm 2023 | x | v | v |
| 3 | Trường đại học kinh tế quốc dẫn có ngành công nghệ thông tin không | x | v | v |
| 4 | Viện công nghệ thông tin và kinh tế số ở đại học kinh tế quốc dân thành lập năm nào | x | v | v |
| 5 | Điểm chuẩn của các khoa trong viện công nghệ thông tin và kinh tế số của đại học kinh tế quốc dân năm 2022 | x | v | v |
| 6 | các tổ hợp xét tuyển của đại học kinh tế quốc dân | x | v | v |
| 7 | viện công nghệ thông tin và kinh tế số ở đại học kinh tế quốc dân có những khoa nào | x | v | v |
| 8 | định hướng tuyển sinh đại học chính quy của đại học kinh tế quốc dân năm 2024 | x | x | x |
| 9 | Điều kiện tuyển sinh đại học chính quy của đại học kinh tế quốc dân năm 2023 | x | v | v |
| 10 | Điểm ngưỡng đầu vào của đại học kinh tế quốc dân năm 2023 | x | v | v |
| 11 | điều kiện xét tuyển kết hợp của đại học kinh tế quốc đân là gì | x | v | v |
| 12 | Điểm chuẩn ngành khoa học máy tính trường đại học kinh tế quốc dân năm 2022 | x | v | v |
| 13 | Điểm chuẩn ngành công nghệ thông tin trường đại học kinh tế quốc dân năm 2023 | x | v | v |
| 14 | Ngành công nghệ thông tin của đại học kinh tế quốc dân hà nội tuyển bao nhiêu chỉ tiêu năm 2022 | x | v | v |
| 15 | Điều kiện tuyển sinh đại học chính quy của đại học kinh tế quốc dân năm 2023 | x | v | v |
| 16 | Tổng số tín chỉ của các khoa trong viện công nghệ thông tin và kinh tế số đại học kinh tế quốc dân | x | v | v |
| 17 | Các học phần chính ngành công nghệ thông tin đại học kinh tế quốc dân | x | v | v |
| 18 | Cơ hội nghề nghiệp ngành công nghệ thông tin trường đại học kinh tế quốc dân | x | v | v |
| 19 | Định hướng chuyên sâu ngành công nghệ thông tin trường đại học kinh tế quốc dân | x | v | v |
| 19 | Định hướng chuyên sâu ngành công nghệ thông tin trường đại học kinh tế quốc dân | x | v | v |
| 20 | Các học phần chính ngành khoa học máy tính trường đại học kinh tế quốc dân | x | v | v |
| 21 | Tổng quy mô đào tạo khối ngành tiến sĩ | x | v | v |
| 22 | quy đào tạo khối ngành tiến sĩ thuộc khối ngành III là bao nhiêu | x | x | x |
| 23 | quy đào tạo khối ngành tiến sĩ khoa tài chính ngân hàng là bao nhiêu | x | x | x |
| 24 | Tổng quy mô đào tạo trình độ đại học chính quy là bao nhiêu | x | x | x |
| 25 | Chỉ tiêu tuyển sinh của khối ngành khoa học máy tính nắm 2019? | x | x | x |
| 26 | Điểm trúng tuyển của khối ngành khoa học máy tính nắm 2019? | x | x | x |
| 27 | Số lượng Hội trường, phòng học lớn trên 200 chỗ là bao nhiêu | x | v | v |
| 28 | Danh trang thiết bị của Phòng thực hành | x | v | v |
| 29 | Số lượng đào tạo khối ngành V | x | x | x |
| 30 | Tình hình việc làm của khối ngành V sau khi tốt nghiệp | x | v | v |

bảng dữ liệu trên:

+ từ câu 1-20:câu hỏi liên quan đến phần nội dung văn bản

* trước khi train thì 100% câu hỏi không trả lời đúng
* Sau khi train: trả lời đúng 18/20 câu =>90%

+ từ câu 21-30: câu hỏi liên quan đến dữ liệu trong bảng

* Trước khi train thì 100% câu hỏi không trả lời đúng
* Sau khi train thì 4/10 câu trả lời đúng =>40%

KẾT LUẬN

Trong quá trình nghiên cứu đề tài **“Cá nhân hóa ChatGPT với dữ liệu đề án tuyển sinh ĐH kinh tế quốc dân”** em đã có những thành tựu đáng kể. Đầu tiên, thông qua dữ liệu được cung cấp em đã xử lí dữ liệu thành các file dữ liệu word, pdf… phù hợp với việc huấn luyện chatGPT. Thứ hai, tạo ra mô hình Chatbot dựa trên chatGPT APIs có thể đọc dữ liệu từ các file dữ liệu và có thể đưa ra câu trả lời dựa trên thông tin đã cung cấp. Với những thành tựu trên thì có thể hoàn toàn ứng dụng chatbot vào những trang website cá nhân của công ty và cung cấp dữ liệu cần thiết để có thể sử dụng chat.

Đề tài cá nhân hóa ChatGPT dựa trên dữ liệu cá nhân trong ngữ cảnh của một công ty là một hướng nghiên cứu có tiềm năng lớn trong việc cải thiện trải nghiệm tương tác giữa ChatGPT và người dùng trong môi trường doanh nghiệp.

Kết luận có thể tập trung vào các lợi ích mà việc cá nhân hóa có thể mang lại cho doanh nghiệp, như tăng cường tương tác và tương tác cá nhân hóa với khách hàng hoặc người dùng nội bộ. Tuy nhiên, cũng cần nhấn mạnh vào việc thiết lập các biện pháp bảo vệ dữ liệu mạnh mẽ để đảm bảo rằng thông tin cá nhân được xử lý và lưu trữ một cách an toàn và tuân thủ các quy định pháp luật.

TÀI LIỆU THAM KHẢO

1. Mô hình LLMs: <https://vinbigdata.com/kham-pha/large-language-model-la-gi-tat-ca-nhung-dieu-ban-can-biet-ve-mo-hinh-ngon-ngu-nay.html#Large_language_model_la_gi>
2. Tìm hiểu về ChatGPT: <https://www.dienmayxanh.com/kinh-nghiem-hay/chat-gpt-la-gi-loi-ich-va-cach-thuc-hoat-dong-1518107#:~:text=ChatGPT%20>(Chat%20Generative%20Pre%2Dtraining,v%C3%A0%20l%C4%A9nh%20v%E1%BB%B1c%20kh%C3%A1c%20nhau.
3. So sánh ChatGPT và chatbot thông thường: <https://vinbigdata.com/chatbot/ai-chatbot-la-gi-va-dot-pha-so-voi-chatbot-thuong.html>
4. Thư viên Langchain: <https://python.langchain.com/docs/get_started/introduction>
5. Framework Django: <https://aws.amazon.com/vi/what-is/django/>
6. Hướng dẫn cài đặt và sử dụng xampp:

<https://www.thegioididong.com/hoi-dap/xampp-la-gi-cach-cai-dat-va-su-dung-localhost-tren-may-1339338>

1. So sánh llama2 và ChatGPT:

<https://careerfoundry.com/en/blog/digital-marketing/llama-vs-chatgpt/#which-is-better>

1. Slide bài giảng phân tích và thiết kế hệ thống – Trần Thị Mỹ Diệp – ĐH kinh tế quốc dân.