**TRƯỜNG ĐẠI HỌC NGUYỄN TRÃI**

**KHOA CÔNG NGHỆ THÔNG TIN**

---------o0o---------

A picture containing text, clipart

Description automatically generated

**BÁO CÁO**

**MÔN : THỰC TẬP CHUYÊN ĐỀ 2**

**Họ và tên : Đỗ Ngọc Quý Dương**

**Lớp :K21CNT1**

**MSSV : 2110900071**

**Giảng viên hướng dẫn : Th.S Trịnh Văn Chung**

**ĐỀ BÀI : "Nghiên cứu công nghệ AI trong việc ứng dụng cho chatbox".**

**Hà Nội , Tháng 9 Năm 2024**

MỤC LỤC NGHIÊN CỨU

[PHẦN I : GIỚI THIỆU VỀ ĐỀ TÀI 3](#_Toc178080487)

[1.TỔNG QUAN VỀ AI 3](#_Toc178080488)

[2.HỆ THỐNG VÀ CÁC NGHIỆP VỤ CƠ BẢN 6](#_Toc178080489)

[3.KHẢO SÁT VÀ THU THẬP CÁC BIỂU MẪU 9](#_Toc178080490)

[4.YÊU CẦU CHỨC NĂNG 11](#_Toc178080491)

[PHẦN II : PHÂN TÍCH VÀ THIẾT KẾ BÀI TOÁN 14](#_Toc178080492)

[1.Use-case Diagram 14](#_Toc178080493)

[2.Thiết kế cơ sở dữ liệu 15](#_Toc178080494)

[A.Mô hình ERD 15](#_Toc178080495)

[B.Database Diagram 15](#_Toc178080496)

[C.Mô tả dữ liệu 16](#_Toc178080497)

[PHẦN III : THIẾT KẾ HỆ THỐNG 17](#_Toc178080498)

[1.Giao diện và các chức năng 17](#_Toc178080499)

[PHẦN IV : KIỂM THỬ 19](#_Toc178080500)

[**1. Kiểm thử chức năng** 19](#_Toc178080501)

[**2. Kiểm thử hiệu suất** 19](#_Toc178080502)

[**3. Kiểm thử bảo mật** 19](#_Toc178080503)

[**4. Theo dõi chất lượng** 20](#_Toc178080504)

[PHẦN V : KẾT LUẬN 21](#_Toc178080505)

[PHẦN V : TÀI LIỆU THAM KHẢO 21](#_Toc178080506)

# PHẦN I : GIỚI THIỆU VỀ ĐỀ TÀI

## 1.TỔNG QUAN VỀ AI

**a.** Khái niệm về AI

* **Định nghĩa AI (Trí tuệ nhân tạo)**: Trí tuệ nhân tạo (AI) là một lĩnh vực khoa học máy tính, với mục tiêu tạo ra các hệ thống máy tính có khả năng thực hiện những nhiệm vụ mà thông thường yêu cầu trí thông minh của con người, chẳng hạn như nhận thức, học tập, lập luận, và giải quyết vấn đề.
* **Phân loại AI**: AI được chia thành ba loại chính:
  + **AI yếu (Narrow AI)**: Là loại AI được thiết kế để thực hiện một nhiệm vụ cụ thể, như chatbot, trợ lý ảo, nhận diện khuôn mặt, v.v.
  + **AI mạnh (General AI)**: AI có khả năng thực hiện nhiều nhiệm vụ phức tạp giống như con người, chưa được phát triển hoàn thiện.
  + **AI tự chủ (Superintelligent AI)**: Loại AI vượt trội hơn cả trí tuệ con người, có khả năng tự học và phát triển không ngừng. Đây là một mục tiêu xa trong tương lai và còn đang được nghiên cứu.
* **LỊCH SỬ PHÁT TRIỂN CỦA AI :**
  + **Giai đoạn khởi đầu (1950-1970)**: Trí tuệ nhân tạo bắt đầu với những nghiên cứu lý thuyết về máy tính và các nguyên lý cơ bản của tính toán. Năm 1956, thuật ngữ "Artificial Intelligence" lần đầu tiên được sử dụng tại Hội nghị Dartmouth bởi John McCarthy, Marvin Minsky, và các nhà tiên phong khác.
    - **Turing Test**: Được đề xuất bởi Alan Turing, một bài kiểm tra để xem liệu một máy có thể bắt chước trí thông minh của con người hay không.
    - **ELIZA (1966)**: Một trong những chatbot đầu tiên, được Joseph Weizenbaum phát triển để mô phỏng cuộc trò chuyện giữa người và máy.
  + **Giai đoạn phát triển chậm (1970-1990)**: Trong giai đoạn này, AI không đạt được nhiều đột phá lớn do hạn chế về sức mạnh xử lý của máy tính và những vấn đề phức tạp liên quan đến việc lập trình trí thông minh.
    - **Hệ chuyên gia (Expert Systems)**: Các hệ thống AI đầu tiên được phát triển để xử lý thông tin và đưa ra quyết định trong các lĩnh vực chuyên biệt như y học, tài chính. DEXTER và MYCIN là hai hệ thống nổi tiếng.
    - **Winter AI (Mùa đông AI)**: Trong giai đoạn này, AI phải đối mặt với sự thiếu hụt kinh phí và quan tâm do không đạt được nhiều tiến bộ thực tiễn.
  + **Sự bùng nổ của AI (1990 đến nay)**: Với sự phát triển của Internet, sự gia tăng của dữ liệu lớn (Big Data), và sự tiến bộ trong công nghệ xử lý đồ họa (GPUs), AI đã có một sự trở lại mạnh mẽ.
    - **Deep Blue (1997)**: Máy tính của IBM đánh bại Garry Kasparov, nhà vô địch cờ vua thế giới, một cột mốc lớn trong lịch sử AI.
    - **Sự phát triển của Học máy (Machine Learning)**: Học máy trở thành một yếu tố quan trọng trong sự phát triển của AI, cho phép máy tính học từ dữ liệu và cải thiện hiệu suất mà không cần lập trình rõ ràng.
    - **AI hiện đại**: Các mô hình AI tiên tiến như DeepMind's AlphaGo (2016) đánh bại kỳ thủ cờ vây số một thế giới, Google’s BERT, OpenAI's GPT series đã thay đổi cách mà AI xử lý ngôn ngữ và thực hiện nhiều tác vụ trí tuệ khác.
* **TẦM QUAN TRỌNG CỦA AI**
  + **Tự động hóa quy trình**: AI đã giúp tự động hóa nhiều công việc phức tạp và tiết kiệm chi phí trong các lĩnh vực như sản xuất, chăm sóc khách hàng, và y tế.
  + **Cải thiện chất lượng dịch vụ**: AI đã nâng cao chất lượng dịch vụ khách hàng bằng các chatbot AI, giúp xử lý hàng triệu tương tác cùng một lúc, hỗ trợ nhanh chóng, cá nhân hóa hơn.
  + **Xử lý dữ liệu khổng lồ**: AI có khả năng phân tích và xử lý khối lượng dữ liệu lớn, từ đó rút ra những thông tin có giá trị để hỗ trợ quyết định trong kinh doanh, y học, và các ngành công nghiệp khác.
  + **Nâng cao trải nghiệm người dùng**: AI đã góp phần phát triển các công cụ trợ lý ảo thông minh như Siri, Alexa, và Google Assistant, giúp người dùng dễ dàng tương tác với các hệ thống máy tính và internet.
* **CÁC ỨNG DỤNG CỦA AI**
  + **Y tế**: AI được sử dụng trong việc phân tích hình ảnh y học, dự đoán các nguy cơ bệnh tật, hỗ trợ phẫu thuật và chăm sóc bệnh nhân. Ví dụ, AI có thể dự đoán nguy cơ mắc bệnh tim hoặc giúp chẩn đoán ung thư sớm hơn.
  + **Tài chính**: AI hỗ trợ trong việc phát hiện gian lận, phân tích rủi ro đầu tư, và quản lý tài sản tự động. Các hệ thống AI có thể phân tích dữ liệu tài chính và đưa ra các khuyến nghị đầu tư chính xác hơn con người.
  + **Giao thông**: AI trong các phương tiện tự hành (như xe tự lái) đang dần thay đổi cách thức di chuyển của con người, giúp giảm tai nạn giao thông và tối ưu hóa lộ trình di chuyển.
  + **Thương mại điện tử**: AI được sử dụng để cá nhân hóa trải nghiệm mua sắm trực tuyến, từ việc đề xuất sản phẩm phù hợp đến việc hỗ trợ khách hàng thông qua chatbot.
  + **Giáo dục**: AI có thể tạo ra các hệ thống học tập tự động hóa, giúp cung cấp tài liệu học tập theo nhu cầu của học sinh và phân tích hiệu suất học tập của họ.
* **AI VÀ SỰ THAY ĐỔI TRONG NGÀNH CÔNG NGHỆ**
  + **Học máy và Deep Learning**: AI hiện đại không chỉ tập trung vào lập trình mà còn dựa vào khả năng học từ dữ liệu, cho phép hệ thống AI trở nên thông minh hơn theo thời gian. Các mô hình deep learning với mạng nơ-ron sâu đã giúp AI trở nên mạnh mẽ trong việc nhận diện hình ảnh, xử lý ngôn ngữ tự nhiên, và thậm chí đánh bại con người trong các trò chơi phức tạp.
  + **AI và Dữ liệu lớn (Big Data)**: AI phát triển mạnh mẽ nhờ vào sự gia tăng lượng dữ liệu khổng lồ từ internet, mạng xã hội, và các thiết bị kết nối. Dữ liệu này cung cấp cơ sở để AI có thể học hỏi và cải tiến không ngừng.

1. **AI TRONG GIAO TIẾP VÀ DỊCH VỤ KHÁCH HÀNG**
   * **VAI TRÒ CỦA AI TRONG GIAO TIẾP VÀ DỊCH VỤ KHÁCH HÀNG**
     + **Tự động hóa quy trình giao tiếp:** AI giúp tự động hóa quá trình trả lời các câu hỏi thường gặp và xử lý các yêu cầu đơn giản của khách hàng, giảm tải cho nhân viên chăm sóc khách hàng. Điều này không chỉ giúp tiết kiệm thời gian mà còn giảm chi phí vận hành cho doanh nghiệp.
     + **Tương tác 24/7:** Một trong những lợi ích lớn nhất của AI trong giao tiếp là khả năng hoạt động liên tục 24/7. Khách hàng có thể liên hệ với doanh nghiệp bất cứ lúc nào mà không cần lo lắng về giờ làm việc. AI giúp duy trì dịch vụ khách hàng liên tục và không gián đoạn.
     + **Cá nhân hóa trải nghiệm:** AI có khả năng phân tích dữ liệu khách hàng và tạo ra các tương tác mang tính cá nhân hóa hơn. Nó có thể sử dụng lịch sử tương tác của khách hàng để cung cấp các giải pháp, đề xuất sản phẩm hoặc dịch vụ phù hợp với từng người.
   * **ĐỀ TÀI NGHIÊN CỨU**
     + Chính vì những vai trò trên và để hiểu rõ và cũng như thấy được tầm quan trọng của AI trong Chatbox hỗ trợ khách hàng và tìm hiểu sau hơn , đưa ra ví dụ thực tế cho các thầy cô và mọi người cùng tham khảo nên tôi đã quyết định chọn đề tài : “**Nghiên cứu công nghệ AI trong việc ứng dụng cho chatbox**” cho báo cáo lần này .

## 2.**HỆ THỐNG VÀ CÁC NGHIỆP VỤ CƠ BẢN**

**1. Hệ thống Chatbot**

**Chatbot** là một hệ thống phần mềm được thiết kế để mô phỏng các cuộc hội thoại của con người qua giao diện văn bản hoặc giọng nói. Hệ thống chatbot thường được tích hợp trong các ứng dụng web, ứng dụng di động hoặc các nền tảng nhắn tin. Mục tiêu của chatbot là tự động hóa các cuộc trò chuyện, giúp giảm tải cho nhân viên hoặc cung cấp trải nghiệm tương tác nhanh chóng, thông minh hơn cho người dùng.

**Các nghiệp vụ cơ bản của hệ thống chatbot:**

1. **Xử lý ngôn ngữ tự nhiên (Natural Language Processing - NLP)**: Đây là một thành phần quan trọng trong hệ thống chatbot. NLP giúp chatbot phân tích và hiểu các câu truy vấn của người dùng, từ đó tìm ra ý nghĩa và mục đích của câu hỏi. Quy trình này bao gồm:
   * **Phân tích cú pháp** (Tokenization): Tách các câu hỏi thành từ hoặc cụm từ.
   * **Nhận dạng thực thể có tên** (Named Entity Recognition): Xác định tên, thời gian, địa điểm, số lượng từ câu hỏi.
   * **Phân tích ngữ nghĩa** (Semantic Analysis): Tìm hiểu ý nghĩa và ngữ cảnh câu hỏi.
2. **Xử lý yêu cầu**: Sau khi hiểu câu hỏi của người dùng, chatbot sẽ gửi yêu cầu đến các thành phần phía server hoặc cơ sở dữ liệu để tìm kiếm câu trả lời phù hợp. Với các chatbot nâng cao, chúng có thể gọi các API bên ngoài để xử lý dữ liệu hoặc tương tác với các dịch vụ khác.
3. **Trả lời**: Hệ thống sẽ tạo câu trả lời từ các dữ liệu đã thu thập, sau đó trả về cho người dùng qua giao diện. Các câu trả lời có thể là câu trả lời tĩnh từ một bộ dữ liệu hoặc câu trả lời động tùy thuộc vào ngữ cảnh.
4. **Học từ phản hồi người dùng**: Một số chatbot, đặc biệt là các chatbot AI và học sâu (deep learning-based chatbot), có khả năng học từ phản hồi người dùng, từ đó cải thiện chất lượng phản hồi trong tương lai.

**2. Khái quát về cách hệ thống chatbot hoạt động**

Hệ thống Chatbot hoạt động qua các bước chính sau :

1. **Nhận và phân tích đầu vào của người dùng**: Chatbot sẽ nhận câu hỏi từ người dùng thông qua giao diện văn bản hoặc giọng nói..
2. **Xử lý ngôn ngữ tự nhiên (NLP)**: Phân tích câu hỏi dựa trên kỹ thuật NLP để hiểu ngữ cảnh, từ khóa và ý định.
3. **Xác định mục đích người dùng (Intent):** Dựa vào mô hình đã được huấn luyện, chatbot sẽ xác định mục đích của người dùng từ câu hỏi.
4. **Tìm câu trả lời:** Chatbot sẽ tìm kiếm thông tin trong cơ sở dữ liệu (FAQ), các quy tắc đã được định sẵn hoặc sẽ sử dụng mô hình AI để tạo ra câu trả lời phù hợp
5. **Gửi phản hồi:** Cuối cùng, chatbot sẽ gửi câu trả lời trở lại người dùng thông qua cùng giao diện.

**3. Phân loại Chatbot**

Chatbot được phân loại dựa trên mức độ phức tạp và cách thức hoạt động. Có hai loại chính:

* **Chatbot dựa trên quy tắc (Rule-based Chatbot)**:
  + **Cách thức hoạt động**: Chatbot này hoạt động dựa trên các quy tắc hoặc kịch bản đã được lập trình sẵn. Mỗi câu hỏi mà người dùng đưa ra sẽ được chatbot xử lý dựa trên các mẫu hội thoại có sẵn.
  + **Đặc điểm**:
    - Không thể học hỏi từ dữ liệu mới.
    - Khả năng xử lý giới hạn chỉ trong phạm vi các quy tắc đã được định nghĩa.
    - Hoạt động hiệu quả với các yêu cầu đơn giản hoặc các quy trình cố định.
  + **Ưu điểm**: Dễ xây dựng và chi phí thấp, phù hợp cho các doanh nghiệp vừa và nhỏ, các trang web thương mại điện tử với khối lượng công việc dịch vụ khách hàng đơn giản.
  + **Nhược điểm**: Không thể xử lý các câu hỏi phức tạp hoặc tình huống không dự đoán trước.
* **Chatbot AI (AI-powered Chatbot)**:
  + **Cách thức hoạt động**: Sử dụng trí tuệ nhân tạo (AI), học máy (Machine Learning) và xử lý ngôn ngữ tự nhiên (NLP), chatbot AI có khả năng học hỏi từ dữ liệu tương tác trước đó và tự động cải thiện phản hồi theo thời gian.
  + **Đặc điểm**:
    - Có khả năng xử lý các câu hỏi phức tạp hơn, linh hoạt trong nhiều ngữ cảnh khác nhau.
    - Tự động cải thiện nhờ vào quá trình học tập từ dữ liệu lịch sử.
    - Có khả năng cá nhân hóa phản hồi dựa trên hồ sơ và lịch sử tương tác của người dùng.
  + **Ưu điểm**: Hiệu quả cao, có thể xử lý nhiều yêu cầu phức tạp, nâng cao trải nghiệm khách hàng nhờ khả năng phản hồi tự nhiên.
  + **Nhược điểm**: Chi phí phát triển cao hơn, cần thời gian huấn luyện mô hình và nguồn tài nguyên tính toán lớn.

**Chatbot dựa trên học sâu (Deep Learning-based chatbot)**

* **Nguyên tắc hoạt động**: Chatbot dựa trên học sâu sử dụng các mô hình mạng nơ-ron sâu (deep neural networks) để tạo ra các phản hồi. Khác với các chatbot AI thông thường, mô hình học sâu có khả năng tự học từ dữ liệu mà không cần lập trình các quy tắc hoặc ngữ nghĩa cụ thể.
* **Cơ chế**: Chúng sử dụng các mô hình như **Transformer** (ví dụ GPT), **RNN**, **LSTM**, giúp chúng hiểu được cả ngữ cảnh dài hạn và ngữ cảnh ngắn hạn trong các cuộc hội thoại.
* **Ưu điểm**:
  + Phản hồi giống con người, tự nhiên hơn, đặc biệt trong các cuộc hội thoại dài.
  + Khả năng tự học mạnh mẽ từ dữ liệu mà không cần lập trình chi tiết.
* **Nhược điểm**:
  + Yêu cầu khối lượng dữ liệu rất lớn để huấn luyện mô hình.
  + Khó dự đoán hành vi vì các mô hình học sâu có thể tạo ra phản hồi không mong muốn nếu không được kiểm soát tốt.

**4. Các thành phần chính của Chatbot**

Chatbot hiện đại, đặc biệt là chatbot AI, được xây dựng từ các thành phần chính sau:

* **Xử lý ngôn ngữ tự nhiên (NLP - Natural Language Processing)**:
  + **NLP** : NLP là một nhánh của AI giúp chatbot hiểu được ngôn ngữ tự nhiên của con người, bao gồm việc phân tích từ vựng, cú pháp, ngữ nghĩa và ngữ cảnh trong một cuộc trò chuyện.
  + **Các bước của NLP trong chatbot**:
    1. **Phân tích ngữ pháp**: NLP phân tích cấu trúc câu để hiểu rõ các thành phần của ngữ pháp.
    2. **Nhận diện ý định (Intent Recognition)**: NLP giúp chatbot hiểu được mục đích của người dùng trong mỗi câu hỏi hoặc yêu cầu.
    3. **Nhận diện thực thể (Entity Recognition)**: Chatbot nhận diện các từ khóa quan trọng như tên, địa điểm, hoặc thời gian trong câu để cung cấp phản hồi chính xác.
  + **Tầm quan trọng**: NLP giúp chatbot giao tiếp tự nhiên và hiệu quả hơn, giảm thiểu các lỗi không hiểu ngữ nghĩa và giúp chatbot có thể tương tác một cách linh hoạt với người dùng.
* **Mô hình AI học sâu (Deep Learning)**:
  + **Mô hình AI học sâu** : Deep Learning là một kỹ thuật học máy sử dụng các mạng nơ-ron nhân tạo (Neural Networks) với nhiều lớp để phân tích và học từ dữ liệu lớn. Trong chatbot, Deep Learning giúp chatbot học hỏi từ các tương tác và cải thiện phản hồi qua thời gian.
  + **Ứng dụng trong chatbot**: Chatbot sử dụng mạng nơ-ron để xử lý các tín hiệu ngôn ngữ, phân tích ý nghĩa ngữ cảnh và học từ các tình huống phức tạp mà các chatbot dựa trên quy tắc không thể xử lý.
* **Cơ chế học máy (Machine Learning)**:
  + **Học máy** : là một phương pháp giúp chatbot học từ dữ liệu lịch sử và tự động điều chỉnh hành vi của mình để cải thiện khả năng trả lời trong tương lai.
  + **Ứng dụng trong chatbot**: Chatbot sử dụng học máy để tự động hóa quá trình phân tích các mẫu hội thoại và cải thiện phản hồi. Ví dụ, chatbot có thể học cách phản ứng với các yêu cầu phức tạp và tinh chỉnh trả lời dựa trên các dữ liệu tương tác thực tế.
  + **Các mô hình học máy phổ biến trong chatbot**:
    1. **Supervised Learning**: Chatbot được huấn luyện với các tập dữ liệu đã được gắn nhãn.
    2. **Unsupervised Learning**: Chatbot tự khám phá các mẫu trong dữ liệu mà không cần dữ liệu đã được gắn nhãn trước.
    3. **Reinforcement Learning**: Chatbot học từ phản hồi (feedback) sau mỗi cuộc trò chuyện để cải thiện hiệu suất tương lai.

**Chúng ta có thể thấy** :

* **Chatbot quy tắc** đơn giản, dễ triển khai nhưng giới hạn trong khả năng xử lý các câu hỏi phức tạp.
* **Chatbot AI** thông minh hơn, có thể học và cải thiện theo thời gian, xử lý các cuộc trò chuyện phức tạp hơn.
* **Chatbot học sâu** mạnh mẽ, tự học từ dữ liệu lớn, nhưng khó phát triển và yêu cầu tài nguyên tính toán cao.

Loại chatbot phù hợp sẽ phụ thuộc vào yêu cầu của doanh nghiệp và tính phức tạp của các cuộc hội thoại mà chatbot phải xử lý.

## 3**.KHẢO SÁT VÀ THU THẬP CÁC BIỂU MẪU**

**A. KHẢO SÁT VÀ PHÂN TÍCH**

Việc khảo sát và phân tích các hệ thống chatbot hiện có như **Google Assistant**, **Siri**, và **ChatGPT** giúp chúng ta hiểu rõ hơn về cách chúng hoạt động, từ đó có cơ sở để phát triển một hệ thống chatbot hiệu quả. Chúng ta sẽ cùng phân tích 1 chatbot **ChatGPT** để lấy ví dụ rõ hơn :

* **ChatGPT**, phát triển bởi OpenAI, là một mô hình ngôn ngữ dựa trên mạng nơ-ron sâu (Transformer), được huấn luyện trên khối lượng dữ liệu lớn để có khả năng tạo ra văn bản giống con người. ChatGPT không chỉ hiểu và trả lời câu hỏi mà còn có khả năng giữ được ngữ cảnh của các cuộc hội thoại dài.
* **Chức năng chính**:
  + **Xử lý ngôn ngữ tự nhiên**: ChatGPT sử dụng các mô hình học sâu để hiểu và phản hồi các câu hỏi từ người dùng với độ chính xác cao.
  + **Giữ ngữ cảnh**: Có khả năng duy trì các cuộc hội thoại dài và phức tạp, có thể trả lời các câu hỏi mang tính tổng hợp thông tin từ nhiều câu hỏi trước đó.
  + **Tạo văn bản phức tạp**: ChatGPT có thể tạo văn bản trong nhiều thể loại khác nhau như tóm tắt, sáng tạo nội dung, hoặc viết mã.
* **Ưu điểm**:
  + Xử lý ngôn ngữ tự nhiên rất tốt, phản hồi giống con người và có thể trả lời các câu hỏi phức tạp.
  + Có khả năng học từ nhiều nguồn và đưa ra câu trả lời thông minh hơn qua thời gian.
* **Nhược điểm**:
  + Đôi khi có thể tạo ra thông tin sai hoặc không nhất quán.
  + Cần tài nguyên tính toán lớn để vận hành.

**B. THU THẬP BIỂU MẪU**

Trong quá trình phát triển chatbot, việc thu thập các biểu mẫu tương tác của người dùng là bước quan trọng để có thể thiết kế hệ thống phù hợp và hiệu quả. Các biểu mẫu tương tác này sẽ cung cấp thông tin về hành vi người dùng, loại truy vấn và yêu cầu từ phía người dùng, từ đó giúp cải thiện hệ thống chatbot.

**a. Dữ liệu đầu vào từ người dùng**

* **Loại truy vấn**: Dữ liệu đầu vào từ người dùng bao gồm các truy vấn dưới dạng văn bản hoặc giọng nói. Các truy vấn này thường thuộc nhiều dạng khác nhau, từ các câu hỏi đơn giản như "Thời tiết hôm nay thế nào?" đến các yêu cầu phức tạp hơn như "Tôi cần tìm kiếm chuyến bay giá rẻ đến New York vào tuần tới."
* **Ngữ cảnh**: Một số cuộc trò chuyện có thể có ngữ cảnh phức tạp, ví dụ người dùng hỏi nhiều câu liên quan đến cùng một chủ đề. Hệ thống cần hiểu ngữ cảnh của các câu hỏi liên tiếp này để đưa ra câu trả lời phù hợp.
* **Biến thể ngôn ngữ**: Người dùng thường sử dụng nhiều cách diễn đạt khác nhau cho cùng một mục đích. Ví dụ: “Tôi muốn đặt vé máy bay” hoặc “Giúp tôi đặt vé bay” đều có cùng mục đích nhưng cách diễn đạt khác nhau. Điều này đòi hỏi chatbot cần xử lý và hiểu đúng ý định của người dùng.

**b. Quy trình xử lý ngôn ngữ tự nhiên (NLP)**

Dữ liệu đầu vào từ người dùng sẽ được xử lý thông qua quy trình **xử lý ngôn ngữ tự nhiên**. Các bước trong quy trình này bao gồm:

* **Phân tích cú pháp**: Tách câu của người dùng thành các từ hoặc cụm từ để dễ dàng xử lý.
* **Nhận diện thực thể**: Xác định và trích xuất các thông tin quan trọng từ truy vấn như tên người, thời gian, địa điểm.
* **Nhận dạng ý định**: Sử dụng các thuật toán học máy hoặc học sâu để xác định ý định chính của người dùng, từ đó chọn ra câu trả lời phù hợp.

Ví dụ: Nếu người dùng nhập câu hỏi "Tôi muốn mua vé xem phim tối nay", hệ thống sẽ nhận diện:

* **Thực thể**: "vé xem phim", "tối nay"
* **Ý định**: "mua vé xem phim"

**c. Thu thập và phân tích các biểu mẫu**

* **Biểu mẫu truy vấn**: Bao gồm các câu hỏi phổ biến từ người dùng, dạng câu hỏi, ngữ cảnh của câu hỏi và cách chatbot trả lời.
* **Biểu mẫu phản hồi**: Ghi nhận phản hồi của người dùng sau khi nhận được câu trả lời từ chatbot. Thông tin phản hồi giúp hệ thống cải thiện và học từ dữ liệu này.

Các biểu mẫu có thể được thu thập thông qua hệ thống logs, phân tích dữ liệu từ cuộc hội thoại và phản hồi của người dùng. Điều này giúp hệ thống chatbot không ngừng cải thiện khả năng xử lý và cung cấp phản hồi chính xác hơn.

## 4.**YÊU CẦU CHỨC NĂNG**

**A. Các chức năng cơ bản**

**a. Nhận dạng ý định (Intent Recognition)**

Nhận dạng ý định là chức năng quan trọng nhất của một chatbot AI. Ý định (intent) ở đây ám chỉ mục đích chính của người dùng khi họ đặt ra một câu hỏi hoặc truy vấn.

* **Chức năng chính**: Chatbot cần có khả năng phân tích và xác định đúng ý định của người dùng từ các câu hỏi hoặc mệnh lệnh. Ví dụ, nếu người dùng hỏi "Tôi muốn đặt vé máy bay", chatbot phải hiểu rằng ý định của người dùng là **đặt vé** và thực hiện các thao tác liên quan đến việc tìm và đặt vé.
* **Cơ chế**:
  + Chatbot sử dụng xử lý ngôn ngữ tự nhiên (NLP) kết hợp với các thuật toán học máy hoặc học sâu để phân tích văn bản và xác định các ý định từ các từ khóa và ngữ cảnh.
  + Hệ thống cần một tập huấn luyện có các câu truy vấn và các nhãn gắn kèm (label) để huấn luyện mô hình hiểu và phân loại ý định của người dùng.
* **Ví dụ**: Người dùng có thể hỏi:
  + “Mở đèn phòng khách” → Chatbot nhận dạng ý định **bật đèn thông minh**.
  + “Cho tôi biết thời tiết hôm nay” → Chatbot nhận dạng ý định **tra cứu thông tin thời tiết**.

**b. Trả lời tự động (Automated Response)**

Sau khi nhận dạng đúng ý định, chatbot cần đưa ra câu trả lời tự động dựa trên dữ liệu hiện có hoặc thông qua việc gọi các dịch vụ bên ngoài.

* **Chức năng chính**: Chatbot sẽ cung cấp phản hồi trực tiếp cho người dùng dưới dạng văn bản hoặc giọng nói, tùy thuộc vào nền tảng và thiết kế.
* **Cơ chế**:
  + **Dữ liệu cứng (hard-coded)**: Chatbot có thể trả lời dựa trên các kịch bản hoặc mẫu câu trả lời đã được lập trình sẵn trong hệ thống.
  + **Dữ liệu động**: Chatbot có thể lấy dữ liệu từ các API bên ngoài (ví dụ như API thời tiết, hệ thống đặt vé máy bay, dịch vụ ngân hàng) để đưa ra phản hồi tương ứng.
  + **Mô hình tạo văn bản**: Sử dụng các mô hình AI (như GPT) để tạo câu trả lời từ các đoạn văn bản ngẫu nhiên dựa trên ngữ cảnh người dùng đưa ra.
* **Ví dụ**:
  + Khi người dùng hỏi: "Thời tiết hôm nay thế nào?", chatbot sẽ gọi API thời tiết để trả lời chi tiết.
  + Nếu người dùng yêu cầu: "Hãy tìm nhà hàng gần đây", chatbot có thể trả lời dựa trên dữ liệu vị trí và danh sách nhà hàng từ Google Maps.

**c. Học từ dữ liệu (Learning from Data)**

Chatbot AI cần có khả năng học từ các tương tác của người dùng để ngày càng trở nên thông minh và hiệu quả hơn. Điều này liên quan đến việc phân tích các cuộc hội thoại trước đó và áp dụng các thuật toán học máy để cải thiện các phản hồi trong tương lai.

* **Chức năng chính**: Học từ dữ liệu giúp chatbot tối ưu hóa cách xử lý câu hỏi và trả lời. Nó có thể ghi nhớ các yêu cầu trước đó của người dùng, đưa ra các gợi ý hoặc câu trả lời chính xác hơn khi cùng một người dùng đặt câu hỏi lại lần nữa.
* **Cơ chế**:
  + **Học tăng cường (Reinforcement Learning)**: Chatbot có thể cải thiện bằng cách ghi nhận các phản hồi từ người dùng như "Có hữu ích không?" và điều chỉnh câu trả lời dựa trên phản hồi đó.
  + **Học dựa trên dữ liệu người dùng**: Chatbot lưu trữ lịch sử các cuộc hội thoại, phân tích dữ liệu đó để hiểu các xu hướng, từ đó tối ưu hóa cách nó hiểu và xử lý các câu hỏi.
* **Ví dụ**:
  + Nếu người dùng thường xuyên hỏi về các nhà hàng kiểu Ý, chatbot sẽ ưu tiên các câu trả lời liên quan đến nhà hàng Ý khi người dùng đặt câu hỏi tương tự lần sau.

**B. Yêu cầu phi chức năng**

Ngoài các chức năng chính của chatbot, cần xác định các yêu cầu phi chức năng quan trọng, bao gồm **hiệu suất**, **tính bảo mật**, và **khả năng mở rộng**.

#### a. ****Yêu cầu về hiệu suất****

Hiệu suất của chatbot liên quan đến khả năng xử lý nhanh chóng và chính xác các câu hỏi của người dùng.

* **Độ trễ thấp (Low Latency)**: Chatbot cần phản hồi nhanh chóng, thường là trong vòng vài giây sau khi người dùng đặt câu hỏi. Độ trễ càng thấp, trải nghiệm người dùng càng tốt.
* **Khả năng xử lý đồng thời (Concurrency)**: Chatbot cần có khả năng xử lý nhiều yêu cầu từ nhiều người dùng cùng một lúc mà không bị giảm hiệu suất. Điều này yêu cầu hệ thống cần được thiết kế để chịu tải tốt (scalable).
* **Tối ưu hóa mô hình**: Đối với các chatbot sử dụng mô hình học sâu, việc tối ưu hóa mô hình để giảm thời gian tính toán là rất quan trọng. Việc triển khai các mô hình nặng như GPT yêu cầu tài nguyên tính toán lớn, do đó cần có phương án để giảm thiểu tài nguyên sử dụng mà vẫn đảm bảo hiệu quả.

#### b. ****Yêu cầu về tính bảo mật****

Vì chatbot thường tương tác với dữ liệu cá nhân của người dùng, các yêu cầu về bảo mật là tối quan trọng.

* **Bảo vệ dữ liệu người dùng**: Dữ liệu cá nhân của người dùng (như thông tin tài khoản, địa chỉ, số điện thoại) cần được mã hóa để tránh bị lộ hoặc khai thác trái phép.
* **Xác thực và quyền truy cập**: Chatbot cần đảm bảo rằng chỉ có những người dùng được ủy quyền mới có thể truy cập vào các dịch vụ nhạy cảm hoặc thực hiện các hành động quan trọng (ví dụ, chuyển tiền qua chatbot ngân hàng).
* **Ngăn chặn tấn công**: Hệ thống cần có các biện pháp bảo vệ chống lại các cuộc tấn công phổ biến như **tấn công từ chối dịch vụ (DDoS)**, **tấn công SQL Injection**, và **tấn công giả mạo (Phishing)**.

#### c. ****Yêu cầu về khả năng mở rộng (Scalability)****

Khả năng mở rộng là yếu tố quyết định đến việc hệ thống chatbot có thể phục vụ cho nhiều người dùng cùng một lúc và trong thời gian dài hay không.

* **Tăng trưởng người dùng**: Hệ thống cần được thiết kế để mở rộng khi số lượng người dùng tăng lên, đồng thời không làm giảm hiệu suất. Điều này đòi hỏi cơ sở hạ tầng phải linh hoạt, có thể mở rộng theo chiều ngang (thêm nhiều máy chủ) hoặc chiều dọc (tăng tài nguyên máy chủ).
* **Triển khai đám mây (Cloud deployment)**: Sử dụng các dịch vụ đám mây (cloud services) như **AWS**, **Google Cloud**, **Azure** để giúp hệ thống mở rộng dễ dàng và nhanh chóng khi có nhu cầu.
* **Chia nhỏ dịch vụ (Microservices architecture)**: Một kiến trúc microservices có thể giúp chatbot dễ dàng mở rộng từng phần riêng lẻ mà không ảnh hưởng đến toàn bộ hệ thống. Ví dụ, dịch vụ xử lý ngôn ngữ tự nhiên có thể được tách riêng và mở rộng độc lập với các thành phần khác.

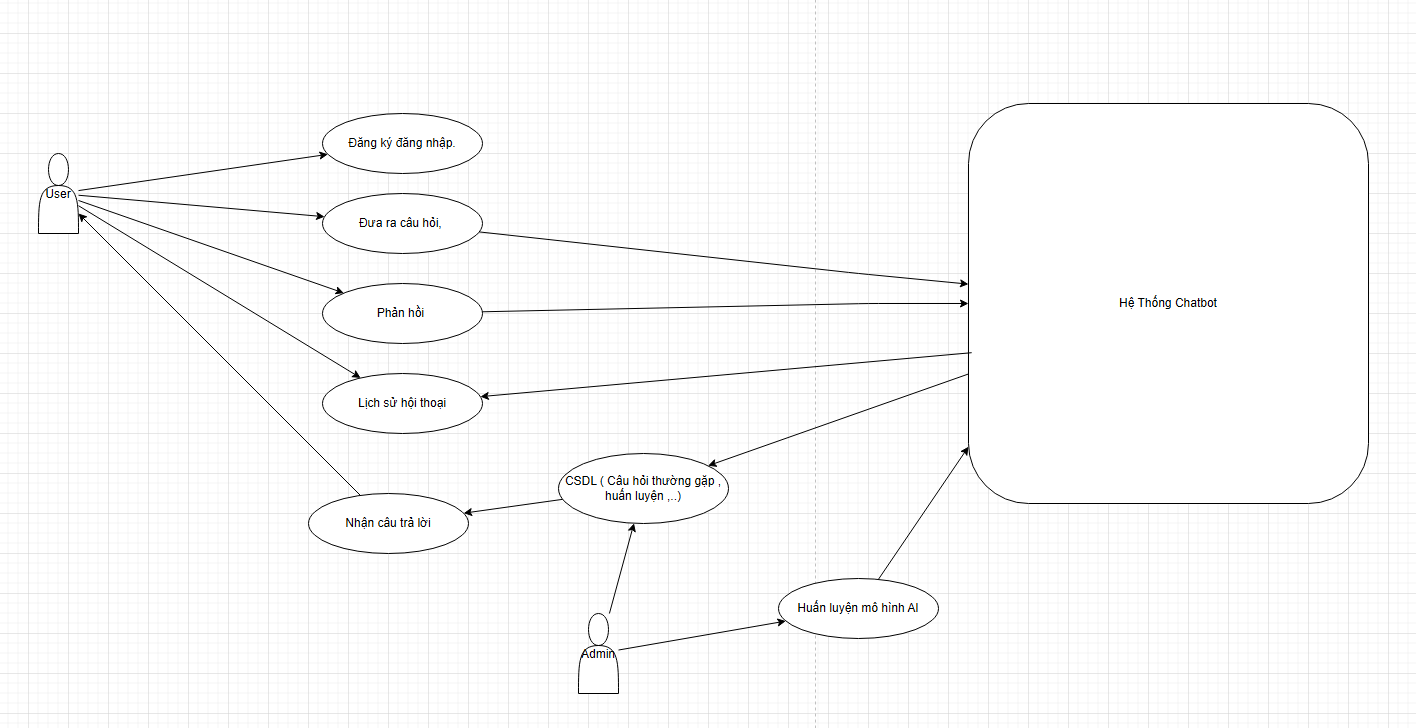
**Từ đó ta có thể thấy như sau** :

 **Chức năng nhận dạng ý định**, **trả lời tự động**, và **học từ dữ liệu** là những chức năng chính của chatbot AI, giúp chatbot trở nên thông minh và đáp ứng yêu cầu của người dùng.

 **Hiệu suất**, **bảo mật**, và **khả năng mở rộng** là những yêu cầu phi chức năng quan trọng, đảm bảo rằng hệ thống chatbot hoạt động ổn định, an toàn và có thể phục vụ số lượng lớn người dùng mà không bị gián đoạn.

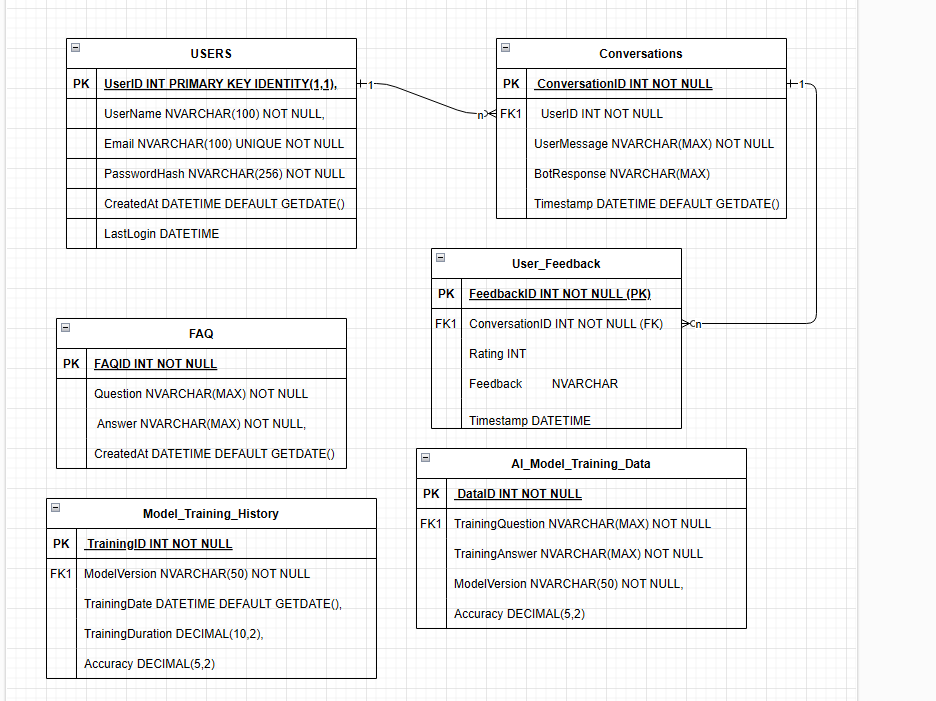
# PHẦN II : PHÂN TÍCH VÀ THIẾT KẾ BÀI TOÁN

## 1.Use-case Diagram

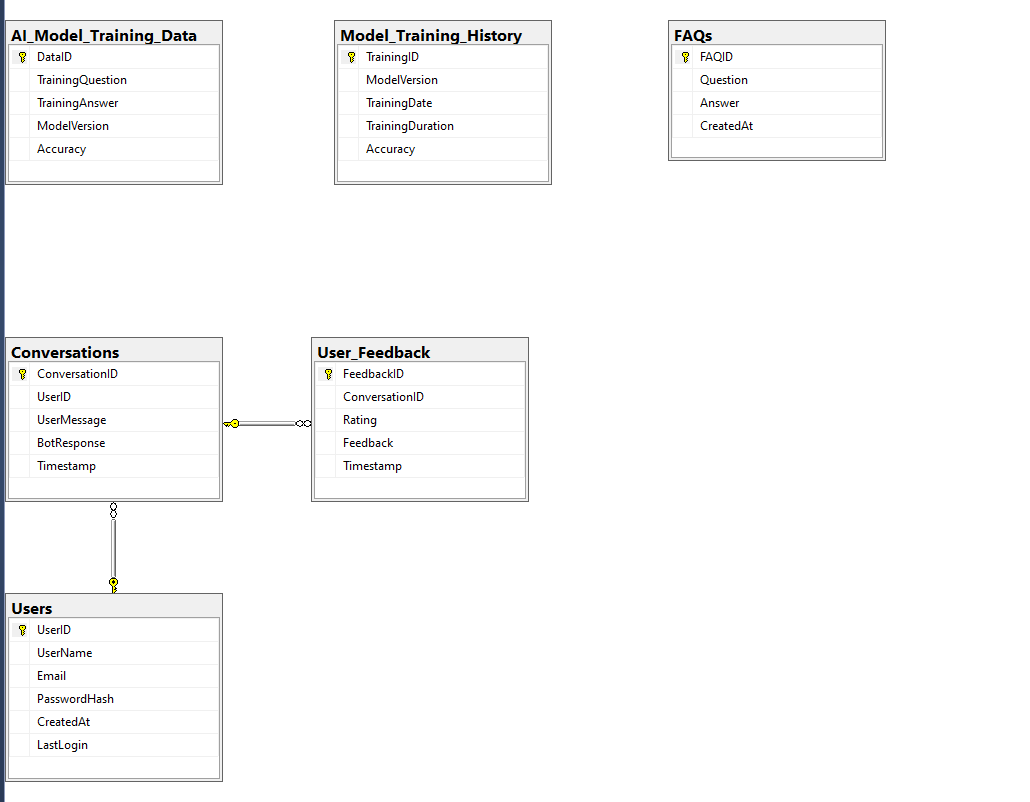
****

## 2.Thiết kế cơ sở dữ liệu

### A.Mô hình ERD

****

### B.Database Diagram



### C.Mô tả dữ liệu

 **Loại dữ liệu sử dụng**:

* **Dữ liệu ngôn ngữ**: Văn bản, giọng nói được sử dụng làm dữ liệu đầu vào cho chatbot.
* **Dữ liệu cảm xúc**: Chatbot có thể sử dụng các mô hình nhận diện cảm xúc từ văn bản hoặc giọng nói để điều chỉnh cách trả lời phù hợp.
* **Dữ liệu ngữ cảnh**: Thông tin về ngữ cảnh (context) từ cuộc trò chuyện trước đó giúp chatbot hiểu rõ hơn nhu cầu của người dùng.

 **Cách xử lý dữ liệu**:

* **Xử lý ngôn ngữ tự nhiên (NLP)**: Giải thích cách dữ liệu ngôn ngữ được xử lý thông qua các bước như phân tích cú pháp, nhận dạng thực thể, và tạo câu trả lời tự nhiên.
* **Học sâu (Deep Learning)**: Sử dụng các mô hình học sâu để chatbot có thể hiểu và xử lý các câu hỏi phức tạp hơn. Trình bày cách mô hình học máy sử dụng dữ liệu đào tạo để tạo ra kết quả tốt hơn.
* **Quy trình làm sạch dữ liệu**: Loại bỏ dữ liệu nhiễu, thiếu sót và tổ chức lại dữ liệu để đảm bảo chatbot hoạt động hiệu quả.

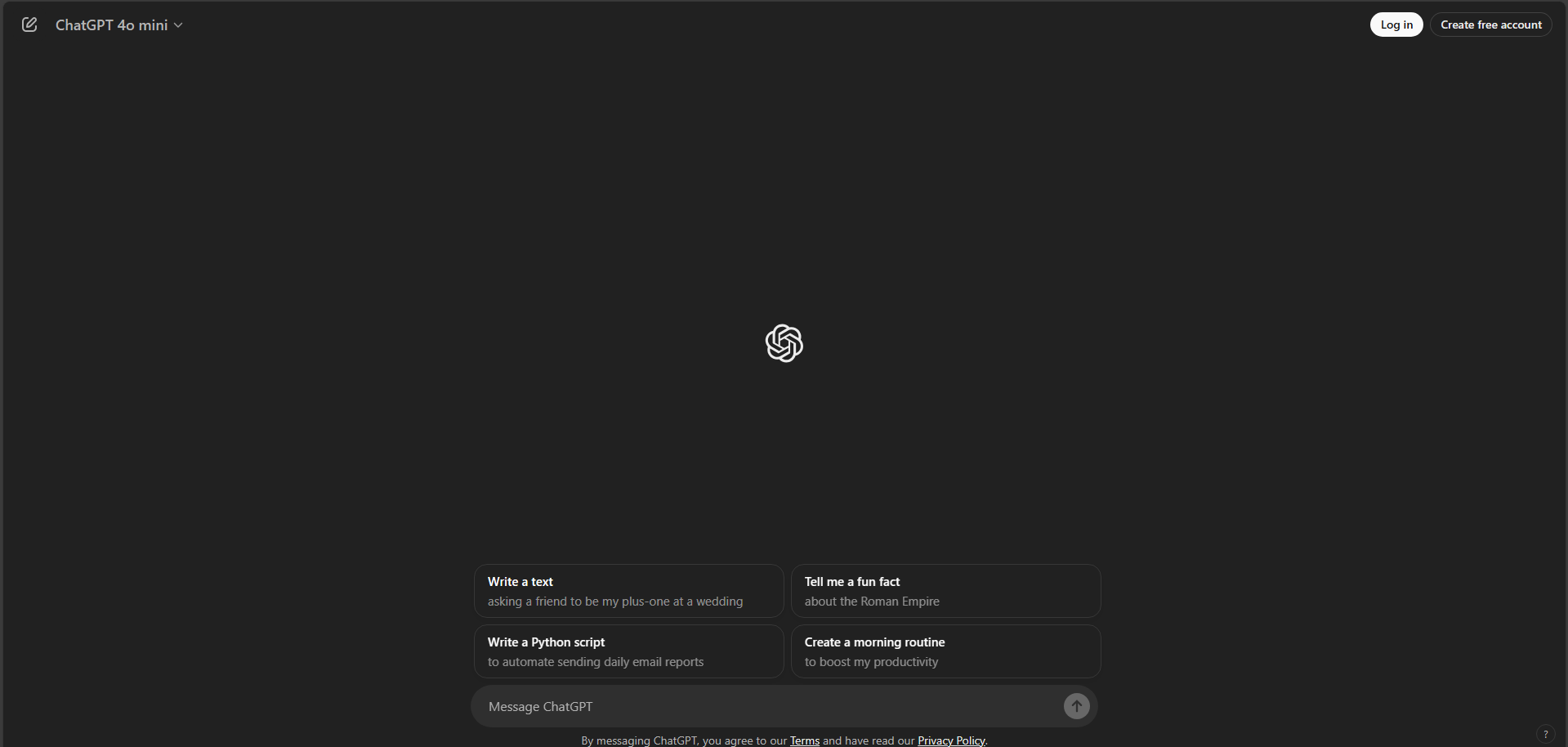
 **Mô hình học máy (Machine Learning Models)**:

* Mô tả các mô hình học máy được sử dụng để phân tích dữ liệu và dự đoán phản hồi, ví dụ: các mô hình GPT, Transformer, hoặc RNN.
* Sử dụng các biểu đồ để minh họa cách dữ liệu được truyền qua các lớp của mô hình học máy, từ đó tạo ra các kết quả đầu ra phù hợp với yêu cầu người dùng.

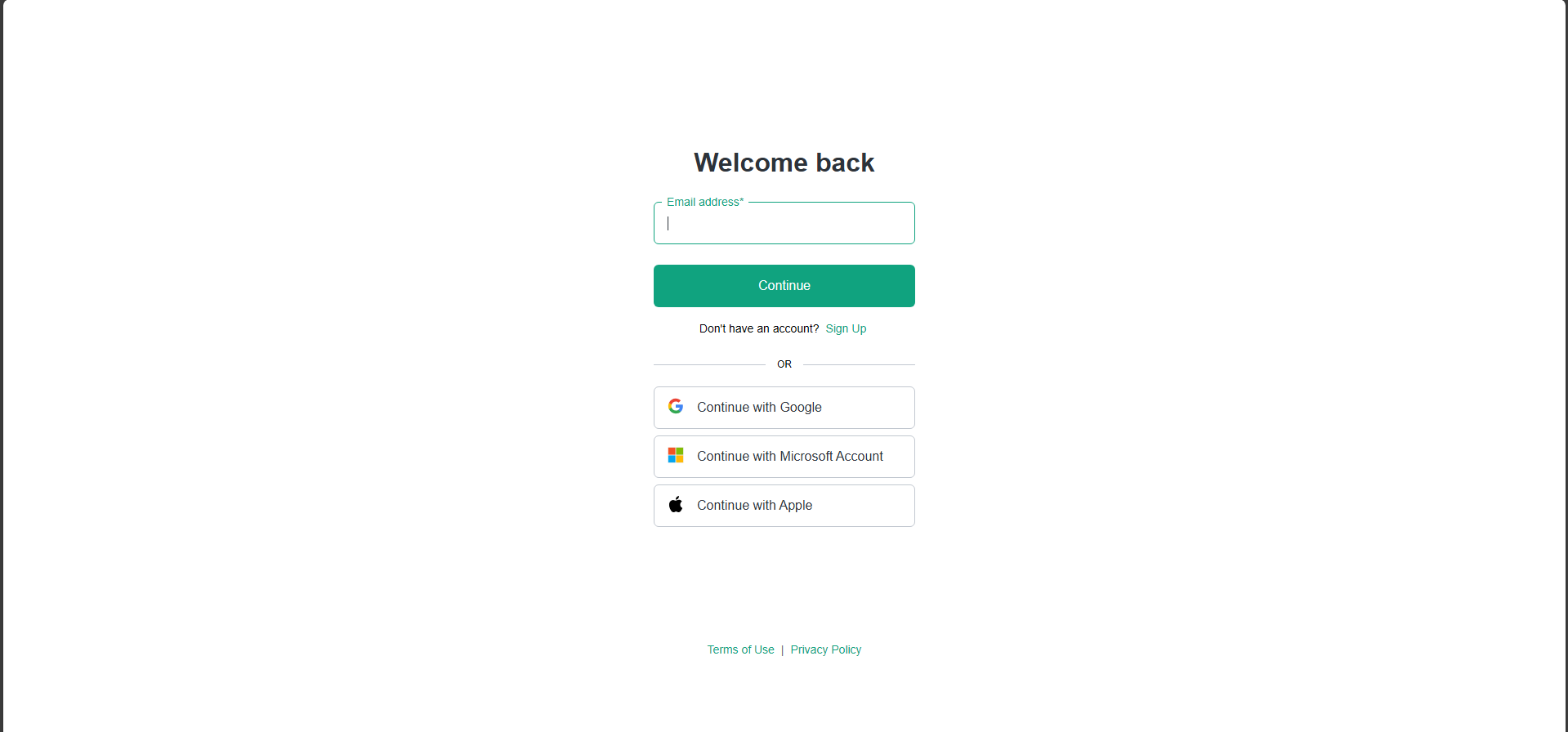
# PHẦN III : THIẾT KẾ HỆ THỐNG

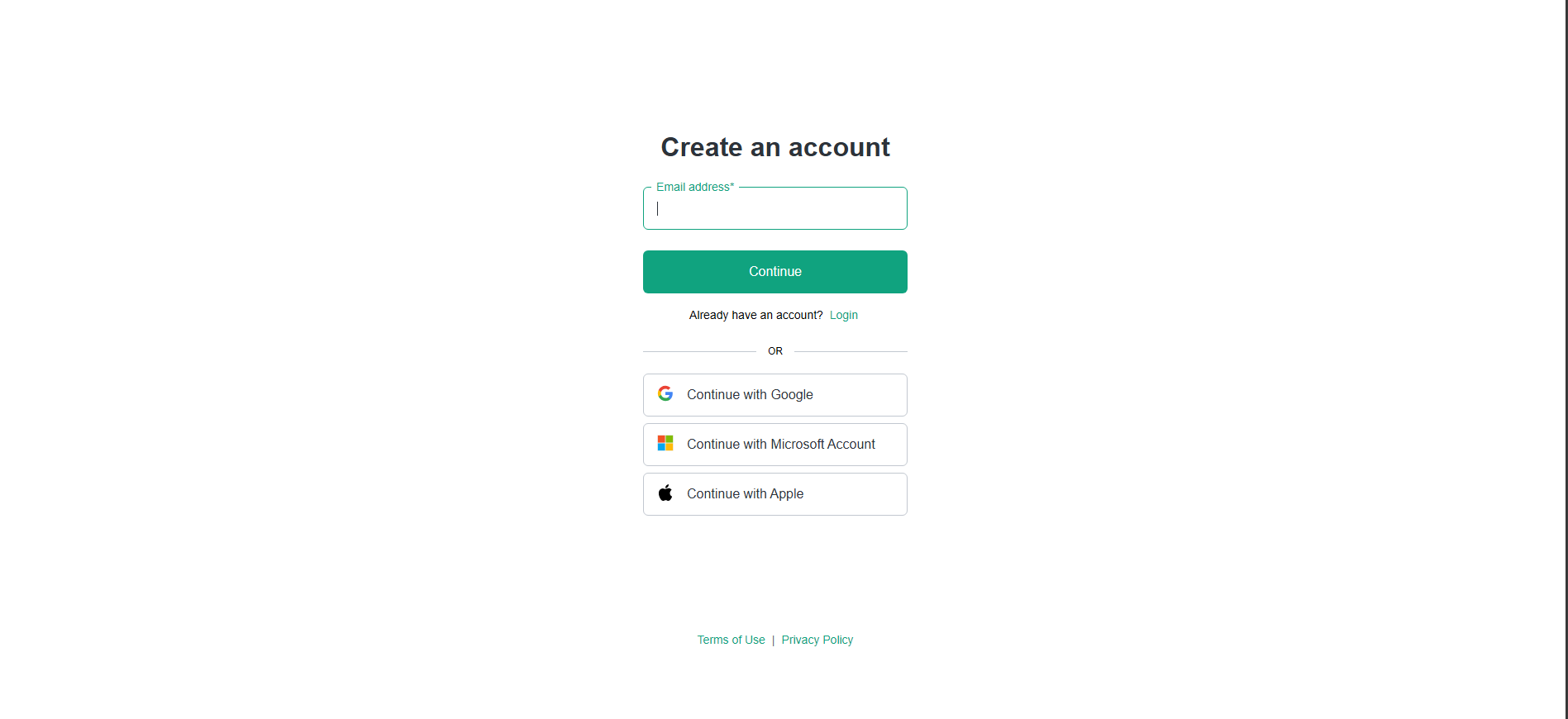
## 1.Giao diện và các chức năng

* 1. **Trang chủ**

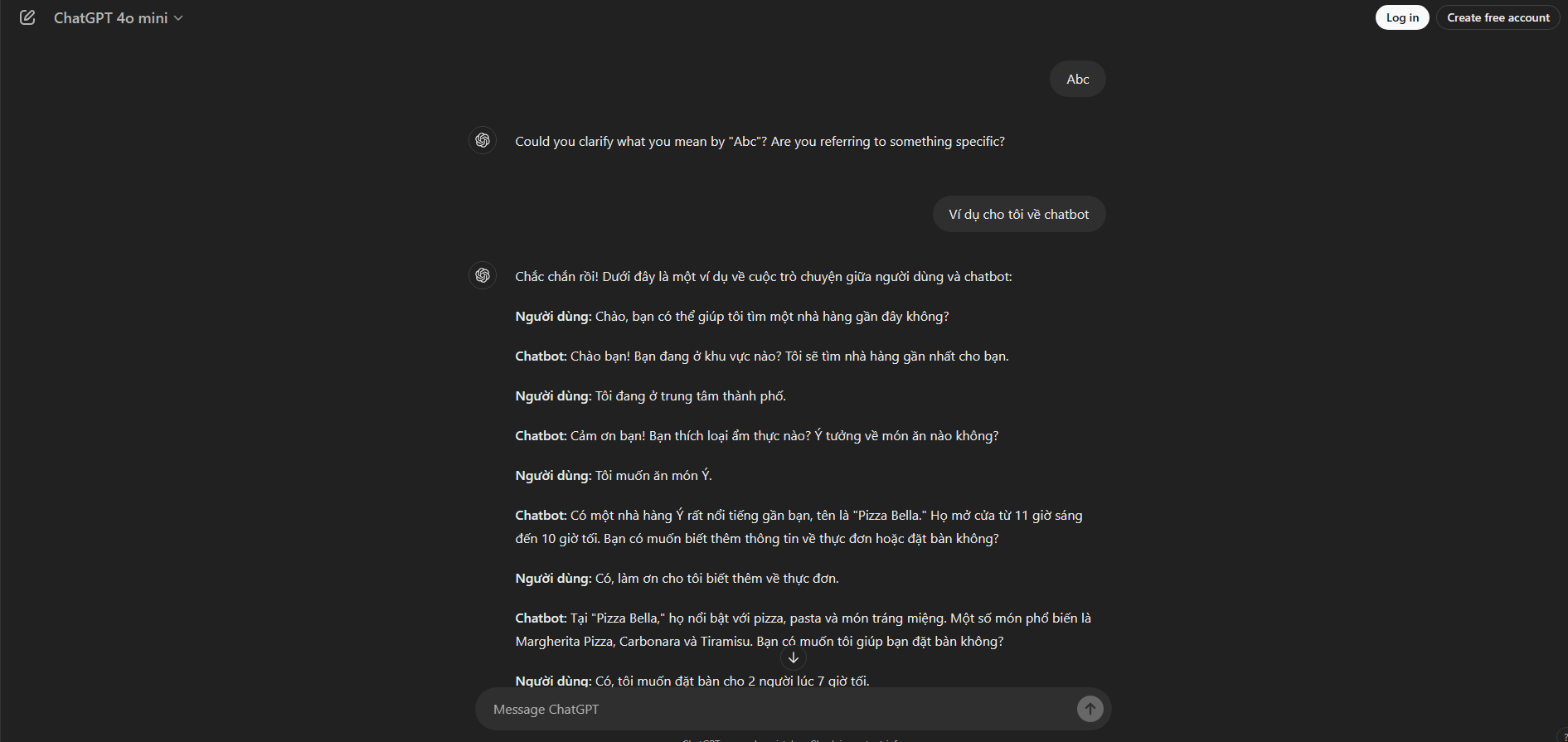
****

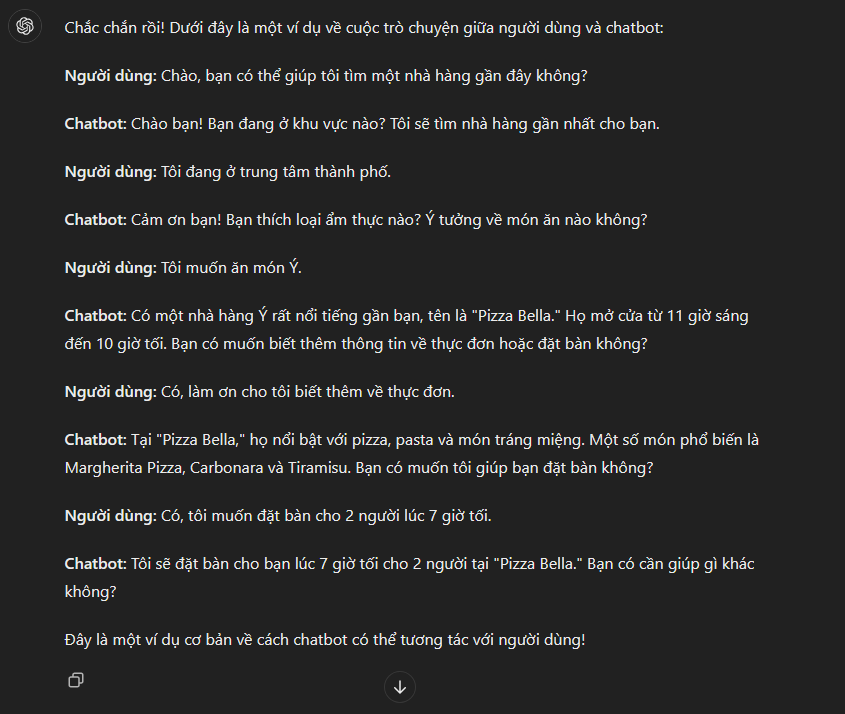
* 1. **Đăng ký đăng nhập**

****

****

* 1. **Sử dụng chatbot**

****

* 1. **Chatbot trả lời **

# PHẦN IV : KIỂM THỬ

**1.** Kiểm thử chức năng

**1.1. Kiểm thử khả năng nhận diện ý định:**

* **Mục tiêu:** Xác định xem ChatGPT có thể hiểu đúng ý định của người dùng từ các câu hỏi khác nhau hay không.
* **Cách thực hiện:**
  + Tạo một loạt câu hỏi mẫu với các ý định khác nhau (ví dụ: hỏi về thời tiết, yêu cầu giải thích một khái niệm, đặt lịch hẹn).
  + Ghi nhận phản hồi và phân tích độ chính xác của chúng.

**1.2. Kiểm thử phản hồi:**

* **Mục tiêu:** Đánh giá tính hợp lệ và sự chính xác trong các phản hồi của ChatGPT.
* **Cách thực hiện:**
  + Đặt các câu hỏi thông thường và kiểm tra xem câu trả lời có phù hợp không.
  + So sánh phản hồi với thông tin chính xác hoặc mong đợi.

**2.** Kiểm thử hiệu suất

**2.1. Đo độ trễ phản hồi:**

* **Mục tiêu:** Đo thời gian ChatGPT cần để đưa ra phản hồi cho người dùng.
* **Cách thực hiện:**
  + Sử dụng nhiều yêu cầu đồng thời từ nhiều nguồn khác nhau.
  + Ghi nhận thời gian phản hồi và phân tích hiệu suất dưới tải cao.

**2.2. Kiểm tra khả năng mở rộng:**

* **Mục tiêu:** Đánh giá khả năng xử lý của ChatGPT khi có nhiều người dùng tương tác đồng thời.
* **Cách thực hiện:**
  + Mô phỏng nhiều người dùng (sử dụng công cụ kiểm thử tự động như Locust hoặc JMeter).
  + Theo dõi hiệu suất và độ trễ phản hồi khi số lượng người dùng tăng lên.

**3.** Kiểm thử bảo mật

**3.1. Kiểm thử bảo vệ thông tin cá nhân:**

* **Mục tiêu:** Đảm bảo rằng ChatGPT không lưu trữ hoặc tiết lộ thông tin cá nhân của người dùng.
* **Cách thực hiện:**
  + Gửi các câu hỏi có chứa thông tin cá nhân và theo dõi cách ChatGPT xử lý chúng.
  + Đảm bảo rằng không có thông tin nhạy cảm nào được ghi lại hoặc chia sẻ.

**3.2. Kiểm tra lỗ hổng bảo mật:**

* **Mục tiêu:** Đánh giá khả năng của ChatGPT trước các tấn công tiềm năng.
* **Cách thực hiện:**
  + Thực hiện các thử nghiệm tấn công thông qua các phương pháp như injection (tiêm mã) để kiểm tra cách hệ thống phản ứng.

**4.** Theo dõi chất lượng

**4.1. Ghi nhận và phân tích dữ liệu:**

* **Mục tiêu:** Theo dõi các chỉ số chất lượng của ChatGPT theo thời gian.
* **Cách thực hiện:**
  + Thiết lập các công cụ theo dõi để ghi nhận các phản hồi, độ chính xác và thời gian phản hồi.
  + Phân tích các dữ liệu này để xác định xu hướng và cải tiến.

**4.2. Cải tiến liên tục:**

* **Mục tiêu:** Tối ưu hóa chatbot dựa trên dữ liệu thu thập được từ quá trình kiểm thử.
* **Cách thực hiện:**
  + Thực hiện các điều chỉnh và cập nhật định kỳ dựa trên phản hồi của người dùng và dữ liệu kiểm thử.
  + Lập kế hoạch cho các phiên bản mới để cải thiện tính năng và hiệu suất.

# PHẦN V : KẾT LUẬN

Việc phát triển công nghệ AI ứng dụng cho chatbot không chỉ mang lại lợi ích trong việc nâng cao hiệu suất làm việc mà còn mở ra nhiều cơ hội mới cho các doanh nghiệp và tổ chức. Để thực hiện điều này, cần có sự đầu tư liên tục vào nghiên cứu và phát triển cũng như sự hợp tác giữa các nhà nghiên cứu, nhà phát triển và doanh nghiệp để tối ưu hóa hiệu quả của chatbot trong các lĩnh vực khác nhau.Qua bài nghiên cứu trên , tôi và các bạn , các thầy cô chắc hẳn đã có thêm hiểu biết , tầm quan trọng của việc ứng dụng AI vào chatbot hiện nay.

# PHẦN V : TÀI LIỆU THAM KHẢO

[**Nghiên cứu các yếu tố tác động đến sự chấp nhận Chatbot AI của khách hàng tại một số ngân hàng thương mại (tapchinganhang.gov.vn)**](https://tapchinganhang.gov.vn/nghien-cuu-cac-yeu-to-tac-dong-den-su-chap-nhan-chatbot-ai-cua-khach-hang-tai-mot-so-ngan-hang-thuon.htm)

[**CVv489S342020032.pdf (thuvienlamdong.org.vn)**](http://thuvienlamdong.org.vn:81/bitstream/DL_134679/19196/1/CVv489S342020032.pdf)

Và một số video , tài liệu khác.