

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



PHAN QUANG THÀNH

**NGHIÊN CỨU XÂY DỰNG CHATBOT TƯ VẤN, HỖ TRỢ
NHẬP HỌC TẠI HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH
VIỄN THÔNG**

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

HÀ NỘI - NĂM 2024

Đề án tốt nghiệp được hoàn thành tại:
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS. TRẦN TIẾN CÔNG
(Ghi rõ học hàm, học vị)

Phản biện 1: PGS.TS. Nguyễn Mạnh Hùng

Phản biện 2: PGS.TS. Đỗ Trung Tuấn

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp
thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: 10 giờ 45 ngày 20 tháng 03 năm 2024

Có thể tìm hiểu đề án tốt nghiệp tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

MỞ ĐẦU

1. Tính cấp thiết của đề tài:

Nhu cầu chatbot hiện nay trong tuyển sinh đang ngày càng tăng cao. Điều này là do những lợi ích mà chatbot mang lại như tiết kiệm thời gian và chi phí bằng cách chatbot có thể tự động trả lời các câu hỏi thường gặp của thí sinh, giúp các cơ sở giáo dục tiết kiệm thời gian và chi phí cho nhân viên tuyển sinh, tăng cường tương tác với thí sinh bằng cách chatbot có thể giao tiếp với thí sinh 24/7.

Dưới đây là một số nhu cầu cụ thể của chatbot trong tuyển sinh hiện nay:

- Trả lời các câu hỏi thường gặp của thí sinh: Chatbot có thể được sử dụng để trả lời các câu hỏi thường gặp của thí sinh về quy chế tuyển sinh, phương thức tuyển sinh, các ngành nghề đào tạo,...
- Hỗ trợ đăng ký tuyển sinh: Chatbot có thể được sử dụng để hỗ trợ thí sinh đăng ký tuyển sinh, bao gồm nhập thông tin, nộp hồ sơ,...
- Tư vấn tuyển sinh: Chatbot có thể được sử dụng để tư vấn tuyển sinh cho thí sinh, giúp thí sinh lựa chọn ngành nghề phù hợp với bản thân.
- Hỗ trợ nhập học: Hướng dẫn hồ sơ, giấy tờ cần chuẩn bị, ngày nhập học...

Với mục đích đưa những tiến bộ công nghệ vào phục vụ cho công tác tư vấn, hỗ trợ nhập học tại Học viện ta, học viên xin chọn đề tài nghiên cứu “*Nghiên cứu và xây dựng chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông*”.

2. Tổng quan về vấn đề nghiên cứu:

Với hướng nghiên cứu và xây dựng chatbot tư vấn, hỗ trợ nhập học, nhiều trường đại học trên thế giới đã thành công và ứng dụng vào trong công tác tuyển sinh của mình như Đại học Harvard, UCAS hay Naviance.

Một số cơ sở giáo dục tại Việt Nam đã triển khai thành công chatbot tuyển sinh, bao gồm:

- Trường Đại học Kinh tế Quốc dân: NEU-Chatbot là chatbot tuyển sinh của Trường Đại học Kinh tế Quốc dân, có thể trả lời hơn 50 loại câu hỏi với độ chính xác lên đến 97,1%.
- Đại học FPT: FPT Chatbot là chatbot tuyển sinh của Đại học FPT, có thể hỗ trợ thí sinh đăng ký tuyển sinh, tư vấn tuyển sinh,...
- Đại học Bách khoa Hà Nội: Bách khoa Chatbot là chatbot tuyển sinh của Đại học Bách khoa Hà Nội, có thể trả lời các câu hỏi về quy chế tuyển sinh, phương

thức tuyển sinh,...

- Ngoài ra còn rất nhiều các nghiên cứu về chatbot tư vấn, hỗ trợ nhập học khác.

Các nghiên cứu trên cho thấy trong tương lai, chatbot tuyển sinh sẽ tiếp tục được ứng dụng rộng rãi hơn nữa, giúp các cơ sở giáo dục nâng cao hiệu quả công tác tuyển sinh.

3. Mục đích nghiên cứu:

Nghiên cứu và xây dựng chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông.

4. Đối tượng và phạm vi nghiên cứu:

Luận văn tập trung vào nghiên cứu phương pháp xây dựng chatbot hiện nay, từ đó đưa ra lựa chọn và xây dựng chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông.

Trong đó, phạm vi nghiên cứu ở đây chỉ tập trung vào xây dựng NLU gồm xử lý ngôn ngữ tự nhiên, phân loại ý định và xây dựng giao diện hỗ trợ xây dựng các luồng kịch bản giữa người dùng và bot.

5. Phương pháp nghiên cứu:

- Tìm hiểu về chatbot, nhu cầu sử dụng hiện nay.
- Nghiên cứu các phương pháp công nghệ, kỹ thuật về chatbot hiện nay.
- Ứng dụng vào xây dựng chatbot hỗ trợ, tư vấn tuyển sinh tại học viện.
- Thực nghiệm, phân tích và đánh giá hệ thống chatbot.
- Triển khai lên facebook, website học viện.

CHƯƠNG 1: TỔNG QUAN VỀ HỆ THỐNG CHATBOT

1.1 Khái niệm chatbot

Chatbot là một dạng ứng dụng phần mềm nhằm mô phỏng giao tiếp với con người trong việc thực hiện một loạt các tác vụ hoặc giải đáp thắc mắc. Chatbot có thể được lập trình sử dụng trí tuệ nhân tạo (AI) để hiểu và phản hồi một cách linh hoạt hơn, trong khi một số khác có thể hoạt động dựa trên một bộ quy tắc cố định.

Có nhiều dạng Chatbot khác nhau, từ những Chatbot đơn giản chỉ có khả năng phản hồi các câu hỏi đã được lập trình sẵn, đến những Chatbot phức tạp sử dụng trí tuệ nhân tạo (AI), học máy (ML) và xử lý ngôn ngữ tự nhiên (NLP) để tiếp nhận thông tin, phân tích câu hỏi và phản hồi chính xác những gì mà người dùng mong muốn, ngoài ra Chatbot còn có khả năng tự học hỏi từ chính những cuộc trò chuyện với người dùng để đưa ra câu trả lời ngoài phân vùng dữ liệu được lập đi lập lại nhiều lần.

1.2 Đặc trưng của công tác tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông

Công tác tư vấn, hỗ trợ nhập học tại Học viện có một số điểm đặc trưng sau:

- Tổ chức nhập học cho tân sinh viên thường diễn ra trong thời gian ngắn (1-2 ngày).
- Cung cấp các bộ tài liệu như sách hướng dẫn nhập học, đến các video giới thiệu chuyên ngành và cơ sở vật chất.
- Tổ chức các buổi định hướng, tour tham quan trường, và buổi giới thiệu, giao lưu với các câu lạc bộ, tổ chức sinh viên để giới thiệu văn hóa và mạng lưới hỗ trợ tại trường.
- Cung cấp thông tin chi tiết về các chương trình học, cơ hội nghề nghiệp sau khi tốt nghiệp và các mối quan hệ với doanh nghiệp, giúp sinh viên hiểu rõ và chọn lựa chương trình học phù hợp với nguyện vọng và khả năng của bản thân.
- Thông tin về học bổng, các chi phí liên quan trong quá trình học tập và hỗ trợ sinh viên trong việc xin học bổng hay chế độ chính sách cho sinh viên.

1.3 Mô tả bài toán

Tổ chức nhập học cho tân sinh viên là một trong những bước quan trọng nhất trong công tác tuyển sinh hàng năm tại Học viện Công nghệ Bưu chính Viễn thông. Tuy nhiên, quá trình này thường gặp nhiều khó khăn và thách thức từ việc kiểm tra các loại giấy tờ, kinh phí nhập học của tân sinh viên tới việc cung cấp thông tin, tư vấn về chương trình học, giới thiệu về Học viện cũng như các câu lạc bộ, tổ chức sinh viên đang hoạt động....

Với số lượng tân sinh viên nhập học đang ngày càng tăng qua các năm, việc xây dựng chatbot tư vấn, hỗ trợ công tác nhập học giúp giảm thiểu sai sót, tiết kiệm công sức và thời gian là việc cấp thiết.

Dưới đây là một số yêu cầu cụ thể của chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông:

- Trả lời các câu hỏi thường gặp của thí sinh: Chatbot có thể được sử dụng để trả lời các câu hỏi thường gặp của thí sinh về các loại giấy tờ, kinh phí cần chuẩn bị để nhập học.
- Hỗ trợ đăng ký chương trình chất lượng cao, liên kết quốc tế, tiếng Anh: Chatbot có thể được sử dụng để hỗ trợ thí sinh đăng ký thi tiếng Anh, đăng ký học chương trình chất lượng cao, bao gồm nhập thông tin, nộp hồ sơ,...
- Tư vấn các hoạt động khác: Chatbot có thể cung cấp thông tin về các câu lạc bộ, hỗ trợ sinh viên ghi danh tham gia câu lạc bộ mong muốn, cung cấp thông tin về nhà cho thuê, quán ăn, cửa hàng tiện lợi... quanh khu vực Học viện.
- Hỗ trợ chế độ chính sách sinh viên: Chatbot có thể cung cấp thông tin về các loại học bổng, chế độ miễn giảm học phí, ký túc xá cho sinh viên...

1.4 Kết luận

Chương này trình bày tổng quan về hệ thống chatbot, các đặc trưng khi ứng dụng trong công tác tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông. Từ đó, nội dung chương chỉ ra những vấn đề còn tồn tại và đề xuất phương án giải quyết các vấn đề này. Nội dung chi tiết phương án giải quyết vấn đề được trình bày tại Chương 2 của luận văn.

CHƯƠNG 2: NGHIÊN CỨU CÁC CÔNG NGHỆ, KỸ THUẬT XÂY DỰNG CHATBOT PHỔ BIẾN HIỆN NAY

2.1 Kiến trúc, thành phần cơ bản của chatbot và những loại chatbot phổ biến hiện nay

2.1.1 Kiến trúc

Có ba kiến trúc chatbot phổ biến hiện nay:

- Kiến trúc dựa trên quy tắc
- Kiến trúc dựa trên tri thức
- Kiến trúc học máy

2.1.2 Thành phần cơ bản

2.1.2.1 Xử lý ngôn ngữ tự nhiên

NLP là công nghệ giúp chatbot hiểu và tạo ra ngôn ngữ của con người. Các kỹ thuật NLP thường được sử dụng trong chatbot bao gồm:

- Phân tích cú pháp
- Tìm kiếm
- Sinh văn bản
- Dịch ngôn ngữ
- Chuẩn hóa
- Nhận dạng thực thể

2.1.2.2 Dữ liệu xây dựng chatbot

Chatbot được đào tạo trên một tập dữ liệu lớn bao gồm các cuộc trò chuyện giữa người dùng và chatbot. Dữ liệu này giúp chatbot học cách hiểu và phản hồi các câu hỏi và yêu cầu của người dùng.

Dữ liệu là yếu tố quan trọng nhất trong việc xây dựng chatbot. Dữ liệu tốt sẽ giúp chatbot hiểu được ngôn ngữ tự nhiên và tạo ra các phản hồi phù hợp với ngữ cảnh của cuộc trò chuyện.

Có hai loại dữ liệu chính được sử dụng để xây dựng chatbot:

- Dữ liệu văn bản: Dữ liệu văn bản bao gồm các văn bản từ sách, bài báo, trang web, mạng xã hội,... Dữ liệu văn bản giúp chatbot học cách sử dụng ngôn ngữ tự nhiên một cách chính xác.

- Dữ liệu hội thoại: Dữ liệu hội thoại bao gồm các cuộc trò chuyện giữa con người với nhau. Dữ liệu hội thoại giúp chatbot hiểu được cách con người giao tiếp và tạo ra các phản hồi phù hợp với ngữ cảnh của cuộc trò chuyện.

2.1.2.3 Giao diện người dùng

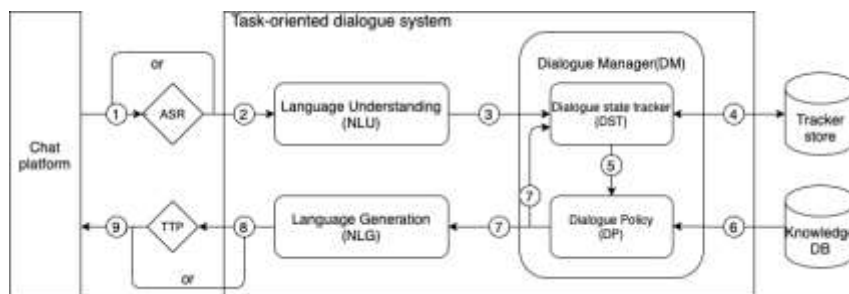
Công cụ giúp người dùng giao tiếp, cấu hình chatbot. Giao diện người dùng có thể là một ứng dụng web, ứng dụng di động hoặc chatbot tích hợp vào trang web hoặc ứng dụng hiện có.

2.1.3 Phân loại chatbot phổ biến hiện nay

2.1.3.1 Hệ thống đối thoại định hướng nhiệm vụ (TODs)

TODs là loại Chatbot được thiết kế để phục vụ một lĩnh vực nhất định như: mua sắm, cung cấp thông tin về thị trường chứng khoán...

Dưới đây là kiến trúc của một hệ thống TODS:



Hình 2.1: Kiến trúc hệ thống TODS

2.1.3.2 Trợ lý cá nhân thông minh (IPAs)

IPAs là loại Chatbot đóng vai trò là trợ lý cá nhân thông minh, là một loại phần mềm được thiết kế để giúp người dùng thực hiện các công việc hàng ngày thông qua việc xử lý ngôn ngữ tự nhiên.

2.1.3.3 Hệ thống đối thoại chit-chat

Chit-chat dialogue systems là những hệ thống trò chuyện thông minh được thiết kế để mô phỏng cách con người chit-chat (trò chuyện phiếm) với nhau. Mục tiêu của hệ thống này không chỉ là cung cấp thông tin hay thực hiện một công việc cụ thể, mà còn là tạo ra một cuộc trò chuyện tự nhiên, thú vị và có khả năng duy trì quan tâm của người dùng.

2.2 Các kỹ thuật sử dụng trong xây dựng chatbot

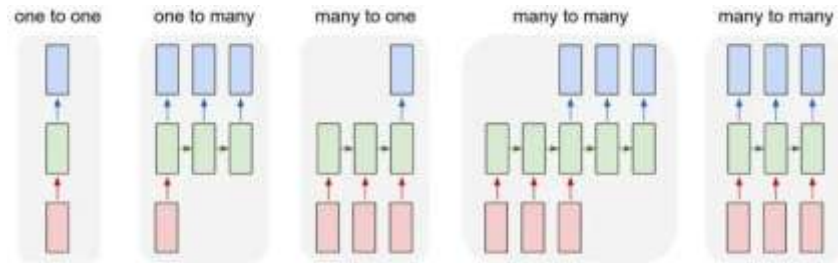
2.2.1 Mạng hồi quy RNN

2.2.1.1 Dữ liệu dạng Sequence

Dữ liệu có thứ tự như các ảnh tách từ video ở trên được gọi là sequence, time-series data

2.2.1.2 Phân loại bài toán RNN

Các mô hình RNN hầu như được sử dụng trong lĩnh vực xử lý ngôn ngữ tự nhiên và ghi nhận tiếng nói. Các ứng dụng khác được tổng kết trong hình dưới đây:



Hình 2.2: Các dạng bài toán RN

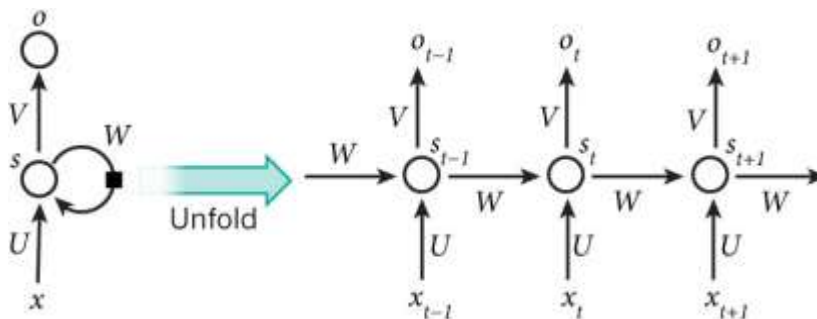
2.2.1.3 Ứng dụng bài toán RNN

Dưới đây là một số ứng dụng của RNN:

- Speech to text: Chuyển giọng nói sang text.
- Sentiment classification: Phân loại bình luận của người dùng, tích cực hay tiêu cực.
- Video recognition: Nhận diện hành động trong video.
- Heart attack: Dự đoán đột quỵ tim.

2.2.1.4 Mô hình RNN

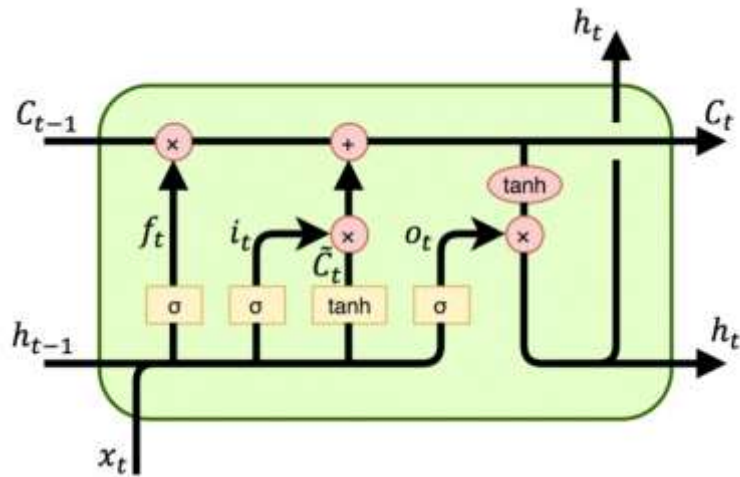
Về cơ bản một mạng RNN có dạng như sau:



Hình 2.3: Mô hình RNN

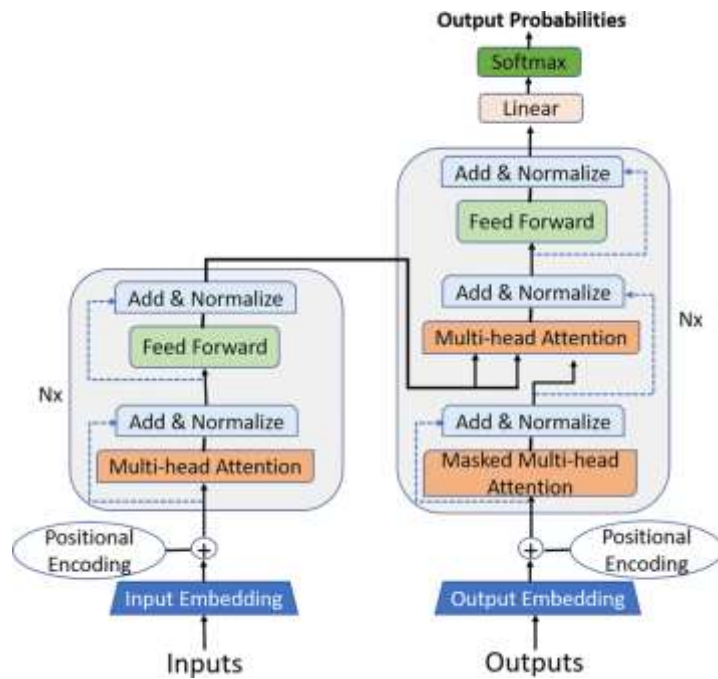
2.2.2 Mạng LSTM

Dưới đây là mô hình mạng LSTM:



Hình 2.4: Mô hình LSTM

2.2.3 Transformer



Hình 2.5: Kiến trúc Transformer

2.2.3.1 Cơ chế Self Attention

2.2.3.2 Cơ chế Multi-head Attention

2.2.4 Điểm cải tiến của Transformer so với LSTM

Kiến trúc transformer cho phép thực hiện các phép tính song song \rightarrow giảm đáng kể thời gian train/inference, tận dụng được sức mạnh tính toán của multi-GPU.

Yêu cầu ít thời gian đào tạo hơn so với các kiến trúc recurrent neural architectures trước đây, chẳng hạn như bộ nhớ ngắn hạn dài (LSTM), và biến thể sau này của nó đã được áp dụng phổ biến để đào tạo các mô hình ngôn ngữ lớn trên các bộ dữ liệu (ngôn ngữ) lớn.

2.3 Trình bày chi tiết công nghệ, kỹ thuật lựa chọn

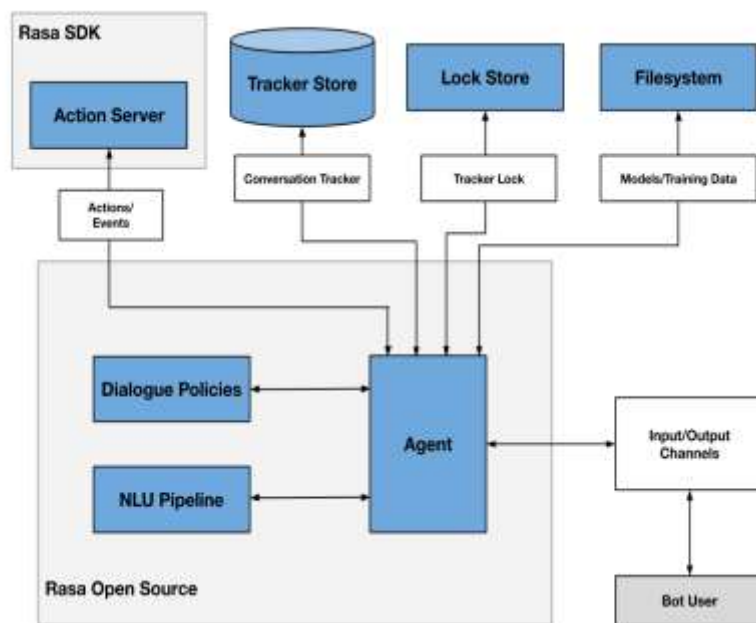
2.3.1 Rasa

Rasa là một framework chatbot mã nguồn mở dựa trên kiến trúc machine learning. Rasa cung cấp một bộ công cụ và thư viện giúp nhà phát triển xây dựng chatbot thông minh và linh hoạt.

2.3.1.1 Thành phần cơ bản

Rasa có hai thành phần chính là Rasa NLU và Rasa Core.

2.3.1.2 Các thành phần hệ thống chatbot RASA



Hình 2.6: Sơ đồ kết nối các thành phần của rasa

2.3.1.3 Tính năng của rasa và mục đích sử dụng

Dưới đây là một số tính năng chính của Rasa:

- Kiến trúc machine learning: Rasa sử dụng kiến trúc machine learning để hiểu và phản hồi các câu hỏi và yêu cầu của người dùng. Điều này cho phép chatbot học hỏi và cải thiện theo thời gian.

- Dễ sử dụng: Rasa cung cấp một bộ công cụ và thư viện đầy đủ tính năng giúp nhà phát triển xây dựng chatbot nhanh chóng và dễ dàng.

- Mở rộng: Rasa là một framework mở và có thể được tùy chỉnh để đáp ứng nhu cầu cụ thể.

2.3.2 Botpress

Botpress là một framework chatbot mã nguồn mở, được xây dựng dựa trên Node.js. Botpress cung cấp một bộ công cụ và thư viện giúp nhà phát triển xây dựng chatbot một cách nhanh chóng và dễ dàng.

2.3.3 Django

Django là một framework web mã nguồn mở, được viết bằng Python. Django cung cấp một bộ công cụ và thư viện đầy đủ tính năng giúp nhà phát triển xây dựng website và ứng dụng web một cách nhanh chóng và dễ dàng.

2.3.4 ReactJS

ReactJS là một thư viện để xây dựng giao diện. React sử dụng một kiến trúc gọi là "components", cho phép nhà phát triển xây dựng các giao diện người dùng phức tạp từ các thành phần nhỏ hơn, đơn giản hơn.

2.4 Kết luận

Chương này trình bày về nghiên cứu các công nghệ, kỹ thuật xây dựng chatbot phổ biến hiện nay trên thế giới. Từ đó đưa ra công nghệ, kỹ thuật được sử dụng để xây dựng chatbot. Trong Chương 3 của luận văn sẽ trình bày về việc áp dụng công nghệ, kỹ thuật đã lựa chọn ở Chương 2 để xây dựng chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông.

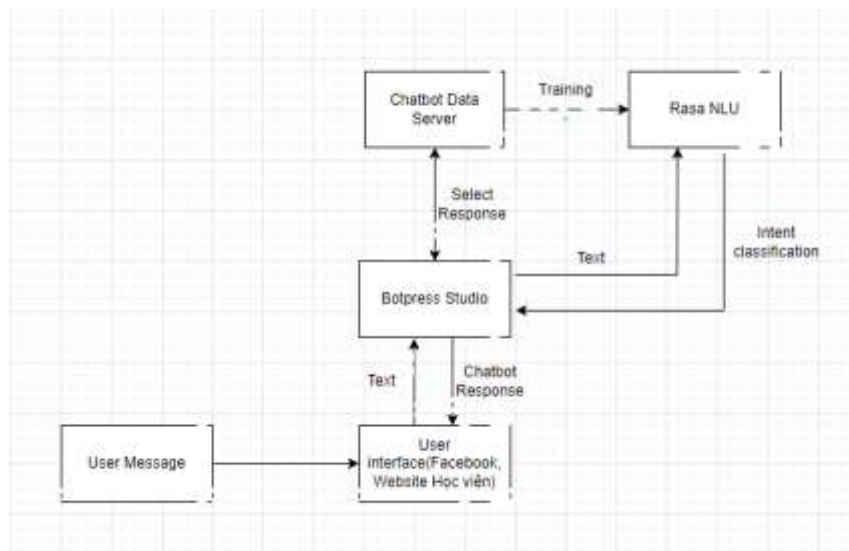
CHƯƠNG 3: XÂY DỰNG CHATBOT TƯ VẤN, HỖ TRỢ NHẬP HỌC TẠI HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

3.1 Thu thập dữ liệu

- 3.1.1 Xử lý và làm sạch dữ liệu
- 3.1.2 Chuẩn hóa dữ liệu
- 3.1.3 Xây dựng bộ câu hỏi, câu trả lời
- 3.1.4 Quy trình thử nghiệm và đánh giá
- 3.1.5 Phân tích và đánh giá kết quả

3.2 Kiến trúc tổng quan hệ thống

chatbot, dữ liệu về bộ câu hỏi câu trả lời, dữ liệu về người dùng, dữ liệu về các truy vấn của người dùng, và dữ liệu về các phản hồi của chatbot.



Hình 3.1: Kiến trúc tổng quan hệ thống

3.3 Xây dựng module quản lý dữ liệu bằng Django và ReactJS

Module quản lý dữ liệu chatbot được xây dựng nhằm mục đích quản lý danh sách chủ đề, danh sách câu hỏi thuộc chủ đề, quản lý câu trả lời chung của các chủ đề, quản lý mô hình và nhiều tính năng khác như quản lý người dùng, phân quyền người dùng hệ thống.

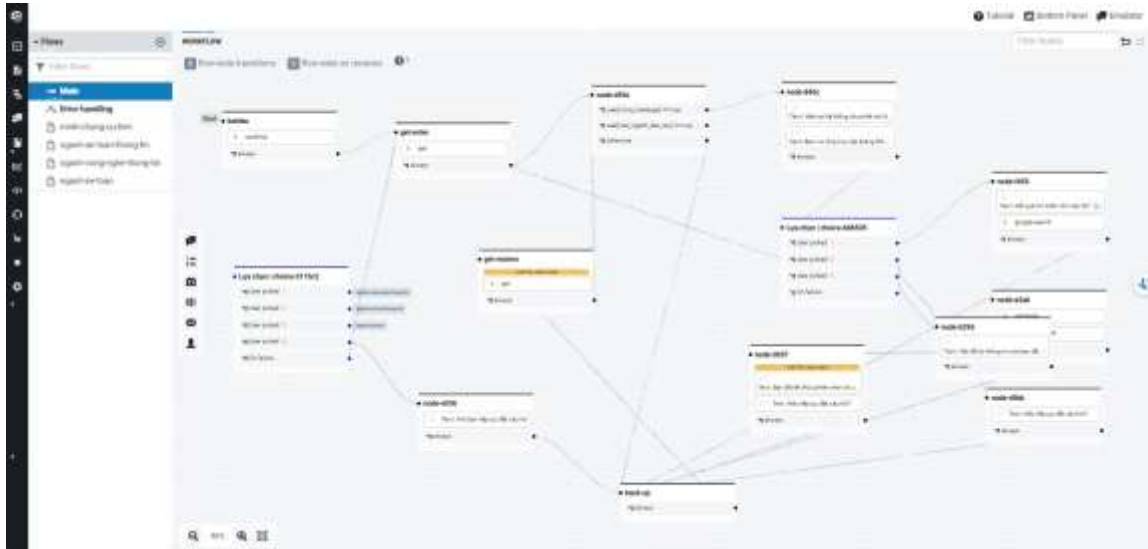
3.4 Xây dựng NLU dựa trên Rasa

Mục đích chính của Rasa NLU là phân tích thông tin do người dùng cung cấp cho chatbot, thông tin này bao gồm các ý định và thực thể cần được trích xuất. Trên nền tảng mã nguồn mở Rasa, tin nhắn gửi đến được xử lý bởi một chuỗi các hàm chức

năng, các hàm này được thực thi lần lượt bên trong quy trình xử lý được định danh trong config.yml . Với quy trình nhúng có giám sát, ta có thể huấn luyện với bất kỳ ngôn ngữ trên thế giới vì công việc này sẽ bắt đầu huấn luyện mọi thứ từ đầu.

3.5 Xây dựng giao diện hỗ trợ quản lý đoạn hội thoại bằng Botpress

3.5.1 Thành phần giao diện quản lý luồng của chatbot



Hình 3.2: Giao diện quản lý luồng hội thoại trên Botpress Studio

3.5.2 Xây dựng luồng kịch bản quản lý hội thoại giữa người dùng và bot

3.6 Kết luận

Chương 3 đã trình bày về cách thức áp dụng công nghệ, kỹ thuật đã lựa chọn ở Chương 2 vào xây dựng chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông. Ở Chương cuối sẽ trình bày về quy trình thử nghiệm chatbot và triển khai trên một số nền tảng phổ biến như facebook hoặc website của Học viện.

CHƯƠNG 4: TRIỂN KHAI THỬ NGHIỆM CHATBOT TƯ VẤN, HỖ TRỢ NHẬP HỌC TRÊN FACEBOOK MESSENGER, WEBSITE CỦA HỌC VIỆN

4.1 Môi trường thử nghiệm và các thước đo đánh giá

4.1.1 Môi trường thử nghiệm

Các thử nghiệm được thực hiện trên máy tính cá nhân với thông số cấu hình của mô trường được mô tả cụ thể trong bảng sau:

Bảng 4.1: Bảng cấu hình môi trường thử nghiệm

Thông tin	Môi trường máy huấn luyện
Vi xử lý	CPU Intel Core I5-12400F
Dung lượng RAM	32GB
Dung lượng bộ nhớ	256GB
GPU	NVIDIA QUADRO RTX 4000 8GB GDDR6
Hệ điều hành	Ubuntu 22.04.3 LTS
Mô hình huấn luyện	bert-base-multilingual-cased
Python	3.10
Rasa	3.1

4.1.2 Thử nghiệm và đánh giá độ chính xác của Rasa NLU

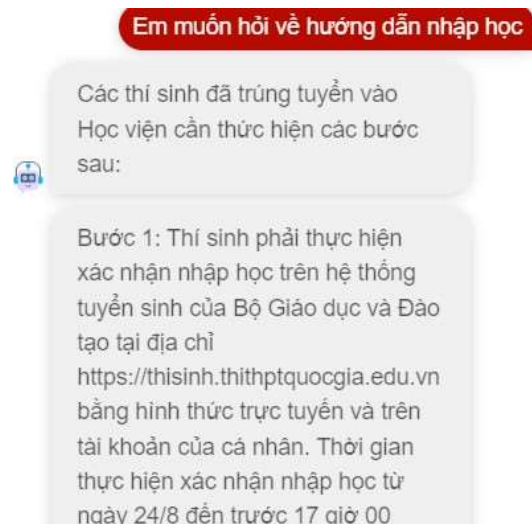
Trong đề án này sẽ sử dụng các phương pháp đo gồm: accuracy score, Precision và Recall, F1 score.

- Accuracy: Cách tính đơn giản và thường được sử dụng nhiều nhất. Tính tỉ lệ giữa số mẫu được dự đoán đúng và tổng số mẫu. Thường được sử dụng cho các bài toán phân loại mà ổn định và không bị lệch hoặc không có sự mất cân bằng của các lớp.
- Precision – là tỷ lệ bao nhiêu cái đúng được lấy ra, Cân nhắc trên tập dữ liệu kiểm soát xem có bao nhiêu dữ liệu được mô hình phán đoán đúng.
- Recall – là tỷ lệ bao nhiêu cái được lấy ra là đúng, Chỉ số này còn được gọi là độ bao phủ, tức là xét xem mô hình tìm được có khả năng tổng quát hóa như nào. Từ hai yếu tố độ chuẩn xác và độ bao phủ người ta đặt ra một chỉ số khác gọi là F1-Score
- F1-Score: Là trung bình điều hòa (harmonic mean) của precision và recall (giả sử hai đại lượng này khác 0). Do đó nó đại diện hơn trong việc đánh giá độ chính xác trên đồng thời precision và recall.

4.1.3 Thử nghiệm trò chuyện với chatbot

Chatbot được em xây dựng với gần 30 chủ đề về tuyển sinh, nhập học như điểm tuyển sinh, chỉ tiêu tuyển sinh, hướng dẫn nhập học, hướng dẫn chuẩn bị hồ sơ nhập học,

thông tin về các vấn đề liên quan nhập học như giấy báo nhập học, sổ đoàn, giấy chuyển nghĩa vụ quân sự, bảo hiểm y tế. Ngoài ra chatbot cũng được trang bị các thông tin cơ bản về học viện như địa chỉ, mã trường hay thông tin liên hệ.



Hình 4.1: Hỏi đáp với chatbot về hướng dẫn nhập học

4.2 Cài đặt và triển khai hệ thống

4.2.1 Cài đặt hệ thống

4.2.2 Các yêu cầu đối với cấu hình máy cài đặt và lưu ý

4.2.3 Triển khai hệ thống chatbot

4.3 Kết luận

Chương này đã trình bày về quy trình thử nghiệm chatbot, cách thức triển khai hệ thống trong thực tế và kết quả thực nghiệm. Từ đó tiến hành đánh giá kết quả thực nghiệm của hệ thống, so sánh với các nghiên cứu đã có và đề xuất phương hướng nghiên cứu, phát triển sau này.

KẾT LUẬN

Đề án tập trung nghiên cứu bài toán Nghiên cứu và xây dựng chatbot tư vấn, hỗ trợ nhập học tại Học viện Công nghệ Bưu chính Viễn thông. Trong đó tập trung chính vào các kỹ thuật xây dựng NLU, xử lý ngôn ngữ tự nhiên, áp dụng các mô hình học máy, học sâu vào quy trình xây dựng phân loại ý định người dùng. Từ NLU đã xây dựng kết hợp với giao diện hỗ trợ xây dựng các luồng kịch bản trò chuyện giữa người dùng và bot. Thông qua thực nghiệm, phân tích và đánh giá các kịch bản, sau đó đánh giá hiệu năng bằng việc triển khai trên các nền tảng như facebook hay website của Học viện. Cuối cùng đưa ra hệ thống chatbot tư vấn, hỗ trợ nhập học. Cụ thể đề án đã được một số kết quả sau:

- Tìm hiểu tổng quan về bài toán và quy trình xây dựng chatbot hỗ trợ hỏi đáp các thông tin tuyển sinh và nhập học.

- Nắm được kiến trúc và nhiệm vụ các thành phần, thuật toán và các kỹ thuật sử dụng trong chatbot như Mạng RNN, mạng bộ nhớ dài-ngắn (Long Short Term Memory networks), Kiến trúc Transformer, Mô hình BERT...

- Xây dựng bộ dữ liệu về nhập học, tuyển sinh 2023 chính xác dựa trên các nguồn thu và thông tin từ Học viện cho bài toán chatbot.

- Nghiên cứu và ứng dụng Framework Rasa, tối ưu hóa các thuật toán khi áp dụng framework này. Tạo mô hình huấn luyện riêng để xác định ý định người dùng. Đồng thời sử dụng mô hình BERT Base Multilingual để cải tiến chất lượng của bài toán xác định ý định. Kết hợp NLU rasa với giao diện người dùng Botpress Studio để quản lý luồng hội thoại giữa người dùng và chatbot dễ dàng hơn.

- Hoàn thành mục tiêu xây dựng chatbot giúp giảm thiểu thời gian và công sức trong công tác hỗ trợ, tư vấn nhập học tại Học viện. Tích hợp trên website Học viện để tương tác với người dùng.

Tuy nhiên đề án tốt nghiệp vẫn còn một số điểm hạn chế như nội dung của dữ liệu hỏi đáp phải được cập nhật thủ công hàng năm trên hệ thống theo đề án tuyển sinh, quy định của Học viện trong công tác tuyển sinh, nhập học. Hiện tại chatbot chỉ mới phân loại theo ý định người dùng và đưa ra câu trả lời chung cho ý định đó. Để cải thiện hệ thống, đề án đưa ra hướng phát triển trong tương lai như cải thiện thêm về nhận diện theo ngữ cảnh để tăng độ chính xác của câu trả lời.