

TP. Hồ Chí Minh, ngày 08 tháng 01 năm 2021

ĐỀ CƯƠNG CHI TIẾT KHÓA LUẬN TỐT NGHIỆP

1. Tên đề tài: Xây dựng giải pháp trả lời tự động (Chatbot) bằng tiếng Việt

2. Sinh viên thực hiện:

- Họ tên: Trần Thị Mỹ Duyên MSSV: 17130045 Lớp: DH17DTA
Khoa: CNTT Niên khoá: 2017 SĐT: 0979006985
- Họ tên: Lương Trung Thành MSSV: 17130216 Lớp: DH17DTB
Khoa: CNTT Niên khoá: 2017 SĐT: 0363998832

3. Giảng viên hướng dẫn: ThS. Nguyễn Đức Công Song

4. Lý do chọn đề tài:

- Do sự phát triển mạnh mẽ của cuộc Cách mạng Công nghiệp 4.0 dẫn đến nhu cầu của xã hội đối với những hệ thống phần mềm thông minh ngày càng cao. Trong đó, việc xây dựng Chatbot AI có khả năng phản hồi chính xác những câu nói của người dùng, là “trợ thủ” đắc lực cho các doanh nghiệp trong quá trình chăm sóc khách hàng cũng là một vấn đề đang được quan tâm. Chatbot AI [1] có thể: tư vấn, bán hàng, Giảm chi phí, tiết kiệm thời gian, chăm sóc khách hàng 24/7,...
- Các Chatbot hiện nay trên thế giới đã phát triển và có khả năng chat gần như con người. Tuy nhiên, điều này chỉ đúng với các con chatbot sử dụng ngôn ngữ khác tiếng Việt. Vì vậy, nhóm muốn thực hiện đề tài xây dựng giải pháp trả lời tự động bằng tiếng Việt bằng giọng nói và văn bản để giao tiếp với người dùng.

5. Mục tiêu của đề tài:

Xây dựng giải pháp trả lời tự động (Chatbot) bằng tiếng Việt bằng việc nghiên cứu ứng dụng các phương pháp:

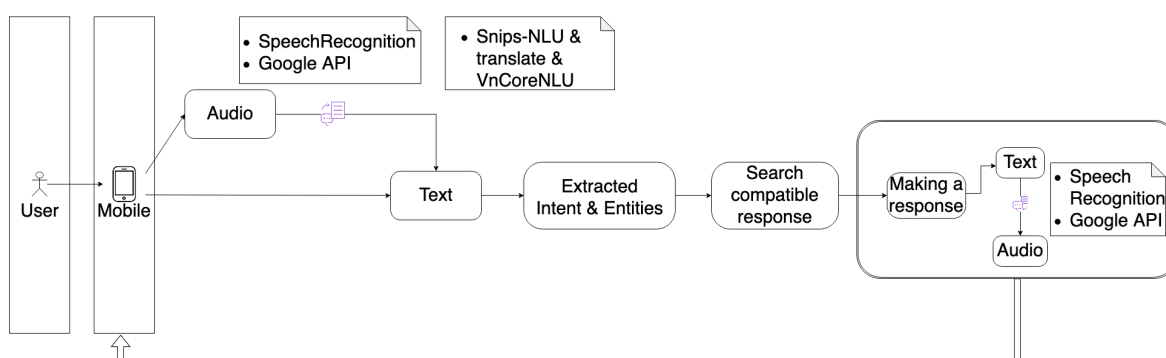
- + Chuyển đổi giọng nói thành văn bản và ngược lại.

- + Detect Intent & Extract Entities: Xác định ý định và trích xuất thực thể từ văn bản nhận được.
- + Search compatible response: Tìm kiếm đưa ra câu trả lời phù hợp.

6. Phạm vi nghiên cứu:

Trong luận văn này, nhóm sẽ giới hạn đề tài trong phạm vi sau: Xây dựng giải pháp trả lời tự động bằng tiếng Việt với dữ liệu đầu vào và kết quả đầu ra là giọng nói và văn bản, dựa vào phương pháp chuyển đổi giọng nói thành văn bản và ngược lại; nghiên cứu áp dụng các phương pháp để nhận diện ý định, thực thể từ người dùng cùng với việc huấn luyện dữ liệu để xây dựng mô hình chatbot, mong muốn chatbot có thể đưa ra câu trả lời phù hợp với câu hỏi của người dùng với lĩnh vực áp dụng: Chỉ đường và hướng dẫn đi đến các địa điểm hành chính nhà nước trong phạm vi Thành phố Thủ Đức.

- Sơ đồ khối dưới đây mô tả tổng quan hệ thống:



Hình 1: Sơ đồ hệ thống

- Nội dung thực hiện: Xây dựng giải pháp trả lời tự động bằng tiếng Việt.
 - + Nghiên cứu ứng dụng phương pháp chuyển đổi giọng nói thành văn bản và ngược lại: Google API, SpeechRecognition python.
 - + Nghiên cứu phương pháp Detect Intent & Extract Entities: Xác định ý định và trích xuất thực thể từ văn bản nhận được.
 - + Nghiên cứu ứng dụng phương pháp Search compatible response: Tìm kiếm đưa ra câu trả lời phù hợp.

7. Cơ sở khoa học và thực tiễn:

- Tìm kiếm giải pháp chuyển đổi giọng nói thành văn bản và ngược lại.

- Để tạo một chatbot, việc đầu tiên chatbot cần phải hiểu những gì người dùng nói (xác định ý định người dùng) bằng cách phân loại ý định và trích xuất các thông tin thực thể quan trọng: địa điểm, ngày, thời gian,...Sau đó để trả lời người dùng cần dựa vào đầu vào hiện tại và bối cảnh của cuộc trò chuyện. Tìm kiếm dữ liệu và huấn luyện mô hình để đưa ra câu trả lời phù hợp.
- Sản phẩm được phát triển trên nền tảng Python.
- Đề tài được phát triển để phục vụ cho nhu cầu sử dụng của người dùng (nhẹ, gọn, tiện dụng). Mặc dù đã có rất nhiều tổ chức lớn thực hiện, tuy nhiên việc phát triển thêm luôn là cần thiết cho sự phát triển của xã hội.

8. Thời gian thực hiện: 6 tháng (bắt đầu từ HKII năm học 2020-2021)

9. Sản phẩm của đề tài:

- Xây dựng giải pháp trả lời tự động bằng tiếng Việt với các chức năng chi tiết như sau:
 - + Ứng dụng chatbot (xác định mục đích và thực thể và xây dựng luồng hội thoại).
 - + Sử dụng ngôn ngữ trò chuyện chính là Tiếng Việt.
 - + Cho phép người dùng cuối tương tác với chatbot bằng giọng nói và âm thanh.
- Bộ dữ liệu dùng để huấn luyện, tài liệu đầu vào mẫu và phương pháp giải quyết các vấn đề về câu trả lời phản hồi.
- Mô hình dùng để phát hiện ý định và trích xuất thông tin từ văn bản.

10. Tài liệu tham khảo:

- Tài liệu tiếng Việt
 1. *5 lý do các doanh nghiệp nên xây dựng chatbot ngay hôm nay*. Truy cập tháng 12/2020, từ <https://fpt.ai/vi/lam-nao-de-cac-doanh-nghiep-tan-dung-loi-cua-chatbot-ai>
 2. Phạm Quang Nhật Minh - FTRI. (2017). *3 vấn đề NLP cơ bản khi phát triển một hệ thống chatbot và một số phương pháp giải quyết điển*

hình. Truy cập tháng 12/2020, từ

<https://techinsight.com.vn/3-van-de-nlp-co-ban-khi-phat-trien-mot-thong-chatbot-va-mot-phuong-phap-giai-quyet-dien-hinh/>

- Tài liệu tiếng Anh

1. Sameera A. Abdul-Kader, Dr. John Woods. (2015). *Survey on Chatbot Design Techniques in Speech Conversation Systems*.

2. Sebastian Schuster, Rushin Shah, Sonal Gupta, Mike Lewis. (2019). *Cross-Lingual Transfer Learning for Multilingual Task Oriented Dialog*.

3. Sangramsing Kayte. (2020). *Natural Language Principles*. Truy cập tháng 12/2020, từ

<https://chatbotslife.com/natural-language-principles-65e88e20b94>

4. *How to Make Conversational AI Work for Low-Resource Languages*.

Truy cập tháng 12/2020, từ

<https://blog.keyreply.com/blog/conversational-ai-low-resource-languages>

Giảng viên hướng dẫn

(Ký và ghi rõ họ tên)

Sinh viên thực hiện

(Ký và ghi rõ họ tên)