TÌM HIỂU VÀ XÂY DỰNG HỆ THỐNG CHATBOT HỖ TRỢ HỌC NHẬP MÔN LẬP TRÌNH

Nguyễn Lê Anh Quân Cát Văn Tài Nguyễn Khả Tiến

19522081 19522147 19522337

Tóm tắt

Link Github: https://github.com/anhquan075/CS519.M11.KHCL

Link YouTube: https://youtu.be/y956Ce2aXAE



Cát Văn Tài



Nguyễn Lê Anh Quân



Nguyễn Khả Tiến

Giới thiệu

Chatbot là một chương trình máy tính có khả năng giao tiếp với con người bằng cách tự động trả lời những câu hỏi hoặc xử lý tình huống.



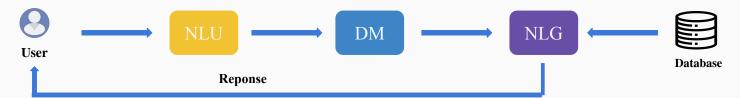
UIT.CS519.ResearchMethodology

Mục tiêu

- Tìm hiểu tổng quan về hệ thống Chatbot và một số framework xây dựng chatbot, trong đó trọng tâm là RASA framework làm cơ sở để xây dựng hệ thống Chatbot.
- Xây dựng bộ dữ liệu để phục vụ cho việc xây dựng hệ thống Chatbot tập trung vào chủ đề
 Nhập môn lập trình.
- Xây dựng ứng dụng minh họa.

Nội dung và Phương pháp thực hiện

Cấu trúc các thành phần cơ bản của hệ thống Chatbot



1. Hiểu ngôn ngữ tự nhiên - NLU

Bao gồm việc xử lý ngôn ngữ tự nhiên (NLP) có nhiệm vụ xác định được ý định câu hỏi của người dùng (intent classification) và trích chọn thông tin (slots filter)

Intent classifiers
Embedder (word2vec, glove, ..)
Bag of words embedder
Slot filler

2. Quản lý hội thoại - DM

Có nhiệm vụ xác định được hành động (action) tiếp theo dựa vào trạng thái hành động trước đó hoặc ngữ cảnh của cuộc hội thoại. Các ngữ cảnh này phải được tham chiếu trong các kịch bản dựng sẵn (history) được đào tạo cho Chatbot.

RNN Dense softmax

3. Mô hình sinh ngôn ngữ - NLG

Là thành phần sinh ngôn ngữ dựa vào chính sách (policy) và hành động được xác định trong DM thông qua các bộ hội thoại. NLG có thể sinh ra câu trả lời dựa vào tập mẫu câu trả lời (pre-defined template) đã đào tạo cho bot.

Template

Nội dung và Phương pháp

RASA Framework

Rasa NLU

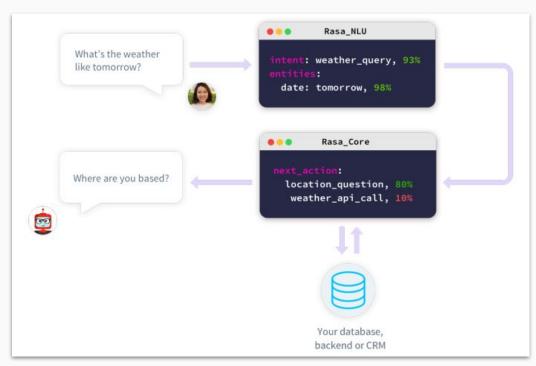
Đảm nhiệm các vấn đề liên quan đến Xử lý ngôn ngữ tự nhiên (NLP) như phân loại ý định (Intent Classification) và trích xuất thực thể (Entity Extraction)... Nói cách khác, Rasa NLU dùng để đọc hiểu tin nhắn từ người dùng.

Rasa Core

Bộ khung xử lý các vấn đề xảy ra trong một cuộc hội thoại, Rasa Core sẽ quyết định Chatbot nên thực hiện hành động như thế nào

Rasa X

Một công cụ hỗ trợ triển khai mô hình Chatbot để xây dựng thành sản phẩm thực tế, giúp tiếp cận nhanh chóng đến người dùng



Ví dụ minh họa về quy trình xử lý của một hệ thống Rasa Chatbot

Nội dung và Phương pháp

Xây dựng bộ dữ liệu phục vụ việc xây dựng chatbot Nhập môn lập trình và Ứng dụng minh họa

- Thu nhập dữ liệu từ các trang web tutorial uy tín về lập trình như là codelearn, geeksforgeeks, kteam, v.v.
- Xử lý dữ liệu để phù hợp với framework xây dựng Chatbot, framework chúng tôi sử dụng ở đây là RASA.
- Xây dựng hệ thống dựa theo các module và các phương pháp đã chọn lọc được.
- Đánh giá hiệu suất hệ thống.

Kết quả dự kiến

- Tài liệu trình bày lại về thiết kế hệ thống Chatbot, thiết kế hệ thống cơ sở dữ liệu cho
 Chatbot, ưu nhược điểm của từng hướng tiếp cận trong việc xây dựng hệ thống Chatbot,
 Framework sử dụng cho việc xây dựng hệ thống Chatbot.
- Tập dữ liệu hội thoại dùng để huấn luyện cho hệ thống Chatbot chuyên về hỗ trợ môn học
 Nhập môn lập trình và bằng tiếng Việt.
- Úng dụng của hệ thống Chatbot.

Kết quả dự kiến

- Úng dụng của hệ thống Chatbot với các khả năng:
 - Giải đáp các thắc mắc về C++.
 - O Đưa ra ví dụ và tìm kiếm những video tutorial liên quan đến thắc mắc của người dùng.
 - Gợi ý cách giải các bài tập sơ đồ khối.



Vấn đáp giữa Chatbot và người dùng về Hàm



Vấn đáp giữa Chatbot và người dùng về bài tập sơ đồ khối

Tài liệu tham khảo

- [1] Ruane, Elayne, Abeba Birhane, and Anthony Ventresque. "Conversational AI: Social and Ethical Considerations", 2019.
- [2] M. Huang, X. Zhu, and J. Gao. "Challenges in Building Intelligent Open-domain Dialog Systems". arXiv:1905.05709, 2020.
- [3] G. Molnar and Z. Szuts. "The Role of Chatbots in Formal Education". In 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY).
- [4] A. Jiao. "An Intelligent Chatbot System Based on Entity Extraction Using RASA NLU and Neural Network". J. Phys.: Conf. Ser., vol.1487, p.012014, 2020.