CSE 4392 Special Topic: Natural Language Processing

Homework 10 - Spring 2024

Due Date: Apr 23rd, 2024, 11:59 p.m. Central Standard Time

Welcome to this week's NLP homework assignment! Your task is to design and implement a real-time communication robot that can talk to you with audio, just like Siri or Google Assistant! This bot consists of three key components: speech-to-text conversion (ASR), dialogue response generation, and text-to-speech synthesis (TTS). **You're encouraged to utilize any API or tool** available for each component, allowing you to explore various NLP techniques and unleash your creativity in crafting an engaging and responsive communication robot. Whether you envision a virtual assistant, chatbot, or conversationalist, this assignment offers the opportunity to experiment with cutting-edge NLP technologies and bring your ideas to life. Happy coding!

Problem 1 - 100%

This assignment is to examine the abilities of:

- Information Retrieval

- Application of existing APIs

- Product design

- Engineering efforts

Real Time System

You need to design a real-time system that can:

1. Once the system is started, it should be able to listen to the user's voice.

2. The system could automatically stop listening when the user stops speaking.

3. A response should be generated in real-time in the form of audio.

4. After the response is generated, the system should be able to listen to the user's voice again.

5. Once the system hears a specific keyword (e.g., "Exit the bot"), it should stop listening and terminate the program.

Speech to Text

For the speech-to-text conversion, you can use any API or tool available. For example:

- SpeechRecognition Library

- OpenAI Speech to Text

- Google Cloud Speech-to-Text

- ...

Response Generation

For the text generation, you can use any API or tool available. For example:

- ChatbotAI Library

- Chatterbot Library

- OpenAI GPT 3 API

- ...

Text to Speech

For the text-to-speech synthesis, you can use any API or tool available. For example:

- OpenAI Text to Speech

- pyttsx3 Library

- text-to-speech Library

- ...

Bonus: If you can design a system with a nice front end, you will receive a bonus of 10%.

If you can design a system that can answer questions about worldly facts, such as "Who is the president of the United States?" or "When did the total eclipse in Dallas happen this year?," you will receive a bonus of 20%.

Attach your codes and a captured video demo that explains how your app works.