VOICE - CONTROL

(AI ASSISTANT FOR SANFRANCISCO BAY UNIVERSITY)

MARYAM ZUBAIR

STUDENT ID : 19709

DESIGN & IMPLEMENTATION - OVERVIEW

- The project involves creating a voice-controlled AI assistant using Python. Key components include audio processing, speech recognition, AI interaction, and audio playback.
- The script imports essential libraries like pydub for audio manipulation, mutagen.mp3 for MP3 file handling, simpleaudio for audio playback, speech_recognition and whisper for speech to text conversion, gtts for text-to-speech, and openai for AI interaction.
- Initialization includes setting up the OpenAI API with an API key and configuring command-line arguments.

DESIGN & IMPLEMENTATION - AUDIO PROCESSING

The system uses pydub and simpleaudio for audio processing and playback.

A continuous audio recording thread captures user speech through the microphone.

The recorded audio is then processed and formatted for further transcription.

DESIGN & IMPLEMENTATION - SPEECH RECOGNITION

Speech recognition is achieved using the whisper and speech_recognition libraries.

The system listens for a specific wake word to initiate processing.

Upon detecting the wake word, the script transcribes the speech into text for AI processing.

DESIGN & IMPLEMENTATION - AI INTERACTION

Text transcribed from user speech is sent to OpenAI's GPT-3 model.

GPT-3 generates a response based on the input query.

This response is then prepared for audio output.

TEST - OVERVIEW

Testing involves verifying the accuracy of speech recognition and the appropriateness of AI responses.

The system is tested with a variety of question related to pdf loaded to check the responses.

RESPONSE:

```
User's question:  are there any coding courses being taught in San Francisco Bay University
AI's response: Yes, San Francisco Bay University offers degree programs in Computer Science and Electrical Engineering that include course
```

```
User's question:  can you please tell me what is the CGP required for Presidential Scholarship
AI's response: The minimum cumulative grade point average (CGPA) required for consideration for the Presidential Scholarship varies depend
```

TEST - PROCESS DETAIL

Audio samples are used to test the system's ability to accurately transcribe and respond.

The system's response time and accuracy in different acoustic environments are also evaluated.

Stress tests are conducted to assess the system's performance under continuous input.

ENHANCEMENT IDEAS

Future enhancements could include integrating more advanced speech recognition algorithms and noise reduction techniques.

Future enhancements will be "Fine Tuning" the model to receive more appropriate results.

1. By training base model on specific task and data
2. By Moderation techniques: to filter out or prevent the generation of harmful, inappropriate, or offensive content by the LLM
3. By prompt Injection: to provide a specific context or instruction to the LLM. Moderation

Expanding the AI's capabilities to understand and respond to a broader range of queries.

CONCLUSION

- The project successfully demonstrates the creation of a voice-controlled AI assistant using Python.
- It highlights the potential of integrating various technologies for natural user interaction.
- The system offers a foundation for further enhancements and applications in AI-driven voice recognition and response systems.

REFERENCES

- References include documentation and resources related to the Python libraries used: pydub, mutagen.mp3, simpleaudio, speech_recognition, whisper, gtts, and openai.
- Additional references may include academic papers or online resources related to speech recognition and AI.

GITHUB LINK

https://github.com/Maryam-Zubair/MachineLearning_Assignment/tree/main/ChatGPT/SFBU-Voice-Controlled-AI-Assistant