Audio transcription and speech generation

Presented By

PRANJAL RATHORE

OBJECTIVE

Research the Whisper API ,Collect a sample audio data set to test your application. Use the Whisper API to convert the audio input to text. Test and measure the accuracy of the transcription against the original audio input.input it into the OpenAI Davinci model to generate a logical response.using the OpenAI Davinci model, convert the text output back to audio format.

Software/libraries/API used

- Open AI's Whisper API
- Open AI's Da Vinci Model of GPT 3
- Numpy
- Pandas
- Wedio- An online TTS software.

Project Repository

The dataset consisted of 10 voice samples with different contexts, emotion sentences, and noise elements.

Accuracy came out to be **94.2957%** for conversion to text.

I have generated the response using the Da Vinci model of Open AI GPT3 .and converted back to the audio using an online TTS CALLED wedio.