Avishkar 2k21
Robomania

Problem Statement: Decode the information

Team: Machine Manav
Ayush Singh Gour 20203044
Anadi Garg 20203016
Saurabh Shukla 20203133
Shubham Kumar 20203144

DECODE THE INFORMATION

Theme

The year is 2030. Pandemic is long gone but there is something else that is disturbing. Aliens have now invaded earth and have been successful in establishing a base here.

Unfortunately, their technologies are far more advanced and superior than that of the human race. Reports suggested that these Aliens lost their planet due to a fatal collision with another celestial body. Ergo they chose Earth because Earth has a lot of energy in its core which would help suffice their need for energy. Additionally, the soothing temperature of Earth is also a boon for them.

Due to this new habitation, these aliens have decided to eradicate their previous habitats i.e. Humans. As a result, there has been a hostile takeover in most of the continents by these Aliens. As of now, South Asia is the only part of the world that has been considerate of the plight of the human race.

Unfortunately, misery has not stopped for us because our only living spy has been spotted and eliminated immediately. However, he was successful in installing an undisclosed microphone at the headquarters of Aliens.

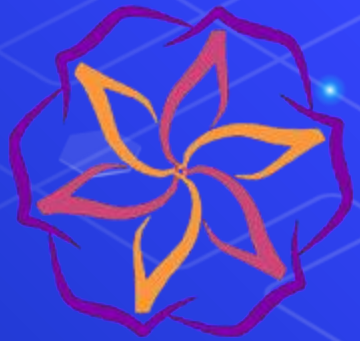
Software,Tech and Libraries used

Language: Python

Neural Network Library: Tensorflow Keras

Audio Processing Library:Librosa

Speech to Text: Google SR Api



TensorFlow

PROJECT PLANNING

Accessing AUDIO DATA From the file system

Extracting Numeric data and converting audio signal into small frames for recognition and finding MFCC and other features of them

Passing the features data to Neural Network an d finding the labels for each frame.

On the basis of the labels use Google Speech to Text API for printing the transcript of the audio file to be tested on the model.

Mel Frequency Cepstral Coefficents

Mel Frequency Cepstral Coefficents (MFCCs) are a feature widely used in automatic speech and speaker recognition.The MFCC feature extraction technique basically includes **windowing the signal, applying the DFT, taking the log of the magnitude, and then warping the frequencies on a Mel scale, followed by applying the inverse DCT**.MFCC has **39 features**.

SPECTRAL CENTROID

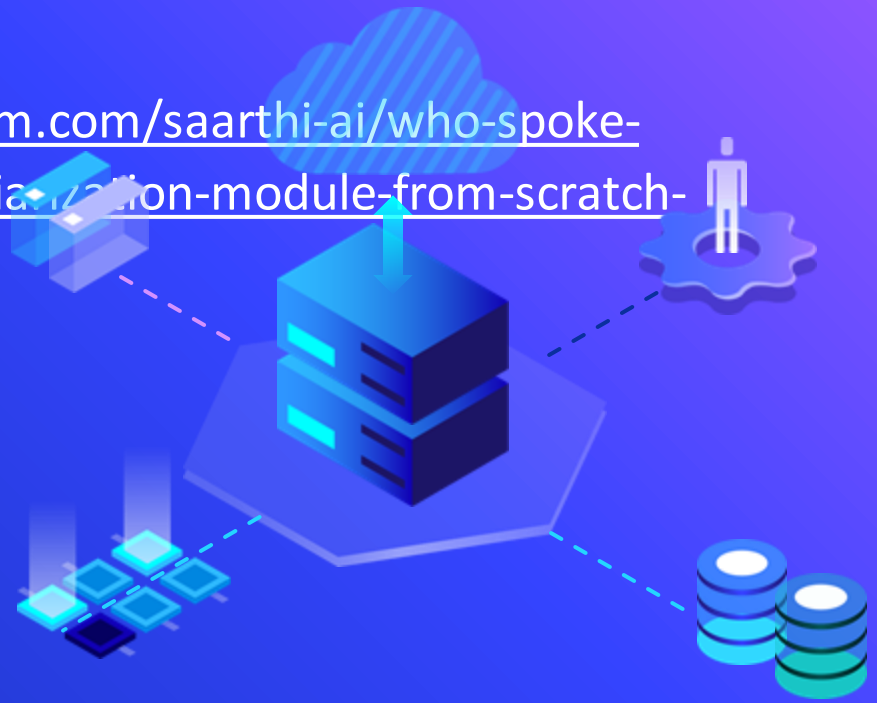
- The spectral centroid is **a measure used in digital signal processing to characterise a spectrum**.
- It indicates where the center of mass of the spectrum is located.
- Perceptually, it has a robust connection with the impression of brightness of a sound.

CHROMA

- Chroma-based features, which are also referred to as “pitch class profiles”, are **a** powerful tool for analyzing music whose pitches can be meaningfully categorize.
- It is a descriptor, **which represents the tonal content of a musical audio signal in a condensed form**.
- Hence it is widely used in musical and harmonic analysis.

References

- Medium articles : https://medium.com/saarthi-ai/who-spoke-when-build-your-own-speaker-diarization-module-from-scratch-e7d725ee279
- Codebasics youtube channel
- Towardsdatascience Website



Thanks!

Any questions?

Email: anshugour2438000@gmail.com