Group 29

B shiv kiran(2000500xx), B Likith Rajesh(2000500xx), D Venkatesh Prasad(2000500xx), L Jitendra(2000500xx)

AI ML Project Report

Problem Statement

Prediciting Gender based on voice by extracting important features of the voice

Dataset Used

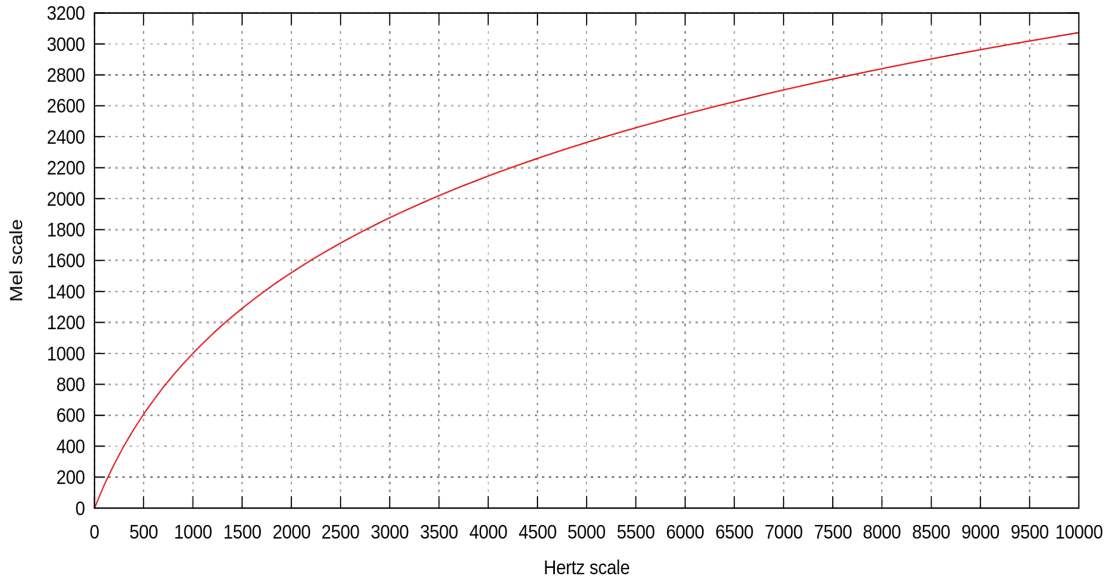
Mozilla Common Voice Dataset which has many voices labelled with the corresponding gender and language used

Feature Extraction

We have used librosa library to extract the following features

Mel Spectogram Frequency

A **mel spectrogram** is a spectrogram where the frequencies are converted to the
mel scale,



Group 29

B shiv kiran(2000500xx), B Likith Rajesh(2000500xx), D Venkatesh Prasad(2000500xx), L Jitendra(2000500xx)

The frequencies are obtained using fourier transform on the signal and the mel spectogram frequency is obtained mapping these frequencies on the mel scale.

Mel frequency cepstrum coefficients(mfcc)

the mel-frequency cepstrum (MFC) is a representation of the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency.
Mel-frequency cepstral coefficients (**MFCCs**) are coefficients that collectively make up an MFC.

Tonnetz

Networks or lattices of tones. The tonnetz tonal centroids — the “central” tones. These are features that help in Detecting Harmonic Change in Musical Audio or variances due to tones in audio.

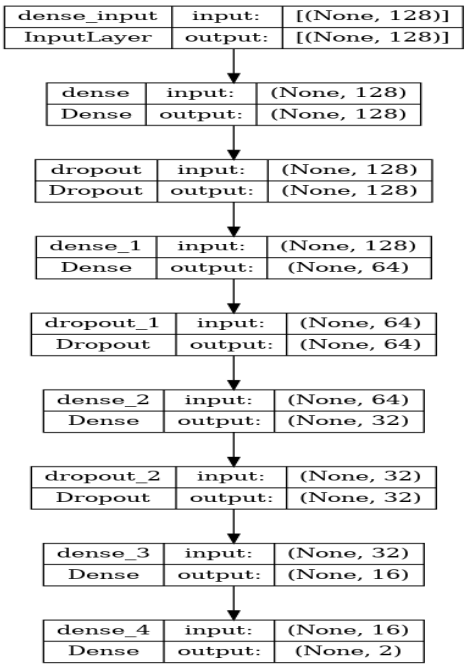
After extracting the above features from the mp3 audio files we store in a corresponding .npy we train our model on these features

Training

We have used the tensorflow.keras to train the model here is the architecture of the model

Group 29

B shiv kiran(2000500xx), B Likith Rajesh(2000500xx), D Venkatesh Prasad(2000500xx), L Jitendra(2000500xx)



Testing

To run a test with a given audio file you can simply run python3 test.py {path_to_audio_file}

Or to record your mic you can simply run python3 test.py