Voice Gender Detection Model

TASK:

Develop a machine learning model that is capable of predicting the gender of a speaker based on a short audio clip. This task involves handling audio data pre-processing, feature extraction, and applying the appropriate machine learning techniques to achieve reliable predictions.

How the Task was implemented:

The type of approach towards this model I used was GMM(Gaussian Mixture Model: Is a parametric probability density function represented as a weighted sum of Gaussian component densities.)

For the above task the data required was Audio Mp3 files which was taken from Kaggle and an labeled CSV file from the same Dataset which consisted of the name of the files and the gender of the speakers. From those Mp3 files the MFCC(Mel-frequency cepstrum coefficients were extracted in order to feed the model and train it from those features.

After Training the model, manipulating its hyperparameters and testing it using the test dataset from Kaggle I was able to predict the gender of the speakers from the Test dataset correctly.

Links

Dataset: Common Voice (kaggle.com)

Audio: Common Voice (kaggle.com)

References: Vol-8-issue-4-M-30.pdf (jncet.org)