

Male and Female Audio Classification with TIMIT Dataset

Anshul Maurya

Florida Institute of Technology

Date: May 03, 2023

AGENDA:

- **Problem Statement**
- **Motivation**
- **Proposed Methodology**
- **Implementation**
- **Experimental Results**

Problem Statement

Problem Statement:

- From the TIMIT dataset extract features information which are close to the pitch.
- Train a Machine Learning model which can discriminate between Male and Female voices based on the extracted features from the dataset.

Motivation

Motivation:

- In the class we discussed speaker identification is a hard problem, because:
 - **Variability in speech** : due to age/environment etc
 - **Background noise** : Speech signals can be corrupted by background noise
 - **Many other factors**

- In this project, I tried to solve a subset of this problem to form a method of more complex speaker identification task.

- Scope of this project is limited to discriminate the male and Female voices.

Proposed Methodology

Proposed Methodology:

- Take a small subset of the data features and analyze which features pairing give the best classification.
 - A NN can take the features and can do the classification without this exercise, but knowing the features is beneficial
- Based on the above observation, plot those features and check with Linear Discriminative Analysis (LDA) and Quadratic Discriminative Analysis (QDA) and check which method fits the data well and reduces the loss.
- Features related to the frequency variation could be a good guess, As we are interested in speaker's Audio Frequency related information to find a pitch related information.

Implementation

Implementation:

Tools Used:



A Python library for audio feature extraction, classification, segmentation and applications

- pyAudioAnalysis is an open source project for audio analysis

Source : <https://github.com/tyiannak/pyAudioAnalysis>



- A industry standard ML library



- Jupyter provides the environment best for the data science related projects.
- This project used pyAudioAnalysis on Jupyter Notebook

Implementation:

- PyAudioAnalysis extracts 138 features from the .wav file, Including MFCC, Chroma, and many other variational features related to frequency analysis.
- After Investigation we found few best pairs very useful in classifying male and female voices:
 - spectral centroid & spectral flux mean
 - Spectral_spread_mean & spectral flux mean ***Best Pair***
 - delta energy_entropy_mean & spectral flux mean
 - delta spectral_centroid_mean & delta energy_entropy_mean

Implementation:

- Pair of 3 feature **spectral_flux_mean**, **spectral_spread_mean**, **energy_mean** to train our classifier.
 1. **Spectral_flux_mean:** A measure of how quickly the power spectrum of a signal is changing, calculated by comparing the power spectrum for one frame against the power spectrum from the previous frame.
 2. **Spectral_spread_mean:** Mean value of the second central moment of the spectrum of current and previous frame.
 3. **Energy_mean:** The sum of squares of the signal values, normalized by the respective frame length.

Source:

<https://journals.plos.org/plosone/article/file%3Fid%3D10.1371/journal.pone.0144610%26type%3Dprintable>

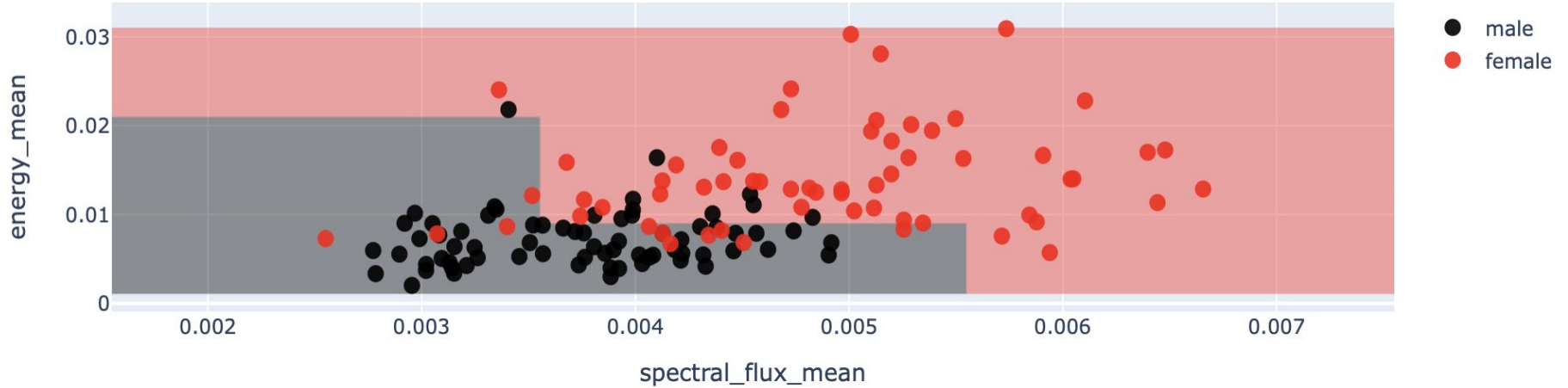
Implementation:

- After extracting the features of the 15 ms frame windows we are ready to use Scikit-Learn's LDA and QDA classifier.
- Below functions from scikit learn will help us to train our classifier:
 1. `scikit.QuadraticDiscriminantAnalysis()`
 2. `scikit.LinearDiscriminantAnalysis()`
 3. `scikit.fit(Data, respectiveLabel)`

Results

Results:

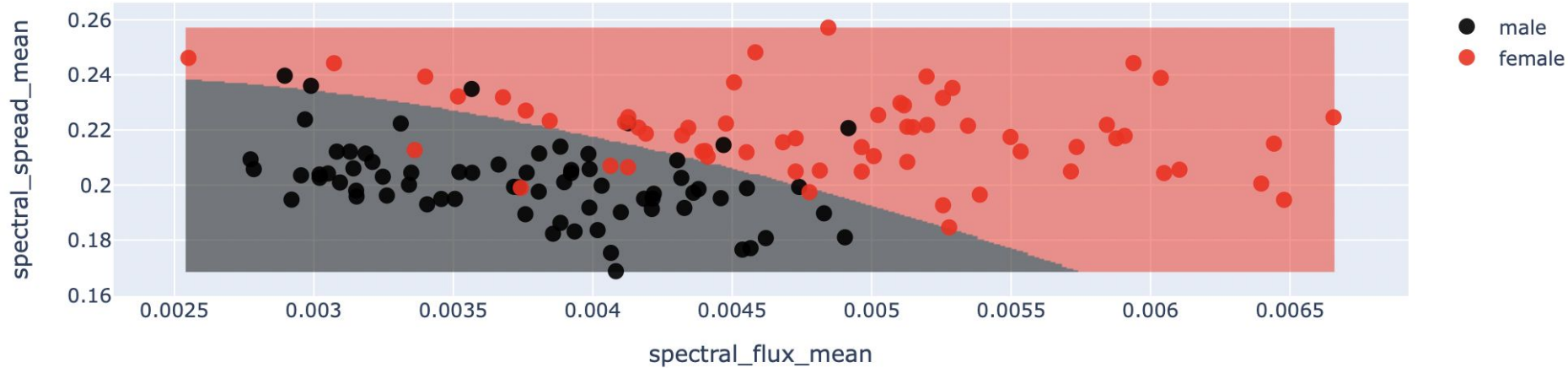
LDA:



- After Training we can see that our model is classifying the Male and Female Audio
- For the sake of simplicity graph only shows two features while model is trained on 3 features as discussed before.

Results:

QDA:



- After Training we can see that our model is classifying the Male and Female Audio.
- For the sake of simplicity graph only shows two features while model is trained on 3 features as discussed before.

Questions?

Thank You