**Folder Description:**

Files attached are:

1. AudioClassifer.jar

To run the jar:

1. Please go to the directory path where jar is placed.

2. run **java –jar AudioClassifier.jar**

**OR**

3. Try double clicking the jar file

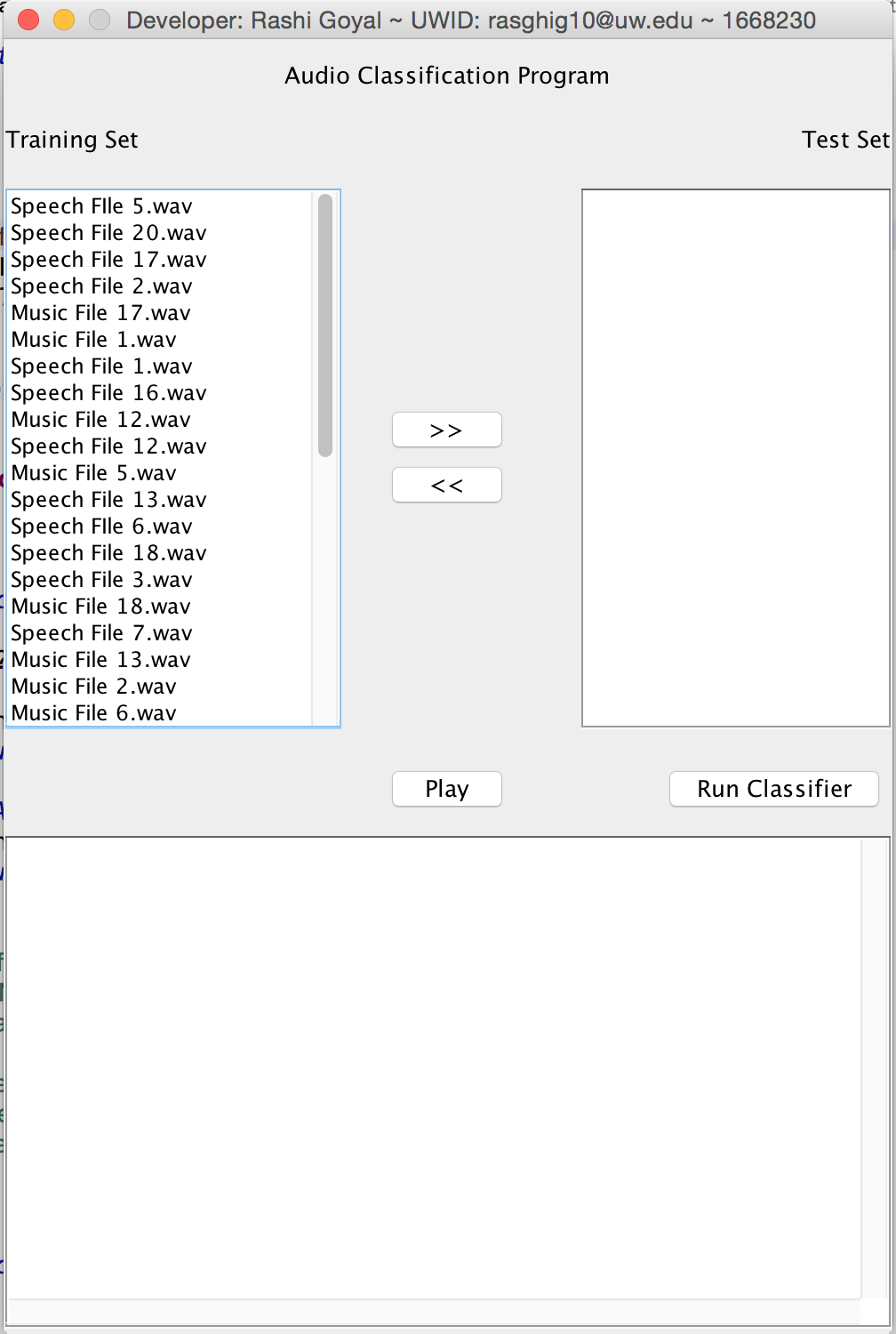
1. MultimediaAssignment4 Folder with Java Code files.
2. Assignment4\_Report\_Rashi Goyal report file.

**Please note:** feature values are being printed in the GUI.

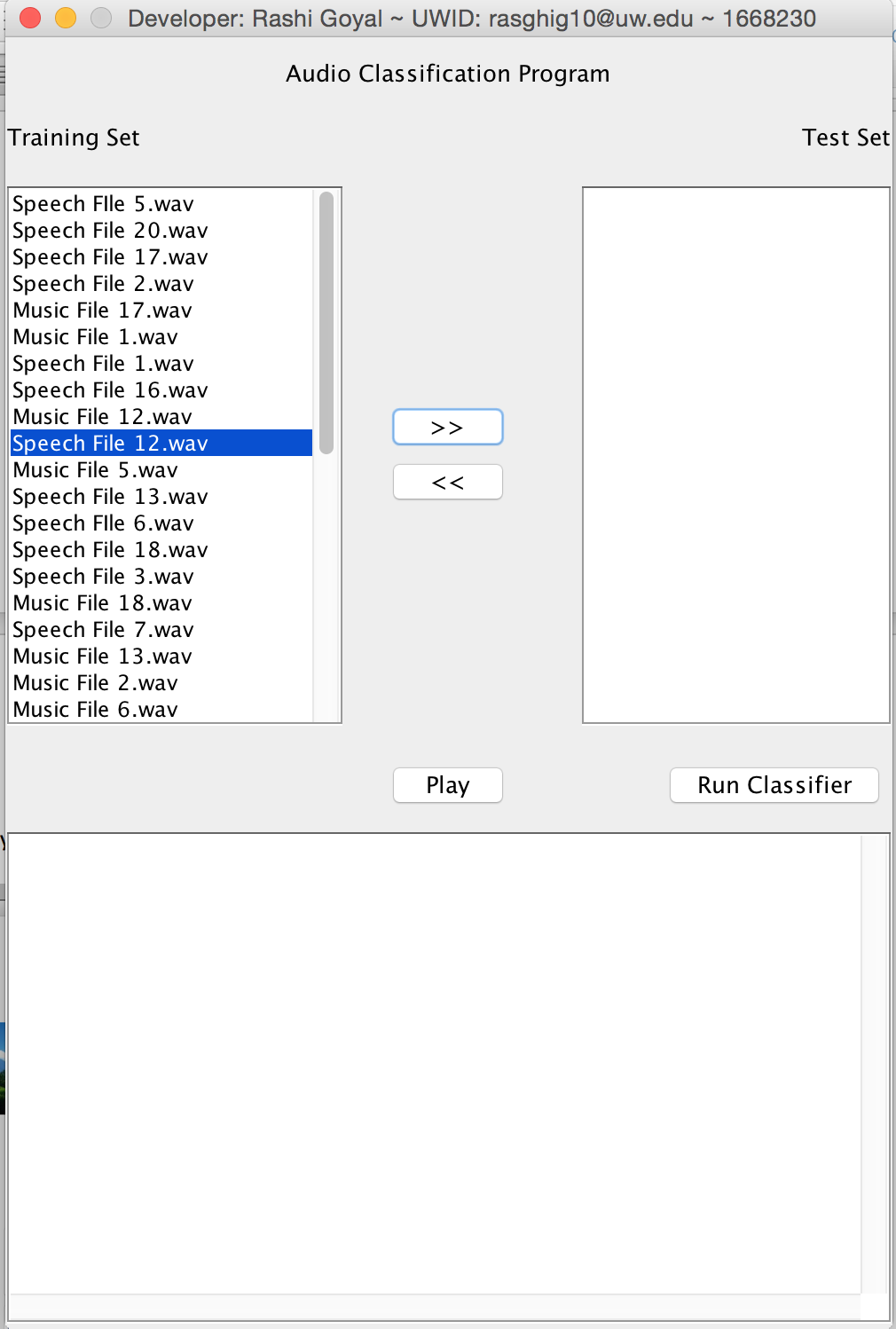
**System Functionality:**

(Please scroll below)

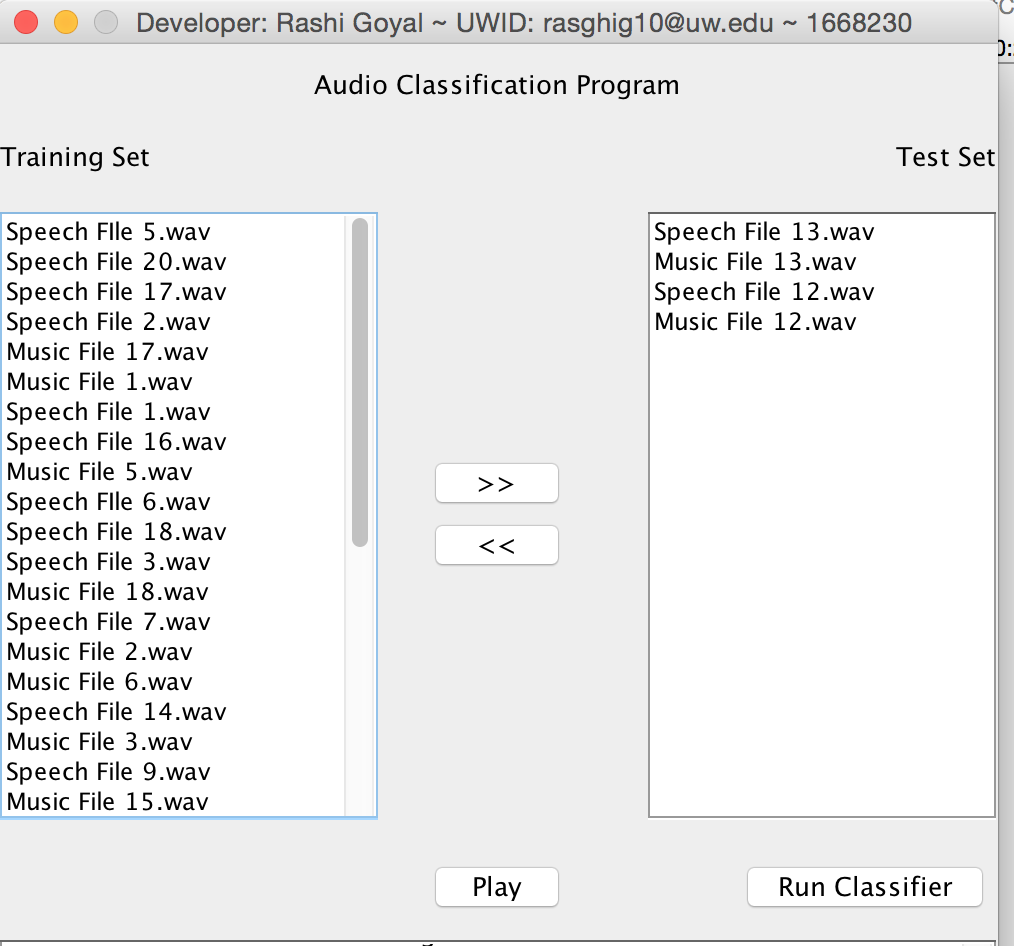
1. On running the jar, a webpage will be displayed as shown below:



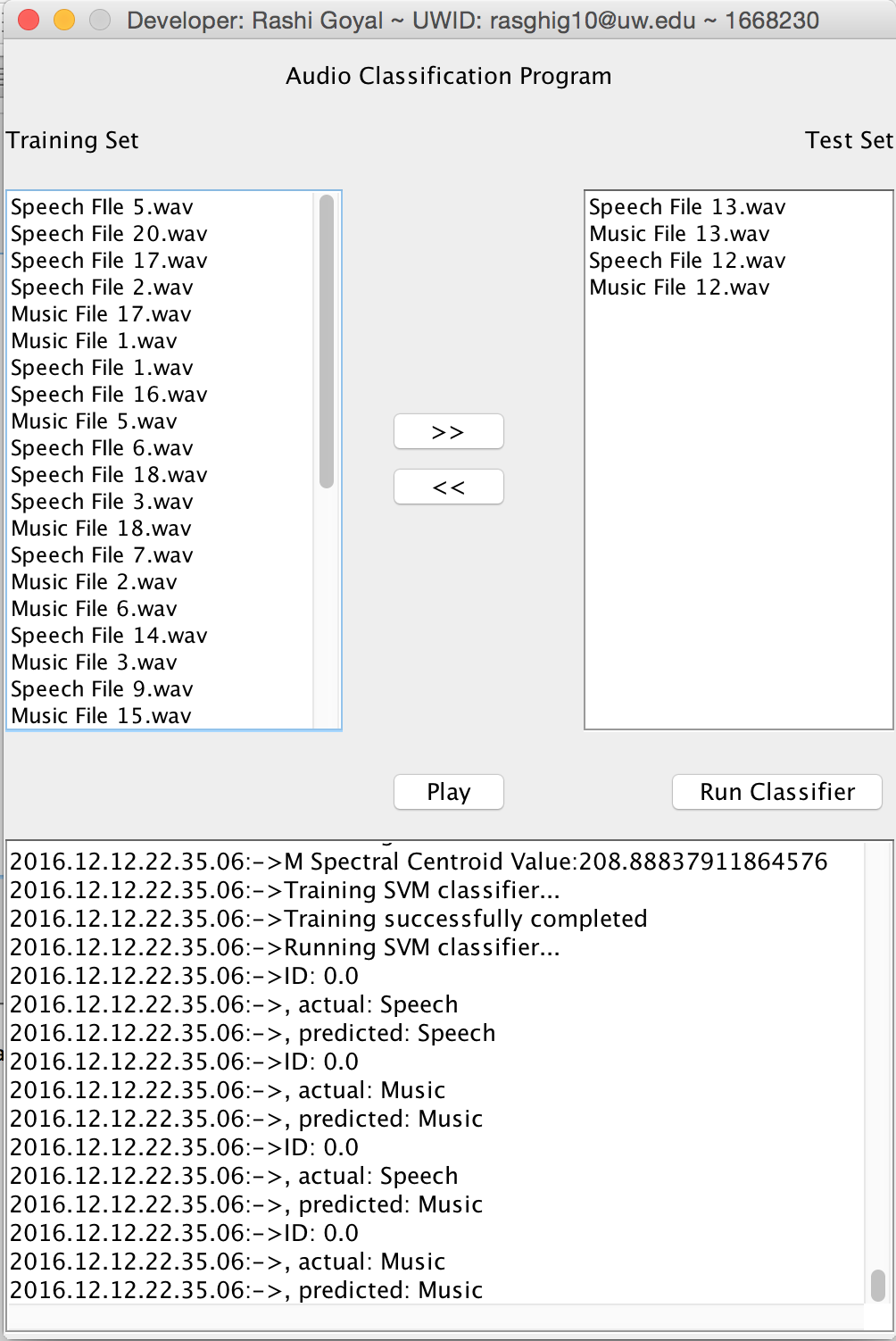
1. Please select a File from the Training Set for Testing. All files are by default added to training set. Please select at least 3-4 files for classifier Testing

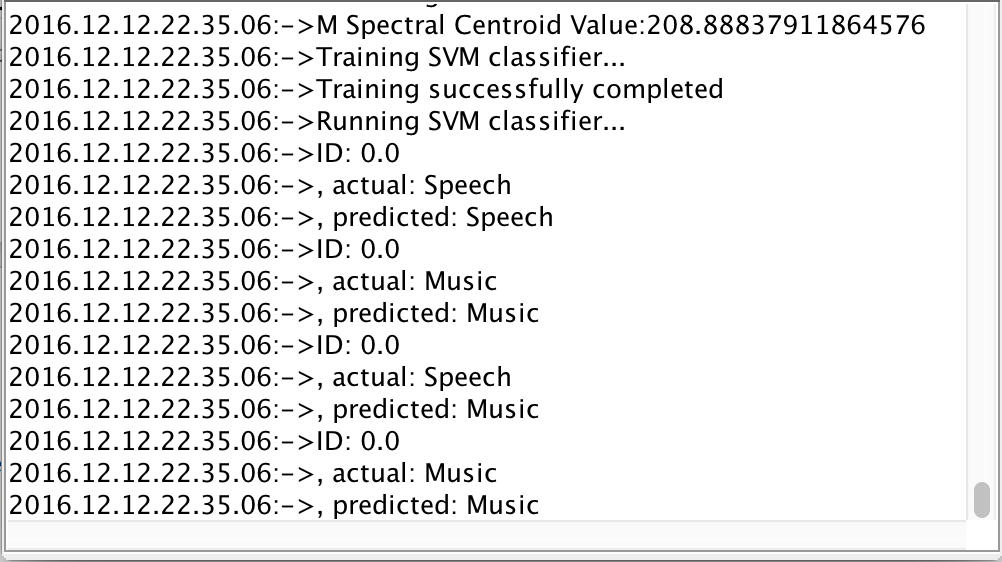


After selecting the files

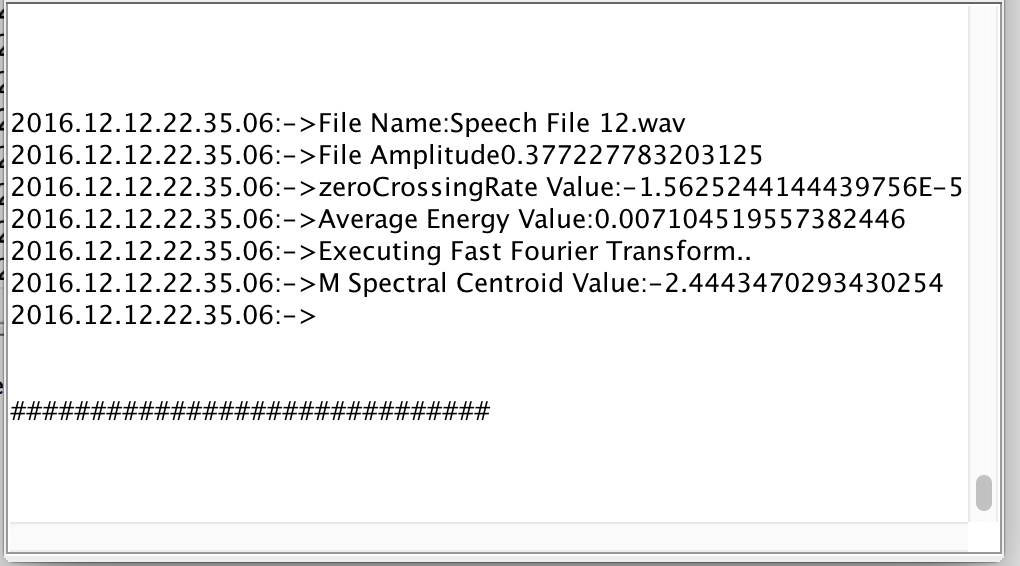


Please not that I am taking 4 files i.e. 2 Music files and 2 Speech files for testing here.

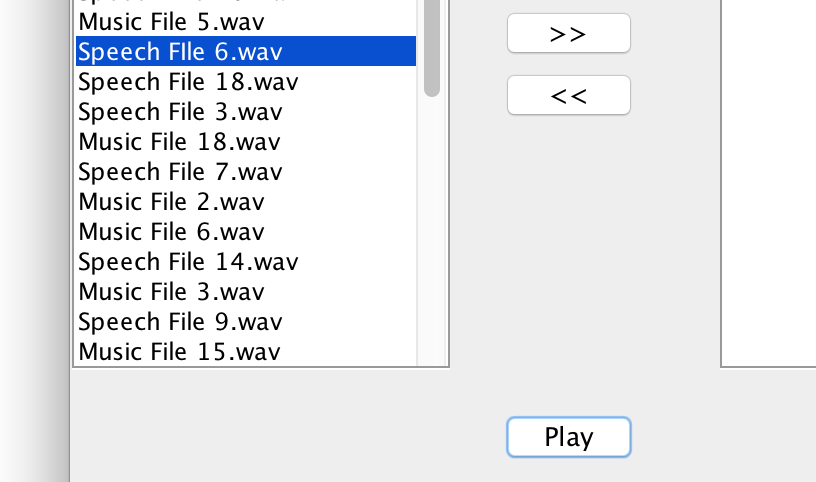
1. Please click Run Classifier Button to run the SVM classification. 
2. You can also notice that the log console at the bottom of the screen will be displaying feature values as well as the classification results.



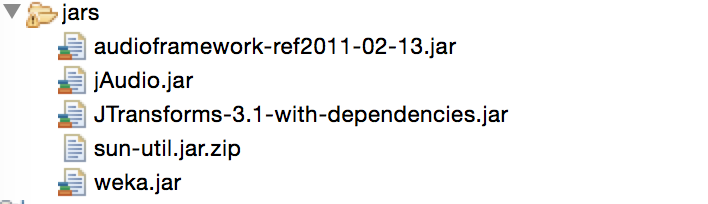
1. You can also notice that if you scroll up the the log console you can see feature values for audio files.



1. You can also play any audio file. Please make sure to select the file you want to play & your speaker volume is up.



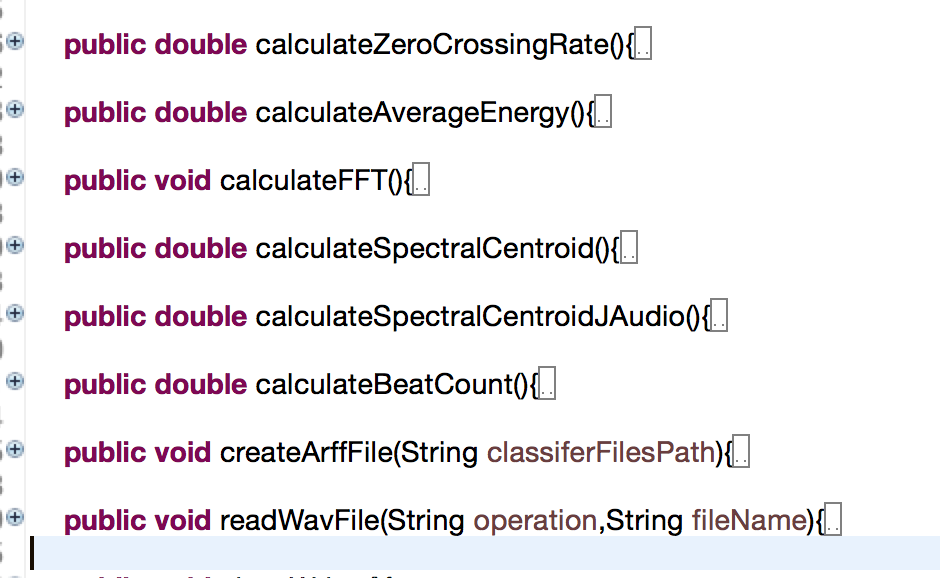
**Libraries Used:**

****

1. **JAudio:** It is used to extract Samples, Sampling Rates, Max Amplitude.
2. **JTransforms:** It is used to extract Fast Fourier Transformation values.
3. **Weka:** It is used to apply Support Vector Machine classifier also known as SMO in Weka

**Please Note:**

1. JAudio was returning more than 64064 FFT values after applying the FFT feature extraction. However, JTransforms provided exactly 64064 fft values which is equal to my sample count. Therefore, I used JTransforms to apply FFT.
2. Zero Crossing Rate & Spectral Centroid values calculated using JAudio were very High. I also looked at the code at Github for them which did not look similar to the formulas explained in class. So, I have manually calculated **Average Energy, Zero Crossing Rate, Spectral Centroid Values**. I have frequency values extracted by JTransforms FFT method.
3. I also experimented with Beat Count, Max Amplitude values extracted using JAudio library. But it negatively affected the classification results.
4. You can look at ReadAudioFile.java class code to review the feature extraction code.



**Features & their Values for all Audio Files:**

Below is the log from the console after running the classifier.

I have used three features to train my model.

1. Average Energy
2. Zero Crossing Rate
3. Spectral Centroid

The values below are printing for each audio file in the format:

File Name:

File Amplitude

zeroCrossingRate:

Average Energy Value:

M Spectral Centroid Value:

Please go through the log below to check the values for individual audio files. At the end of the log, Results are displayed.

2016.12.12.22.35.00:->File Name:Speech FIle 5.wav

2016.12.12.22.35.00:->File Amplitude0.14239501953125

2016.12.12.22.35.00:->zeroCrossingRate Value:0.0

2016.12.12.22.35.00:->Average Energy Value:7.194257057271898E-4

2016.12.12.22.35.00:->Executing Fast Fourier Transform..

2016.12.12.22.35.00:->M Spectral Centroid Value:5.405069726395503

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 20.wav

2016.12.12.22.35.01:->File Amplitude0.62310791015625

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.01583643425212111

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-35.040790246861526

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 17.wav

2016.12.12.22.35.01:->File Amplitude1.0

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.06188370646098337

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-104.90330787487864

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 2.wav

2016.12.12.22.35.01:->File Amplitude0.12591552734375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:4.5428834724589253E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-157.81002648804304

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 17.wav

2016.12.12.22.35.01:->File Amplitude1.0

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.07671875950264803

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:784.2031000840129

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 1.wav

2016.12.12.22.35.01:->File Amplitude0.656585693359375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.013580961924918637

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-116.10562145282039

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 1.wav

2016.12.12.22.35.01:->File Amplitude0.164642333984375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:5.281566308112815E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:47.65292443603363

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 16.wav

2016.12.12.22.35.01:->File Amplitude0.999969482421875

2016.12.12.22.35.01:->zeroCrossingRate Value:1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:0.035685484402129077

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:138.78245077607124

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 5.wav

2016.12.12.22.35.01:->File Amplitude0.125946044921875

2016.12.12.22.35.01:->zeroCrossingRate Value:1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:8.042437358817551E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-28.26415588443338

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech FIle 6.wav

2016.12.12.22.35.01:->File Amplitude0.147064208984375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:3.178433641926483E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-0.8068365590248734

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 18.wav

2016.12.12.22.35.01:->File Amplitude0.409149169921875

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.560963426626914E-5

2016.12.12.22.35.01:->Average Energy Value:0.0035985503325520964

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-1.954669155448461

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 3.wav

2016.12.12.22.35.01:->File Amplitude0.236907958984375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:5.023559251130791E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:4.357909966317283

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 18.wav

2016.12.12.22.35.01:->File Amplitude0.30572509765625

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.003944127021080931

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:0.11204822182252519

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 7.wav

2016.12.12.22.35.01:->File Amplitude0.41754150390625

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.008307291079458082

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:22.859274116772262

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 2.wav

2016.12.12.22.35.01:->File Amplitude0.583953857421875

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625488296509265E-5

2016.12.12.22.35.01:->Average Energy Value:0.009680068761833372

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:67.93819313792795

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 6.wav

2016.12.12.22.35.01:->File Amplitude0.505096435546875

2016.12.12.22.35.01:->zeroCrossingRate Value:1.5625488296509265E-5

2016.12.12.22.35.01:->Average Energy Value:0.01100800370500622

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:7.323168814118677

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 14.wav

2016.12.12.22.35.01:->File Amplitude0.5279541015625

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.006607371829483275

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:1.0631182010842073

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 3.wav

2016.12.12.22.35.01:->File Amplitude0.44927978515625

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625488296509265E-5

2016.12.12.22.35.01:->Average Energy Value:0.007852619107573464

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:270.1265134820597

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 9.wav

2016.12.12.22.35.01:->File Amplitude1.0

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:0.1503126584312122

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-1.9532606650131725

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 15.wav

2016.12.12.22.35.01:->File Amplitude0.174407958984375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.0020078544451243943

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-0.09607657584184574

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 19.wav

2016.12.12.22.35.01:->File Amplitude0.500152587890625

2016.12.12.22.35.01:->zeroCrossingRate Value:1.560963426626914E-5

2016.12.12.22.35.01:->Average Energy Value:0.003410028306289123

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:6.889205184806167

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 4.wav

2016.12.12.22.35.01:->File Amplitude1.0

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.1063265164035256

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:20.524599326025804

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 19.wav

2016.12.12.22.35.01:->File Amplitude0.2928466796875

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.001811424343395629

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:9.505730358094112

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 14.wav

2016.12.12.22.35.01:->File Amplitude0.881683349609375

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625488296509265E-5

2016.12.12.22.35.01:->Average Energy Value:0.03666144637375401

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-13.207466232600888

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music FIle 20.wav

2016.12.12.22.35.01:->File Amplitude0.220794677734375

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:6.723543111729668E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:25.31922381755154

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 10.wav

2016.12.12.22.35.01:->File Amplitude0.385223388671875

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625488296509265E-5

2016.12.12.22.35.01:->Average Energy Value:0.00612919615867232

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-13.08855677064656

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 7.wav

2016.12.12.22.35.01:->File Amplitude0.122772216796875

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625488296509265E-5

2016.12.12.22.35.01:->Average Energy Value:4.711084564748068E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-5.671823250011399

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 10.wav

2016.12.12.22.35.01:->File Amplitude1.0

2016.12.12.22.35.01:->zeroCrossingRate Value:1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:0.139020341253432

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:1.2081322090113622

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 4.wav

2016.12.12.22.35.01:->File Amplitude0.369903564453125

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.006939856863414269

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-325.5800562224756

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 11.wav

2016.12.12.22.35.01:->File Amplitude0.4307861328125

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.006398781227301685

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-26.49899406401108

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 16.wav

2016.12.12.22.35.01:->File Amplitude0.612274169921875

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.017265332418450272

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-31.40751302357776

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 9.wav

2016.12.12.22.35.01:->File Amplitude0.175750732421875

2016.12.12.22.35.01:->zeroCrossingRate Value:1.560963426626914E-5

2016.12.12.22.35.01:->Average Energy Value:0.0022246069743382425

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-53.36276883650913

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 8.wav

2016.12.12.22.35.01:->File Amplitude0.572235107421875

2016.12.12.22.35.01:->zeroCrossingRate Value:1.560963426626914E-5

2016.12.12.22.35.01:->Average Energy Value:0.016456336317529605

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:30.74088201389756

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 15.wav

2016.12.12.22.35.01:->File Amplitude0.622833251953125

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.0038153758800326615

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-45.183757398753805

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 8.wav

2016.12.12.22.35.01:->File Amplitude0.147247314453125

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:8.666588370863872E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:9.800333515230758

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 11.wav

2016.12.12.22.35.01:->File Amplitude0.188995361328125

2016.12.12.22.35.01:->zeroCrossingRate Value:1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:5.541150277858833E-4

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-0.9691409680612026

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 13.wav

2016.12.12.22.35.01:->File Amplitude1.0

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.17705465030105552

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-4.119740197579374

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 13.wav

2016.12.12.22.35.01:->File Amplitude0.7344970703125

2016.12.12.22.35.01:->zeroCrossingRate Value:0.0

2016.12.12.22.35.01:->Average Energy Value:0.016870478262704297

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-1.0864845890995734

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Speech File 12.wav

2016.12.12.22.35.01:->File Amplitude0.377227783203125

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:0.007104519557382446

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:-2.4443470293430254

2016.12.12.22.35.01:->

##############################

2016.12.12.22.35.01:->File Name:Music File 12.wav

2016.12.12.22.35.01:->File Amplitude0.96759033203125

2016.12.12.22.35.01:->zeroCrossingRate Value:-1.5625244144439756E-5

2016.12.12.22.35.01:->Average Energy Value:0.010393799339537508

2016.12.12.22.35.01:->Executing Fast Fourier Transform..

2016.12.12.22.35.01:->M Spectral Centroid Value:208.88837911864576

2016.12.12.22.35.01:->Training SVM classifier...

2016.12.12.22.35.01:->Training successfully completed

2016.12.12.22.35.01:->Running SVM classifier...

2016.12.12.22.35.01:->ID: 0.0

2016.12.12.22.35.01:->, actual: Speech

2016.12.12.22.35.01:->, predicted: Speech

2016.12.12.22.35.01:->ID: 0.0

2016.12.12.22.35.01:->, actual: Music

2016.12.12.22.35.01:->, predicted: Music

2016.12.12.22.35.01:->ID: 0.0

2016.12.12.22.35.01:->, actual: Speech

2016.12.12.22.35.01:->, predicted: Music

2016.12.12.22.35.01:->ID: 0.0

2016.12.12.22.35.01:->, actual: Music

2016.12.12.22.35.01:->, predicted: Music

2016.12.12.22.35.05:->

**Comparison between Model Output and Ground Truth Label:**

Confusion Matrix:

(Considering the 2 Music files and 2 audio files I tested for this report as shown in the results above)

|  |  |  |
| --- | --- | --- |
| N=4 | Predicted Label Music | Predicted Label Speech |
| Actual Label Music | 2 | 0 |
| Actual Label Speech | 1 | 1 |

Accuracy:

* (True Positive + True Negative)/Total
* (2+1)/4
* 75%

Precision:

* True Positive/Predicted Total Positive
* Precision for Music = 2/3
  + .66
* Precision for Speech= 0/1
  + 0

Recall:

* True Positive/Ground Truth
* Recall for Music=2/2
  + 1
* Recall for Speech=1/2
  + .5