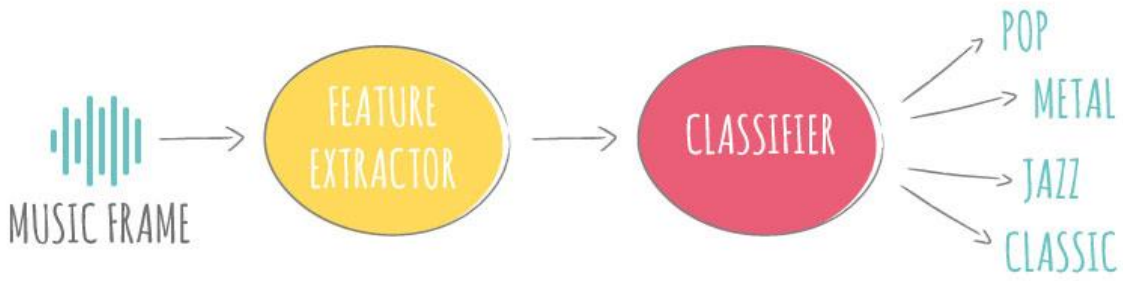
Music Genre Classification

Many companies nowadays use music classification, either to be able to place recommendations to their customers (such as Spotify, Soundcloud), or simply as a product (for example Shazam). Determining specific music genres is a first step towards this goal.

The basic flowchart for music genre classification is shown below:



The following features were used for performing classification:

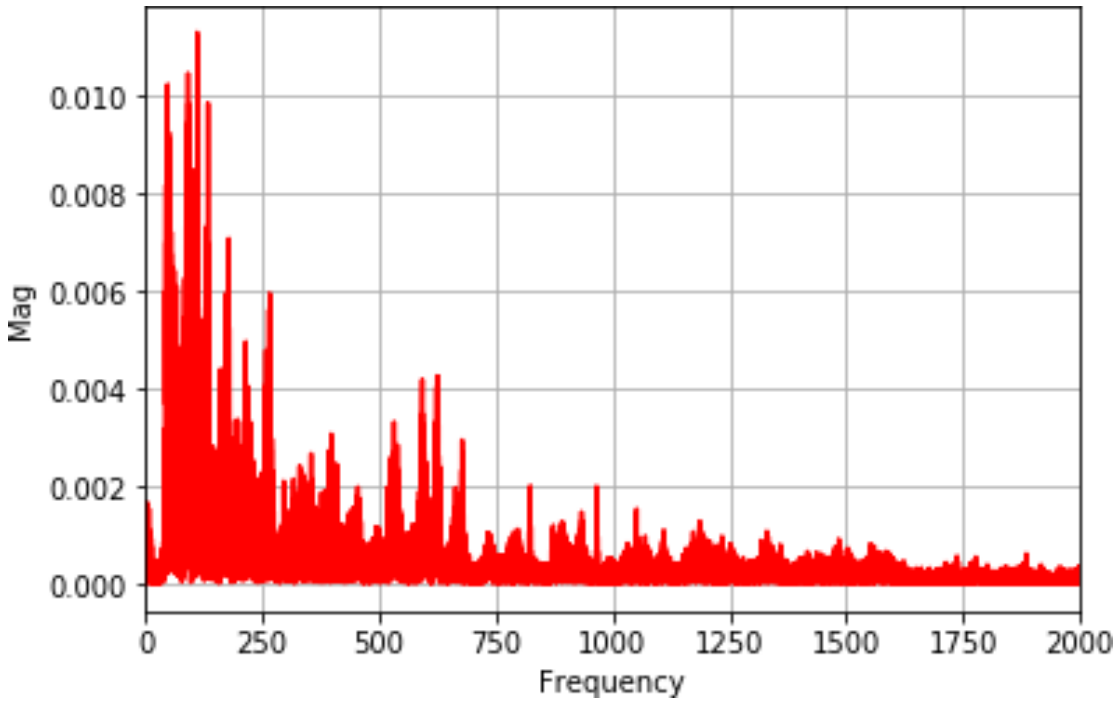
1) Chroma Frequencies
2) RMSE (Root Mean Squared Energy)
3) Spectral Centroid
4) Spectral Bandwidth
5) Spectral Roll-off
6) Zero Crossing Rate
7) Mel Frequency Cepstral Coefficients (MFCCs)

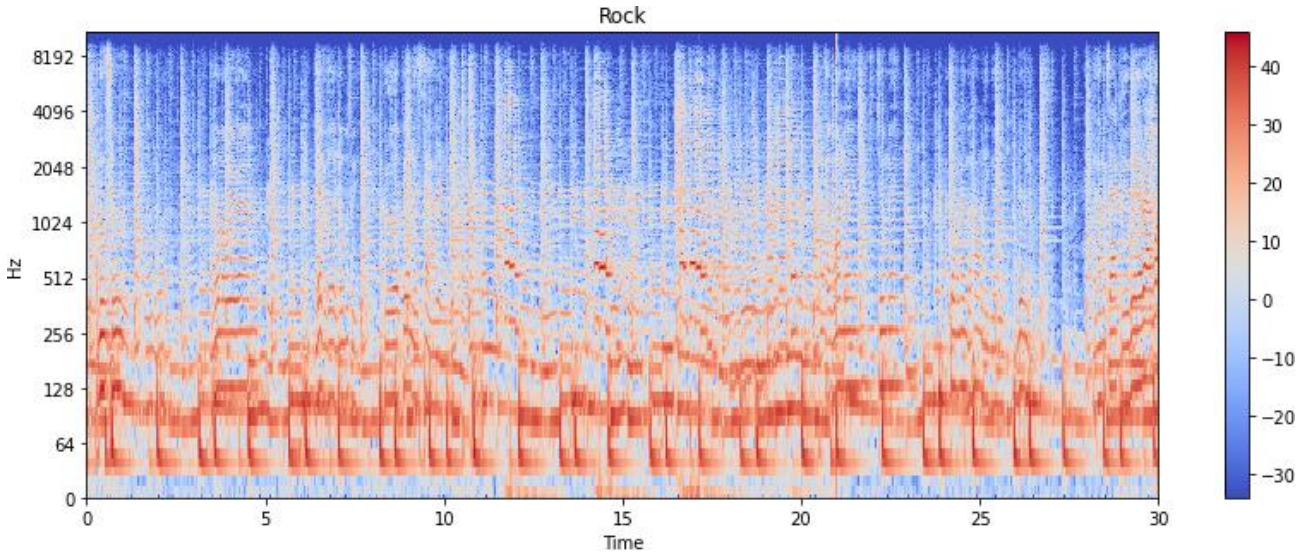
INFERENCES FROM THE FOURIER TRANSFORM OF THE SONGS

The following observations can be made from the graph of Fourier Transform of the songs. In the case of classical songs, the frequencies are uniformly distributed over the range of 20-2000 Hz while in the case of rock song most of the frequencies are between 20-250 Hz.

Thus, using frequency-based features such as spectral centroid, spectral roll-off and MFCCs would help us in correctly classifying the song.

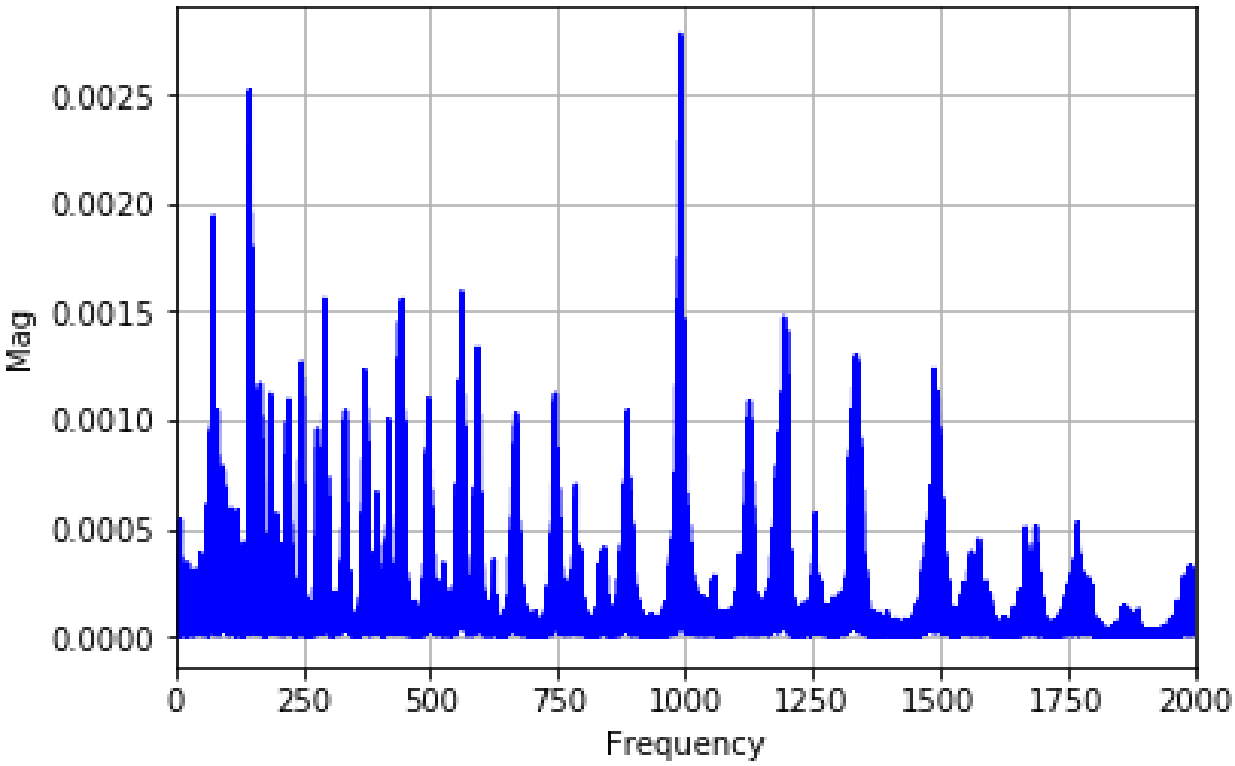
Fourier Transform of Rock Song

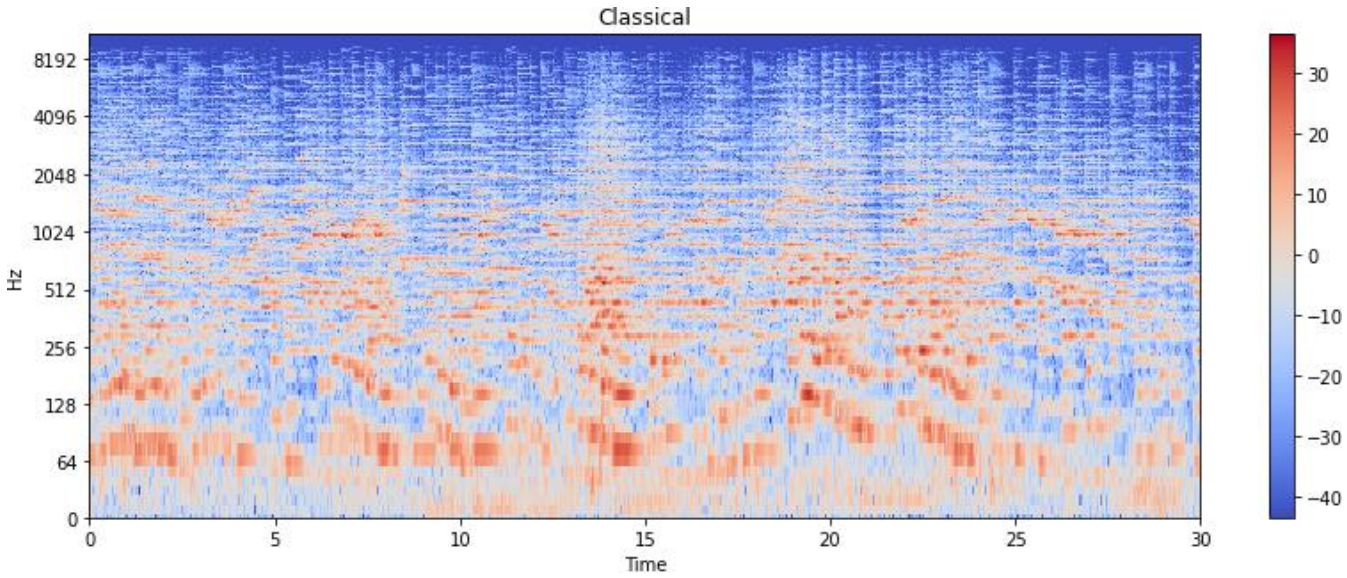




Spectrogram

Fourier Transform of Classical Song

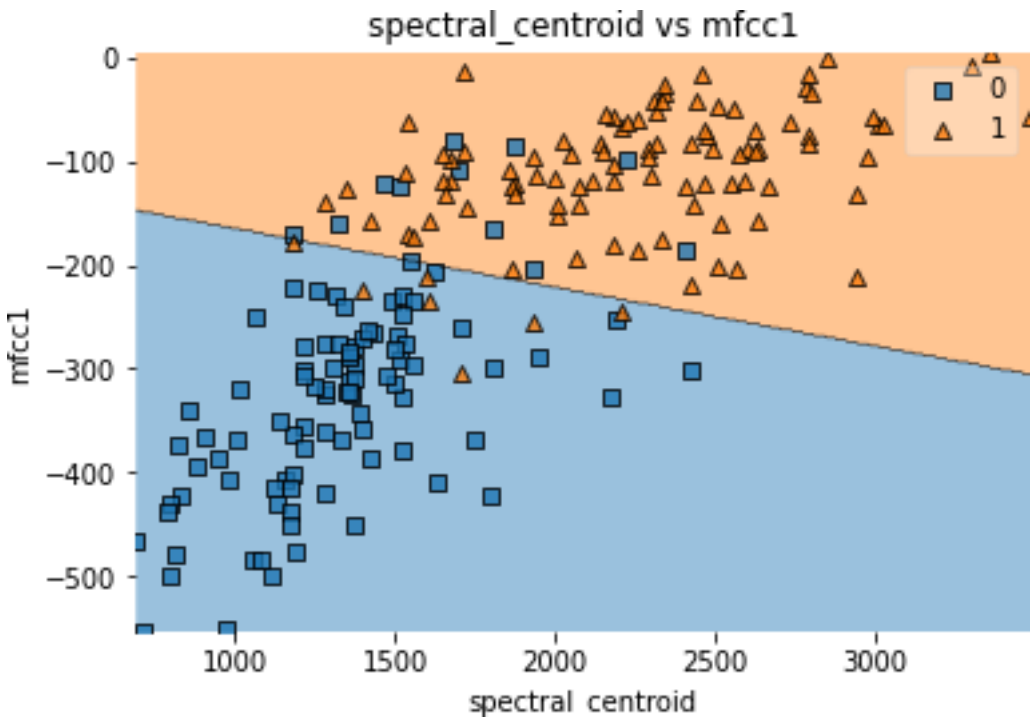


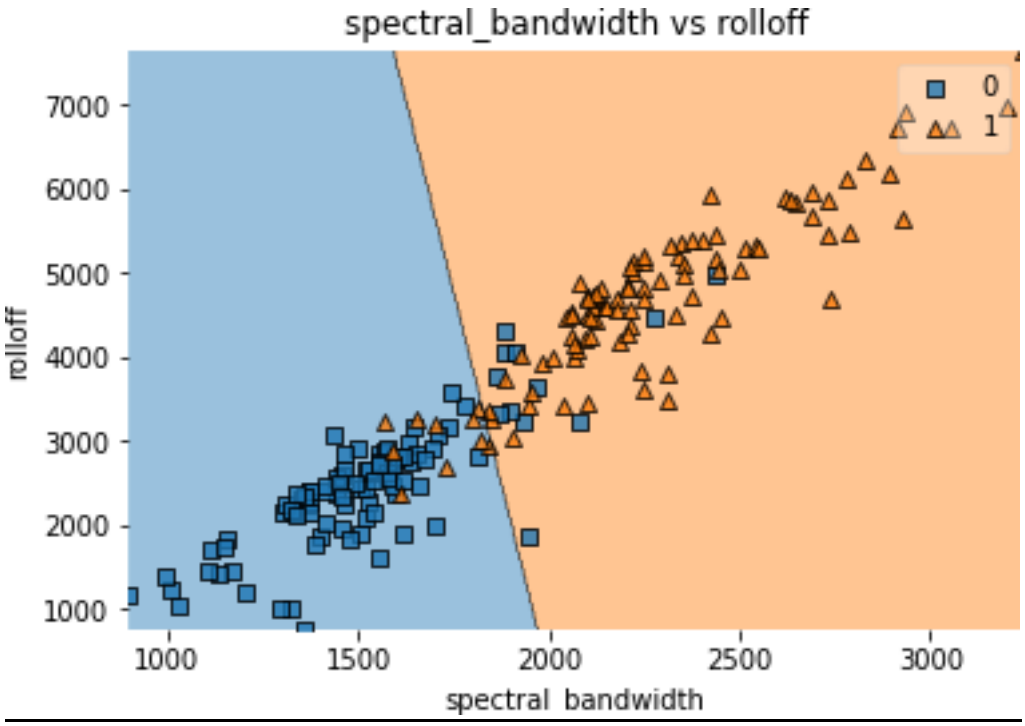


Spectrogram

FEATURE SELECTION FOR THE MODEL

To see which features to use for training the Neural Network, we implemented an SVM and plotted the decision boundary to check how well the selected features were performing.





The linear decision boundary does a pretty good job at separating the two genres of music. Thus, these two features can be considered for building our model. Similarly, graphs for other features can be plotted and selected in our model.

BUILDING OUR MODEL

We perform feature scaling on the input data to normalise it. Also, we encode the output as 0 as classical and 1 as rock. We also split the data into train data and test data.

To build a neural network we import it from the Keras library and train it on our train data. We have also implemented regularisation to prevent over fitting. We see an accuracy of 0.9625 on our train data after 20 epochs of training. On our test data we see an accuracy of 0.95 which is quite close to our accuracy of train data.

CONCLUSIONS

As we had to classify between only two genres, our simple Neural Network performed quite well but if we had to classify between a lot of genres, we would have to use CNNs-RNNs or ensemble techniques for more accuracy.

Predictions of the model on test files

Name of File	Classical	Rock	Result
test1.au	0.472	0.527	rock
test2.au	0.0721	0.922	rock
test3.au	0.0655	0.934	rock
test4.au	0.052	0.947	rock
test5.au	0.030	0.969	rock