

---

# Music Genre Classification

Mandeep Singh (19534)  
Parimal Gajhbhiye (19584)

MLSP Final Project 2021

---

- Music genre classification aims to classify the music files into certain categories of genres it belongs to.
  - Automatic categorization of music into different genres is done using Machine learning / Deep learning tools.
  - In this project, we are extracting features from the music data (GTZAN) and then using different learning models to perform the classification.
-

### Features:

- Spectral centroid
- Spectral flux
- Root Mean Square Energy (RMSE)
- Zero-crossing rate (ZCR)
- Spectral contrast
- Spectral bandwidth
- Spectral flatness
- Spectral roll-off
- Mel-Frequency Cepstral Coefficients (MFCC) (20 in number)
- Chroma
- Tonnetz
- Tempo
- Pitch

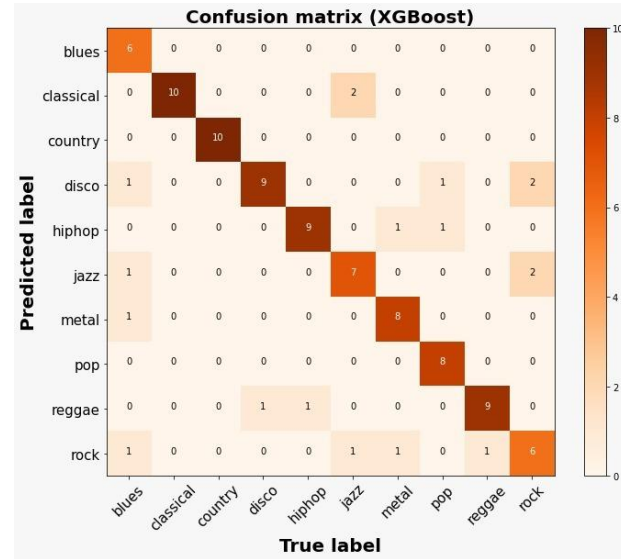
### Dataset Details :

- ❑ 1000 Music Files
- ❑ 10 Different Genres
- ❑ Each genre has 100 Music Files
- ❑ Each Music File is 30 seconds long with a sample rate of 22050 Hz
- ❑ Genres: Blues, Classical, Country, Disco, Hip-Hop, Jazz, Pop, Metal, Reggae, and Rock.
- For Frequency based features, we have used window length = 1024 and hop length = 512 for Short-time Fourier transform(STFT).
- For each feature (except tempo) we have taken 6 statistics such as maximum, minimum, mean, standard deviation, kurtosis, and skewness.
- Total number of features for each sample becomes 31 (20 MFCCs + 11 other features) \* 6(statistics) + 1(tempo)= 187
- Size of data = 1000\*187
- For CNN and LSTM, we divide each file into 10 segments and then get MFCC features and use that dataset for modelling.
- We use 10% of total data as test data. For CNN, we use 10% of training data as validation data

# Results

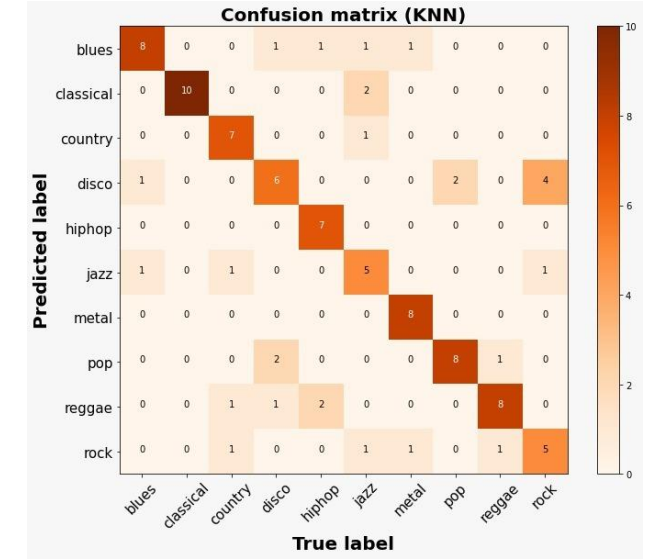
## XGBoost

- Max\_Depth = 15
- Min\_Child\_Weight = 0.5
- Estimators = 100
- Accuracy =  $80 \pm 2\%$
- Except Rock, Jazz and Blues, all were well classified



## KNN

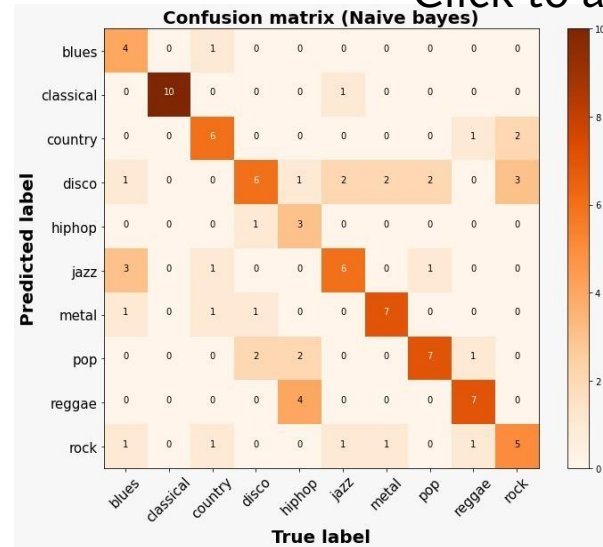
- Number of Neighbours = 5
- Leaf Size = 20
- Metric = Minowski
- Accuracy =  $72 \pm 4\%$
- Classical, Blues, Metal, Pop and Reggae were well classified



Click to add text

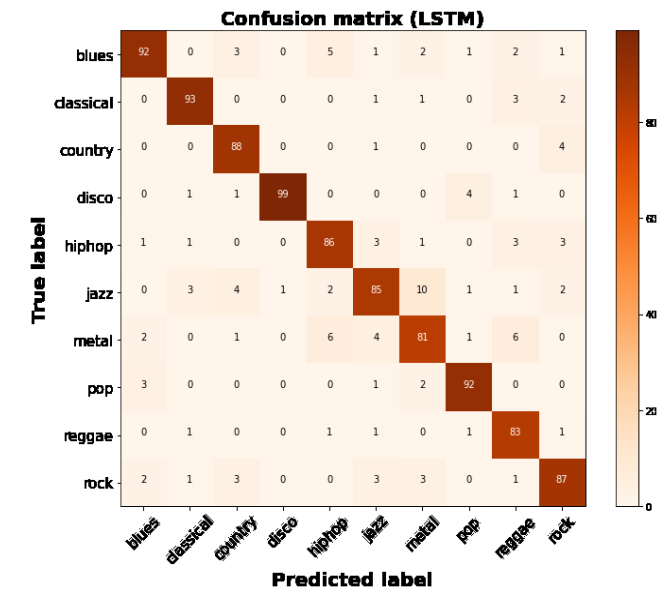
## Naïve Bayes

- Accuracy =  $61 \pm 5\%$
- Only Classical was well classified



## LSTM

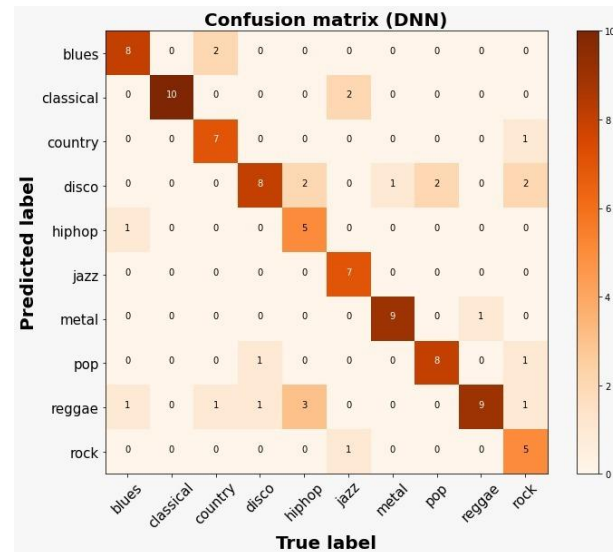
- LSTM layers = 2
- Dense layers = 1
- Activation Function = Relu, Softmax
- Optimizer = Adam
- Loss Function = Sparse Categorical Cross Entropy
- Accuracy =  $89 \pm 1\%$
- All were well classified



# Results

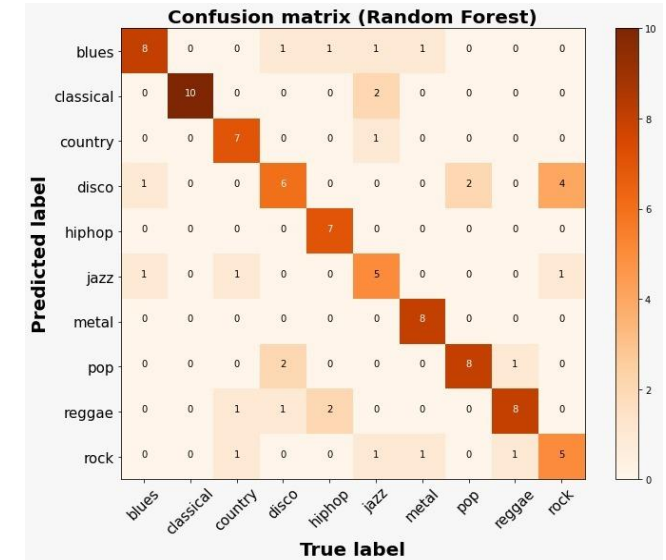
## DNN

- Hidden Layers = 6
- Activation Function = Relu,Softmax
- Optimizer = Adam
- Loss Function = Categorical Cross Entropy
- Accuracy =  $77 \pm 3\%$
- Except Rock, Country, Hiphop and Jazz, all were well classified



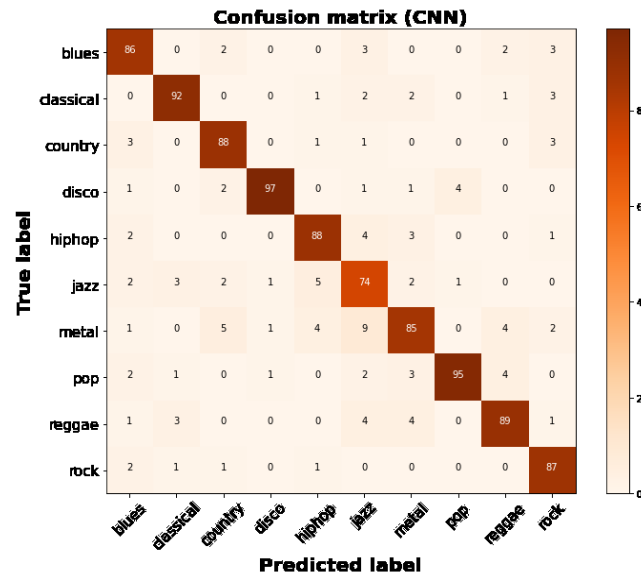
## Random Forest

- Criterion = entropy
- Max\_Depth = 8
- Estimators = 500
- Accuracy =  $74 \pm 4\%$
- Classical, Blues, Metal, Pop and Reggae were well classified



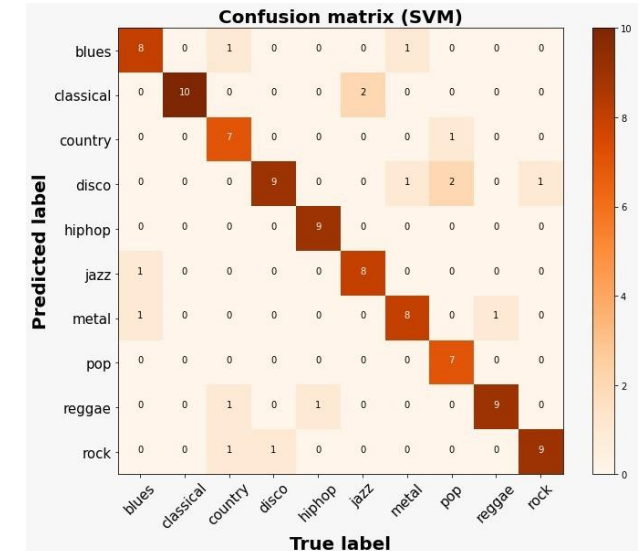
## CNN

- Convolution Layers = 3
- Number of Filters per Convolution Layer = 256
- Dense Layers = 1
- Activation Function = Relu,Softmax
- Optimizer = Adam
- Loss Function = Sparse Categorical Cross Entropy
- Accuracy =  $88 \pm 2\%$
- Except Jazz, all were well classified



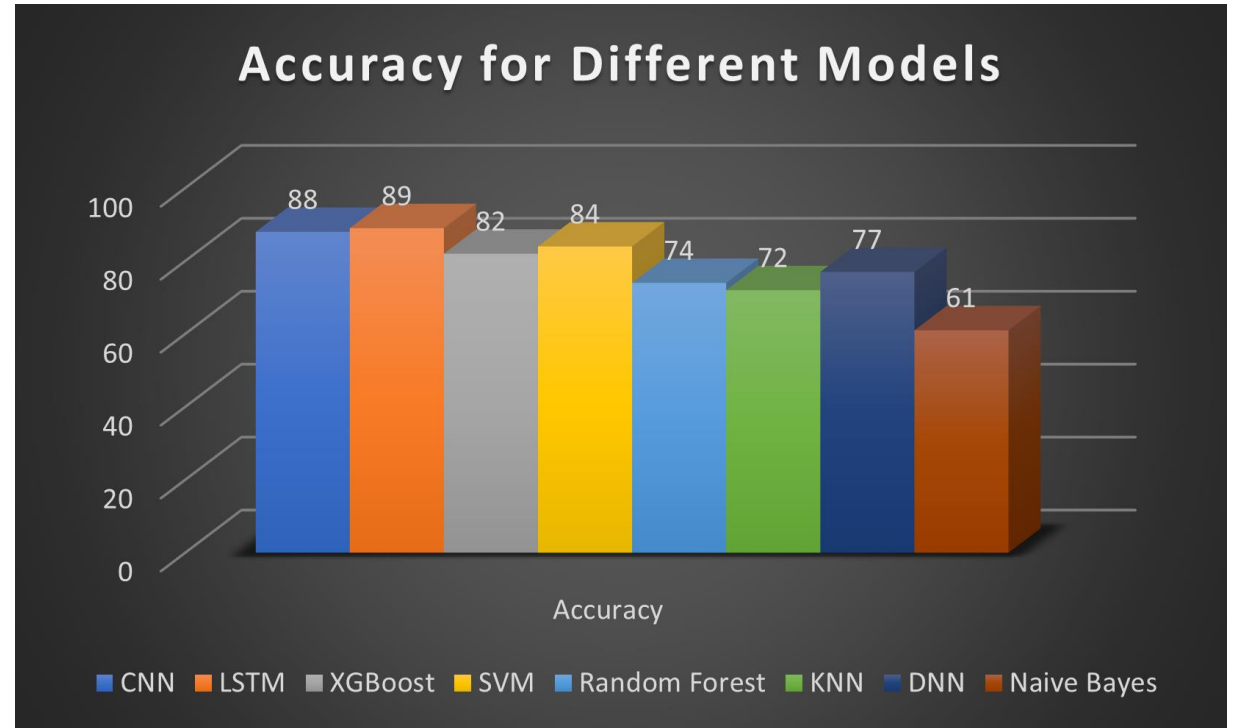
## SVM

- C = 10
- Degree = 1
- Kernel = 'rbf'
- Accuracy =  $84 \pm 2\%$
- Except Pop and Country, all were well classified



## Results

- LSTM and CNN are giving best performance with 89 and 88 percent.
- Performance of XGBoost was improved by tuning the hyperparameters using Randomized-SearchCV.
- GridSearchCV was used to tune the hyperparameters for KNN, SVM and Random Forrest



Using Research Paper [1], we studied the feature extraction from an audio sample for music genre classification purpose. We implemented paper [2], used some additional audio features and implemented additional models such as CNN, LSTM, DNN, XGBoost.

We divided the task amongst ourselves as follows:

- Mandeep Singh : CNN, SVM, DNN, Random Forest, PPT
- Parimal Gajbhiye : Feature Extraction, XGBoost, LSTM, KNN, Naive Bayes, Project report

Tools Used :

The libraries that assisted us in obtaining the results are as follows:

- Numpy
- Matplotlib
- Pandas
- Seaborn
- Scikit-learn: Naive Bayes, SVM, Random Forest, KNN
- Keras: CNN, LSTM

---

[1] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," in IEEE Transactions on Speech and Audio Processing, vol. 10, no. 5, pp. 293-302, July 2002, doi:10.1109/TSA.2002.800560.

[2] A. Elbir, H. Bilal C, am, M. Emre Iyican, B. O`ztu`rk and N. Aydin, "Music Genre Classification and Recommendation by Using Machine Learning Techniques," 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), 2018, pp. 1-5, doi: 10.1109/ASYU.2018.8554016.

---