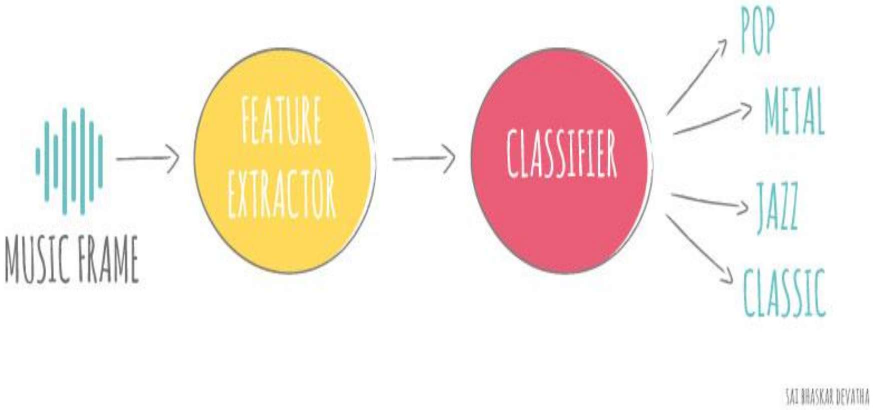
MUSIC GENRE CLASSIFICATION
Interim Project Report

Priyanka singh[2018174]
Manish [2018047]
Aditya Singh [2018378]

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Aim:

Music forms are a very crucial part of our lives. Genre classification is an essential task with many real-world applications. Since the quantity by which music is getting released daily reaches the mountain peaks, the need for a proper and accurate genre classification rises in proportion. In today's era, internet services have a massive amount of multimedia exchanges and browsing. So the need for an effective way to organise and categorise these data rises in proportion. Searching a query song in a database containing millions of songs can become a cumbersome and time-consuming task. Categorising songs into different genres can reduce the search space for these query songs. The main objective of our project is to come up with an effective and accurate machine learning model that can automatically classify the music based on the genre.



Dataset And Features

- The dataset GTZAN used for building music genre classification has been taken from Kaggle . It consists of 1000 samples of songs.The dataset has 10 genre classes:

  1. Reggae, 2. Jazz, 3. Disco, 4. Rock , 5. Metal , 6. Pop , 7. Country, 8. Blues, 9.Classical, 10. Hiphop.Each of the 10 genre classes have 100 examples of songs. Thus the dataset is balanced.

- We used the songs examples present in the dataset to extract 63 features using the librosa module available in python.

  The extracted features includes:

  1. Mel Frequency Cepstral Coefficients(mean and variance)
  2. Spectral bandwidth(mean and variance)
  3. Spectral centroid(mean and variance)
  4. Zero crossing rate(mean and variance)
  5. Spectral roll-off(mean and variance)
  6. Beats location (mean and variance)
  7. Estimated global tempo(mean and variance)
  8. Chroma short time fourier transform (mean and variance)
  9. Root mean square energy for each frame(mean and variance)
  10. Harmonic component(mean and variance)
  11. Percussive component(mean and variance)

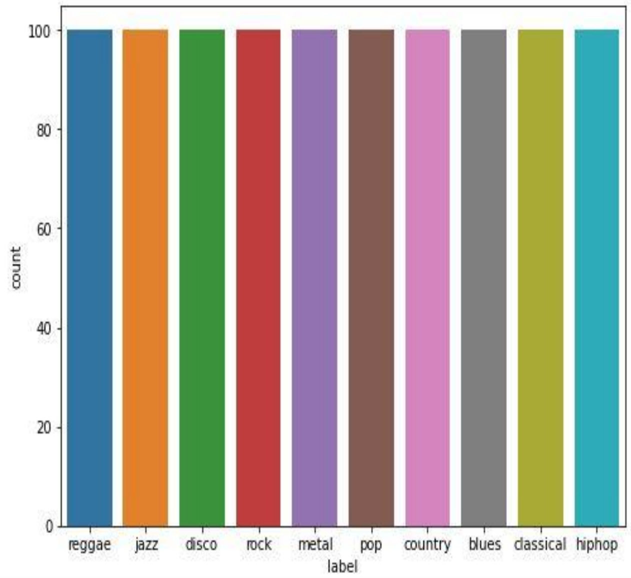


Fig1: Distribution of classes in dataset

EXPLORATORY DATA ANALYSIS

While performing exploratory data Analysis the following things are observed:

1. No null values were present in the dataset
2. The target variable is the music genre which is categorical in nature. Hence label encoding was performed to obtain better results.
3. We analysed that the ranges of features differ by a significant amount. So in order to avoid the algorithm to give more importance to the feature with higher range, the feature set was standardised.
4. Relevant features were selected using the insights drawn from variance and correlation. A feature with low variance doesn't have much predictive power and doesn't contribute much. So features like harmonic_var(variance=1.357712e-04),percussive mean (variance=1.170194e-06), percussive var(variance=4.226615e-05) and harmonic mean(variance=2.835795e-06) were removed. If two features are highly correlated then the feature with highest correlation with the target label was kept in relevant features and the other with lower variance with the target variable was removed.
5. We observed that there is a significant difference between mean, 75th percentile and maximum value for several features thus indicating presence of outliers. We removed the outliers using percentile capping method where anything greater than 99th percentile and lesser than 1st percentile has been removed. After removing the outliers and important feature selection the final dataset contains : 988 training examples and 56 feature sets.

	spectral_bandwidth_var
count	1000.000000
mean	137110.238662
std	96478.361350
min	10787.185064
25%	67376.554428
50%	111977.548036
75%	182371.576801
max	694784.811549

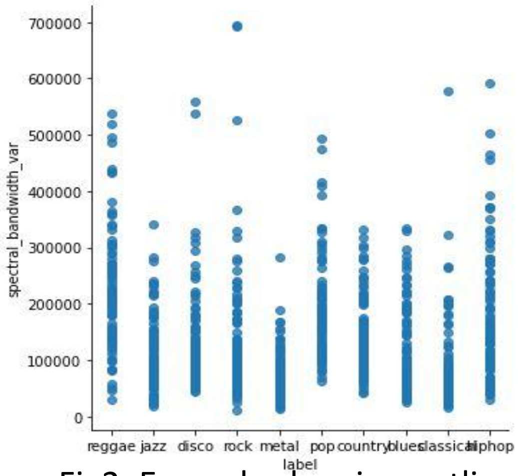


Fig2: Example showing outliers in dataset

DESIGN CHOICES:

- So we started with SVM as our baseline and then we looked for other classifiers Which includes KNN, NN, LDA, QDA, Naive Bayes and Logistic regression.
- We used accuracy as a evaluation metric to analyse the performance of different model.
- We have used Grid Search method along with k- fold cross validation resampling technique to choose the best parameter from given set of parameters. The chosen parameters gives best performance on the dev set and hence these parameters were used to finally estimate the genre of test data.

FURTHER ANALYSIS:

The dataset is not linearly separable as we observed that the train accuracy obtained on SVM model was not 100% on the best hyperparameter evaluated but was 0.73422. This shows that the data is not linearly separable.The same has been illustrated by taking example of two features in fig.

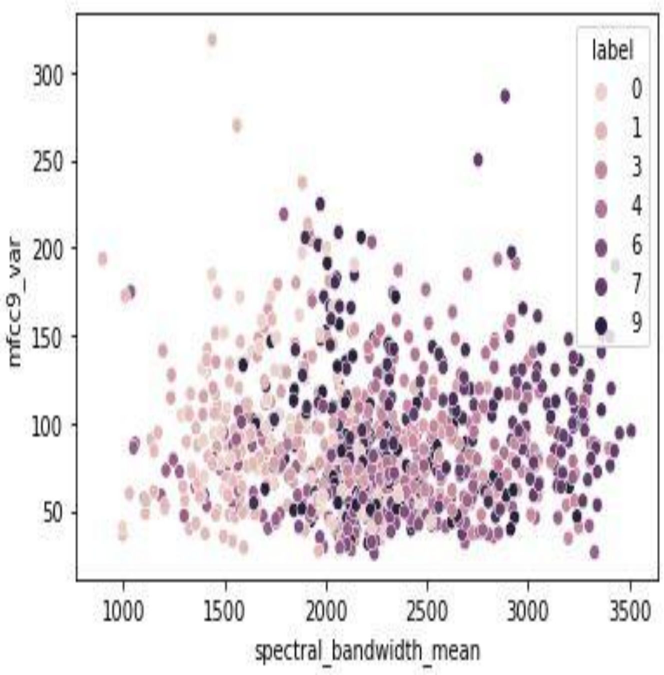


Fig3: Example showing non linear separability in data

RESULT OBTAINED:

- The learning model we have explored so far and the corresponding accuracies obtained have been mentioned in Table 1.
  The highest accuracy obtained is by using Poly Kernel SVM model with accuracy as 73.4%.
- We got to know that SVM and neural networks performed well and 8 of our 10 models have accuracy 70% however random forest also performed well with accuracy of 72%. It means that all these models similar in performance since difference in accuracy is not much.
- Gap between between our best model and state of the art model is about 20% where our poly kernel SVM gives about 73% accuracy the state of the art model is at 91%

Learning Model	Accuracy obtained
Poly Kernel SVM	73.4%
KNN	72%
NN	71%
Linear Kernel SVM	71%
Random Forest	72%
LDA	71%
RBF Kernrel SVM	72%%
Sigmoid Kernrel SVM	70%

Table 1. Results obtained

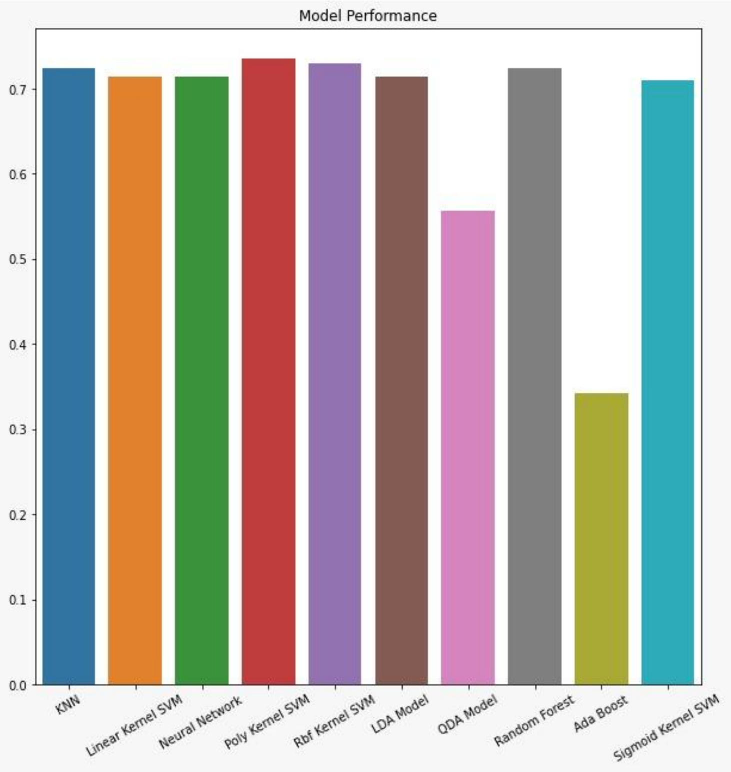


Fig4: Accuracies obtained using different models

ANALYSIS OF RESULT:

We checked the training and testing error of our models and

following were the results received-

As we can see that there is not a case where training error is too low and test error is too high means our model is not overfitting and also there is not a case where both training and test error are high means our model is not underfitting as well.So it shows that the final tuned hyperparameter through grid search is right.

Model Name	Test Error	Training Error
KNN	0.27	0.20
Linear Kernel SVM	0.28	0.19
Neural Network	0.28	0.17
Poly Kernel SVM	0.26	0.16
Sigmoid kernel SVM	0.29	0.25
RBF Kernel SVM	0.27	0.25
LDA	0.28	0.24
Random Forest	0.27	0.20

Fig 5: Train and test error analysis

Individual contribution and future work distribution

Member Name	Individual contribution	Work assigned for future
Priyanka singh[2018174]	Feature selection, EDA, Outlier Detection, Application of SVM , KNN and Neural Network,Hyperparameter Tuning.	work on Ensemble learning techniques,RNN .
Manish	Hyperparameter tuning,Application of Random forest, LDA,QDA,Ada Boost, SVM and model comparison.	Drawing conclusion from plots and improving the accuracy, hyperparameter tuning, model comparison
Aditya Singh	Feature extraction and application of SVM, KNN and Neural Network	Convolution neural network,ANN.