Music Genre Classification

Name: Kavya Sree Kaitepalli
dept. name of organization
ECE
Name of organization
IIIT SriCity
S20200020270
mailto:kavyasree.k20@iiits.in

Name: Himaja.A
dept. name of organization
ECE
Name of organization
IIIT SriCity
S20200020264
mailto:himaja.a20@iiits.in

Name: Jahnavi.CH
dept. name of organization
ECE
Name of organization
IIIT SriCity
S20200020253
mailto:jahnavi.c20@iiits.in

***Abstract*— The availability of many sites providing music downloads, categorizing the various downloaded music becomes a challenge and many-a-times quite frustrating too. One cannot think about paying attention to each single MP3 file before golf shot it within the right storage bin. If we are able to train the machine for classifying these files to cluster them in bound pre-defined classes, it'll be one thing nice to own.**

Music is often divided into genres. We use this categorization for grouping our MP3 downloads. The classes distinguish every style of music by its kind and magnificence. this type of machine classification isn't simply aiming to be helpful for your own purpose, however, conjointly has its actual use within the music business

Keywords—component, formatting, style, styling, insert (key words**)**

I. Introduction (*Motivation*)

Genre classification is very important for several real-world applications, because the amount of music being discharged daily continues to sky-rocket, particularly on net platforms like Soundcloud6 and Spotify. Spotify releases Forty thousand tracks per day, that is that the equivalent of 280,000 songs per week, or around one.2 million tracks per period of time. In a year, this volume would add up to a walloping fourteen.6 million. several corporations today use music classification, either to position recommendations to their customers (for example Spotify, Soundcloud) or as a product (for example Shazam). crucial the music is that the opening in this direction. To classify songs in any given listing or library by genre is a very important practicality for any music streaming/purchasing service.

Methodology

Based upon the inputs (attributes) we classify the music genre by using Decision tree, Random Forest, KNN classifier, Naïve bayes classifier, Multiclass support vector machines, sequential feature selection, correlation matrix, confusion matrix.

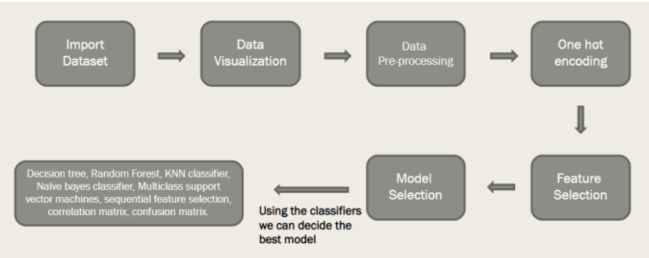

fig (1): Block diagram

The attributes that are taken into consideration are artist name, track name, popularity, acousticness, danceability, duration_ms, energy, instrumentalness, key, liveness, loudness, mode, speechiness, tempo, obtained date, valence (these are input features), music genre (output feature).

1. DATA PRE-PROCESSING

A. *CORRELATION MATRIX*

The correlation matrix is used to find how strong the relation is between two features/attributes. To reduce the redundancy, the relation of the features is compared and if it is found that these two features have the correlation value near to 1 then one of these features can be dropped without resulting in information loss.

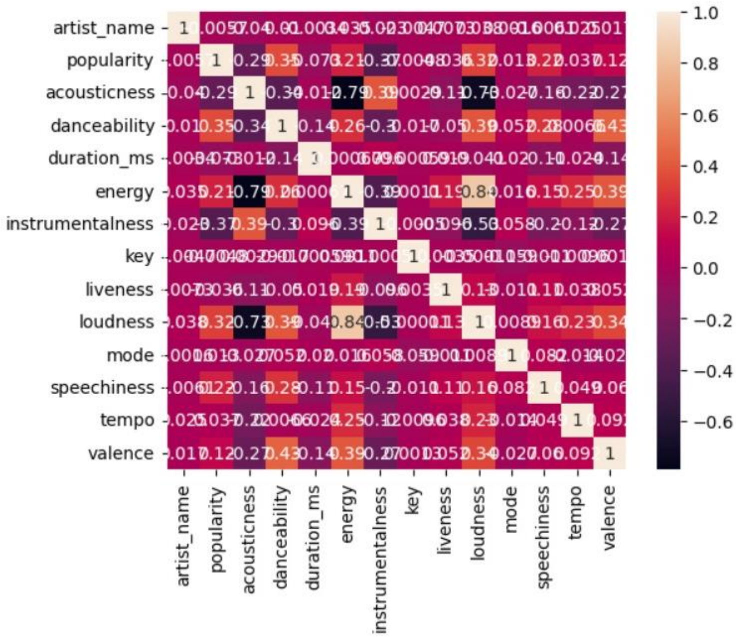

Fig. (2) correlation matrix

B. *LABEL ENCODING*

Label encoding is defined as a process of converting the categorical labels into a numeric form to convert them into machine-readable language. Machine learning algorithms can then decide in a better way how those labels must be operated. It is an important pre-processing step for the structured dataset in supervised learning.

2. CLASSIFICATION ALGORITHMS

C. DECISION TREE

A decision tree is a type of supervised machine learning used to categorize or make predictions based on how a previous set of questions were answered. The model is a form of supervised learning, meaning that the model is trained and tested on a set of data that contains the desired categorization. Decision trees are useful for categorizing results where attributes can be sorted against known criteria to determine the final category.

D. RANDOM FOREST CLASSIFICATION

Random forest algorithm builds decision trees on different samples and takes their majority vote for classification and average in case of regression. One of the best features of the Random Forest classification is that it easily deals with the categorical data as well besides numerical. This algorithm is useful for both classification as well as regression. It performs well in classification problems as compared to that of regression.

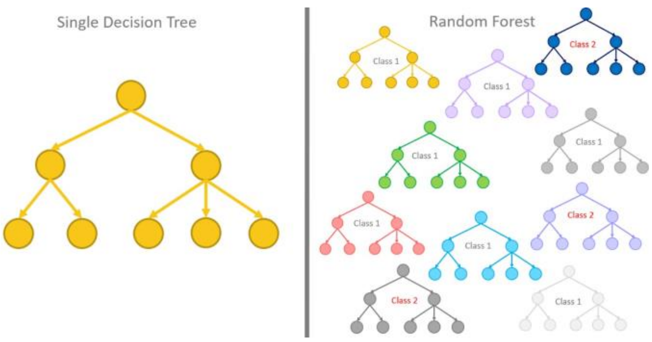

Fig. (3)random forest classifier

E. KNN

K-Nearest Neighbours (KNN) algorithm is a non-parametric algorithm. The principle is that known data are arranged in a space defined by the selected features and when a new data is supplied to the algorithm, the algorithm will compare the classes of the k closest data to determine the class of the new data. This is a supervised learning algorithm and uses proximity measures in the classification process.

F. GAUSSIAN NAIVE BAYES CLASSIFIER

Naïve Bayes classifier is a statistical classifier. It performs probabilistic prediction which means it predicts the class membership probabilities. This classifier takes the foundation of the Bayes' theorem. Assumptions considered in the classifier: The classes are mutually exclusive and exhaustive. The attributes are independent given the class.

G. SVM (Support Vector Machine)

SVM is a traditional computing paradigm. SVM aims to find a hyperplane that provides maximum separation between the classes. Usually, it is used for the binary classification. However, there are ways to modify the algorithm so that it can also be applied in multi-class classification. This is a supervised learning algorithm.

H. LOGISTIC CLASSIFICATION

Logistic regression is a classification technique borrowed by machine learning from the field of statistics. Logistic Regression is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The intention behind using logistic regression is to find the best fitting model to describe the relationship between the dependent and the independent variable.

I. MULTI LAYER PERCEPTRON

Multi-layer perceptron is one of the most common neural network models in the field of deep learning. A multi-layered perceptron consists of interconnected *neurons* transferring information to each other, much like the human brain. This consists of three layers namely, input layer, hidden layer and the output layer. The input layer takes the data as input and passes it to the hidden layer where it does all the computations to produce a meaningful output at the output layer.

J. GRADIENT BOOSTING ALGORITHM

The idea of gradient boosting is that we will mix a gaggle of comparatively weak prediction models to create a stronger prediction model. As every weak learner is additional, a new model is fitted to provide a more accurate estimation of the response variable. it's an awfully powerful technique for building prophetic models.

K. SEQUENTIAL FEATURE SELECTION

Feature selection is one of the most important tasks performed in the field of machine learning. Feature selection is defined as selecting the most meaningful features from the available features. The goal of feature selection algorithms is to automatically choose a subset of features that are most relevant to the problem at hand. By eliminating unimportant information or noise, feature selection aims to increase computing efficiency and decrease the generalization error of the model. This Sequential Feature Selector adds (forward selection) or removes (backward selection) features to form a feature subset in a greedy fashion. This estimator selects the best feature to add or remove at each step based on the estimator's cross-validation score. This Sequential Feature Selector solely considers the features (X) rather than the desired outputs in the case of unsupervised learning (y).

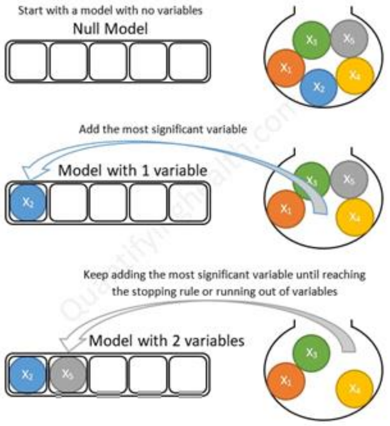

Fig (4) forward sequential feature selection

L. *MRMR*

MRMR (acronym for *Maximum Relevance — Minimum Redundancy*) is a feature selection algorithm that has gained new popularity.

Maximum Relevance - Minimum Redundancy" is so called because — at each iteration — we want to select the feature that has maximum relevance with respect to the target variable and minimum redundancy with respect to the features that have been selected at previous iterations. The redundancy (denominator) is computed as the average (Pearson) correlation between the feature and all the features that have been selected at previous iterations.

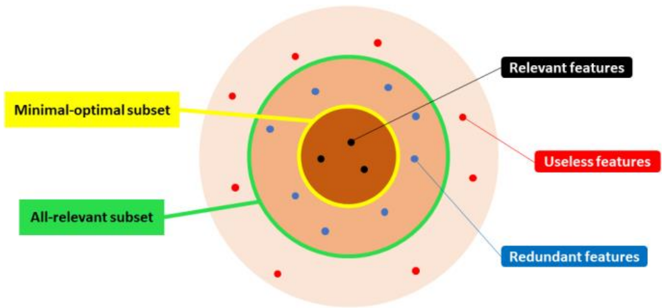

Fig (5) Maximum Relevance — Minimum Redundancy

II. RESULTS

In this section we obtain the confusion matrices for different classifiers. Once we get the confusion matrices, we can calculate the performance estimation metrics such as accuracy, F1 score, recall, precision.

The confusion matrices for some of the classifiers are being listed:

CONFUSION MATRICES

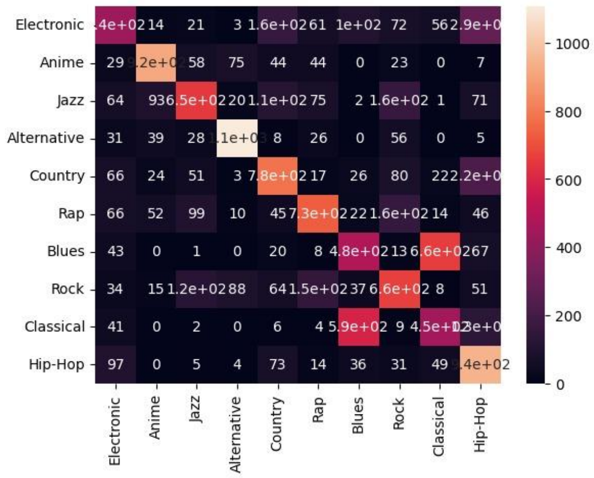

Fig. (6)
confusion matrix for random forest classifier

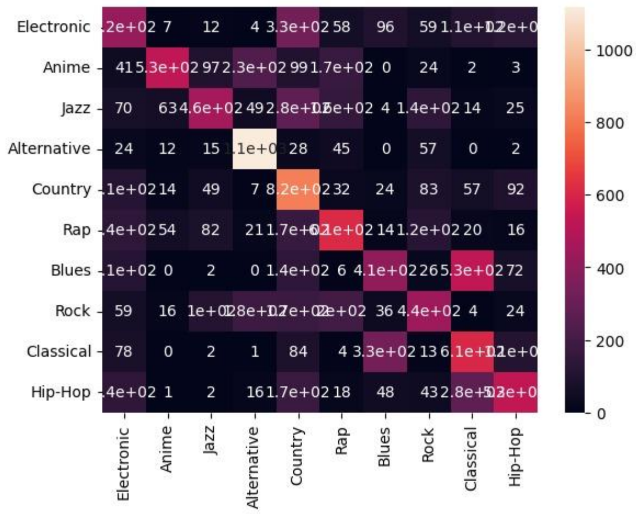

Fig. (7)
Confusion matrix for naïve bayes classifier

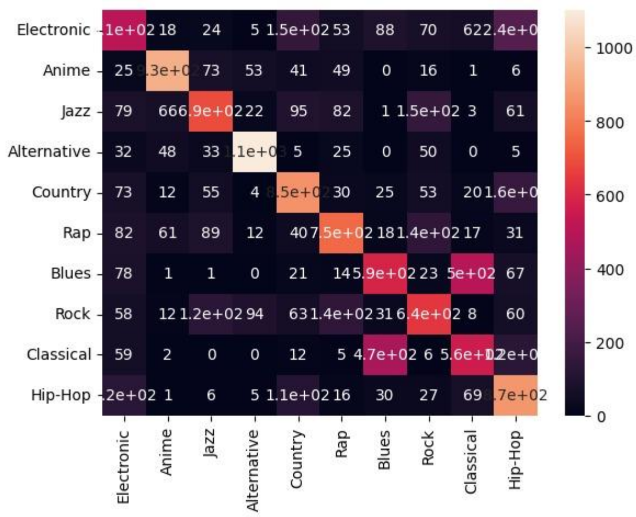

Fig. (8)
Confusion matrix for gradient boosting

PERFORMANCE ESTIMATION METRICS

- KNN classifier

Accuracy: 0.16368
F1 score: 0.1483
Recall: 0.164
Precision: 0.1593

- Decision Tree

Accuracy: 0.444
F1 score: 0.4442
Recall: 0.443
Precision: 0.4456

- Random forest classifier

Accuracy: 0.5729
F1 score: 0.5704
Recall: 0.5723
Precision: 0.5761

- Naïve Bayes

Accuracy: 0.4759
F1 score: 0.4715
Recall: 0.4738
Precision: 0.4963

- SVM

Accuracy: 0.1969
F1 score: 0.1645
Recall: 0.1953
Precision: 0.1923

- Gradient boosting classifier

Accuracy: 0.5996
F1 score: 0.5982
Recall: 0.5989
Precision: 0.6008

- Linear classification model

Accuracy: 0.2722
F1 score: 0.2488
Recall: 0.2735
Precision: 0.2538

These are the results obtained before applying feature selection.

We use Maximum relevance Minimum redundancy (MRMR) algorithm and Forward Sequential Feature Selection (Forward SFS) algorithm to select the most relevant features.
MRMR algorithm selects a subset of features having the most correlation with the class (output) and the least correlation between themselves. It ranks the attributes accordingly.
Forward SFS is used to find the set of features that gives maximum accuracy.

III. CONCLUSION

From the above results, we observe that accuracy obtained for the Random Forest model, Gradient boosting algorithm and the Naïve Bayes classifier are better compared to other classifiers.
The highest accuracy obtained is for the Gradient Boosting Classifier which is 59.96 percent.
Accuracy for the random classification turns out to be ten percent since we have ten unique music genres (classes). The obtained accuracy for the Gradient Boosting Classifier is about six times the random accuracy which leads us to a conclusion that it performed better than other classifiers for this particular data set.

IV. APPENDIX

Link to access code:
https://drive.google.com/drive/folders/1HezMN1hi5aK4h9AHpD8LIP_JbX4VWxEe?usp=share_link

Individual contributions:

Kavya Sree Kaitepalli: Gradient boosting classifier, Decision tree, SFS algorithm, parts of report

Jahnavi.CH: KNN classifier, SVM, Linear classification, parts of report

Himaja Anchuri: Random Forest classifier, Naïve Bayes classification, MRMR algorithm, parts of report