Analysis of Genre Detection Accuracy

based on Audio and Sentiment Score

Rishita Bansal, Parth Padalkar

Erik Jonsson School of Engineering and Computer Science,

University of Texas at Dallas

1. Introduction

Songs can be classified into different genres and often users have their own preferences of songs based on genre. Users like to hear songs based on their genre preference. Thus it is required to classify the songs into the correct genre for a good user experience. So most of the existing songs are already classified into genres, but the problem arises when we have some new unclassified songs and we want to classify them into their correct genre. Thus we are using the audio features of the songs and sentiments based on the key idea of the song obtained from the album description and the song title . We are classifying genres based on these features and seperately identifying if there is an increase in accuracy when considering the sentiment score of the key idea of the song.

2. Problem Definition and Algorithm

2.1 Task Definition

We have metadata from FMA (Free Music Archive) dataset that gives us all extracted audio features and song information and genres. We need to classify these songs based on their genres. Our algorithms would train on this data and then we can predict genre for new songs twice.
Once based only on the audio features and the second time based on the audio features and sentiment score of the key idea.

2.2 Algorithm Definition

We are using SVM classifier, Random forest classifier, MLP classifier, Gaussian Naive bayes classifier, KNN classifier from sklearn.

3. Experimental Evaluation

3.1 Methodology

We have taken the FMA Dataset from UCI Dataset Repository. Dataset includes audio and metadata from Free Music Archive. The metadata gives information about artists, tracks, genres and features. There are a total of 161 genres out of which 16 are root genres and others are sub genres. There are 9 audio features that consist of 518 attributes. The data also contains user experience.

Our aim was to use these features to predict the genre of the song and to see if there is any increase in the accuracy if the key idea of the songs in the form of their Album description and song title are considered as another feature.

We are using the SVM classifier, Random forest classifier, MLP classifier, Gaussian Naive bayes classifier, KNN classifier from sklearn for both datasets (with sentiment score and without).

We tuned the hyperparameters and tested the performance of the above classifiers for the dataset containing only the audio features. Then we performed sentiment analysis on the song title and album description of each song and recorded the sentiment score on a scale of 0 to 2. 0 is most negetive and 2 is the most positive.
We combined the list of sentiment scores obtained and treated this as a new feature along with the existing audio features and ran the classification algorithms on this new dataset again.
We used the values of hyperparameters that gave the maximum accuracy on the test set for the previous dataset.

3.2 Results

Without Sentiment Score

MLP

Activation	**Solver**	**beta_1**	**beta_2**	**Hidden_layers**	**Accuracy**
Relu	lbfgs	0.9	0.999	5,2	0.28436592449177156
Relu	adam	0.5	0.5	600,500,300,200	0.4674491771539206

Tanh	adam	0.4	0.4	500,500,300, 200	0.284365924 49177156
Relu	adam	0.5	0.5	530,500,400, 300,200	0.478702807 357212

SVM

Kernel	**C**	**Max_iter**	**Accuracy**
sigmoid	1.0	-1	0.284
rbf	0.5	500	0.3335
rbf	0.3	500	0.33355
rbf	0.8	1000	0.33355
rbf	0.8	1000	0.3297

KNN

N neighbours	**Accuracy**
200	0.42231
20	0.52087
10	0.55003
5	0.56401
3	0.56455

Random Forest

Max_depth	**n_estimator**	**Accuracy**
200	100	0.55227
150	100	0.55227
250	100	0.55227
100	300	0.62578
500	300	0.62578

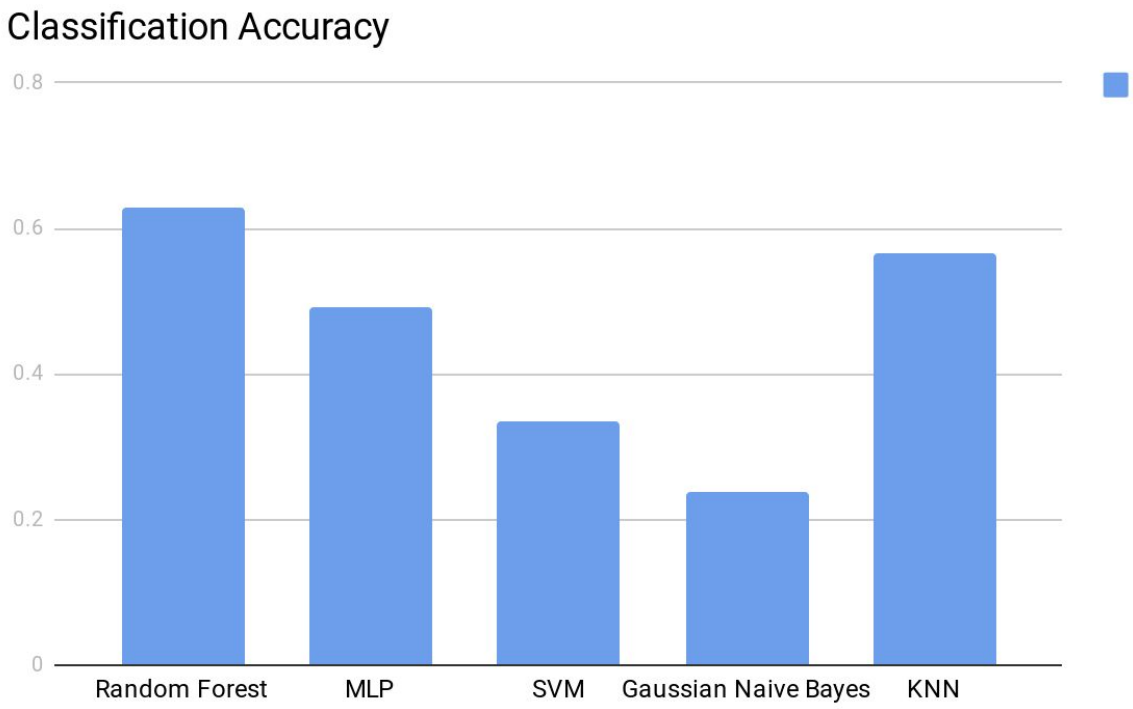
Gaussian Naive Bayes accuracy - 0.23904

After adding sentiment score to data

After adding sentiment score we ran all five classifiers with the hyperparameters that gave us highest accuracy with the previous data.

Following are the accuracies for all the classifiers:

Random Forest	MLP	SVM	Gaussian Naive Bayes	KNN
0.62760	0.49146	0.33379	0.23916	0.56455



3.3 Discussion

The performance ranking of the classifiers on the dataset is as follows:

1) Random Forest Classifier : Winner
2) KNN Classifier : Second Position
3) Random forest : Third Position
4) MLP classifier : Consolation Prize
5) Gaussian Naive Bayes : Honorable Mention

After adding the sentiment score to the dataset, the accuracy increased slightly. Thus we can say sentiment is a good feature to get an idea of the genres. We can add some more relevant features to increase the accuracy even further.

4. Related Work

There is a lot of work where sentiments can also be detected using audio data. We can also find out the instruments being used in the songs. Using all the features we can also determine the mood of the song, so this can be very useful when users want to hear songs according to the mood. For example, if a user is sad and wants to become cheerful, he can get access to a playlist of songs that makes him/her cheerful. Similarly there can be songs for party mood, workout mood, study mood, work environment mood. Until now mood classification does not have a good accuracy.

5. Future Work

There is a lot of scope to increase the accuracy of genre detection. This can be done in 2 ways

1) Adding more relevant features like
2) Performing some more feature engineering and eliminating some audio related features that may not be relevant for genre detection.
3) We can also get sentiment score using audio features rather than song key idea.

6. Conclusion

Thus we have seen that in addition to the audio data, sentiment score helps us improve classification score for genre classification. Thus we can say sentiment is quite an important feature for genre classification. Also we need to implement more feature engineering to increase accuracy furthermore.

Bilbiography

1. Michael Deffarard, Kirell Benzi, Pierre Vandergheynst, Xavier Bresson, FMA:A dataset for music analysis,EPFL, 2017

2. Tyler Dammann and Kevin Haugh, Genre Classification of Spotify Songs using Lyrics, Audio Previews, and Album Artwork, Stanford University, 2017

3. Emiru Tsunoo, Taichi Akase, Nobutaka Ono and Shigeki Sagayama, MUSIC MOOD CLASSIFICATION BY RHYTHM AND BASS-LINE UNIT PATTERN ANALYSIS, Graduate School of Information Science and Technology, The University of Tokyo, 2010

4. George Tzanetakis, Musical Genre Classification of Audio Signals, IEEE, 2002