Pattern Recognition

Pattern Recogniton Project 2

Shashank Reddy Sama
S20180020244
Pavan Yachamaneni
S20180020261
P.V Pavan Choudary
S20180020231

Contents

1 Introduction

Pattern recognition is the automated recognition of patterns and regularities in data.It has applications in statistical data analysis, signal processing, image analysis, information retrieval, bioinformatics, data compression, computer graphics and machine learning.

With the rapid development of digital technology, the amount of digital music content increases dramatically everyday. To give better music recommendations for the users, it's essential to have an algorithm that could automatically characterize the music. This process is called **Musical Information Retrieval (MIR)** and one specific example is **music genre classification**.

However, music genre classification is a very difficult problem because the boundaries between different genres could be fuzzy in nature. For example, testing with a 10-way forced choices task, college students could achieve 70 percent classification accuracy after hearing 3 seconds of the music and the accuracy doesn't improve with longer music . Also, the number of labeled data often is much smaller than the dimension of the data. For example, GTZAN dataset 2 used in the current work contains only 1000 audio tracks, but each audio track is 30s long with a sampling rate 22,050 Hz.

2 Problem Description

2.1 Aim

To classify music into respective categories.

2.2 Data Description

Dataset has 10 genre's and our ultimate goal is to have an accurate classification algorithm.**Blues,Country,Classical,Disco,HipHop,Jazz,Metal,Pop,Reggae,Rock.**

Total audio tracks are having equal number of data samples of each genre's. After dataset is collected and extracted, a classification algorithm is used on entire training dataset. There needs to be a trade-off between final result and classifier used. In general Recurrent Neural Networks gives better results.

Note: We are using SVM and its kernels as classifiers.Testing accuracies are shown in confusion matrix.

3 Methodology

Raw music has less information in time domain but in frequency domain gives precise results.Therefore feature extraction has to be carried out. But,the issue is what features are to be considered. For this we used most traditional techniques which constitutes.
1.STFT(Short Time Fourier Transform)
2.Spectral Centroid
3.Spectral Bandwidth
4.Zero Crossing Rate
5.Spectral Roll Off
6.Zero Crossing Rate
7.MFCC's (Mel Frequency Cepstral Coefficients)

From all above features MFCC's are the most dominating features which constitutes more information. In addition to it, MFCC helps in better understanding of human hearable music therefore results in better music recommendation.

4 Implementation and Results

We used python and scikt Learn framework for doing our classification.
Dataset repository is mentioned in I-python Notebook linked in master folder.
At first we have to iterate through every music file and take down the audio magnitude and sampling rate.
Then for feature extraction we used the inbuilt function provided by librosa library.
After finding each feature we initiated a CSV file for storing information in chronological order.
Finally feature data is in data.CSV file which will be created during run time of program.
We used Gaussian and Polynomial kernel of degree 3 as main classifiers and results are analyzed accordingly.
Data is split into training and testing data randomly by a fraction of 4:1 ratio.

Note: For our project Polynomial Kernel has edge over Gaussian kernel.

Also, so far the better accuracies with Neural Networks are around 75 percent but in our project Only Training data is used for finding suitable parameters and model is trained ubntil required constraints are successfully converged.
Gaussian Kernel shows around 70 percent accuracies. Accuracies are attained from testing data.

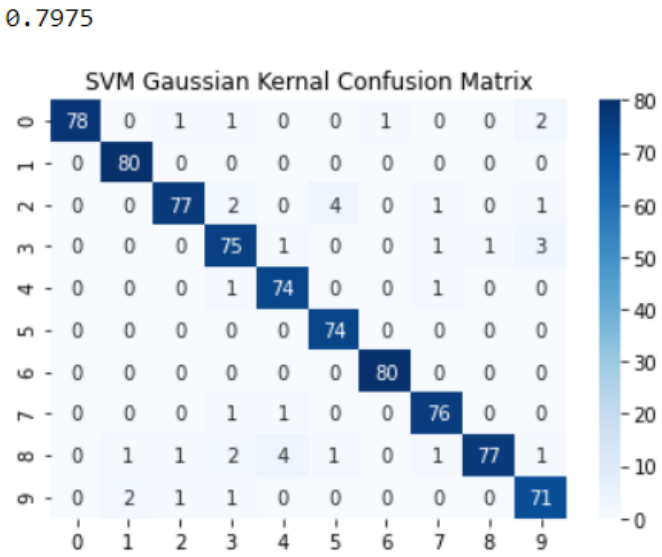

Figure 1: Training Accuracies of SVM Gaussian Kernal

0.9525

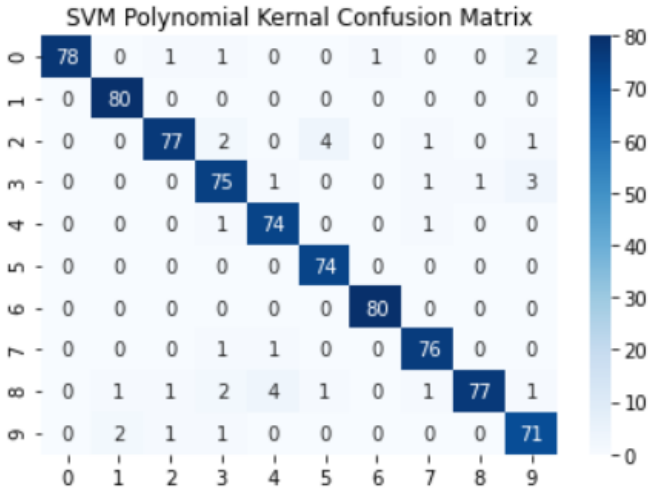

Figure 2: Training Accuracies of SVM Polynomial Kernal

0.695

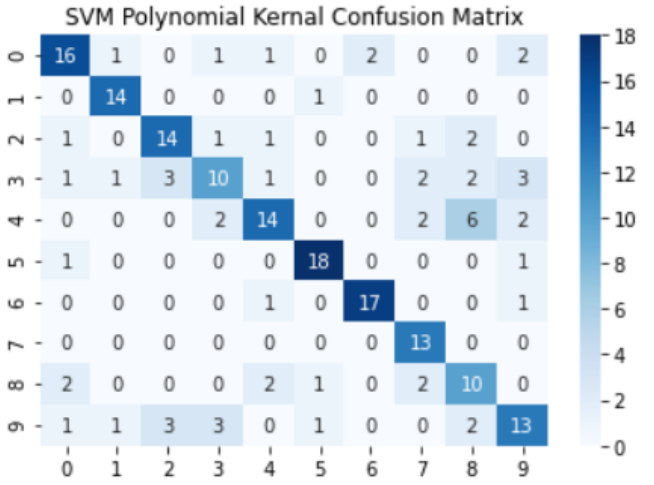

Figure 3: Testing Accuracies of SVM Gaussian Kernal

0.64

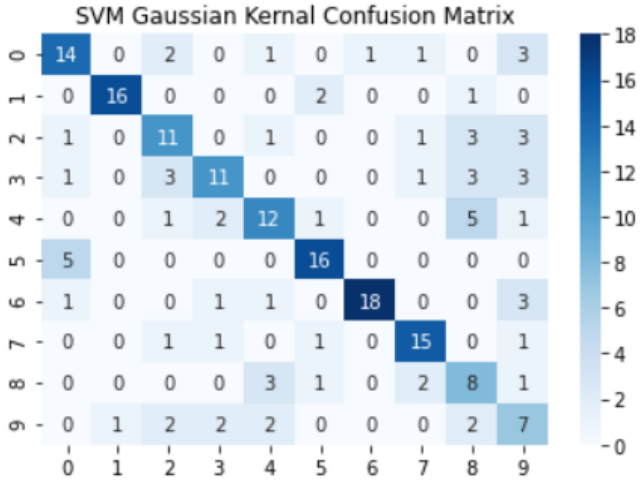

Figure 4: Testing Accuracies of SVM Polynomial Kernal

5 Analysis

Even though accuracies are quite small but, human level perception of music classification is around 70 percent. Main reason for this is use of electric equipments and using signal processing techniques on music which affects its ultimate genre.

6 Contributions

1.Shashank Reddy Sama - Algorithms and visualization of Data
2.Pavan Yachamaneni - Feature Extraction
3.PV Pavan Choudary - Resource Collection,Report Making and Data Preprocessing

7 Resources and References

1. https://towardsdatascience.com/musical-genre-classification-with-convolutional-neural-networks-ff04f9601a74
2. https://www.kaggle.com/c/music-genre-classification/leaderboard
3. http://cs229.stanford.edu/proj2018/report/21.pdf