Music Genre Classification

DS303

Aziz Shameem
Kalp Vyas
Tanmay Dokania
Nishant Thakre

IIT bombay

July 2021

Introduction

Music genre prediction is one of the topics that digital music processing is interested in. In this study, acoustic features of music have been extracted by using digital signal processing techniques and then music genre classification and music recommendations have been made by using machine learning methods. In addition, convolutional neural networks, which are deep learning methods, were used for genre classification and music recommendation and performance comparison of the obtained results has been. In the study, GTZAN database has been used and the highest success was obtained with the Neural Networks.

In this implementation, we have used several Machine Learning / Deep Learning models, in an effort to classify music based on its genre. The data set used contains extracted features (57 in number), on which the training was performed, and the data was classified into 10 classes, representing different genres that the corresponding music belonged to.

Feature Extraction

For each song, the raw music signal was divided into multiple features using a few digital processing techniques.

Features extracted

- **Zero Crossing Rate:** no of sign changes of a signal in a certain period of time.
- **Spectral Centroid:** Centre of gravity of the frequencies in the frequency bin.
- **Spectral Bandwidth:** Weighted average amplitude difference between frequency magnitude and brightness
- **Spectral Rolloff:** Frequency value corresponding to a certain ratio of the distribution in the spectrum (about 85 percent).
- **Mel Frequency Coefficient of Cepstrum (MFCC):** Describes the overall shape of a spectral envelope

Few other features like chroma_stft, harmony, perceptr, rms and tempo were also used.

Models : Decision Trees

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a data set, branches represent the decision rules and each leaf node represents the outcome.

Hyper-Parameters Tuned

- Criterion : gini or entropy
- Splitter : best or random
- Maximum depth of tree

Models : Random Forest Classifier

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression.

Hyper-Parameters Tuned

- Number of Estimators
- Maximum Depth of Trees
- Criterion : Entropy or Gini

Models : Support Vector Machines

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.
The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

Hyper-Parameters Tuned

- C : Regularization parameter
- gamma : Kernel coefficient
- Kernel : Linear or rbf

Models : Naive Bayes

Naive Bayes classifiers are a collection of classification algorithms based on Bayes' Theorem. It is not a single algorithm but a family of algorithms where all of them share a common principle, i.e. every pair of features being classified is independent of each other. We have used 2 forms of naive Bayes classifiers namely, Gaussian and Bernoulli.

Hyper-Parameters Tuned

Gaussian

- variance smoothing factor

Bernoulli

- alpha: Additive (Laplace/Lidstone) smoothing parameter
- fit_prior : to decide whether to learn class prior probabilities or not

Models : K-Nearest Neighbours

The KNN (K-Nearest Neighbours) algorithm is one of the simplest algorithms in supervised ML and is mainly used for classification type of problems. It finds the distance between rows of test and train data using different criteria to measure distance like Hamming, Euclidean, Manhattan etc. It uses these distances to decide the class the set point belongs to. One interesting thing about this algorithm is that it always gives 100 percent accuracy on the train data so if the train data and test data are very similar, the algorithm works great.

Hyper-Parameters Tuned

- N - no of neighbours being considered
- p - to decide which distance to be used (p=1 is manhattan (L1 norm),p=2 is euclidean(L2 norm))
- weights - to decide what weight functions to be used in prediction

Models : Neural Networks

Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.
Artificial neural networks (ANNs) are comprised of a node layers, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, connects to another and has an associated weight and threshold.

Hyper-Parameters Tuned

- Number of neurons in hidden layer
- Number of epochs
- Activation Function(used in hidden layers)
- Optimizers : Functions/Algorithms that modify the attributes of the NN, such as weights and learning rate

Results

Table summarizes the accuracies obtained from the models fitted on the data :

.

Table: Final Accuracies

Model	Accuracy
Decision Tree	67
Random Forest	79.8
Neural Nets	81.1
Naive Bayes	51.1
SVM	58.7
KNN	32

Based on the result obtained, we can see that Neural Networks have given the best result in terms of accuracy, however, Random Forests are not far behind, giving similar accuracies with much lower training times.

Contributions

The publication and data searching was done by Tanmay and Kalp.
Pre-processing of the data was done by Kalp.
Out of the six models trained, the distribution of work was as follows :

- Nishant Thakre : DT, RF
- Tanmay Dokania : SVM (PCA)
- Kalp Vyas : NB, KNN
- Aziz Shameem : ANN

The report drafting and compilation was accomplished jointly by Aziz, Kalp and Tanmay.

Contributions

The publication and data searching was done by Tanmay and Kalp.
Pre-processing of the data was done by Kalp.
Out of the six models trained, the distribution of work was as follows :

- Nishant Thakre : DT, RF
- Tanmay Dokania : SVM (PCA)
- Kalp Vyas : NB, KNN
- Aziz Shameem : ANN

The report drafting and compilation was accomplished jointly by Aziz, Kalp and Tanmay.

References

- one "Music Genre Classification and Recommendation by Using Machine Learning Techniques" https://ieeexplore.ieee.org/document/8554016
- two "Feature Extracted Data sets" https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification item For the in-depth analysis of the training, along with the results obtained using various sets of hyper-parameter choices, visit https://github.com/kalp121212/DS303_Project