CS 454

Introduction to Machine Learning and Artificial Neural Networks

Prof. Dr. Ethem Alpaydın

Fall 2021

Progress Report

Music Genre Classification

Team A

Burcu Arslan - S018468

Aleyna Nur Ölmezcan - S015338

Emin Sadikhov - S017026

Abdullah Saydemir - S014646

Muhammed Esad Simitcioğlu - S018039

December 5, 2021

What have we done so far?

We downloaded the GTZAN dataset that includes 100 songs from 10 different genres. We separated the dataset to training, test and validation samples. We extracted 76 different features out of each song using the librosa library in Python. Extracted features are saved to training and validation csv files alongside the genres.

Librosa library provides features as numpy arrays within specific timeframes. To get a constant we used mean and standard deviation functions as it is done by N. Ndou et al.[1] (2021). We analyzed the contribution of each column to variance using PCA but did not delete the columns that do not contribute much, from the dataset.

So far, we implemented the K Nearest Neighbor algorithm using the Scikit-learn library. Hyperparameters we tested are distance metrics (euclidean, minkowski, manhattan, chebyshev and hamming distance), number of neighbors (1 to 20), and number of components PCA holds (1 to 50).

Currently, our results for the K-Nearest Neighbor algorithm change between 30% to 50% percent. We recognize that the accuracy is very low compared to the articles we have found. According to Ndou et al.[1] (2021) this value is reported as 72.80% for hyperparameters linear search for distance metric, number of neighbours = 1, using absolute error for cross-validation; 69.70% for hyperparameters manhattan distance for distance metric, k = 7, weighting = distance and 92.69% for nearest neighbours = 1.

What are we planning to do next?

We are planning on implementing Convolutional Neural Networks and Support Vector Machines to provide the features extracted from the data from GTZAN dataset into the algorithms. We are then going to analyze the classifications for each algorithm and how close to truth they are while calculating their respective confusion matrices and accuracies. The data obtained will be compared to previous work in the field.

We are not satisfied with the results. Though most of these features are used in the previous works, we would like to specifically include some other features. We had to exclude the features that are not available in libraries, because they require low level audio analysis.

References To Related Work

[1] Ndou, N., Ajoodha, R., & Jadhav, A. (2021). Music Genre Classification: A Review of Deep-Learning and Traditional Machine-Learning Approaches. IEEEXplore. https://ieeexplore.ieee.org/document/9422487