NAML Project Report - Group 14

Music Genre Classification using
k-Nearest Neighbours
Nearest Centroid
Multiclass SVM
-

by Silvia Marino (codice persona) and Francesco Panebianco (10632465)
DEIB - Politecnico di Milano

February 2022

Contents

1 Introduction

Scope

The scope of this project is to create a music genre classifier using the machine learning algorithms *k-Nearest Neighbours* and *Multiclass SVM*, which were assigned to our group. The project is part of the evaluation of the "Numerical Analysis for Machine Learning" course, which is part of the first semester of the first year of Master's Degree in Computer Science and Engineering at Politecnico di Milano. Given the similarities between *k-Nearest Neighbours* and *Nearest Centroid*, we chose to implement the latter as well, comparing its performance to the former, even though it is outside the specification of the project.

The Dataset

The dataset assigned to our project is the notorious *GTZAN Genre Collection*[2], which contains 100 different extracts from 10 different music genres, provided in .wav (Waveform Audio File Format). The genres considered are:

- Blues
- Classical
- Country
- Disco
- Hip Hop
- Jazz
- Metal
- Pop
- Reggae
- Rock

As it can be seen, genres that share similarities are included in the dataset (e.g. Blues and Jazz), but also dramatically different types of music such as Rock and Classical, which we expect the algorithms to classify with higher precision.

The *.wav* file format used by the dataset is the most common uncompressed audio file format in Microsoft Windows systems for. It was developed by IBM and Microsoft, for storing an audio bitstream on PCs[1].

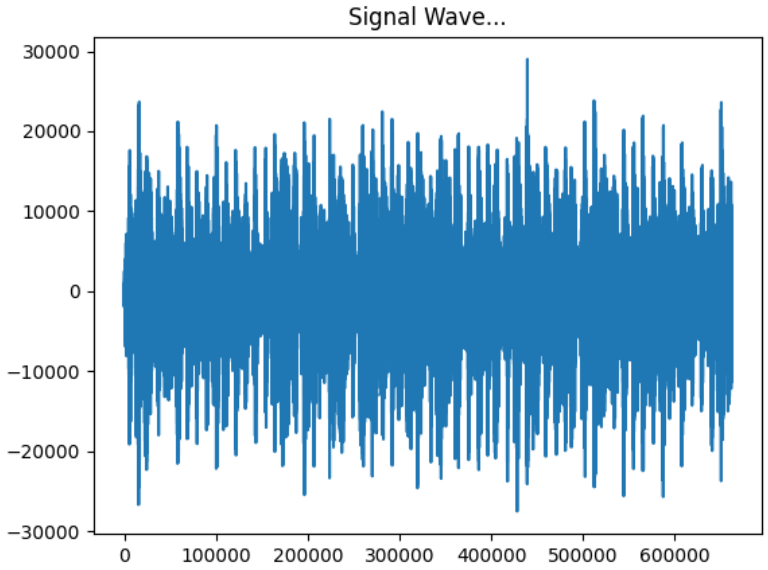

Figure 1: Plotted waveform of blues.00000.wav

This kind of files are managed in Python by using the *wave* module, which provides convenient functions to work with the WAV sound format. In our case, the audio samples are provided as 22050Hz Mono 16-bit 30 second tracks[2], which are easily transformed into a numpy array of *int16*.
The picture above is the result of the following code:

```
import wave
import matplotlib.pyplot as plt
import numpy as np

test_file = wave.open('genres/blues/blues.00000.wav', 'rb')
# Extract Raw Audio from Wav File
signal = test_file.readframes(-1)
signal = np.frombuffer(signal, dtype='int16')
test_file.close()

plt.figure(1)
plt.title("Signal Wave...")
plt.plot(signal)
plt.show()
```

Feature Extraction

2 k-Nearest Neighbours Classifier

In this section it will be explained which methodologies have been used to implement the music genre classifier and how the code has been structure to achieve the results exposed at the end of the report

3 Nearest Centroid Classifier

4 Multiclass SVM Classifier

References

[1] Fleischman E. *WAVE and AVI Codec Registries*. 1998. URL: `https://datatracker.ietf.org/doc/html/rfc2361`.

[2] Leben Jakob. *Music Analysis, Retrieval and Synthesis for Audio Signals*. URL: `http://marsyas.info/downloads/datasets.html`.