Statistical Learning Project

1st Milestone

Group 19

Members: ***Alice Schirina'****, Valerio Antonini, Maria Luisa Croci, Daniele Sanna, Giuseppe Stefanelli*

Research Title

Guessing it in a few seconds

Abstract

The goal of the following project is to classify music genres using a neural network.

Main research aim & framework

In order to reach our goal we would like to use machine learning techniques that we are studying applying it in a field of common interest (the music). The idea came up taking the cue from known app that identify songs. Finally, as an extension of the project, we are going to try to compare our outcomes with others obtained with a clustering algorithm.

Data collection & source(s)

As already said previously, we used 3 sensors: PitchSensor (Hz), DecibelSource (db) and BrightnessEV (lumen); using Science journal app we have actually recorded 5 songs for each of the 4 music genres, for a total of 20 songs. Our dataset is composed by 3 columns (features), one for each sensor described above. Each index of the row represents the timing. Once collect these observations, we gathered information about how to transform our data into something that is more specific in order to analyze the sound and to obtain a genre classification; sound could be totally described using the pitch, frequency and time, through these three variables it could be built a spectrogram. The spectrogram is extremely useful to reach our goal, since it is less sensitive to noise and it is indirectly possible to extract other features (for example we can find the exactly the center of mass of the sound, that is the weighted average of the frequencies).

Other than the spectrogram, we used the brightness sensor too; our idea is to use a strobo in order to extract a new feature useful for differentiate better the genres. The strobe reacts at the variations of the tones, generating bright each time that an activation bound is overcame: we are NOT interested at the value of the lumen, but how many times we have a specific value as the time varies. We expect that a rock track has more lumen's values than a classic track and we get as value the average. In case we are able to do that,

our final dataset will be different from the actual one. This will be a necessary evolution since our dataset has only 3 features for making the training on the neural network and we think that are not enough to reach a good genre classification. Moreover, we will try to implement a clustering techniques too for classifying the genres.

Models & Methods

Given our knowledge we expect to use the spectrogram from which we detect most of our feature, some clustering techniques and CNN (convolutional neural network).

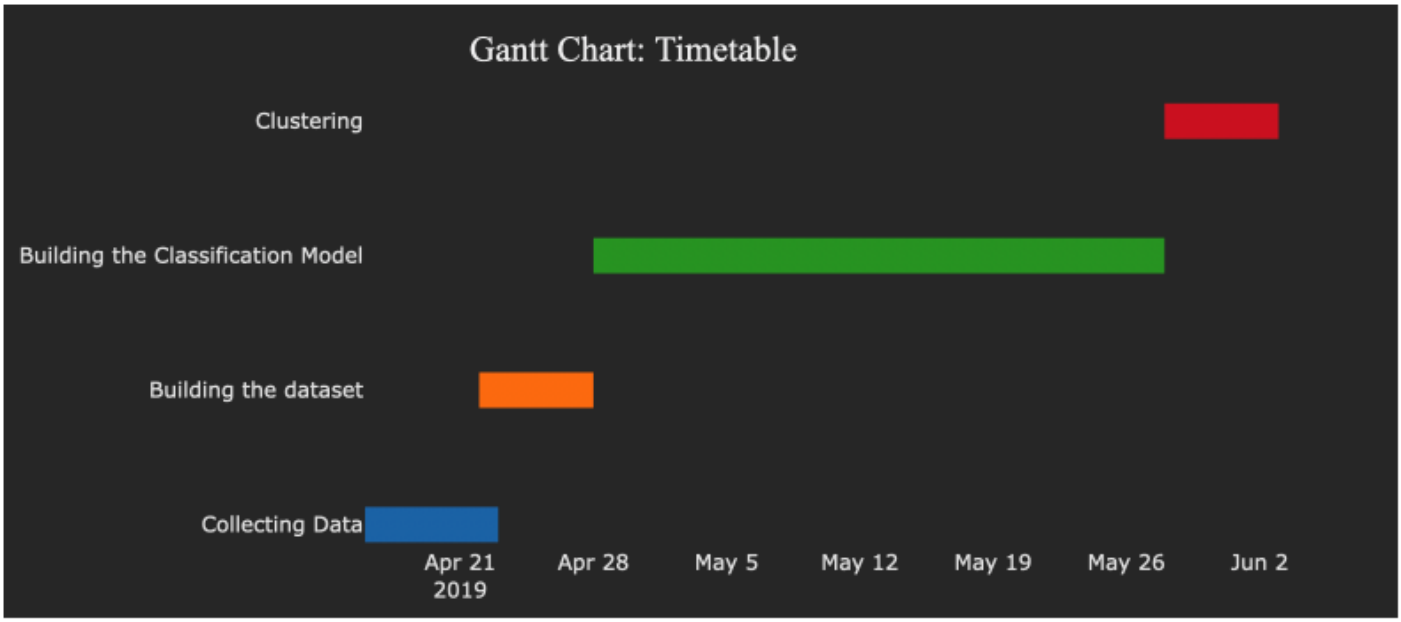
Software/Hardware Toolkit *For our task we are going to write our codes in Python, using several libraries and packages as Pandas, Scikit-Learn, Numpy, LibROSA, Keras and others.*

Problems so far...

The first problem was gathered information about Fourier transformation, in order to use the power and the frequency. We learned that for get a spectrogram we need to divide the total interval of time into equal subintervals and then to calculate the Fourier transformation for the wave form contained in each interval and at the end merge all the subintervals in order to get the final spectrogram. This operation is executed with the python library "Librosa", using a function called "short time fourier transformation". We tried to apply that transformation to the frequencies, but we didn't get any relevant outcome, so we thought that it happens because for passing from the frequency's domain to the time's domain it's needed apply the reverse fourier transformation (and the feature given by science journal is not good for get a wave form to working on). Next, we get linear values and then we applied the fourier transformation to the given decibel values; in this case we manage how to get a spectrogram, but actually it doesn't seem to be significant because all of them are quite similar for all the genres. Then we thought that, since we already have time, pitch and frequency, we could obtain the spectrogram using only science journal without using any transformation, since the plot is a heat map, but the outcome is almost useless because the spectrogram is very sparse and it is not possible to extract information. Now we are working on a new way that it's giving us good results for getting the spectrogram that we want: in case it is not possible, we would change strategy for classify the genres. Another issue is that switching on more sensors together, we obtain 1/3 of the data with Nan values. For now, since we don't have more than 2 nan values consecutively, we replaced them with the previous values (but we are also considering other techniques).

Project Timeline

Hoping that everything will go as we expect, the following chart shows our plans:



References

- Automatic music classification and the importance of instrumnet identification
- Exploring Pitch with the Science Journal App
- Music Genre Classification using Machine Learning Techniques

---