Mateusz Bąk

Distinguishing and classifying percussion elements using machine learning

Introduction

Data from a set of 7371 samples of 200DrumMachines drum machines located on the ismir.net website was used for the analysis. The base used contains samples of kicks, snares, hats and claps.

1. Purpose of the project

The goal of the project is to train a model that recognizes the percussion elements.

2. Materials and methods

● Description of the population

The data consists of 137 kicks, 211 snares, 74 claps and 58 hats, i.e. a total of 480 samples of rhythm instruments. Before programming, they were selected to contain only wav files with a sampling frequency of 44.1 kHz and the full name of the instrument.

● Classification methods

The collected data were classified using the k-nearest neighbors method and the random forest method, using 13 melo-frequency cepstral coefficients, abbreviated as MFCC, i.e. mean value, standard deviation, median, 1st and 3rd quartiles, dispersion between the 10th and 90th percentile, kurtosis, skewness, minimum value, maximum value.

The model was trained using the Python programming language and the Jupyter environment.

3. Results

● k-nearest neighbors method

The number of neighbors was chosen empirically for the best results and is 10.

```
Dokładność (ang. accuracy):  0.875
Czułość (ang. recall):  0.8609126984126984
Precyzja (ang. precision):  0.8956243032329988
F1:  0.8751945556190839
```

Pic. 1. Success metrics of the k-nearest neighbors method.

Accuracy, i.e. the sum of correct predictions by the total number of predictions, is quite high, but the parameter itself is not sufficient to evaluate the classifier. Sensitivity, i.e. the ratio

of correctly recognized elements to all that should be recognized by the classifier, and precision, i.e. the ratio of correctly classified elements to all elements that the classifier placed in a given class, are also quite high, as they are close to 90%. F1, i.e. the harmonic averaging of the remaining success metrics, is 87.5% and is also the most reliable coefficient.

- **Random forest method**

The random_state parameter in the RandomForestClassifier function was selected empirically so that the results were as good as possible and is 10.

```
Macierz pomyłek (ang. confusion matrix):
[[ 0  0  0]
 [52 10  0]
 [57  1  0]]
Dokładność (ang. accuracy):  0.08333333333333333
F1:  0.091324200913242
```

Pic. 2. Random forest error matrix and success metrics.

The error matrix consists of 3 rows, indicating a positive result of classification into a given subgroup, and 3 columns, indicating a true result of classification into a given subgroup. The fourth subgroup was used in the algorithm as a test set. You can calculate accuracy, precision, sensitivity and F1 from it, but there is no need to do that, because you can see that the accuracy and F1 of the whole operation are very low, because they are below 10%.

4. Summary

In classifying percussion elements, the k-nearest neighbors algorithm turned out to be much more effective and the use of a model trained in this way would be quite reliable, while the random forest method is less effective than guessing, so in this case it is not recommended. Perhaps it is a matter of too small a group of data, which is 480.

5. Links

- https://ismir.net/resources/datasets/ – link to the sample base used.