Logistic regression and Bayes-classifier study of classification of songs to genres based on timbre, pitch and rhythm of the music signal

Ville Virkkala, Jarno Leppnen
Ekahau Research
(Dated: March 11, 2019)

A genre of a song can be estimated based on its music signal's characteristics . In this work we use two classifiers, Bayes Classifier and Logistic classifier, to classify songs into one of ten possible genres. The two classifiers are trained against the training data and their performance is compared against each other. In this work we show that both classifiers perform much better compared to random guess. However their capability to classify all songs is clearly limited the accuracy for both classifiers being around 60%.

I. INTRODUCTION

An automatic music transcription, *i.e.*, notating a piece of music to a speficic genre, *e.g.*, Blues, dates back to 1970s when first attempts towards automatic music transcription were made[?] . Since then interest in automatic transcription of music has grown rapidly and various approaches, statistical methods, modelling human auditory system, have been applied to music transcription problem. However even today an expert human musician often beats a state-of-the-art automatic transcription system in accuracy.

Characteristics of music signal that are useful in classification of a song are *timbre*, *rhythm*, *pitch*, *loudness* and *duration*[?] from which the three first one, described below are used in this work.

- The timbre of the music can be most easily described as the factor which separates two sources of music from each other. For example if the same song is played by violin or a guitar the timbre is called the character which separates the violin from the guitar.
- The pitch is related to frequency scale of a song a can be defined as the frequency of the sine-wave fitted to target sound by human listener.
- The rhythm of the music can be described as arrangement of sounds as time flows.

In classification problem the object is classified into a certain class based on it's characteristics called features. A linear classifier does the classification by making a linear combination of the features and converting the resulting value into a class or a probability that the object belongs to given class. In logistic regression the feature vector of the object is transformed into a probality by taking a linear combination of features and mapping the result into interval $[0, 1]$ using a sigmoid function. The Bayes-classifier in contrast assumes that the feature vector is drawn from a multidimensional-Gaussian distribution. The posterior probability of the object belonging to a certain class is then obtained as a product of the prior of the class and the probability to sample the given feature vector from the multidimensional Gaussian distribution.

The paper is organized as follows. The used data-set and the computational methods are described in detail in Sec. II. In Sec. **??** the results for the both logistic regression- and Bayes-classifier are given. Sec. III is a summary of the results and the differences between the two classifiers are discussed.

II. USED DATA-SET AND COMPUTATIONAL METHODS

A. Used data-set

The data-set consisted of 4363 songs and was divided into training and test data sets including every third song to test set and rest of the songs to training set. Each song contained 264 features and the songs were labeled to 10 different categories. The gatecories were: 1 Pop Rock, 2 Electronic, 3 Rap, 4 jazz, 5 Latin, 6 RnB, 7 International, 8 Country, 9 Reggae and 10 Blues. The musical characteristics of the songs were packed to a feature vector of length 256. The first 48 elements in the feature vector can be associated to timbre, the next 48 elements to pitch and the final 168 features to rhythm. The distribution of the features resembled in most cases a Gaussian distribution or a skew symmetric distribution. This is illustrated figures **??**a and **??**b.

B. Computational methods

In this work two different methods were used to classify the songs to different genres. First method is logistic-regression method in which the logistic-loss is minimized iteratively using the gradient descent method. The other method used is the Bayes-classifier which classifies the song to certain category that gives the maximum posterior probability with respect to label i. Both methods are described below in detail. In addition we studied the effect of feature extraction and for that purpose we used principal component analysis method to exclude features with little impact.

TABLE I: Confusion matrix corresponding to classification obtained using logistic regression. The column direction indicates the true value and the row direction is the predicted value. The labels 1 . . . 10 are the ten music genres specified in section II A.

	1	2	3	4	5	6	7	8	9	10
1	652	35	7	9	4	10	1	3	2	3
2	57	130	9	4	2	1	0	2	2	0
3	14	9	82	4	2	1	0	0	1	0
4	25	3	0	43	0	2	0	1	0	2
5	38	4	1	5	11	1	1	0	3	0
6	37	6	12	8	4	25	0	0	0	0
7	34	6	1	2	2	4	3	1	0	0
8	50	0	0	1	2	1	0	9	0	0
9	2	5	0	2	1	0	0	8	0	27
10	21	0	0	6	1	2	0	1	0	3

1. *Logistic-regression*

2. *Principal component analysis*

III. CONCLUSIONS

In this work we used logistic-regression and Bayes-classifier to classify songs to different genres based on the music signal's characteristics. For logistic regression the obtained accuracy for test set was 0.67 and logistic-loss 0.27. For the external data set used the obtained accuracy and logistic-loss were 0.65 and 0.178 respectively in the case of logistic-regression. For the Bayes-classifier the obtained accuracy and logistic-loss were 0.53 and 0.33 for the test-data and for external data set 0.32 and 1.17 respectively. According to obtained results both classifiers performed clearly better than random guess, but remained far from perfect classification. From the two classifiers used the logistic-regression classifier performed clearly better. The logistic classifier also generalized much better to completely new data giving nearly equal performance for test data set and external data set.