Report of Project 4: Words for Music

Yifei Hu(yh2781)

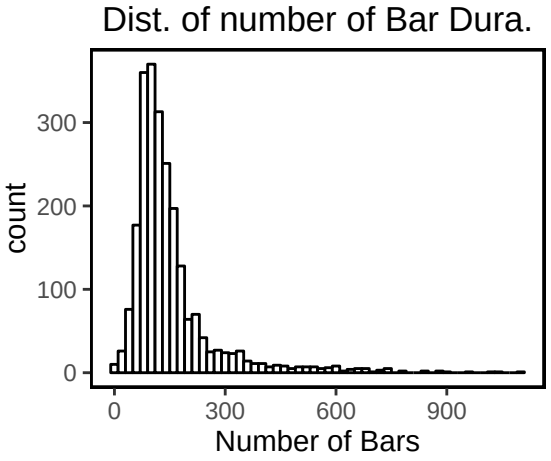
New York, 2016-11-18

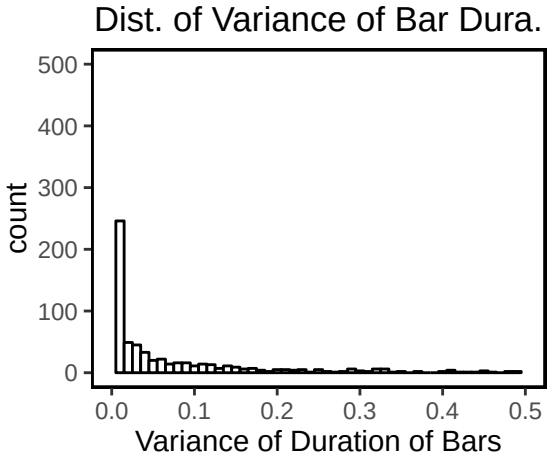
Steps:

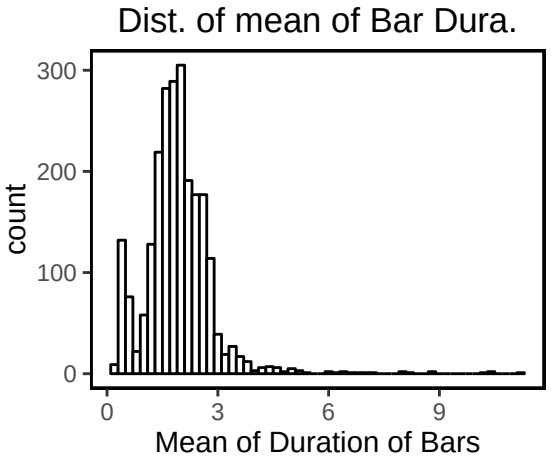
1.Given the frequencies table of the lyrics, I deleted the columns 2,3,6:30 since the words in the columns will be given 0 probability at last.

2.Assign Latent Dirichlet allocation model (LDA) on the frequencies table of the lyrics, choose the number of topic you want and then get the log of conditional probabilities of the words in the dictionary given each topic and the conditional probabilities of the topics given each song.

3.For the feature part, though there are 15 features which we can use for each song, I choose the features as the number, mean and variance of lag1 differences of bars_start, beats_start, sections_start, segments_start and tatums_start to represent the number, mean duration and variation of duration of the subsections. These features give us a sense of the change of the rhythm of the song. For figure of number of bars and mean of duration of bars we can see that different songs really have different number, mean duration of the bars. And for figure variance of duration of bars we can see that the variances of the songs are similar which shows the correctness of the measure of the provider of the h5 file.







For segments_pitches and segments_timbre, the mean and variance of each row of the two matrices

give a sense of the change of the attributes of the 12 musical scales. For segments_loudness_max, segments_loudness_max_time and segments_loudness_start, the mean and variance of each of them give a sense of the change of the max loudness of the song. And now we have total 69 features for each song.

4.We assign the label (the topic) to the songs which has the largest conditional probabilities of the topics given the song.

5.For the model part, I tried 3 models, multinomial logistic regression, random forest and multiple classes support vector machines with radial kernel. Then input the features and the labels of the songs in the training set into the three models.

6.For test sets, for multinomial logistic regression, the output is the probabilities assigned to each topic of each song. By using the formula

$$P(word|features) = \sum_{topics} P(word|topic) \times P(topic|features)$$

we get the probabilities of each word in the dictionary appearing in the song. For random forest and SVM the output is the topic assigned to each song. Then we find the conditional probabilities of the words in the dictionary given the topic which we already attain in the LDA. And these probabilities of the words are the probabilities of each word in the dictionary appearing in the song. The evaluation formula is:

$$Evalutaion_{testsong1} = \frac{1}{m}\frac{1}{\bar{r}}\sum_{i=1}^{m} r_{w_i},$$

where m is the number of different words in the test song, $\bar{r}$ is the summation of the rank of the predictied lyrics, r_{w_i} is the rank of word w_i that occurs in the true lyrics. And we take the mean among all the test songs.

After using 5 folders cross validation to choose the number of topics (1,3,5,10,15,20 topics), the predictive rank sum for multinomial logistic regression (actually 1 topic can not be done by this model, I present it here is just for comparison) is

Number of Topics	1(mean)	3	5	10	15	20
Predictive Rank Sum	0.2518	0.2445	0.2489	0.2514	0.2525	0.256

For random forest, the predictive rank sums are all about 0.7, which is worse. For SVM, the probability of correctly assigning label to each songs in test set is only about 10% which is really low.

7.At last I choose 3 topics in LDA and use multinomial logistic regression as the finally model.

Remarks

Last but not least, when I choose 3 topics, it's a little better than considering all the songs as one topic, and under other choice of number of topics the predictive rank sums are not far from considering one topic. So maybe the model I choose is not good or maybe the features provided do not have an enough high correlation with the lyrics which we can use to predict the rank of lyrics in a song by the features of the song.