Restaurant Grading And Recommendation

-- Based On The Sentiment Analysis In Twitter

TEAMMEMBER: YIZHI GU   A20334298

                JIE   HUANG   A20279698

PROBLEM

We consider if we can recommend the restaurant to the twitter users by analysis the tweets about the restaurant.

Our hypothesis is that the people's tweets can help us rating the different restaurant.

APPROACH

We use the machine learning method

First we training the Yelp reviews and choose the higher accuracy by testing the different parameter

Second we use the training data to predict the positive or negative reviews of tweets about the restaurant

Third we find positive review user and recommend the restaurant to his friends

DATA

We collect the data from Yelp —*yelp_academic_dataset.json*

Take 1000 positive reviews and 1000 negative reviews to be the training files.

Take another 1000 positive and 1000 negative reviews to be the testing

Second we request the tweets by using Twitter API

Set a restaurant file including 3 sample restaurants
Request 100 recent tweets of different restaurants

RESULTS

Based on the experiments, we set the parameter:

binary=True, min_df=4, max_df=.5, 5-fold cross-validation, and the estimated accuracy is 83.7%.

After we test the testing files, we got accuracy 85.9% which are pretty close to the estimated accuracy.

.

CONCLUSIONS

Limited by data scale and request limitation, we can't perform too much analysis on twitter users. And the limitation may affect the accuracy of the result.

If possible, I really hope to build a system with all twitter data. And perform a Cascades and Clusters analysis on a specific user's relation. Meanwhile provide a popular restaurant in their circle.