# Measuring the value of an online-review

As we known, online-review plays an important role in e-commerce that they have a major impact on the users’ purchase decisions. However, the increasing number of reviews instead of helping the customers better understanding the features of the products, it makes more bias and misunderstanding. In order to solve this problem, most of the e-commerce industries have their ranking and filtering system for users to extract the reviews they want. But the extant ranking system either can’t represent the value of the reviews or could be easy to manipulated by attackers. Our study focus on finding a method by using text mining to explore the real quality of a review, which not only reflect the value of the review, but also be non-vulnerable.

General Terms

Algorithm, Experimentation, text mining

Keywords

Review quality, labels, sentiment, CNN

1. **Instruction**

In the world of abundant online-reviews, one of the main challenges is to identify the high-quality ones from others. Nevertheless, because of the diversity of natural language, it’s difficult to compute the quality directly by the review contents. Instead, the common practice in e-commerce industries is to rank by the review update time or the review helpfulness. Unfortunately, due to the limitation of these ranking methods [4], it’s hard to reflect the real value of a review[3]. Thanks to the development of text mining, we discover a method to reflect the quality of a review. In the remainder of the report, we will talk about why do we use this method? What is this method? And how can we use this method.

1. **Motivation and objectives**

For many modern consumers, they make the purchase decision highly depends on the reviews of product. A real, objective review with detail help them draw better decision. For the sellers, the reviews are their word of mouth. They hope to have real and detailed feedbacks. For the e-commerce platform, high quality reviews help them increase the number of active users. Therefore, this is a triple win proposal.

1. **The value of a review and the measurement**

The value of a review is an abstract concept. Typically, measurement around reviews is made of a mix of Qualitative and Quantitative data. It's impossible to simply say 'A Review is Worth £X.XX' because all manner of things from tone to language to placement will affect the value of the review.

Generally speaking, a review has two kinds of value. One comes from the review itself, such as the update time or the level of the reviewer. The other one is given by people, such as how helpful the review is or the number of the comments. The measures of a single kind of value above is simple and straightforward that you can directly get it from the application. We develop a measurement which can reflect the two value at the same time.

Our approach is to analyze the amount of information in a review. In our consideration, more information higher value of a review. For example, “the steak is crispy and juicy” is better than “the food is good here”, because the first sentence has more information that it told us what food is good and why it is good. And we found that more information is actually more details and more objective.

To represent the details of a sentence, we manually score how related a sentience and the labels we have. In our case, we have 6 labels: amenities, environment, food, location, price, service. [0.5,0,0,0,0,0] means the sentence is about 50% probability related to amenities. For the sentiment, we use a constant between 0~1 to represent how objective the review is. To represent the value of a review, we sum up the product of each details and sentiment. The higher value, the higher quality the review has.

1. **The limitation of common practice in review ranking**

The reviews always ranked by update time and user’s votes. But both ranking methods have their limitations. The disadvantage of ranking by update time is that it doesn’t reveal how people think of the review. In other words, it doesn’t show how popular the review is and how people like the review. For ranking by user’s vote, such as rank by review helpfulness or the number of comments, it is vulnerable that the ranking position can be manipulated. And what’s more, the popular review will become more popular, because of its high ranking. To break these limitation, we introduce a new ranking method – rank by the value/quality of a review [3], which not only inherit the advantage of common ranking method, but also give up the shortcomings of them. Next, we will present the details of implementation of our ranking method.

1. **Methodology**

The main challenge of our study is to quantificat the amount of information of a review. First of all, we need to quantificat the review content into a embedding matrix [5.2]. And then we need to make a regression to change the matrix into the vector of labels or the value of sentiment by using Convolution Neural Network [5.3,5.4]. At last, we apply a formula to compute the quality of the review**.**

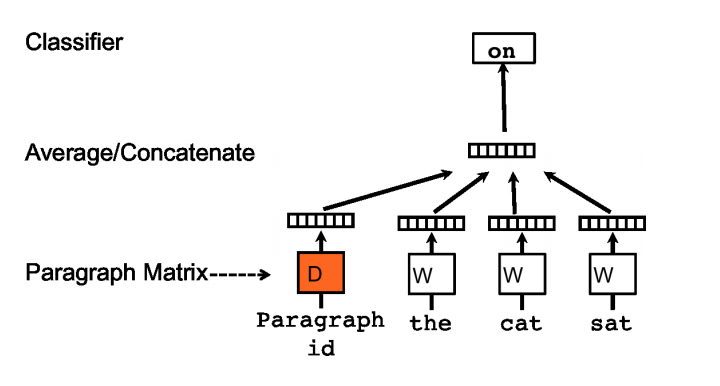


5.1 Data Set

We use the subset of the yelp official data set in this study. The Yelp dataset is a subset of their businesses, reviews, and user data for use in personal, educational, and academic purposes. It contains 4,700,000 reviews about business attributes like hours, parking, availability, and ambience over 12 metropolitan areas.

* 1. Document Vector

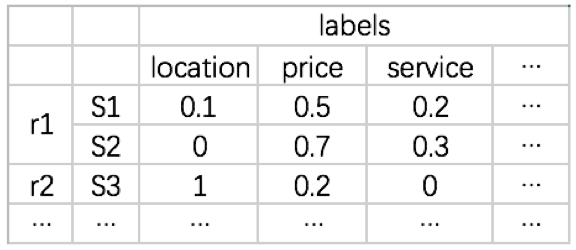
In our case, we need to represent the review text input as a fixed-length vector. When it comes to texts, one of the most common fixed-length features is bag-of-words (Harris, 1954). Despite their popularity, bag-of-words features have two major weaknesses: they lose the ordering of the words and they also ignore semantics of the words. To overcome the weaknesses of bag-of-words, a concept called Paragraph Vector is proposed (Tomas, 2014), which is an unsupervised algorithm that learns fixed-length feature representations from variable-length pieces of texts. The algorithm represents each document by a dense vector which is trained to predict words in the document.



In our experiment, we split each review into sentences. Each sentence is a document/paragraph. Every paragraph is mapped to a unique vector, represented by a column in matrix D and every word is also mapped to a unique vector, represented by a column in matrix W. The paragraph vector and word vectors are averaged or concatenated to predict the next word in a context. After doing that, we transfer each sentences into the fixed-length vector.

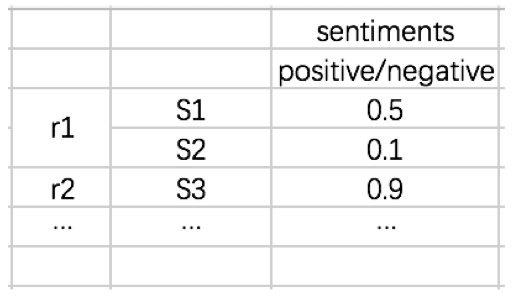
* 1. Measuring details of a review

In order to represent that a sentence is relative to some particular areas, the common practice is to manually label the sentences with tags. For example, “The steak is juicy and crispy” can be labelled as “food”. “The sweet potato is cost $10” can be labelled as “price”. When tagging the sentence, we create a vector base on the list of total tags. In the beginning all the initial numbers in the vector are 0. If the sentence hit one of the tag of the list, the relative element of vector becomes 1. For instance, if we have total 6 tags: amenities, environment, food, location, price, and service. Given the sentence “The steak is juicy and crispy”, the output vector would be [0,0,1,0,0,0]. These output vector of training set is used to predict the next labels. Comparing the performance with other methodologies [6], we use Convolution Neural Network as our model to train the data. In addition, instead of generating binary values to the vector, we generate a constant between 0 and 1 to represent how related are the sentence and the particular areas. The value is closer to 1 if there are more details about the particular area. Besides, we still use CNN as the regression model.



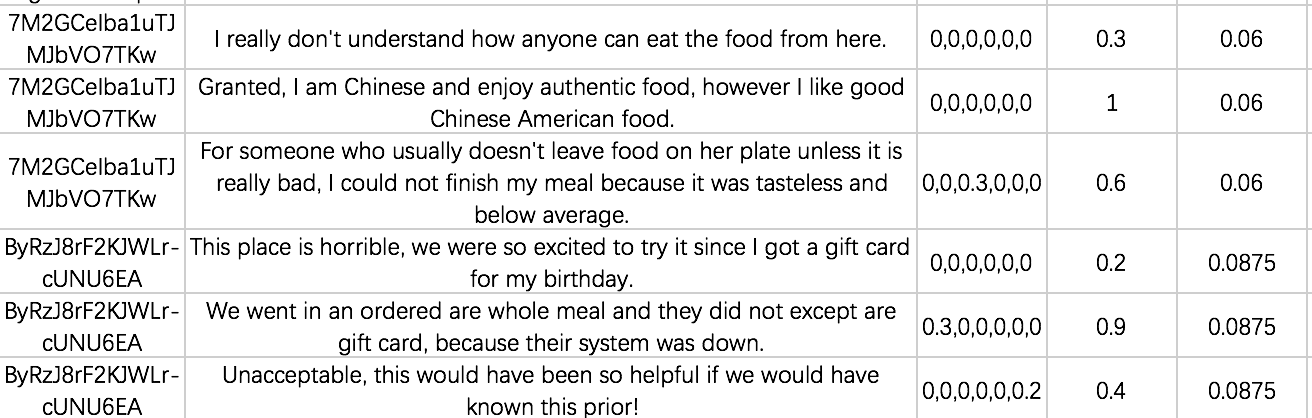
* 1. Measuring sentiment of a review

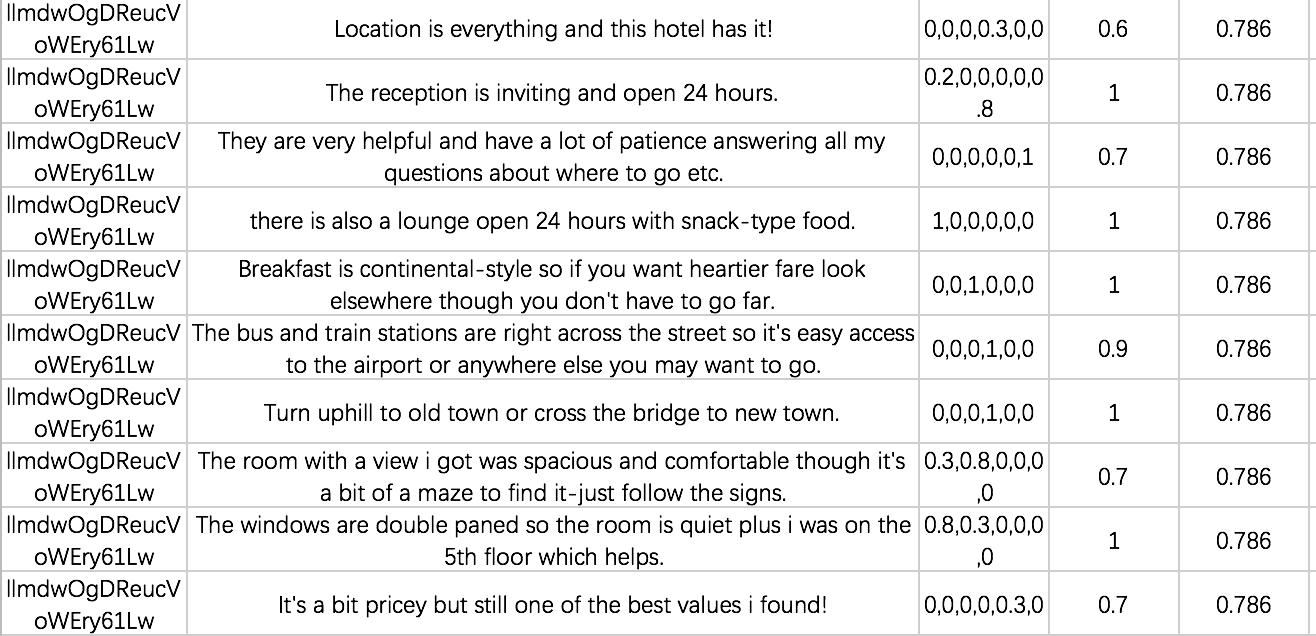
Traditionally, the sentiment in a review can be positive, negative or neutral. And we only want to know how objective the review is. So we can mark neutral as 1 and positive/negative as 0. But it is still hard to do the agreement because of the bias of the review. Therefore, we generate a constant between 0 and 1 to represent the level of objectivity. More neutral, the value is closer to 1. We train another Convolution Neural Network to predict the sentiment of the next review base on the training data we labeled. And finally, we can predict the level of a review being objective.



* 1. Calculating the quality of a review

Intuitively, the rank of the quality of the review can be: objective review with details > subjective review with details >= objective reviews without details > subjective review without details. For example, “the beef steak is crispy and juicy” > “the beef steak is delicious” >= “I am a meatatarian” > “I hate the food here”. Therefore, we use a formula to represent the quality Q of a review: Q = (S\*L) / n = (Sj \* Lij ) / n = ((S1\*L11+S1\*L12+…+S1\*L1m)+(S2\*L21+S2\*L22+…+S2\*L2m)+ (Sn\*Ln1+Sn\*Ln2+…+Sn\*Lnm))/n. S refers to the vector of the sentiment[5.4]. L points to the vector of labels[5.3]. n means there are n sentences in the review. m is the total number of the tags.

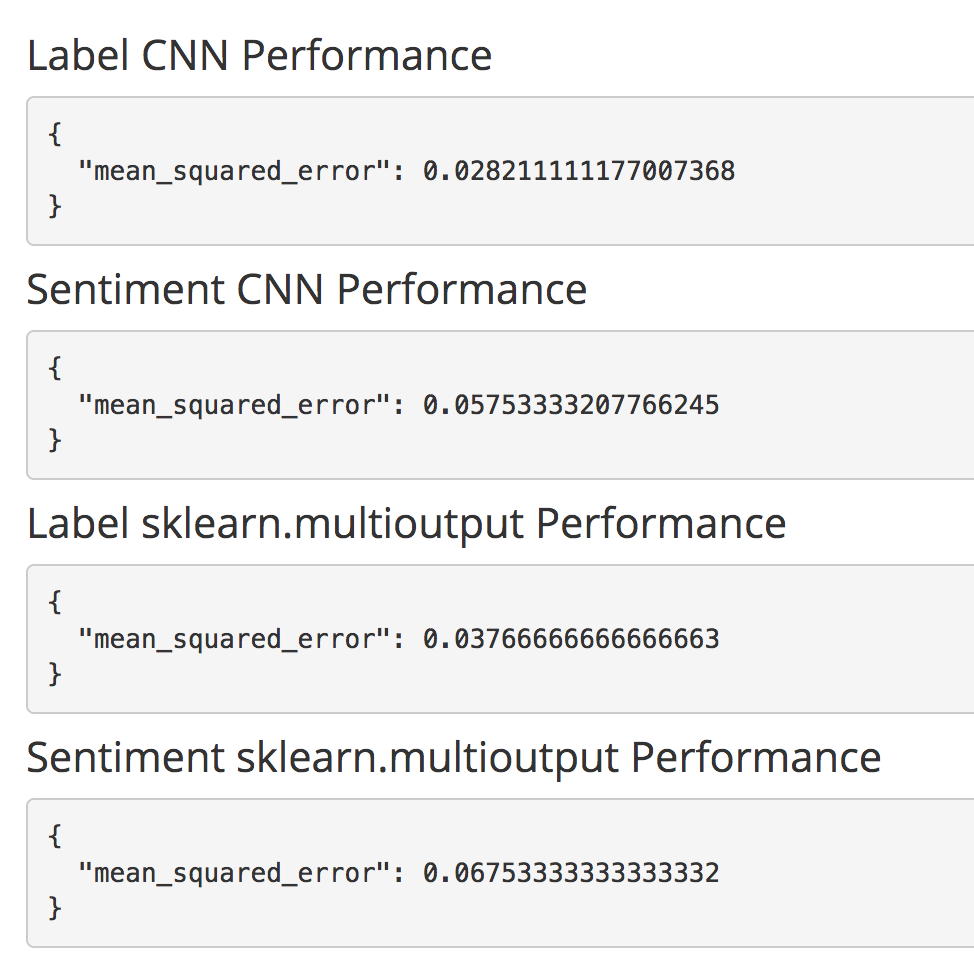




The last column is the quality of the review

low quality review vs high quality review

1. Evaluation Comparison



We use 350 training data and 150 validation data to train the model. The result shows the performance of CNN is better than the performance of multiclass regression in sklearn in our case.

**7. Deploying**

7.1 Install the dependencies

python 2.7

sklearn

matplotlib

keras

gensim

nodejs (admin)

npm (admin)

word\_sample.json

data\_sample2.csv

7.2 Training the samples

1. open jupyter

2. modify all retrain=0 to retrain=1 on the bottom of ReviewAnalyser.ipynb

3. run the cells

7.3 Start Restful API server

python RestfulAPI.py

7.4 Run the backend program

1. cd admin

2. npm install

3. npm start

7.5 Access the admin system

Access <http://localhost:3001> in the browser

**Reference**

Theodoros Lappas, Gaurav Sabnis, Georgios Valkanas The Impact of Fake Reviews on Online Visibility: A Vulnerability Assessment of the Hotel Industry

Quoc Le, Tomas Mikolov Distributed Representations of Sentences and Documents

Rob Thomas What is the value of an online review?

Amazon official website https://www.amazon.com/review/guidelines/top-reviewers.html/