Sentiment Analysis Based Online Restaurants Fake Reviews Hype Detection

Xiaolong Deng[1] and Runyu Chen[2]

[1] Key Laboratory of Trustworthy Distributed Computing and Service,
Beijing University of Posts and Telecommunications
10 Xitucheng Road, Beijing, 100876, China
shannondeng@bupt.edu.cn

[2] International School, Beijing University of Posts and Telecommunications
10 Xitucheng Road, Beijing, 100876, China
4007821@qq.com

Abstract. In our daily life, fake reviews to restaurants on e-commerce website have some great affects to the choice of consumers. By categorizing the set of fake reviews, we have found that fake reviews from hype make up the largest part, and this type of review always mislead consumers. This article analyzed all the characteristics of fake reviews of hype and find that the text of the review always tells us the truth. For the reason that hype review is always absolute positive or negative, we proposed an algorithm to detect online fake reviews of hype about restaurants based on sentiment analysis. In our experiment, reviews are considered in four dimensions: taste, environment, service and overall attitude. If the analysis result of the four dimensions is consistent, the review will be categorized as a hype review. Our experiment results have shown that the accuracy of our algorithm is about 74% and the method proposed in this article can also be applied to other areas, such as sentiment analysis of online opinion in emergency management of emergency cases.

Keywords: Sentiment analysis, Hype review, Multi-dimension analysis, Bayes judgment.

1 Introduction

In nowdays, most consumers have the habit of scanning the online reviews before purchasing the commodity. Products with many positive reviews often make good impact to customers. Therefore, for merchants in e-commerce time, it is important to maintain the online reputation of their products. However, some of the merchants, in order to enhance the popularity of their products, they hired some groups of people who is called internet "Water Army" to post positive reviews of their products online, and which can greatly affect the choice of consumers.

In this article, we have established collaborative relationship with dianping.com and have obtained fake review dataset collected by the company. Through serious observation and analysis, we concluded that most of the fake reviews are hypes. And they are very similar to authentic reviews in many aspects including length, tone and

W. Han et al. (Eds.): APWeb 2014 Workshops, LNCS 8710, pp. 1–10, 2014.

wording and so on which is the reason that is why they are most misleading to consumers.

We have done some further investigation on fake reviews of hype we sort out. This kind of reviews is most distinctive in its text, i.e. totally positive or negative. Some reviews of this kind, although not suspicious according to analysis on other attributes, possess little reference value because of extreme sentiment.

Aiming at the problem described above, we have applied sentiment analysis methods to detection of such kind of fake reviews. To obtain a desirable accuracy, we first find out all the subject words by matching the sentiment words we extracted from all the fake reviews data set with the corresponding features. We manually judged the subject words and divided them into four dimensions: taste, environment, service, overall attitude. To evaluate the authenticity of a review about restaurant, our algorithm will conduct sentiment analysis from the four dimensions of it based on Bayes classifier. If the analysis results of the four dimensions are the same (all positive or negative), the review is hype. And to our surprise, the algorithm has an good accuracy of about 74%.

2 Related Work

Relevant research is always focus on spam detection [1] and rubbish website diction [2]. And in recent years, researchers started to identify spam reviews.

Researchers have concentrated on some different characteristics of spam reviews. Jindal proposed the algorithm concerns unusual score [3][4].Wang and Xie paid more attention on store reviews to detect spam.[5][6].Arjun considered a number of indicators to find fake reviewer groups[7]. Some mainly considered relevant, Song defined the relevance of their basic characteristics and features among other comments [8]. Myle proposed the combination of language features SVM modeling [9].

In this paper, a innovative method is proposed to classify the content itself from the marked spam reviews, for which account the largest number and most influential speculation reviews. Then we draw the appropriate model program to solve this problem. By using sentiment analysis in more special areas of multi-dimensional analysis, our method gained greater accuracy of identification of such spam reviews. In comparison, our algorithm requires much more manual handling, but the recall rate reflects better performance.

3 Sentiment Analysis

3.1 Bayes Classifier

Through investigation on the possibility of a specific event happening in the past, Bayes Theorem can calculate the approximate the probability of the event happening in the future. And Bayes theorem only requires a few parameters for estimation and is not sensitive to the lost data. Algorithms that implemented from this theorem can run more faster due to the simplicity of Bayes theorem [10].

Bayes classifier is the application of Bayes Theorem to classification of text. By calculating the probability of each category, the algorithm categorizes the text to the one with largest probability.
Bayes equation:

$$P(c_j \mid x_1, x_2, \dots, x_n) = \frac{P(x_1, x_2, \dots, x_n \mid c_j)}{P(c_1, c_2, \dots, c_n)} \tag{1}$$

Since each condition is independent and identically distributed, we have:

$$c_{NB} = P(c_j) \prod_i P(x_i \mid c_j) \tag{2}$$

Because there is no apparent feature in the labeled text, we apply multi-event model based on word frequency here:[11]

$$P(t_i \mid c = spam) = \frac{1 + Count(t_i)}{n + Counts(\sum_1^n t_i)} \tag{3}$$

3.2 Training Set Obtainment

Training data is always needed in calculating Bayes prior probability, and we manually selected 500 totally positive and negative reviews respectively. We have proved that the training set is enough since by enlarging the size from 300 to 500 we did not find noticeable change in the result. When selecting the sample of training set, we followed the principle that the review is as comprehensive as possible.

Table 1. Training set data

Positive	Negative
三人行骨头锅太给力了，冬天啃绝对过瘾，料很足，好多肉，人气爆满，装修美观，去晚了就要排队，钵钵鸡，个人觉得不错，喜欢吃辣的朋友可以尝试下，味道十足。	服务态度极差提前一周打去说只能提前两天预订等到提前两天打去又没位子了还说没人接过电话说要提前两天预订领班态度也差还挂客人电话。
昨天请朋友来吃饭，听说这里的蟹晏菜挺有名的，于是就点了几只大闸蟹，点了个蟹粉豆腐。还有几道特色菜，片皮鸭，5A牛柳粒，古法局鱼。朋友都说不错我很高兴。这里环境也挺舒服的，服务态度很好，注重细节，我们想要的在未开口之前服务员都想到了。	第一次去就被咖喱牛腩的牛腩震惊了，那么老的牛腩，都快赶上牛肉干了，牛腩也完全不入味...询问服务员还说他们的牛腩一直是这样的，那请问前面点评里面说牛腩嫩、烂的吃的是什么...强烈不推荐！服务员态度很不好，下次绝对不会去了。蟹粉小笼十分一般，根本不値这个价格。

Table 1. (*Continued.*)

同事们商量着一起去吃点东西，大家一致赞同，冬天里吃火锅应该是再好不过的选择了吧，就这样我们一伙人就过来了，店里的生意很好的，但还是井然有序的，我们找了个地坐下，点了个火锅，一会的功夫就上来了，菜很丰富，话不多说，大家就开动了，有菜有肉，真是一顿营养大餐呢，汤鲜味美啊，大家都吃的很高兴，一边聊着天，一边吃着饭，真是享受呢，这才是生活啊，以后有时间还要常来啊，朋友们有空也可以过来尝一下，相信不会让你失望的。	锅贴的品种不多，也没什么可挑选的，贡丸汤一份，偌大一只碗，里面晃悠着两个贡丸；中午时分也没有坐满，三两个服务员在休班闲聊，毫不避讳的高谈阔论着；我们点的锅贴好了，得自己去台面上找， 就几个单就有点混乱了，服务水平可见一斑。
...	...

3.3 Sentiment Evaluation

To find out the sentiment of a review, we use the method of multiplying the conditional probability of each word. Whether a review is positive or negative depends on the value of the result of multiplying process. We multiply the probability of each word with 10 first in case the result is too small.

4 Multidimensional Discrimination

4.1 Establishment of Sentiment Word Library

First we sorted out 56483 reviews about restaurants from our original data set. With the help of ICTCLA50 algorithm [12], we divided all the Chinese contents into words. We extracted all the sentiment words and set up our sentiment word library. In the same time, we calculated the frequency of each adjective, which enables us to get rid of some undesirable words with low frequencies. The final version of our sentiment library contains 1590 adjectives.

Table 2. Emotional word library

Words/Frequencies		Words/Frequencies	
很好/a	12.389356882973637	实在/a	7.453503647044743
好吃/a	11.14669772771281	舒适/a	7.406438314516372
不错/a	11.052315097990974	丰富/a	7.3550685280099195
实惠/a	9.742061863915502	亲切/a	7.350785294189377
干净/a	9.651343822312056	安静/a	7.289251208561696
值得/a	9.466860974306062	确实/a	7.262434541770648
周到/a	9.096459212424548	可口/a	7.18616605729332
特别/a	9.000203638575638	贴心/a	7.141585285545526
舒服/a	8.995655552806497	独特/a	7.044336876237392
便宜/a	8.929070356945344	合理/a	7.03207042683553
最好/a	8.40933391245834	一样/a	6.988856149668119
优惠/a	8.213934707582403	美味/a	6.957181112804157
优雅/a	8.193907482484725	重要/a	6.9435164471041055
温馨/a	8.078017307503831	合适/a	6.795440558011954
地道/a	8.023967758424863	有味/a	6.752448672847059
划算/a	8.008506960447109	适中/a	6.7334198176058955
一般/a	7.982753540338287	整洁/a	6.654735396670123
方便/a	7.9191048016013506	一流/a	6.63665330359303
热情/a	7.905370323918099	主动/a	6.636631314797421
鲜美/a	7.887444208296432	过瘾/a	6.61751253180918
地方/a	7.806704408558014	失望/a	6.612146670170864
入味/a	7.788232816199374	随便/a	6.533682077002941
精致/a	7.6268743931156555	耐心/a	6.46625147932886
其实/a	7.513344086828133	开心/a	6.452121735931351
…		…	

4.2 Feature Finding

Feature refers to the corresponding none phrase of existing sentiment words. We managed to find them by searching words before the adjectives that match those in our library; the nearest noun phrase is the matching result. Finally, we obtained 2303 subject words of reviews about restaurants.

4.3 Classification of Subject Words

After finding all the subject words, we analyzed the dimension each of them belongs to and finally limited the dimension into four types: taste, environment, service and overall. Then we classified these subject words. There are 1314 words belong to dimension “taste”, including those that describe the look, flavor and taste of foods. 222 words describing the environment, geographical conditions and traffic conditions of the restaurants are categorized into “environment”. 286 words are categorized into “service”, most of them describe the quality of service, prices. And there are 151 words classified as “overall”, including name of restaurants, names of places and holistic description etc.

Table 3. Main body of word library

Taste	Service	Environment	Overall
口味	阿姨	摆设	安溪
阿拉斯加蟹	按摩	包房	安阳
爱尔兰咖啡	按摩师	包房环境	澳门豆捞
安格斯牛肉	包装	包房装饰	澳门街
安徽菜	包装盒	包间	澳洲
安排的菜品	保安	包间风格	八佰伴
八宝辣酱	杯子	包间环境	巴贝拉
八宝年糕	菜的价格	包厢	巴蜀风
八宝养生茶	菜价	包厢环境	斑鱼府
八宝鱼	菜价格	背景音乐	北京烤鸭
霸王蛙	菜肴的服务	布局	餐厅
霸王鱼头	菜肴的价位	布置	大酒店

Table 3. (*Continued.*)

Taste	Service	Environment	Overall
白菜	菜肴的质量	餐馆生意	店
白菜粉丝	菜肴服务	餐厅	店气氛
白肉丝	餐厅的厨师	餐厅的布置	店人气
白水鱼	餐厅的服务人员	餐厅的环境	店生意
白汤	餐厅的老板	餐厅的设计	店味道
白汤牛奶	操作过程	餐厅印象	东西总体
白糖	茶水	厕所	饭店
…	…	…	…

5 Hype Identification

5.1 Determination of Dimension

By observing the sample data we find that spans of reviews are usually very large, thus using period to divide the sentences brings many errors. To avoid this, we use comma as monitoring sign to divide the sentences first.

When categorizing each sentences, we first set the default dimension as "overall". Then we matching each subject word in the sentence with our library established before, the sentence will be categorized into the dimension with most matching words.

5.2 Detection of Hype

In our process of detecting reviews of hype, our model analyzed the sentiment of all sub sentences of each dimension. If the results of four dimensions are same (all positive or negative), we will label the review as hype. In the following statistics, we find that positive reviews of hype are far more than negative reviews. Thus we assume that the default sentiment of a review is positive when there is information loss in 1 or 2 dimensions. Result shows that this preprocessing brings error less than 0.1%.

6 Experiment Result

6.1 Algorithm Difficulty Analysis

Through the analysis and category on the original data provided by dianping.com, we obtain 17681 fake reviews of hype about restaurants, and we applied our model to test these data. With the increase of reviews being tested, the result accuracy remained stable around 68%.

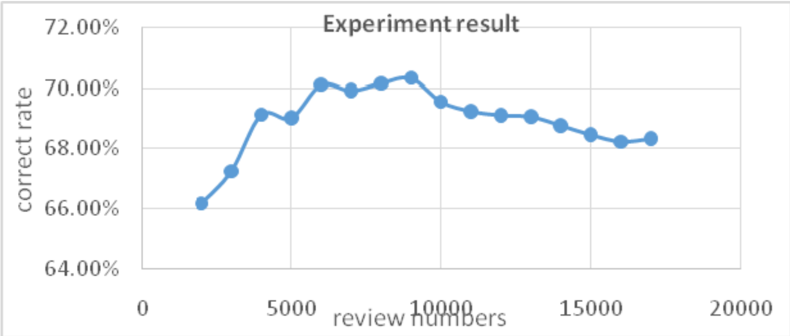

Fig. 1. Experiment result on raw data

Afterwards, we randomly sampled 2000 reviews to analyze. We found some error in labeling among the original data by checking the content of each review. Possible reasons are summarized below: (1) reviews are not about restaurants; (2) reviews are apparently not hype. After correcting the label of the 2000 reviews one by one, we witnessed the accuracy increased from 67% to 73%. Therefore we can estimate that same labeling errors exist in other parts of original data set, and the accuracy of our algorithm could reach around 74%.

Table 4. A random sample of 2,000 reviews

Review numbers	**Original correct rate**	**Non-restaurant reviews**	**Available reviews**	**Error mark**	**Revised correct rate**
1-500	68.4%	34	466	7	**73.2%**
501-1000	62.8%	29	471	11	**70.3%**
1001-1500	65.6%	38	462	9	**71.4%**
1501-2000	71.8%	21	479	5	**76.2%**
All	67.15%	122	1878	32	**72.79%**

In Figure 2, sample 1 to 6 is respectively stands for review numbers 1-500,501-1000,1001-1500,1501-2000,1-2000 and 1-17681. It is hard for us to clean all the error marked data, but the ultimate accuracy can be predicted about 74%.

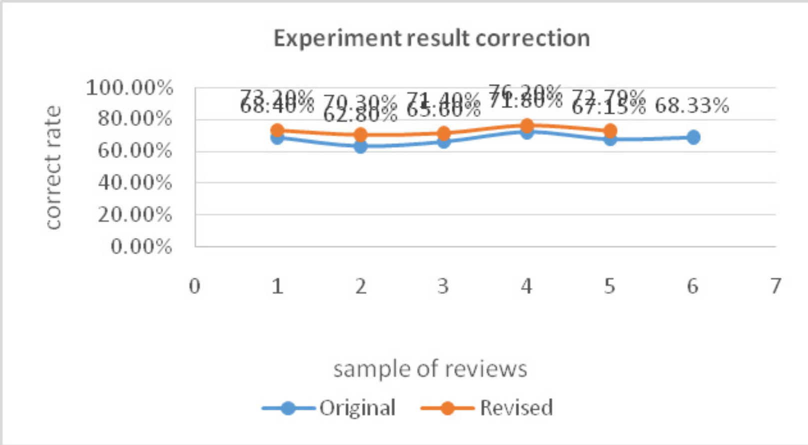

Fig. 2. Experiment result on revised data

6.2 Algorithm Difficulty Analysis

During the error analyzing process, we found that nearly 60% of errors occur for mainly two reasons. (1) Using comma as delimiter can cause incomplete sentence-breaking. (2) Some neutral reviews can be sorted wrongly by different training set, and it is difficult to judge whether they are hype or not.

7 Application of Emergency Management

The method proposed in this paper can also be used in emergency management for sentiment analysis. We can divide comments into different dimensions to get more specific sentiment analysis. We can also base on individuals, considering some of its comments together, to discover some associations with highly consistent in the emotion.

8 Conclusion

In this paper, we proposed a method to detect fake reviews of hype based on sentiment analysis. Reviews of hype take large part of fake reviews and are influential. The model we established can detect this type of review with an accuracy of around 74%. We set up our own sentiment word and multi-dimensional subject word library focus on reviews about restaurants as well. This method can also be applied to field other than reviews about restaurants.

Acknowledgement. Thanks to the support of National Natural Science Foundation of China (NNSF) (Grants No.90924029), National Culture Support Foundation Project of China (2013BAH43F01), and National 973 Program Foundation Project of China (2013CB329600, 2013CB329606).

References

1. Drucker, H., Wu, D., Vapnik, V.N.: Support vector machines for spam categorization. IEEE Transactions on Neural Networks 10(5), 1048–1054 (2002)
2. Ntoulas, A., Najork, M., Manasse, M., et al.: Detecting spam web pages through content analysis. In: Proceedings of the 15th International World Wide Web Conference (WWW 2006), Edinburgh, Scotland, May 23-26, pp. 83–92. ACM, New York (2006)
3. Jindal, N., Liu, B., Lim, E.P.: Finding Unusual Review Patterns Using Unexpected Rules. In: The 19th ACM International Conference on Information and Knowledge Management (CIKM 2010), Toronto, Canada, October 26-30, pp. 625–628 (2010)
4. Jindal, N., Bing, L.: Opinion spam and analysis. In: Proceedings of the 1st ACM International Conference on Web Search and Data Mining (WSDM 2008), February 11-12, pp. 137–142. ACM, New York (2008)
5. Wang, G., Xie, S., Liu, B., Yu, P.S.: Review Graph based Online Store Review Spammer Detection. In: 2011 IEEE 11th International Conference on Data Mining (ICDM), Vancouver, BC, December 11-14, pp. 1242–1247 (2011)
6. Xie, S., Wang, G., Lin, S., Yu, P.S.: Review Spam Detection via Temporal Pattern Discovery. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, August 12-16, pp. 823–831 (2012)
7. Mukherjee, A., Liu, B., Wang, J., Glance, N., Jindal, N.: Detecting Group Review Spam. In: Proceedings of the 20th International Conference Companion on World Wide Web (WWW 2011), New York, NY, USA, pp. 93–94 (2011)
8. Haixia, S., Xin, Y., Zhengtao, Y., et al.: Detection of fake reviews based on adaptive clustering. Journal of Nanjing University: Natural Sciences 49(4), 38–43 (2013)
9. Ott, M., Choi, Y., Cardie, C., Hancock, J.T.: Finding Deceptive Opinion Spam By Any Stretch of the Imagination. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL HLT), Portland, OR, United States, June 19-24, pp. 309–319 (2011)
10. Wang, J., Wang, L., Gao, W., Yu, J.: A research on the keywords extraction of naive Bayes. Computer Applications and Software 31(2), 174–181 (2014)
11. Zhang, F., Wu, Z., Yao, F.: Research of spam filter based on Bayes. Journal of Yanshan University 33(1), 47–52 (2009)
12. Xia, T., Fan, X., Liu, L.: Implementation of ICTCLAS system based on JNI. Computer Application 24(z2), 945–950 (2004); Deng, J.: Control problems of grey system. Systems & Control Letters 1, 288–294 (1982)