A STUDY ON THE REVIEW ANALYSIS FOR THE AUTOMATIC EVALUATION OF THE PRODUCT

[1]ZHUORAN SUN, [2]MINORU YOSHIDA, [3]KAZUYUKI MATSUMOTO, [4]KENJI KITA

[1,2,3,4]Faculty and School of Engineering, University of Tokushima, Tokushima, 7708506, Japan
E-mail: [1]c501437042@tokushima-u.ac.jp, [2]mino@is.tokushima-u.ac.jp, [3]matumoto@is.tokushima-u.ac.jp, [4]kita@is.tokushima-u.ac.jp

Abstract- Nowadays, e-business has been never more popular and more convenient. Online merchants have flourished and merchandize have diversified. However, it's still difficult for users to quickly locate their desired products on online markets. Often times, an online shopper will evaluate ratings and comments from other users to speculate the current product. In particular, the easiest and the most intuitive way is to see the scores to determine the quality of the product. Therefore, if we can automatically identify and analyze all the reviews and accurately and objectively give a score to a particular product, it will be more efficient for users to search merchandize and for online business owners to manage the products they sell.

Keywords- E-business, Reviews, Analysis, Automatic Evaluation, Manage

I. INTRODUCTION

Nowadays,online reviews for goods and services sold and provided on the online marketplace have shown a practical phenomenon of Big Data. For example, Amazon.co.jp*1of Japan, Taobao marketplace of China*2, Rakuten marketplace of Japan*3 all usea 1-5 ratingschemefor users to rate the products they purchased.To give another example, "GURUNAVI" *4allows user to post reviews on restaurants. From a statistical point of view, almost all websites use the aforementioned 1-5 rating scheme..Note that this rating scheme does not limit commentators, who may comment on the pros and cons of the provided goodsand services. Of course, positive reviews will most likely benefit them.

For consumers, looking at the ratings is the most straightforward way to make a purchase decision. However, research has shown that current rating mechanism on most website is based on consumers' personal opinions. The problem is with this is that a user's rating fails to align with the reviews. For example, the review with a 5-start rating doesn't meet the 5-star standard. Another problem is the polarization of ratings where a user will either give a 5 or a 1. Users rarely usescores in the middle range or they claim that they don't know how to apply the rating rubrics.

Therefore, I wonder if it is possible to rate a merchandize from analyzing user reviews and summarizing the number of favorable reviews. As a result, ratings will vary with the number of user reviews and will be more objective and practical.Also it can make a comprehensive judgment with commentswhich are written by people in different country on the same commodity such as books, movies, CDs, etc.

II. DETAILS EXPERIMENTAL

2.1. Data Collection

In this research, we have gathered database from the online shopping platform, Amazon for Japan and China. Specifically, we extracted two types of data, user evaluation data and user review data. Either type is extracted with 209 records (129 positive reviews and 80 negative reviews) for Japanese,and 382 records (200 positive reviews and 182 negative reviews) for Chinese which will be used as raw database for this experiment.But we collect more data for research.

But, in this paper, I will show a simple example about comparison between Japan and China, about the Harry Potter series of comments, and get a composite score for this book.

Table 1Data base for Japanese comments

Number	Comment of Japanese
1	日本では京極夏彦が妖怪小説の様式を使い本格ミステリーを書いて度肝を抜かせたがハリーポッターシリーズはファンタジーの様式で書かれたミステリーである。…
2	とっても面白いです。最近このたぐいの小説に出会えなかったのでうれしいです。シリーズ化しているので、さらに楽しさ倍増です。…
3	UK 版を友達から借りて読みました。1章目は読みにくかったのですが、2章目からは、本を置くことが出来ないくらいはまりました。…
…	…
128	非常に話にのめりこみやすく、何度読んでもあきません！！！ぜひ、一回読む事をお勧めします！！
129	私は本が大好き！『魔法』物の本には特に目がなく、この本ももちろん即ご購入

Table 2Data base for Chinese comments

Number	Comment of Chinese
1	开始买书一直担心质量，其实七册都很好，绝对是正版！快递很快，比当当快了一天！我真不是拖，不过都是13年加印个的，我想买第一次的，真的很难！
2	拿到书一口气全看完了，再来评价的，纸页质量轻有韧性，印刷很好，没有错别字。应该是正版。故事内容也超级吸引人。
3	全7本 都看过。给4年级的女儿看的 我自己倒是特别的喜欢。
…	…
199	不错！纸质非常好！字非常清晰！物流非常快！
200	不错，这本书与预想的一样好。纸面的质量也不错。

2.2. Data analysis

Inthis experiment, we use a phrase score SS(S) for string s defined as:

$$SS(s) = \frac{\alpha_1 + freq(s,C)}{\alpha_2 + freq(s,C) + freq(s,C')}$$

This score can be considered as the Bayesian a posterior probability of s being in corpus C after observing freq(s,C) and freq(s,C').When extract good reviews, the freq(s,C) is for 5 score, and freq(s,C') is for below 4 score. On the contrary, when extract poor reviews, the freq(s,C) is below4 score, and the freq(s,C') is for 5 score.

The edge score of string s is defined by below.

$$ES(s) = \frac{(freq(s) - 1) \cdot len(s)}{DL(s)}$$

Using this equation, we can get the most coherent meaning, and the most complete expression. So we use there score to sort the results. We put the number set to 1000 for results.

Table 3DataAnalysis for Good Expression of Japanese

Number	Result of Data Analysis
1	この本は(0.838)
2	の中で(0.834)
3	す！(0.834)
…	…
999	ます。何(0.681)
1000	ラシ…ンも(0.681)

Table 4 Data Analysis for Poor Expression of Japanese

Number	Result of Data Analysis
1	残念(0.719)
2	。ただ(0.691)
3	より(0.680)
…	…
999	り面白く(0.486)
1000	をつけ(0.486)

Table 5 Data Analysis for Good Expression of Chinese

Number	Result of Data Analysis
1	呵呵呵呵呵呵呵呵呵呵(0.887)
2	很快(0.758)
3	非常喜欢(0.720)
…	…
999	！希望(0.603)
1000	，发货也很快(0.603)

Table 6 Data Analysis for Poor Expression of Chinese

Number	Result of Data Analysis
1	字体(0.790)
2	小(0.750)
3	盗版(0.732)
…	…
999	错误(0.480)
1000	本书不(0.480)

III. RESULTS AND DISCUSSION

3.1 Results Analysis

From the result of data analysis according chapter 2, we can see that we get a lot of phrases, some can distinguish whether good or poor expression keywords, but also have some irrelevant phrases. So we must get rid of noise. We put results in the second filter, remove noise, punctuation, emoticons, and so on irrelevant information. The secondary screening for the four results (praise and bad review of Japanese, praise and bad review of Chinese) as following tables.

Table 7The Second Filter for Get Useful Good Expression of Japanese

Number	The Second Filter
1	おもしろい(0.755)
2	何度も読(0.727)
3	読みたい(0.709)
…	…
267	面白いの(0.488)
268	好きなので(0.488)

Table 8The Second Filter for Get Useful Poor Expression of Japanese

Number	The Second Filter
1	残念(0.719)
2	退屈(0.637)
3	ではない(0.614)
…	…
74	読めない(0.488)
75	足りない(0.488)

Table 9The Second Filter for Get Useful Good Expression of Chinese

Number	Result of Data Analysis
1	很喜欢(0.797)
2	很快(0.758)
3	非常喜欢(0.720)
…	…
273	值得购买(0.480)
274	喜欢 爱不释手(0.480)

Table 10The Second Filter for Get Useful PoorExpression of Chinese

Number	Result of Data Analysis
1	字体(0.790)
2	小(0.750)
have3	盗版(0.732)
…	…
127	劣(0.480)
128	买到盗版(0.480)

3.2 Comprehensive Score Calculation

Using the results of second filter, we can get the comprehensive score by follow equation.

$$CS = \frac{\sum_{i=1}^{n} GE}{\sum_{i=1}^{n} GE + \sum_{i=1}^{m} PE}$$

$\sum_{i=1}^{n} GE$:means the total number of Good Expression, in the same way $\sum_{i=1}^{m}$ PEmeans the total number of Poor Expression. We want to get the composite score, so we need get the sum of the total number of good expression of Japanese and Chinese, in a similar way get the sum of the total number of poor expression of Japanese and Chinese. So we get the final score as following table.

Table 11The Comprehensive Score

Form	G.E.	P.E.	Total	Rate	Score
Japan	*268*	*75*	*343*	*0.78*	*3.9*
China	*274*	*128*	*402*	*0.68*	*3.4*
Total	*542*	*203*	*745*	*0.73*	*3.7*

There are many possible ways to rate items using the extracted results. Among this, we used the following score assuming that the smaller the number of noises in the extracted expressions is, the more coherent the set of reviews is.

CONCLUSIONS

For now users are free to write comments on the internet is increasingly popular, and the amount is also more and more hugeness. So the analysis for big data such as this is also more and more important and useful. However, natural language parsing is not particularly perfect so far.

For example, in this experiment, phrases in the database for feature extraction, sometimes the results are not very ideal. When dividing words sometimes is divided into a mess, cannot understand the meaning completely, or it can dismantle a complete word, lead to word meaning is not complete, or have polysemy. So it is hard to judge the word is good or poor.

REFERENCES

[1] The web site of Amazon.co.jp: http://www.amazon.co.jp/
[2] The web site of Amazon.cn: http://www.amazon.cn/
[3] The web site of Taobao: http://www.taobao.com/
[4] The web site of Rakuten: http://www.rakuten.co.jp/
[5] The web site of GURUNAVI: http://www.gnavi.co.jp/

[6] Minoru Yoshida, Kazuyuki Matsumoto, Qingmei Xiao, XielifuguliKeranmu, Kenji Kita, Hiroshi Nakagawa, "Extracting Corpus-Specific Strings by Using Suffix Arrays Enhanced with Longest Common Prefix", Information Retrieval Technology, 10th Asia Information Retrieval Societies Conference, AIRS 2014, Kuching, Malaysia, pp.360-370, December 3-5, 2014

[7] Tsuyoshi Katasho, "Analysis of evaluation data on the Internet" Mejiro University, The Graduate School of Psychology

[8] Miyuki Yamasawa, Hiroki Yoshimura, Hiroshi Masuichi, "Distinguishing usefulness among Amazon's review sentences" IPSJ SIG Technical Report 173, 15-20, 2006-05-19

[9] Yuki Uchida, Tomohiro Yoshikawa, Takeshi Furuhashi, EijiHirao, Hiroto Iguchi, "A Proposal of Text Mapping Method to Analyze Web User Reviews" 24th Fuzzy System Symposium, Osaka, September 3-5, 2008

[10] Bo Pang, Lillian Lee, "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales" ACL '05 Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, pp.115-124, Association for Computational Linguistics Stroudsburg, PA, USA ©2005

★★★