INTELLIGENT AUTOMATION & SOFT COMPUTING, 2017
http://dx.doi.org/10.1080/10798587.2016.1267244

Analyzing and Assessing Reviews on JD.Com

Jie Liu[a,b,c,d], Xiaodong Fu[d], Jin Liu[a,b,c] and Yunchuan Sun[a,e]

[a]State Key Laboratory of Software Engineering, School of Computer, Wuhan University, Wuhan, China; [b]Institute of Automation, Chinese Academy of Sciences, Guilin, China; [c]Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Beijing, China; [d]Information Engineering and Automation, Kunming University of Science and Technology, Beijing, China; [e]Business School, Beijing Normal University, Kunming, China

ABSTRACT

Reviews are contents written by users to express opinions on products or services. The information contained in reviews is valuable to users who are going to make decisions on products or services. However, there are numbers of reviews for popular products, and the quality of reviews is not always good. It's necessary to pick out reviews, which are in high quality from numbers of reviews to assist user in making decision. In this paper, we collected 21,501 reviews flagged as good from 499,253 products on JD.com. We observed the level of users is an important factor affects the quality of reviews, and users prefer to post short reviews containing the description of the quality and price of the product. We proposed a system to assess the quality of reviews automatically in this paper. We achieved that by applying SVM classification based on two kinds of features; reviews and reviewers that would help users find out high quality reviews and useful information from massive reviews. We evaluated our system on JD.com. The accuracy of our experiments for reviews quality assessing reached to 87.5 percent.

KEYWORDS

Online reviews; Quality of reviews; SVM classification; JD.com

1. Introduction

With the development of the Internet, more and more opportunities are provided for users to write reviews to share their experiences and opinions on products or services by online retailers like Amazon.com, movie sites like imdb.com and other websites. Users tend to read reviews to know the quality of products or get other useful information from reviews before they are choosing something, such as buying a product or picking a movie. The researches done by Chevalier and Mayzlin (2006) and Dabholkar (2006) have showed that the reviews played an important role on the decision-making. And websites are encouraged to make reviews be better to enhance user experience. The quality of reviews is also the measure of service quality.

A star rating is allowed to assign to show the satisfaction of users to products or services. The time when users are posting the review is always recorded by the website. Then reviews are ranked in the order of star ratings or the time on most websites. Unfortunately, the star ratings do not convey enough information for users. The content of reviews has to be read to get detailed information. The reviews of the best-selling products can be thousands. It's impossible for users to read all of these reviews. Some reviews post by the product manufacturer may speak highly of the product to promote the selling. Some reviews post by the competitor may say bad words to the product to suppress each other. These reviews will mislead users to make wrong decisions. Users are encouraged to vote the review written by others to confirm whether it is useful, then reviews are ranked in the order of useful votes, but participation is low.

Because the number of reviews is large and the quality of reviews is various, traditional review ranking and assessing methods could not deal with that well. We analyzed the reviews on JD.com to better understand features, which have influence on a reviews quality. By analyzing the meta-data, the level of users, which is the representation of user's reputation, is in line with the quality of reviews. Based on the analysis of reviews content, we found that the length of reviews is short and the number of words in reviews is not large, because reviews are mainly about the description of products features and the price. Then we proposed a system to assess reviews quality automatically using a machine learning approach. The main contributions of this system are: (1) a system for automatically assessing quality of reviews using SVM classification. We evaluated our system on the data-set collected from JD.com by assessing its reviews quality. (2) an analysis of different features helpfulness on assessing reviews quality. These features include reviews features and reviewers feature. This system would help users assess the quality of reviews and pick out reviews, which are in high quality from massive numbers of reviews and, which are in various qualities.

2. Related Work

There were more and more researches about assessing and ranking reviews. Reviews were assessed only according to text features by Yang, Yan, Qiu, and Bao (2015). Because authors believed that review quality was only related to review text, and they chose a user's manual label as the ground truth to compare with. Reviews on Amazon.com were assessed by Zheng, Zhu, and Lin (2013) based on not only review text features, but also reviewers feature. The assessing performance was improved when user's social feature was considered. They also found that the products type, such as digital or physical, has influence on

CONTACT Jin Liu jinliu@whu.edu.cn; Yunchuan Sun yunch@bnu.edu.cn

assessing reviews quality. Similar to Zheng et al. (2013), both review text features and reviewer's features were considered when Ngo-Ye and Sinha (2014) were assessing reviews quality. The reviewer's features were discussed in more detail. They also found that reviewer's features could improve the reviews quality assessing performance. Reviews of Yelp were analyzed by Bakhshi, Kanuparthy, and Shamma (2015). They also considered both reviews and reviewer's features, and they found that the quality of reviews was related to reviewers and reviews length. Book reviews on Amazon.com were assessed by Chua and Banerjee (2015) based on reviewer reputation, review rating and review depth, and they found an interesting result that the quality of reviews was positively related to reviewer reputation and review depth, but was negatively related to review rating. User reviews were analyzed by natural language processing and machine learning to get QoS (Quality of Service) in which users were interested (Liu, Kale, Wasani, Ding, & Yu, 2015). Reviews were assessed based on the combination of latent topics and star ratings (Krestel & Dokoohaki, 2015). Then the reviews were ranked to help users get information in quick. The sentiment contained in reviews was studied by two steps: Reviews text informativeness analysis and sentences structural study (Fang, Qian, Huang, & Zhu, 2014). Crowdsourcing was introduced to analyze the emergency events (Xu et al., 2016). Sentiment contained in reviews was analyzed mainly based on lexicon features, and they analyzed reviews from different domains, including books, hotels and electronics (Mao, Niu, Wang, Wang, & Qiu, 2015). A similar work was also mentioned (Shi, Zhan, & Li, 2015). They used CRF model for text fine-grained sentiment analysis. The nouns and adjectives in game reviews were analyzed (Zhu & Fang, 2015) to assist in better understanding user experiences and requirements of computer games. App reviews were analyzed to help understand user needs and experiences (Guzman, Aly, & Bruegge, 2015), which was similar to research (Zhu & Fang, 2015). They mined conflict opinions contained in reviews, which made developers know both the advantages and shortcomings of apps.

Recently, others investigated reviews summarization and extraction. Reviews summarization was achieved by exploiting user-labeled reviews helpfulness (Xiong & Litman, 2014). The helpfulness ratings were used to assist in scoring sentences. The problem of generating semantic concepts was studied (Xu, Liu, Mei, Hu, & Chen, 2014). A similar research work was represented (Hu et al., 2014). The knowledge contained in Chinese reviews was extracted by building knowledge space and retrieving knowledge (Zhao, Zhu, Jin, & Qiang, 2015). One interesting research was showed by Xu, Chen, and Santhanam (2015). They analyzed the relationship between reviews quality and reviews presentation formats (text, image and video), especially video. At the same time, the product type was also considered to analyze its influence on reviews quality. Another interesting study (Kwon, Kim, Duket, Catalán, & Yi, 2015) found that users preferred to read negative rated reviews to collect weakness of products. There were also researches about spam reviews detection. Spam reviews mainly of movies were detected based on user features and user-product relations features (Wu et al., 2015). Group spam reviews detection was studied by Xu and Zhang (2015).

3. Datasets

Our data was obtained by crawling JD.com to collect reviews. The crawler collected reviews starting from a set of the product IDs we set. According to Nelson (1970, 1974), products are divided into search products and experience products. There are reviews flagged as good by JD.com only on search products. It won't be labor intensive if we choose these reviews to investigate. So, we crawled search products reviews on JD.com. For each product review, we collected the following data: The tag assigned to the product, the pros and cons of the product, the content of the product review, the star rating assigned to the product, the number of people who think the review is helpful and reply to the review, the name and the category of the product and the name, the level of the user.

The crawler we used is Heritrix. The reason why we choose it is that we can get the whole web page of reviews on JD.com stored in HTML format. Then we got reviews by HTML Parser, a library used to parse HTML files. Finally, reviews were stored in database for analyzing later.

We collected reviews from 499,253 search products, of which 6,022 have 21,501 reviews flagged as good by JD.com and post by 14,250 users. We will analyze the meta-data and the content of reviews in the following section. The meta-data could be gotten directly, because it's independent of reviews text. In order to get some statistic data from the content of reviews, we will separate reviews into separate words by ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System). The source codes can be downloaded from http://ictclas.nlpir.org. We also analyzed the sentiment of reviews to get a better understand of reviews based on a lexical resource which contains sentiment annotation.

Table 1. Star Ratings Distribution.

Star	Frequency	Percent	Cumulative Percent
1	7	0.0	0.0
2	8	0.0	0.1
3	80	0.4	0.4
4	3327	15.5	15.9
5	18079	84.1	100.0
Total	21501	100	

4. Analysis of Reviews

4.1. Meta-data Analysis

Meta-data are independent of the text of reviews. The distribution of star ratings assigned to the product by users is showed in Table 1.

We discover that most reviews flagged as good by JD.com are assigned 4 or 5 stars. It means that most users are satisfied with the product they have bought. We will analyze the content of reviews later.

The distribution of helpful votes and reply numbers is showed in Fig. 1. We observe that more than 60 percent of reviews get no helpful vote or reply from users. Reviews whose reply numbers are 1 to 4 are less than 30 percent. It means that users are inactive to vote on or reply to reviews. Other disadvantages of helpful votes ranking were mentioned by Liu, Cao, Lin, Huang, and Zhou (2007), such as imbalance vote bias, winner circle bias and early bird bias.

We believe that there is a correlation between useful votes and reply numbers. The most familiar measure of dependence between two quantities is the Pearson's correlation coefficient. The Pearson product-moment correlation coefficient is a measure of the linear dependence between two variables X and Y, giving a value between + 1 and −1 inclusive, where 1 is total

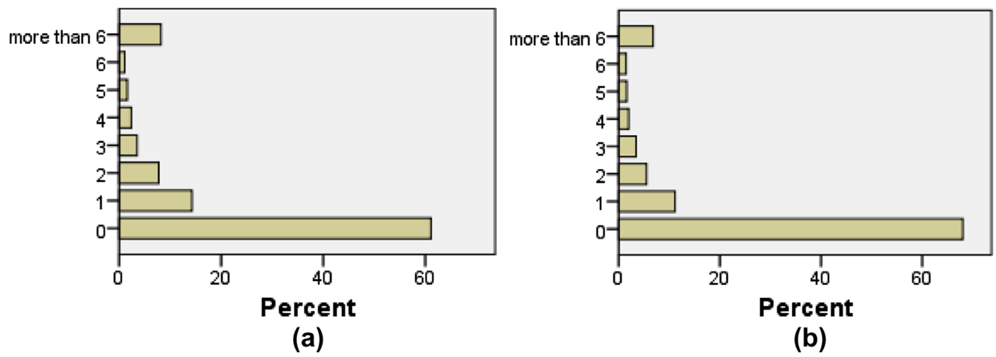


Figure 1. (a) Distribution of Useful Votes. (b) Distribution of Reply Number.

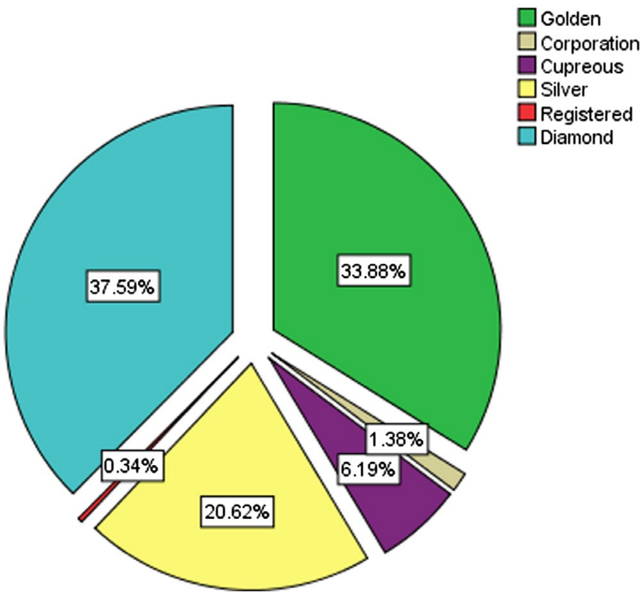


Figure 2. Distribution of User Levels.

positive correlation, 0 is no correlation, and −1 is total negative correlation. We test the data and result is showed in Table 2. The result confirms that the useful votes are correlated to the reply numbers.

Table 2. Correlation Test on Useful Votes and Reply Numbers.

		Reply	Useful
Reply	Pearson Correlation	1	0.686
	Sig. (2-tailed)		0.000
	N	21501	21501
Useful	Pearson Correlation	0.686	1
	Sig. (2-tailed)	0.000	
	N	21501	21501

Table 3. Textual Features of Reviews.

Feature Type	Feature Name	Feature Description
Statistic	Sentences Number	The number of sentences in each review
Statistic	Words Number	The number of words in each review
Syntactic	Nouns Percent	The percent of nouns
Syntactic	Adjectives Percent	The percent of adjectives
Syntactic	Verbs Percent	The percent of verbs
Syntactic	Adverbs Percent	The percent of adverbs
Statistic	Strings Percent	The percent of words which are not Chinese

The distribution of user levels is showed in Fig. 2. The user levels from high to low are diamond, golden, silver, cupreous and registered on JD.com. We discover that most users whose review is flagged as good are in high level. It means that experienced users will provide reviews in high quality.

4.2. Content Analysis

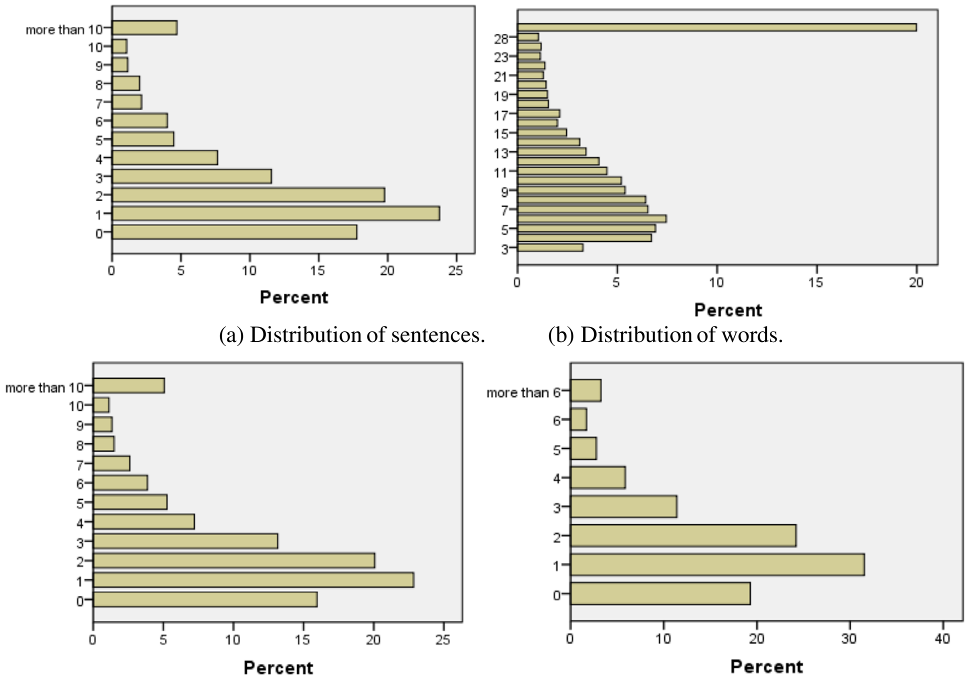
On one hand, we can get to know user's overall attitude to the product from star ratings assigned to the product. On the other hand, we can infer the usefulness of reviews from helpful votes and reply numbers. However, we can get nothing detailed about the product from star ratings, helpful votes or reply numbers. In order to know more detailed information about the product, we have to dig the content of reviews. We will collect and analyze the textual features of reviews. The detailed features description is showed in Table 3.

The distribution of sentences and words numbers in each review is showed in Fig. 3. In Fig. 3(a) the number of sentences is zero, because there is no punctuation in some reviews. We find that most reviews contain 1 to 3 sentences. That may be, because users who have bought the product want to share their opinions in a few sentences and users who are going to buy the product want to read few sentences to aid decisions.

The number of words in each review showed in Fig. 3(b) is the number of words excluding prepositions, conjunctions and pronouns. We observe that there are about 4 to 14 useful words for most reviews. It means that not only the number of sentences but also the number of words are few in reviews. It confirms that users want to provide or get information about the product in short reviews again.

The distribution of nouns and adjectives numbers in reviews is also showed in Fig. 3. We notice that most reviews contain 1 to 3 nouns and about 15 percent reviews with no noun. There are 1 to 2 adjectives in reviews and about 20 percent reviews without adjective.

In order to know what nouns and adjectives are in reviews, we collected high frequency nouns and adjectives. They are listed in Table 4. We discover that the quality and the price are the most concerned factors for users. Users also pay attention to the function and other characteristics of the product, such as speed, color, size and volume. The logistics and the package are also mentioned because; the product has to be delivered from the warehouse to users. It's part of user experience of buying products. Another interesting finding is that users like to mention their relatives or friends comments in reviews. As for adjectives, they together with nouns are used to describe the product features.



(a) Distribution of sentences. (b) Distribution of words.

(c) Distribution of nouns. (d) Distribution of adjectives

Figure 3. The Distribution of Nouns and Adjective Numbers in Reviews

In order to know what verbs and adverbs are in reviews, we also collect high frequency verbs and adverbs. The analysis result of verbs, adverbs and String Percent is omitted as the limited pages.

The sentiment is contained in reviews to show users attitudes, feelings and opinions. We analyzed the sentiment of reviews based on HowNet, which is a lexical resource containing words sentiment. It can be downloaded from http://www.keenage.com. There are mainly 4 kinds of words in HowNet, including words describing positive feelings, negative feelings, positive comments and negative comments. We analyze reviews words over 2 axes; positive sentimental words (positive feelings and comments) and negative sentimental words (negative feelings and comments). We discovered that there are 1 to 5 positive words in reviews to present users positive feelings and comments. It's in line with the star ratings of reviews, which were analyzed above. Positive sentimental words are contained in high star rated reviews to express user's feelings.

5. Assess Reviews Quality

5.1. Features Indicating Quality

In order to analyze the helpfulness of different features on assessing reviews quality, we evaluated our system with lots of features divided into two kinds; reviews features and reviewers feature.

Star Ratings (SR): A number assigned to the product to express the users overall attitude towards to the product. The number ranges from 1 to 5. Number 1 means the negative attitude towards to the product. On the contrary, number 5 means the positive attitude towards to the product. In order to scale the feature between [0, 1], we define 20 percent of the original number to be the star ratings SR.

Helpful votes and Reply Numbers (HRN): Helpful votes are numbers of people who think the review is helpful. It's a measurement of reviews quality, which assessed by users manually. And reply numbers are numbers of people who reply to the review. Usually, users reply to reviews, which they are interested in and useful, then it also could be regarded as a measurement of reviews quality which assessed by users manually. We define the total number of a review's helpful votes and reply numbers to be the helpful votes and reply numbers HRN, because helpful votes are correlated to reply numbers and both of them are not large.

Words numbers and Sentences Numbers (WSN): Words numbers are numbers of words contained in the review. Sentences numbers are numbers of sentences in a review. The more words and sentences are contained in reviews, the more information can be gotten by users.

These features are statistic or syntactic features, which makes some of them be correlated to others. Then we did a principal component analysis on these features to make new components be independent to each other. Principal component analysis is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

Before the principal component analysis, we perform the KMO (Kaiser-Meyer-Olkin) and Bartlett's test to check whether it's suitable for the analysis. The KMO value is close to 1 when the simple correlation coefficient is larger than the partial

Table 4. Top-25 High Frequency Nouns and Adjectives.

Nouns				
Goods	Quality	Price	Feeling	Effect
Speed	Appearance	Brand	Certified goods	Workmanship
Function	Products	Problem	Friends	Mobile phone
Voice	Package	Taste	Color	Time
Mouse	Disadvantage	Wife	Earphone	Card
Adjectives				
Good	Nice	Big	Small	Fast
Cheap	Beautiful	Many	High	Convenient
Practical	Comfortable	Good looking	Fit	Expensive
Cute	Exquisite	Long	Tough	Clear
General	Simple	Old	Light	Easy

correlation coefficient. It means that there is a strong correlation between variables. The primary hypothesis of Bartlett's test of sphericity is that the correlation coefficient matrix is the identity matrix. If the Sig. value is less than the significant level, we reject the primary hypothesis, indicating that there is a strong correlation between variables. The result is showed in Table 5. The KMO value is 0.7 close to 1 and the Sig. value in Bartlett's test of sphericity is less than 0.05, which means the correlation between variables is strong and it's suitable for principal component analysis.

The result was showed in Table 6. We can infer that there are two principal components. The first component is different words and sentences numbers component, and the second component is strings numbers component. In fact, more than 95 percent reviews have no strings in our data-set, which means we can ignore the second component. As for the first component, different words numbers and sentences numbers are correlated to each other. Then we let words numbers stand for other words numbers, because other words numbers are contained in words numbers. Above all, we define the total numbers of words numbers and sentences numbers to be the words numbers and sentences numbers WSN.

Sentimental Words Numbers (SWN): Sentimental words numbers are numbers of positive and negative sentiment words contained in reviews, and it's much more precise and valuable than star ratings. Then we define the total numbers of positive and negative sentiment words numbers to be sentimental words numbers SWN.

Description Words Numbers (DWN): Description Words Numbers are numbers of words, which describe the quality of products and other things users concerned. It means the more those words are contained in reviews, the more useful reviews are. Reviews, which contain lots of useful information, are in high quality. We define the numbers of words, which describe products quality and other user concerned to be the description words numbers DWN.

In order to analyze the description words contained in reviews, we collected nouns, which are separated from pros and cons of products. We also collected high frequent nouns contained in reviews content. We assumed that both the nouns contained in pros and cons and high frequent nouns contained in reviews content are words, which describe the product quality and other things users concerned, then the description words set was generated.

Table 5. KMO and Bartlett's Test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.700
Bartlett's Test of Sphericity	Approx. Chi-Square	263401.641
	df	153
	Sig.	0.000

Table 6. Component Matrix.

	Component	
	1	2
Words Numbers	0.335	−0.154
Nouns Numbers	0.905	−0.019
Adjectives Numbers	0.787	0.033
Verbs Numbers	0.926	−0.037
Adverbs Numbers	0.846	0.068
Strings Numbers	0.077	0.961
Sentences Numbers	0.990	−0.012

Features for assessing reviews quality are based on not only reviews themselves, but also reviewers who post them. Reviewer Feature is the feature, which is related to reviewers.

User Levels (UL): The level of users is raised as they buying products and posting reviews. The higher the user level is, the more experienced the user is. We believe reviews written by experienced users are in high quality because experienced users are good at sharing their opinions and experiences. As there are five user levels, we make the user levels become numbers 1to 5. And we define 20 percent of user level numbers to be user levels UL. It means the more the UL close to 1; the higher the user is leveled.

5.2. Experiments and Results

We evaluated our system by applying SVM classification based on features defined above on JD.com. Below, we described the evaluation procedure and experimental results.

As there are reviews, which are flagged as good by JD.com, we assumed that reviews, which are flagged as good, are in high quality. On the contrary, reviews, which are not flagged as good are assumed to be in low quality. Therefore, assessing reviews quality on JD.com was defined to be a binary classification task. We compiled a feature vector for each review according to features definition mentioned in Section 5. The feature values, which were not normalized by definition, were scaled between [0, 1]. In order to classify the reviews, we used a C-SVM with a radial basis function (RBF) kernel as implemented by LibSVM. We tuned the RBF kernel parameters C (the penalty parameter) and γ (the kernel width hyperparameter) performing full grid search following Hsu, Chang, and Lin (2003). We performed ten-fold cross validation for performance evaluation. We experimented with various combinations of feature sets.

The average cross validation accuracy for all combinations of feature was showed in Table 7. We can infer that the WSN feature performs best. The WSN feature contains information, which is related to words and sentence numbers. It shows the useful information richness from which users can get. The SR, SWN and UL features also perform well in assessing reviews quality. The HRN feature performs worse than other features. The HRN feature is about user's feedback, which should have been important. However, there is not enough feedback from users for the product, which makes it perform worse. The performance gets improved when different reviews features are combined to assess reviews quality. However, when all reviews features are combined, it performs worse than meta-data or content-data reviews features are combined. The reason may be that the attitude of review content is not always same as the attitude of review meta-data. The performance improves significantly when both reviews and reviewers features are combined for reviews quality assessing.

Table 7. Accuracy with Different Feature Sets.

Feature Combinations	RBF Kernel
SR	82.6351%
HRN	65.0103%
WSN	83.4258%
SWN	82.3464%
DWN	73.4523%
UL	81.3245%
SR+TP+HRN (meta-data of reviews)	84.5672%
WSN+SWN+DWN (content-data of reviews)	85.5278%
SR+TP+HRN+WSN+SWN+DWN (reviews feature)	83.6955%
SR+TP+HRN+WSN+SWN+DWN+UL (reviews and reviewers feature)	87.5343%

6. Conclusion

Reviews are widely supported by different kinds of websites to enhance user's online experience. However, reviews vary in quality and can be in large amount for popular products. Thus, assessing and ranking numbers of reviews turn out to be necessary.

We analyzed reviews flagged as good on JD.com to better understand how different factors affect the quality of reviews. We observed that the quality of reviews is related to the level of users. We also observed that the number of both words and sentences is few, making reviews be concise. We found the features and the price of products are the focus of reviews for search products.

We also proposed a system for automatically assessing reviews quality to solve that problem. We achieved that by applying SVM classification based on several reviews and reviewers features. We trained our system and assessed reviews quality on JD.com. The accuracy reached to 87.5 percent for our system evaluation. We also analyzed the usefulness of different features on assessing review quality. The performance of combining reviews and reviewers features to assess review quality is significantly better than those of other features combination.

Newly post reviews, which contain more useful information in recent are much more helpful for users. Then reviews, which were labeled in high quality by our system, ought to be listed in chronological order on JD.com.

Our system, which automatically assesses review quality, could help users deal with large numbers of reviews, which are in various qualities. It could pick out high quality reviews and assist user in making decisions.

As further future work, we plan to experiment with all reviews belonging to one category, including reviews, which are not flagged as good. We will analyze the reviews especially few stars rated to investigate the characteristic of them. We will also perform an experiment on reviews of experience products to find the differences between search products and experience products.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was partially supported by the National Natural Science Foundation of China [grant number. 61572374, U163620068, U1135005, 71161015, 61462056] and Opening fund of State Key Laboratory of Software Engineering (SKLSE2012-09-17) Guangxi Key Laboratory of Trusted Software (No.kx201607) and the Academic Team Building Plan for Young Scholars from Wuhan University.

Notes on contributors

Jie Liu received a BS degree in Computer Science and Technology from Kunming University of Science and Technology in 2013, and an MS degree in ComputerTechnology from Kunming University of Science and Technology in 2016. He is currently a Ph.D. student in State Key Lab. of Software Engineering in Wuhan University, China. His research interests include software data analysis.

Xiaodong Fu is a professor in the Faculty of Information Engineering and Automation, Kunming University of Science and Technology, China. He received his MS in Computer Science in 2000 and his Ph.D. in Management Science in 2008. He has authored more than 20 publications on services computing topics. He is a member of the IEEE Computer Society and the ACM. His current research interests include services computing and intelligent decision.

Jin Liu received a Ph.D. degree in Computer Science from Wuhan University, China, in 2005. He is now a professor in State Key Laboratory of Software Engineering, Computer School, Wuhan University. He has published more than 50 papers in well-known conferences and journals, such as ICSE, SANER, TSE, FGCS, COMPUT J, CCPE, etc. His current research interests include software service engineering and software repository mining.

Yunchuan Sun received a Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Science, Beijing, China, in 2009. He is now an associate professor at Beijing Normal University, Beijing, China. He has successfully organized several special issues in some international journals like Springer Personal and Ubiquitous Computing, Automatika, Elsevier Journal of Networks Computer Applications, and so forth. His current research interests include Internet of Things/Web of Things, semantic link network, knowledge representation, cyber-physical-socio intelligence, and business models for the Internet of Things.

References

Bakhshi, S., Kanuparthy, P., & Shamma, D.A. (2015, February). Understanding online reviews: Funny, cool or useful? In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 1270–1276). ACM.

Chevalier, J.A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43* 345–354.

Chua, A.Y. & Banerjee, S. (2015). Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. *Journal of the Association for Information Science and Technology, 66*, 354–362.

Dabholkar, P.A. (2006). Factors influencing consumer choice of a "rating Web site": An experimental investigation of an online interactive decision aid. *The Journal of Marketing Theory and Practice, 14*, 259–273.

Fang, L., Qian, Q., Huang, M., & Zhu, X. (2014, November). Ranking sentiment explanations for review summarization using dual decomposition. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (pp. 1931–1934). ACM.

Guzman, E., Aly, O., & Bruegge, B. (2015, October). Retrieving diverse opinions from app reviews. In Empirical Software Engineering and Measurement (ESEM), 2015 ACM/IEEE International Symposium on (pp. 1–10). IEEE.

Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). A practical guide to support vector classification (pp. 1–16).

Hu, C., Xu, Z., Liu, Y., Mei, L., Chen, L., & Luo, X. (2014). Semantic link network-based model for organizing multimedia big data. *IEEE Transactions on Emerging Topics in Computing, 2*, 376–387.

Krestel, R. & Dokoohaki, N. (2015). Diversifying customer review rankings. *Neural Networks, 66*, 36–45.

Kwon, B.C., Kim, S.H., Duket, T., Catalán, A., & Yi, J.S. (2015). Do People really experience information overload while reading online reviews? *International Journal of Human-Computer Interaction, 31*, 959–973.

Liu, J., Cao, Y., Lin, C.Y., Huang, Y., & Zhou, M. (2007, June). Low-quality product review detection in opinion summarization. In EMNLP-CoNLL (pp. 334–342).

Liu, X., Kale, A., Wasani, J., Ding, C., & Yu, Q. (2015, June). Extracting, ranking, and evaluating quality features of web services through user review sentiment analysis. In Web services (ICWS), 2015 IEEE International Conference on (pp. 153–160). IEEE.

Mao, K., Niu, J., Wang, X., Wang, L., & Qiu, M. (2015, August). Cross-domain sentiment analysis of product reviews by combining lexicon-based and learn-based techniques. In High Performance Computing and Communications (HPCC) (pp. 351–356). IEEE.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy, 78*, 311–329.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy, 82*, 729–754.

Ngo-Ye, T.L. & Sinha, A.P. (2014). The influence of reviewer engagement characteristics on online review helpfulness: A text regression model. *Decision Support Systems, 61*, 47–58.

Shi, H., Zhan, W., & Li, X. (2015). A supervised fine-grained sentiment analysis system for online reviews. *Intelligent Automation & Soft Computing, 21*, 589–605.

Wu, Z., Wang, Y., Wang, Y., Wu, J., Cao, J., & Zhang, L. (2015, November). Spammers detection from product reviews: A hybrid model. In Data Mining (ICDM), 2015 IEEE International Conference on (pp. 1039–1044). IEEE.

Xiong, W., & Litman, D.J. (2014). Empirical analysis of exploiting review helpfulness for extractive summarization of online reviews. In COLING (pp. 1985–1995).

Xu, C., & Zhang, J. Towards. (2015, November). Collusive fraud detection in online reviews. In Data Mining (ICDM), 2015 IEEE International Conference on (pp. 1051–1056). IEEE.

Xu, Z., Liu, Y., Mei, L., Hu, C., & Chen, L. (2014). Generating temporal semantic context of concepts using web search engines. *Journal of Network and Computer Applications, 43*, 42–55.

Xu, P., Chen, L., & Santhanam, R. (2015). Will video be the next generation of e-commerce product reviews? Presentation format and the role of product type. *Decision Support Systems, 73*, 85–96.

Xu, Z., Liu, Y., Yen, N., Mei, L., Luo, X., Wei, X., & Hu, C. (2016). Crowdsourcing based description of urban emergency events using social media big data. *IEEE Transactions on Cloud Computing*. doi: 10.1109/TCC.2016.2517638

Yang, Y., Yan, Y., Qiu, M., & Bao, F.S. (2015). Semantic analysis and helpfulness prediction of text for online product reviews. In *ACL, 2*, 38–44).

Zhao, F., Zhu, H., Jin, H., & Qiang, W. (2015). A skip-gram-based framework to extract knowledge from chinese reviews in cloud environment. *Mobile Networks and Applications, 20*, 363–369.

Zheng, X., Zhu, S., & Lin, Z. (2013). Capturing the essence of word-of-mouth for social commerce: Assessing the quality of online e-commerce reviews by a semi-supervised approach. *Decision Support Systems, 56*, 211–222.

Zhu, M., & Fang, X. (2015). A lexical analysis of nouns and adjectives from online game reviews. In Human-computer interaction: Interaction technologies (pp. 670–680). Springer International Publishing.