ARTICLE IN PRESS
JBR-09222; No of Pages 10
Journal of Business Research xxx (2016) xxx–xxx
Contents lists available at ScienceDirect
Journal of Business Research

Predicting the "helpfulness" of online consumer reviews

Jyoti Prakash Singh [a,*], Seda Irani [b], Nripendra P. Rana [b], Yogesh K. Dwivedi [b], Sunil Saumya [a], Pradeep Kumar Roy [a]

[a] National Institute of Technology Patna, India
[b] School of Management, Swansea University Bay Campus, Swansea SA1 8EN, UK

ARTICLE INFO

Available online xxxx

Keywords:
Online user reviews
Helpfulness
Product features
Text mining
Product ranking

ABSTRACT

Online shopping is increasingly becoming people's first choice when shopping, as it is very convenient to choose products based on their reviews. Even for moderately popular products, there are thousands of reviews constantly being posted on e-commerce sites. Such a large volume of data constantly being generated can be considered as a big data challenge for both online businesses and consumers. That makes it difficult for buyers to go through all the reviews to make purchase decisions. In this research, we have developed models based on machine learning that can predict the helpfulness of the consumer reviews using several textual features such as polarity, subjectivity, entropy, and reading ease. The model will automatically assign helpfulness values to an initial review as soon as it is posted on the website so that the review gets a fair chance of being viewed by other buyers. The results of this study will help buyers to write better reviews and thereby assist other buyers in making their purchase decisions, as well as help businesses to improve their websites.

1. Introduction

Online user reviews have become today's word of mouth for the current generation of customers and business managers. Hence, understanding the role of online user reviews in e-commerce has attracted the attention of both academics and practitioners around the world (Duan et al., 2008a; Forman et al., 2008; Li & Hitt, 2008). Online user reviews influence both product sales via consumer decision-making and quality improvement via business firms (Duan et al., 2008b). With the rapid penetration of the Internet into society and e-commerce business, the amount of user reviews is increasing rapidly. Such a large volume of data constantly being generated can be considered as a big data challenge (Chen et al., 2013) for both online businesses and consumers.

Online reviews in the form of unstructured big data have both negative and positive impacts on consumers. First of all, the consumers are getting the real experience of their peers about a product, which helps them make intelligent decisions about the product or service. But at the same time, the large amount of reviews can cause information overload. In some cases, it is not possible for any customer to go through all the reviews and then make decisions. For example, an average-ranked book on Amazon.com can have more than several hundred reviews, whereas for a popular product such as the latest mobile phone, the number of reviews can be in the thousands. In such situations, it is virtually impossible for consumers to read all the reviews before making purchase decisions, especially for products that have been reviewed by hundreds and sometimes thousands of customers with their inconsistent opinions. Chen et al. (2013) classify such a large volume of unstructured data (i.e., big data) in the form of user generated content, which clearly poses a big data management challenge. It would be more useful for customers if they had a higher level of visibility of helpful user reviews that reflect the overview of the product or services. That would encourage websites to evaluate the helpfulness of reviews written by other users. This is traditionally done by asking a simple question such as "Was this review helpful to you?" and putting "thumbs up" and "thumbs down" buttons.

The usefulness of reviews is generally assessed and their rank assigned by websites based on the helpfulness voting. For example, by default, user reviews are sorted by their helpfulness on Amazon.com. This is very useful to consumers as they can see the most helpful reviews on top. This also makes the website more user-friendly and hence attracts more consumers. Reviews that are perceived as helpful to customers bring considerable benefits to companies, including increased sales (Chevalier & Mayzlin, 2006; Clemons et al., 2006). It is estimated that this simple question "Was this review helpful to you?" brings in about $2.7 billion additional revenue to Amazon.com (Spool, 2009). However, the helpfulness voting is not a silver bullet and does not solve all problems. The reasons for this are (i) very few user reviews receive helpfulness votes, and without helpfulness votes, the helpfulness voting mechanism does not work effectively; and (ii) recent

* Corresponding author.
E-mail addresses: jps@nitp.ac.in (J.P. Singh), sedatuna@hotmail.com (S. Irani), nrananp@gmail.com (N.P. Rana), ykdwivedi@gmail.com (Y.K. Dwivedi), sunils.cse15@nitp.ac.in (S. Saumya), pradeep.cse15@nitp.ac.in (P. Kumar Roy).

http://dx.doi.org/10.1016/j.jbusres.2016.08.008
0148-2963/© 2016 Published by Elsevier Inc.

reviews have yet to get votes, and hence their helpfulness cannot be decided. Given that reviews are posted so rapidly, the useful reviews are likely to get buried beneath less useful reviews before attracting helpfulness votes.

Since most helpful reviews get higher exposure to consumers, they normally become more helpful whereas less helpful reviews become less attractive to consumers due to less exposure. As a result, the reviews with fewer helpfulness votes' are ignored by customers whereas reviews with more helpfulness votes get more visibility and readership. The result of this is that consumer decision-making is mostly influenced by the helpfulness votes and is skewed without considering when the review was posted and what the context was.

Although online reviews have helped consumers in deciding the pros and cons of different products, which ultimately helps in deciding the best product for an individual's needs, they introduce a challenge for consumers to analyze this huge amount of data because of its volume, variety, and velocity. Review data is getting big day by day, at a very fast pace. Some users have started putting up pictures/images of the product to make their reviews more attractive and user-friendly. Hence, the review dataset may be seen as a big data analytics problem. It is interesting for businesses to dig into those review data to get insight about their products. Chen et al. (2013) suggested that “machine learning is possibly a feasible way to improve traditional data reduction techniques to process or even pre-process big data”. However, such an approach (i.e., machine learning) is yet to be tried and tested for enhancing the value of online user reviews.

Considering the discussion presented above about using a machine learning approach for big data analysis, this research investigates the helpfulness of online consumer reviews. We propose a system where the website itself would be able to perform the initial evaluation of the review using the model put forward by this research. That would help in prioritizing the better reviews in an appropriate order so that they can be viewed by other users. This will mitigate the Matthew effect, which implies that the top reviews gain more helpfulness votes as they are more visible and the lower reviews get fewer helpfulness votes as they are buried inside the review heap (Wan, 2015). A recent study by *brightlocal.com* (BrightLocal, 2016) suggests that 87% of buyers read 10 or less than 10 reviews before trusting a business. Hence, if a review is really helpful, but it is not put in the top 10 list, then it will lose its purpose. The proposed approach ensures that this helpful review is ranked appropriately in the review.

We chose the Indian context for this study because e-commerce businesses have just started flourishing here. People have started buying online and writing reviews for the related products. The reviews on Indian e-commerce sites are very different from the reviews from other parts of the world where e-commerce is very popular. Reviews written by Indian buyers are mainly in English, but they contain some Hindi text (written in English script only) as well. Some of the most widely used Hindi words, such as *bahut achha*, *bakbas*, and *pesa wasool*, are found in a number of reviews (Singh et al., 2015).

The rest of the article is structured as follows: Section 2 reviews the related literature and is followed by the methods of data collection and analysis in Section 3. The results are presented in Section 4. Next, Sections 5 and 6 report the discussion and conclusions respectively.

2. Literature review

Various research studies have been done on helpfulness of reviews (Kim et al., 2006). Some researchers have used regression techniques to show the most helpful reviews while others have used neural networks. Ghose and Ipeirotis (2006) proposed two ranking mechanisms for ranking product reviews: a consumer-oriented ranking mechanism that ranks the reviews according to their expected helpfulness, and a manufacturer-oriented ranking mechanism that ranks them according to their expected effect on sales. They used econometric analysis with text mining to make their ranking work. They found the reviews that tend to include a mix of subjective and objective elements are considered more informative (or helpful) by the users.

Liu et al. (2007) considered the product review helpfulness problem as a binary classification problem. They performed manual annotation to check review comments on many products using ‘favorable’ and ‘unfavorable’ as the classification targets but they did not use the original helpful feedback for their study. However, (Liu et al., 2008; Otterbacher, 2009) proposed a model for predicting the helpfulness of reviews using many features, such as length of reviews and the writing style of the reviewers. Out of these, the three most important factors named and used for prediction are the reviewer's expertise, the writing style of the reviewer, and the timeliness of the review. Radial basis functions are used to model expertise and writing style. The training data is taken from a tally present in the reviews itself, called a helpfulness vote.

Forman et al. (2008) suggested that in the context of an online community, the reviewer's disclosure of identity-descriptive information is used by consumers to supplement or replace product information when making purchase decisions and evaluating the helpfulness of online reviews. They found that the online community member's rate reviews containing identity-descriptive information more positively. Danescu-Niculescu-Mizil et al. (2009) found a new correlation between proportion of helpful votes of reviews and deviation of the review ratings from the average ratings of products. They report that helpful votes are consistent with average ratings. Mudambi and Schuff (2010) undertook an analysis on reviews collected from Amazon.com to determine which properties of a review make it useful for the customers. Three hypotheses were formulated and verified. On analyzing the hypotheses, they found that the impact of review extremity was dependent on product type. Products were grouped into two categories: (i) search product and (ii) experience product. A search product is one that customers can easily acquire information about concerning its quality before interacting with the product directly and where it does not require much customer involvement to evaluate the key quality attributes of the product, which are objective and easy to compare. An experience product is one that customers have to interact directly with to acquire information about its quality. With an experience product, the customer's involvement is required in order to evaluate the level of quality as key attributes are subjective or difficult to compare.

For experience products, extreme reviews were found to be less helpful as compared with moderate reviews. However, for search products, extreme reviews were more helpful than moderate ones. The review length also had an impact on helpfulness but was dependent on product type. For search products, review length had a greater positive impact as compared to experience products. So, it was concluded that helpfulness depended on star rating and review length but was also dependent on product type.

Ghose and Ipeirotis (2011) analyzed many characteristics of review texts, such as spelling errors, readability, and subjectivity, and examined their impact on sales. Linguistic correctness was found to be a vital factor affecting sales. There is a feeling that reviews of medium length with fewer spelling errors are more helpful to naive buyers as compared to reviews that are very short or very long and have spelling errors.

To analyze the impact of various characteristics of online user reviews, (Cao et al., 2011) used text mining on the helpfulness as indicated by the number of votes a reviewer receives. They found that helpfulness is more affected by semantic features as compared to other features of reviews. They also found that reviews expressing extreme opinions are more impactful than reviews with neutral or mixed opinions.

Korfiatis et al. (2012) explored the interplay between online review helpfulness, rating score, and the qualitative characteristics of the review text as measured by readability tests. They constructed a theoretical model based on three elements: conformity, understandability, and expressiveness. They investigated the directional relationship between the qualitative characteristics of the review text, review helpfulness, and the impact of review helpfulness on the review score. To validate

Please cite this article as: Singh, J.P., et al., Predicting the “helpfulness” of online consumer reviews, *Journal of Business Research* (2016), http://dx.doi.org/10.1016/j.jbusres.2016.08.008