Decision Support Systems 50 (2011) 755–768

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

# Quality evaluation of product reviews using an information quality framework

Chien Chin Chen *, You-De Tseng

*Department of Information Management, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan, ROC*

ARTICLE INFO

Available online 27 August 2010

*Keywords:*
Text mining
Classification
Opinion mining
Opinion retrieval

ABSTRACT

The ubiquity of Web2.0 makes the Web an invaluable source of business information. For instance, product reviews composed collaboratively by many independent Internet reviewers can help consumers make purchase decisions and enable enterprises to improve their business strategies. As the number of reviews is increasing exponentially, opinion mining and retrieval techniques are needed to identify important reviews and opinions to answer users' queries. Most opinion mining and retrieval approaches try to extract sentimental or bipolar expressions from a large volume of reviews. However, the process often ignores the quality of each review and may retrieve useless or even noisy documents. In this paper, we propose a method for evaluating the quality of information in product reviews. We treat the evaluation of review quality as a classification problem and employ an effective information quality framework to extract representative review features. Experiments based on an expert-composed data corpus demonstrate that the proposed method outperforms state-of-the-art approaches significantly.



## 1. Introduction

When making customer-related decisions, business managers frequently ask questions like “How do people feel?” “What opinions do people have?” To answer such questions, enterprises usually need to conduct laborious surveys in order to gather enough opinions for analysis. However, with the advent of Web2.0, many online collaborative tools, e.g., weblogs and discussion forums, are being developed to allow Internet users to express opinions and share valuable knowledge. One consequence is that the Web has become an invaluable source of information for business enterprises. Roed [25] observes that Internet users are often willing to divulge personal information and are forthcoming in presenting their personal viewpoints honestly. This kind of behavior has an indirect *word-of-mouth*[1] effect on marketing because users' opinions posted on the Web have a huge impact on consumer decisions [3]. Many e-commerce companies and websites, such as Amazon,[2] are aware of the word-of-mouth effect and offer users a platform to post their product reviews. However, as the number of reviews is growing exponentially, companies and customers are finding it increasingly difficult to find desired information. To alleviate this information overload problem, *opinion mining* and *retrieval* techniques [5,11,15,20,24] have been devised to extract and retrieve meaningful opinions from reviews.

A major task of opinion mining and retrieval involves identifying sentimental (or bipolar) text units in review documents. A text unit can be a word, a sentence, a paragraph, or even the whole document, depending on the granularity of opinion mining and retrieval. While many opinion mining and retrieval approaches try to identify and analyze opinions extracted from reviews, comparatively few works consider the quality of reviews. As Web2.0 encourages knowledge sharing, there are no constraints on review writing; hence, the quality of reviews varies enormously. For example, many reviews posted on Amazon simply contain emotional expressions, such as “*I just don't like this camera.*” Since such reviews lack constructive expressions, they should not be included in the opinion mining and retrieval process. Some websites do consider the quality of reviews, and provide a review quality evaluation mechanism that ranks reviews based on votes submitted by readers. Fig. 1 shows a review quality evaluation on Amazon, where 121 out of 128 users thought the review was helpful. However, even under a voting mechanism, the quality evaluations are still affected by *imbalanced vote bias*, *winner circle bias*, and *early bird bias*, which make the system impracticable [21]. Thus, there is an urgent need for effective quality evaluation mechanisms to help opinion mining and retrieval algorithms identify informative reviews.

In this paper, we propose a method for evaluating the quality of information in product reviews. We treat the quality evaluation of product reviews as a classification problem and employ a multiclass support vector machine (multiclass SVM) [31] model to categorize reviews. In addition, we adopt a mature information quality (IQ) framework, which has been widely used in many domains over the last twenty years, to define review features for classification. Rowley and Hartley [26] define raw data as unprocessed observations,

* Corresponding author.
*E-mail addresses:* paton@im.ntu.edu.tw (C.C. Chen), r96725044@ntu.edu.tw (Y.-D. Tseng).

[1] http://en.wikipedia.org/wiki/Word_of_mouth.
[2] http://www.amazon.com/.

0167-9236/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2010.08.023

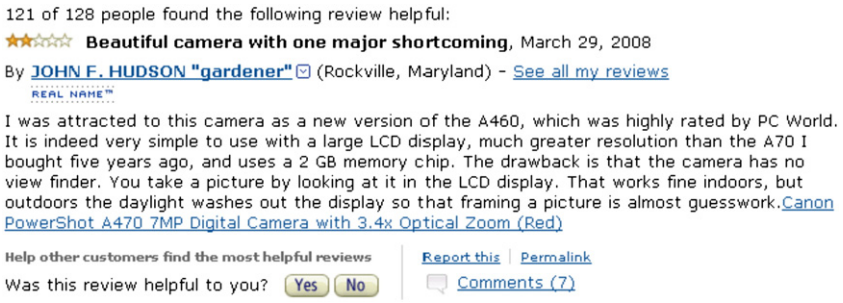


**Fig. 1.** A review quality evaluation on Amazon.

whereas information is structured and organized data. Thus, information is valuable for a specific purpose or application. In addition to raw textual units (i.e., unigrams), the proposed IQ features can comprehend opinions and polarity information embedded in review content; hence, the feature set is information-oriented. A review retrieval system based on the proposed method is also implemented. The system ranks a review according to the quality of the information it contains as well as the content's relevance to a user query. We hypothesize that the proposed IQ-based features are effective in classifying review quality, and review retrieval systems that incorporate the proposed classification model could retrieve informative and query-relevant reviews. We conduct various experiments based an expert-composed data corpus to validate the hypotheses. The results demonstrate that the proposed method outperforms state-of-the-art review quality evaluation approaches significantly and the reviews retrieved by the proposed retrieval system are query-relevant and informative. In addition, the learned model helps identify the factors that are critical for compiling high-quality reviews.

The remainder of this paper is organized as follows. Section 2 contains a review of related works. In Section 3, we introduce the information quality framework and apply it to the problem of review quality classification. We present the review retrieval system in Section 4 and evaluate the system performance in Section 5. Then, in Section 6, we summarize our conclusions.

## 2. Related work

In this section, we consider a number of review quality evaluation, information quality, opinion retrieval, and opinion mining approaches.

### 2.1. Review quality evaluation

Most review quality evaluation approaches adopt sets of review features to evaluate the quality of reviews. However, many adopted feature sets are lexical or syntactically oriented, so they hardly reflect the diverse characteristics of reviews. Zhang and Varadarajan [34] and Kim et al. [16] measure the quality of a review in terms of the helpfulness votes submitted by readers. The quality of a review is high if it receives many helpfulness votes. Zhang and Varadarajan collected a set of product reviews posted on Amazon along with the corresponding helpfulness votes given by readers. Based on the votes, the authors quantified the quality of a review as the ratio of helpfulness votes submitted by readers and employed SVM regression to approximate the quality of the reviews. The resulting regression function was then used to estimate the quality of new reviews. The reported experiment results show that the shallow syntactic features, e.g., the counts of proper nouns, modal verbs, and comparative adjectives in a review, are highly correlated with review quality estimation. Kim et al. also employed SVM regression to predict the helpfulness vote ratio of a review, and used five categories of review features, namely, structural, lexical, syntactical, semantic, and meta-data features, to construct a regression function. They found that the lexical features, i.e., the unigrams of a review, together with review's length and product rating have a significant impact on users' assessments of the review's quality.

Liu et al. [21] conducted a detailed survey of reviews posted on Amazon and found that users' votes were influenced by three types of bias: imbalanced vote bias, winner circle bias, and early bird bias. The imbalanced vote bias means that Internet users tend to rate others' opinions positively rather than negatively. The winner circle bias indicates that reviews awarded a lot of helpfulness votes will continue to attract many votes. This is because the reviews are top ranked, so they are easily accessed by Internet users. The early bird bias suggests that the earlier a review is posted, the more votes it will receive. Because of these biases, the methods [16,34] that employ users' votes as training examples are unreliable. Rather than making quality evaluations based on biased votes, Liu et al. treat review quality evaluation as a classification problem. They also use an expert-composed data set to train an unbiased SVM classifier, which then categorizes reviews as either high-quality or low-quality.

In this paper, we adopt an effective framework for assessing the quality of information. It considers various aspects of reviews to derive information-oriented features for evaluating the quality of the reviews. Using information-oriented features improves the evaluation performance, and resolves important issues related to the composition of informative reviews.

### 2.2. Information quality

The objective of information quality (IQ) research (also known as data quality research) is to determine the characteristics of information items that are important to, or suitable for information consumers [30]. Over the past decade, many IQ studies were developed to assess the information quality of various information technologies [1]. Most IQ studies consider information quality as a multi-dimensional framework in which each dimension represents a single aspect or construct of information items and is described by a set of features [17]. Zhu and Gauch [35] assessed the information quality of a Web page in terms of an IQ framework comprised of six quality dimensions, namely, currency, availability, information-to-noise ratio, authority, popularity, and cohesiveness. The authors measure the dimensions through the properties of Web pages and combine the measures linearly to calculate the quality score of a Web page. The reported experiment results show that the precision of Web searches is increased significantly by incorporating the framework into an information system. Hence, quality evaluation of Web pages is important for Web search engines. Chae and Kim [2] proposed an IQ framework for assessing the information quality of a mobile Internet service. The framework consists of four dimensions, namely, connection quality, content quality, interaction quality, and contextual quality. The authors conducted a large-scale survey to examine the effectiveness of the dimensions and found that all the dimensions help increase customer satisfaction and loyalty. However, the relative