A Mixed-Methods Approach to Disclose the Influence of Twofold Information Usefulness on Sales

Libo Ivy LIU
Department of Information Systems and Logistics,
Faculty of Business and Enterprise,
Swinburne University of Technology, Australia
ivy.liboliu@gmail.com

Jimmy SJ. REN
Lenovo Research and
Department of Information Systems,
City University of Hong Kong
jimmy.sj.ren@gmail.com

Long SONG
Department of Information Systems,
City University of Hong Kong
songlong.thu@gmail.com

Kristijan MIRKOVSKI
Department of Information Systems,
City University of Hong Kong
kmirkovsk2@gmail.com

Abstract

In the current study, we examine the relative effects of the two types of consumer reviews (i.e., positive and negative eWOM) on consumers’ purchase decisions, and the moderating roles of the two types of information usefulness (i.e., explicit usefulness and implicit usefulness). Analyzing a large-scale panel data collected from an online shopping site, we found that consumers’ purchase decisions are indeed influenced by both positive and negative reviews. In addition, a SVM classifier is built to identify the implicit useful reviews. Our results show that information usefulness, including explicit and implicit useful information, has an important moderating role in consumers’ purchase decisions. This study contributes to the existing literature by explaining how information usefulness (i.e., explicit and implicit usefulness) moderates the influence of consumer review on consumers’ purchase decisions, and providing a classifier for consumer reviews through sentiment analysis in online social shopping sites. The results offer important and interesting insights to IS research and practice.

Keywords: social commerce, positive electronic word of mouth, negative electronic word of mouth, information usefulness, explicit usefulness, implicit usefulness, consumer purchase decision, mixed-methods.

1. Introduction

Online reviews have become an important source of information that initiate and/or simplify consumers’ purchase decisions. Many online shopping sites have already integrated design features that provide rich set of information about products’ quality to potential customers. Consumer reviews are often described as the electronic word-of-mouth (eWOM) communication. This experience information can contain online positive and/or negative statements about a product/service or a company from potential, actual, and/or former consumers [1]. Previous studies have shown that online product opinions significantly influence consumers’ purchase decisions [2]. A study conducted by the Nielsen Company 1 in 2013 revealed that nearly 85 percent of the offline consumers electronically gather information about a local business prior to the actual purchase. Given the significance of consumer reviews for consumers’ purchase decision, it is of a great importance for academics and practitioners to understand how online reviews, including both positive and negative ones, influence the consumers’ purchase decisions in online shopping sites.

The rise of social media contributed to an increasing use of online social platforms to communicate opinions about products and exchange purchase experiences [3, 4]. Consequently, academics and practitioners have been investing considerable efforts to determine how to identify useful reviews among the millions of reviews available. Similarly, a sound classification of useful reviews would be essential for the development of e-marketing strategies that are targeted to effectively manage online consumer reviews.

To address the identified gaps in the literature, we draw on the dual-process theory to investigate the impact of positive and negative reviews on consumers’ purchase decisions. We examine how the information usefulness moderates the impact of consumer reviews
on consumers’ purchase decisions. Moreover, we conduct sentiment analysis to classify consumer reviews in online social shopping sites. The research questions for this study are as follows:

1) How positive and negative reviews influence consumers’ purchase decisions?

2) How information usefulness (i.e., explicit and implicit usefulness) moderates the relationship between positive and negative reviews, and consumers’ purchase decisions?

3) How to identify implicit useful reviews that help online shopping sites further enhance their e-marketing strategies?

We also notice that most of the existing studies on online reviews used an objective approach to explore how general online reviews influence sales [1, 5, 6]. In other words, prior literature obtained objective data from websites. To address the flaws of the research methodologies considered in prior literature, a mixed-method approach was used. We combined a large-scale quantitative panel data and a qualitative sentiment analysis to explore the difference in the influence of positive and negative reviews on consumers’ purchase decisions, and investigate how reviews usefulness moderates the relationship between consumer reviews and consumers’ purchase decision.

To summarize, we believe this study makes three fold contribution. First, existing studies have only empirically shown that in general online consumer reviews influenced consumer’s purchase decisions. We found that negative customer reviews received little attention in the IS literature. Thus, we consider negative reviews in the context of social shopping sites and examine how they influence consumers’ purchase decisions. Second, we notice that the moderating effect of reviews usefulness on the relationship between consumer reviews and their purchase decisions is underexplored in the IS literature. Therefore, this study attempts to investigate the moderating role of reviews usefulness (i.e., explicit and implicit usefulness) on consumers’ purchase decisions. Third, the conclusions in previous studies has been primarily derived from data collected through an objective approach. In this study, we empirically test our research model by a mixed-method approach, which is based on large-scale quantitative panel data, preceded by a qualitative sentiment analysis.

We have organized the rest of this paper as follows. In the first section, we present the conceptual model and elaborate on development of the hypotheses. After briefly describing our data source, we explain our empirical strategy and go through the results of our data analysis (i.e. quantitative panel data and qualitative sentiment analysis). Finally, we conclude this study with a discussion of the implications for theory and practice.

2. Research Model and Hypotheses Development

In this section we lay down the foundations of our research model and elaborate on its constituting hypotheses related to eWOM and information usefulness (see Figure 1).

2.1. Positive eWOM and Negative eWOM

eWOM is a form of interpersonal communication among consumers concerning their personal experiences about some products or services in an online context [7]. There are types of eWOM according to its direction: positive eWOM and negative eWOM. Positive eWOM is to share consumers’ satisfactory experiences towards products or services, whereas negative eWOM is to share dissatisfactory experiences in an online context. The relationship between eWOM and consumers’ purchase behavior is well explored in the prior literature [8-10]. The volume of peer consumer reviews is found significantly affect consumer’s purchase decisions [11]. Opinion-based social information has also been shown to be an important factor that influence product/service sales [12]. In the current study, we focus on the impact of consumer reviews (i.e. positive and negative eWOM) on consumers’ purchase decisions in the context of the online social shopping sites. Duan et al. [13] indicated that eWOM drives product awareness. Thus, we expect that positive eWOM will stimulate potential consumers’ purchases – the greater score of product positive reviews by peer consumer, the more likely a consumer will make a purchase:

H1: Consumer positive review will positively influence consumer purchase decision.

Similarly, negative eWOM will decrease potential consumers’ purchases – the greater score of product negative reviews by peer consumer, the less likely a consumer will make a purchase:

H2: Consumer negative review will negatively influence consumer purchase decision.
2.2. Information Usefulness

Information usefulness refers to the degree to which the information is perceived to be valuable, informative and helpful [14]. Information usefulness levels alter consumers’ elaboration likelihood. In this study, we identify two types of useful information: explicit and implicit useful information. Explicit useful information refers to information obviously marked as “useful” by other users. Implicit useful information refers to information that provides useful and has not been marked as “useful”. Researchers have used the dual-process theory of human information processing, to study how information processing behavior can lead to decision outcomes [14, 16], such as elaboration likelihood model (ELM) [15]. ELM was used to explain information processing strategies through two routes, including central routes and peripheral routes. The ELM allows us to make predictions about the relative impact of various factors on information adoption under different levels of elaboration likelihood. High levels of information usefulness tend to motivate increased elaboration on the review: as high levels of useful information becomes increasingly important to the consumer, he or she is more likely to undertake the cognitive effort to thoughtfully consider the review. Consumers that are highly involved with the useful review are likely to engage in high elaboration. Consequently, for less useful information, the impact of information (both positive and negative eWOM) on consumer purchase decision will be weakened. Thus:

H3: Information usefulness (i.e., explicit and implicit useful information) will positively moderate the relationship between peer consumer positive review and consumer purchase decision.

H4: Information usefulness (i.e., explicit useful information and implicit useful information) will positively moderate (negatively moderate the negative relationship) the relationship between peer consumer negative review and consumer purchase decision.

3. Research Method

A mixed-method approach was used, in which a large-scale quantitative panel data preceded by a qualitative sentiment analysis. In particular, firstly we used men’s sandal to conduct quantitative panel data analysis. And then we used women’s sandal to conduct qualitative semantic analysis.

3.1. Data Collection

The data for this study were collected from Zappos (http://www.zappos.com), a popular social shopping site in US. Zappos provides an online shopping platform where users can buy their favorite products and share experience in form of comments with others. Launched in 1999, Zappos has become one of the world's largest online shoe stores. Based on Alexa.com report, it has a traffic rank of 785 globally and 250 in US, as of May 2014. Zappos provides a platform for consumers to learn about shoe and clothing products, to share their experience related to shoe and clothing products, to interact with other consumers, and to buy shoe and clothing. Consumers on the forum can post comments related to the experience with the purchased shoe or clothing, provide a rating (from 1 to 5) about the product, and “like” others’ review. We aim to explore how a consumers’ purchase decision is influenced by other consumers’ product reviews and how such influence is moderated by the usefulness of review on the social shopping site.

We collected all reviews on man sandal shoe in Zappos before May 2014. There are 1,823 pairs of shoes in the type of man sandal shoe. To ensure the sample includes shoes with reviews, we selected shoes which include at least 1 review. We keep 51,543 reviews on 884 pairs of shoes from man sandal in Zappos before May 2014.

3.2. Operationalization of Constructs

Consumer Positive Review. A consumer can rate a product while sharing her or his experience about the product by providing an overall rating (from 1 to 5) about the product. If the overall rating is above 3, the review is considered as positive one. Therefore, consumer positive review is operationalized as the total score of all positive ratings (on particular products) provided by peer consumers.

Consumer Negative Review. The same as consumer positive review, if the overall rating is below 3 (i.e., 1 or
the review is considered as negative one. Therefore, consumer positive review is operationalized as the total score of all negative ratings (on particular products) provided by peer consumers.

**Information Usefulness: Explicit Usefulness.** Consumers on the forum can mark others’ reviews as “useful”. Individual consumers can mark as “useful” other consumers’ posts or ratings if they are found useful. Therefore, explicit usefulness is operationalized as the total number of “useful” (on particular products) provided by peer consumers. Based on rating score, we divided review usefulness into positive and negative usefulness reviews.

**Information Usefulness: Implicit Usefulness.** We developed a classifier below. The classifier can automatically mark “useful” others’ review. Therefore, implicit usefulness is operationalized as the total number of “useful” (on particular products) provided by our classifier.

**Consumer Purchase Decision.** Zappos provides sales ranking for all products which based on the consumers shopping history. Thus, consumers’ purchase decision is operationalized as the total number of sales on particular products.

## 4. Results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \beta )</th>
<th>Std. Error</th>
<th>Hypothesis Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer Positive Review</td>
<td>.111</td>
<td>.191</td>
<td>1 .002</td>
</tr>
<tr>
<td>Consumer Negative Review</td>
<td>-.319</td>
<td>5.958</td>
<td>1 .000</td>
</tr>
<tr>
<td>Consumer Positive Review * Explicit Usefulness</td>
<td>.311</td>
<td>.427</td>
<td>1 .000</td>
</tr>
<tr>
<td>Consumer Negative Review * Explicit Usefulness</td>
<td>-.172</td>
<td>11.426</td>
<td>1 .000</td>
</tr>
</tbody>
</table>

**R Square = 0.441 (df = 4)**

### 4.1. Quantitative Panel Data Analysis

To test our research model, we examined the research model in a regression framework as below. As this study focuses on the moderating effects of the review usefulness, we didn’t include explicit usefulness and implicit usefulness as independent variable in the model.

\[
\text{Consumer Purchase Decision} = \beta_0 + \beta_1 \text{Positive Review} + \beta_2 \text{Negative Review} + \beta_3 \text{Positive Review} \times \text{Explicit Usefulness} + \beta_4 \text{Negative Review} \times \text{Explicit Usefulness} + w \quad (1)
\]

The regression analysis results are summarized in Table 1. The results indicate that consumer positive review has a significant positive effect (\( \beta = 0.111, p < 0.001 \)) on consumers’ purchase decision. The results support H1. The regression results suggest that consumer negative review has a significant negative effect (\( \beta = -0.319, p<0.001 \)) on consumers’ purchase decision. The results support H2. Table 1 reveals that, as hypothesized, explicit usefulness exerts a significant positive moderating effect on the relationship between the consumer positive review and consumers’ purchase decision (\( \beta = 0.311, p < 0.001 \)). The results support H4. The negative interaction coefficient (\( \beta = -0.172, p < 0.005 \)) indicates that, the drop in consumers’ purchase decision is stronger for more explicit useful negative review, comparing to less explicit useful negative review.

### 4.2. Qualitative Semantic Analysis

In order to identify how information usefulness can moderate the effect of eWOM on consumers’ purchase decision, the very first step is to correctly identify the useful reviews from our panel dataset. Intuitively, we could have taken the number of people who found one review helpful as an indicator of information usefulness. However, this practice had its disadvantages since there could be a few potentially useful reviews with no people finding them helpful while these reviews indeed influence consumers’ purchase decision. Thus, there was a pressing need to perform a semantic analysis to dig out those undiscovered useful reviews through text mining and machine learning techniques before we began to analyze our dataset in our empirical model.

Based on 313,686 reviews from 6,984 pairs of women’s sandals, we conducted semantic analysis of these reviews in order to find what kind of words can effectively convey a signal of “usefulness” to consumers and further to use these discriminative words as text-based features plus some consumer-based features to train a classifier to classify the reviews as useful or useless. What’s more, we used another dataset containing 25149 reviews about men’s sandals to further test the sensitivity of our classifier.

#### 4.2.1. Feature Extraction and Selection

After stemming and segmenting the words for the reviews, we entered the feature extraction part. Here, we...
mainly focused on two kinds of features—text-based features and consumer-based features as shown in Table 2.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Features</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text-based</td>
<td>Tf-idf weights</td>
<td>The product of TF (normalized term frequency) and IDF (inverse document frequency), as shown in (2) and (3)</td>
</tr>
<tr>
<td></td>
<td>Review Length</td>
<td>The total number of words in each review</td>
</tr>
<tr>
<td>Consumer-based</td>
<td>Overall Rating</td>
<td>The consumer’s overall rating about this product</td>
</tr>
<tr>
<td></td>
<td>Comfort Rating</td>
<td>The consumer’s rating about the comfort from the product</td>
</tr>
<tr>
<td></td>
<td>Style Rating</td>
<td>The consumer’s rating about style of the product</td>
</tr>
<tr>
<td></td>
<td>Size Rating</td>
<td>The consumer’s satisfaction with the size of the product</td>
</tr>
<tr>
<td></td>
<td>Width Rating</td>
<td>The consumer’s satisfaction with the width of the product</td>
</tr>
<tr>
<td></td>
<td>Support Rating</td>
<td>The consumer’s satisfaction with the feeling of support from the product</td>
</tr>
</tbody>
</table>

Table 2. Proposed Feature Set for Classification

In text-based feature set, Tf-idf weights and Review Length are included. For Review Length, we found that long review often contain larger amount of information so that consumers are more inclined to mark it as a useful one.

Tf-idf scheme is a very common way to represent a document by computing a weight for each word [18]. Besides, these weights often play an important role in different classifiers like Rocchio, KNN, and SVM [19, 20]. Here, the normalized tf value of word $i$ in a review $d$ is computed by

$$ntf_{i,d} = \alpha + (1 - \alpha) \frac{tf_{i,d}}{\max_i tf_{i,d}} \tag{2}$$

where $\alpha$ is a value between 0 and 1. Thus, the weight of word $i$ in a review $d$ is denoted by the tf-idf scheme as follows.

$$w_{i,d} = ntf_{i,d} \times \log \frac{N}{df_i} \tag{3}$$

$N$ is the total number of reviews in our dataset and $df_i$ is the number of reviews that contain word $i$.

For consumer-based features, we included other ratings which are some numbers ranging from 1 to 5 attached with the review such as the consumer’s rating in terms of the comfort, style, size, width, support of the products as well as the overall rating since we believe some rating patterns can attract consumer’s attention and further obtain a “useful” label. For example, some people care more about the drawbacks of the products thus those with low ratings may have better chance to get a “useful” label from these people.

After analysis, we found that there are 13398 words in our lexicon of our review dataset. Considering the efficiency of classification, we adopted the Chi-Square test to select discriminative words from our lexicon to generate tf-idf features. Chi-square test is one of the most commonly used feature selection methods in text mining area [21, 22]. Finally, we selected the top 3000 discriminative words in the tf-idf weight feature part.

Meanwhile, we divided our review dataset into two categories—positive and negative reviews indicated by the overall rating. After performing Chi-square test for each category, we successfully found some discriminative words that could convey a signal of usefulness in both positive and negative reviews as shown in Table 3. We could see these words often expressed a contagious emotion that is more capable of gaining a sense of empathy from people. That is why these words are discriminative for useful reviews.

<table>
<thead>
<tr>
<th>Words Indicating Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Soft, cute, great, wow, smart, wonderful, fantasy, enjoy, advantage, awesome</td>
</tr>
<tr>
<td>Negative</td>
</tr>
<tr>
<td>Hurt, sad, heartbroken, tear, scrape, warn, uncomfortable, risk, hate, poor, disappoint, mad, honest</td>
</tr>
</tbody>
</table>

4.2.2. Classification and Evaluation

As our goal was to classify the useful reviews, we considered 2/3 of the useful and useless reviews as our training set and the remaining useful reviews as the test set. For a reviews with at least one person finding it helpful, we labeled them as a useful reviews. However, for the useless reviews, we couldn’t simply regard them as useless just because nobody considered them useful. In fact, part of the useless comments were not considered as “useful” just because consumers didn’t have chance to read or notice them. Usually, the reviews posted at a very early stage are possible to be neglected by consumer relative to those that are at the very bottom.
of the page. Thus, these reviews at the bottom of the page might lose the potential to become a “useful” ones no matter how well-written those reviews are. On the contrary, if the top reviews, which had a bigger chance to be seen by consumer, are not marked as “useful”, then it was very possible that they were useless indeed. As a result, when we were selecting the useless reviews for the training set, we preferentially selected the most updated reviews without a mark of “useful” based on the posting time sequence to approximate the useless reviews. Usually, for a common consumer, s/he would go beyond one page to see more reviews on average. We considered those reviews that appeared on the first or second page without a “useful” mark to approximate the genuine useless reviews and added them to the training and test set proportionally.

For the classification, we used the LibLinear package—an implementation of SVM (Support Vector Machine) classifier [23] that is a supervised learning algorithm based on Machine Learning technique to achieve a high performance. And we selected the function with L2-regularized L2-loss form as our objective function. And we selected precision (PRE), accuracy (ACC), recall (REC), F1-measure (F1), receiver operating characteristic curve (ROC) together with the area under the curve (AUC) as the evaluation metrics for our experiment. The results on the test set are shown in Table 4 and Figure 2.

### Table 4. Performance Results (%) on Women’s Sandals Dataset

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text-based</td>
<td>ACC 70.72</td>
</tr>
<tr>
<td>Text- and Consumer-based</td>
<td>ACC 71.82</td>
</tr>
</tbody>
</table>

From the results, we can see the group with both the text- and consumer-based features achieves a slightly better performance than the group with text-based features only. This implies that the patterns of consumers’ rating could explain for the usefulness of a review to some degree. And also those emotionally contagious words also have played a significant role in the classification for useful reviews.

#### 4.2.3. Sensitivity Analysis

To test the robustness of our classifier, we applied our classifier to another dataset which contained 25149 reviews about men’s sandals on the same online shopping website. The results are shown below in Table 5 and Figure 3.

### Table 5. Performance Results (%) on Men’s Sandals Dataset

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text-based</td>
<td>ACC 70.33</td>
</tr>
<tr>
<td>Text- and Consumer-based</td>
<td>ACC 70.51</td>
</tr>
</tbody>
</table>

From the results above, we can see that the overall performance on men’s sandals dataset (new) is slightly
worse than it on the women’s sandals dataset (old). But in terms of the recall and precision, the values are very stable and the recalls on the new dataset are even higher than before. In a conclusion, our classifier is capable of discovering useful reviews even on a freshly new dataset.

To test our research model, we reran research model using explicit usefulness information and implicit usefulness information that identified by the classifier. Similar as model (1), we didn’t include explicit usefulness and implicit usefulness as independent variable in the model. Because this study focuses on the moderating effects of the review usefulness rather than direct effects.

\( \text{Consumer Purchase Decision} = \beta_0 + \beta_1 \text{Positive Review} \times \text{Implicit Usefulness} + \beta_2 \text{Negative Review} \times \text{Implicit Usefulness} + \beta_3 \text{Positive Review} \times \text{Explicit Usefulness} + \beta_4 \text{Negative Review} \times \text{Explicit Usefulness} + w \) (4)

The results are presented in Table 6 (Table 6 summarizes all regression results). The results indicate that implicit usefulness exerts a significant positive moderating effect on the relationship between the consumer positive review and consumers’ purchase decision (\( \beta = 0.381, p < 0.005 \)) and a significant negative moderating effect on the relationship between the consumer negative review and consumers’ purchase decision (\( \beta = -0.479, p < 0.005 \)).

5. Discussion and Implications

5.1. Major Findings

In this work, we drew on the dual-process theory to explore the relationship among positive eWOM, negative eWOM, online information usefulness and customer purchase decision. With our theoretical research model and the large scale real world e-commerce dataset collected from one of the world’s largest online shoe stores, we can show that consumer positive reviews will positively influence consumers’ purchase decisions; consumer negative reviews will negatively influence consumers’ purchase decision. We also prove that information usefulness plays a moderating role in influencing the customer purchase decisions.

Treating this preliminary conclusion as a solid baseline, we took one step further. We aimed to disclose the connection between the potentially useful reviews, namely the reviews would be tagged useful by a future customer, and the customer purchase decisions. To effectively achieve this goal, we build a text classifier using a state-of-the-art text mining technique to explicitly model the useful reviews and non-useful reviews. The performance of our text classification model was validated in a carefully prepared test set containing both useful and “potentially useful” reviews.

By applying the result of this text classifier in our theoretical research model, our result shows that the information usefulness formed by the explicitly useful reviews as well as the potentially useful reviews plays a significant role in moderating the purchase decisions. On one hand, this finding reaffirms the theoretical hypotheses in our research model; on the other hand it strongly suggested that the implicit component of information usefulness in the e-commerce context should not be overlooked. Several important implications can be generated from our findings which we will further discuss in the later sections.

5.2. Theoretical Contributions

This paper makes several important contributions to the existing literature. First, this study contributes to e-commerce literature by investigating the moderating effect of information usefulness in influencing consumers’ purchase decisions. Unlike most of the previous studies which only focused on the effects of the positive eWOM, this study generalizes the finding to both positive and negative eWOM. Second, to our knowledge this paper is the first attempt to use large scale real world data to disclose how the explicit as well as the implicit component of information usefulness influences customer purchase decisions in the e-commerce context. Our results strongly suggested that the explicit component of information usefulness is not the only factor to influence sales. On the contrary, the implicit part of information usefulness may as important and should not be overlooked. While this finding consolidates our theoretical model, it also gives the
model more explanatory power when applying it the real world e-commerce setting.

5.3. Managerial Implications

Our theoretical model and findings may find applications in e-commerce in the area of sales prediction and review information management. First of all, while previous studies suggested that customer reviews can be used as a cue to predict product sales [24, 25], the results in the paper strongly suggest that when predicting the sales of a product one should not only consider the explicit part of information usefulness, the implicit part of information usefulness in the customer reviews should also be used. For instance, given a set of product with similar amount of explicit reviews, we can show that it’s the information usefulness hidden in the implicit useful reviews that determine the sale ranking of those products. Secondly, the plethora of customer reviews may impose a heavy cognitive burden on consumers, resulting in a decrease in perceived informativeness of online shopping sites. To prevent consumers from experiencing information overload, it’s reasonable to suggest that e-commerce sites should expose useful (explicit and implicit) reviews, which can be retrieved by our text classification method, in the first several pages of the content.

5.4. Limitations

However, there are several limitations related to our research. First, the definition of useless reviews is very rough since the useless reviews in the training set for the classifier are selected based on the assumption that on average consumers usually browse one or two pages of reviews before purchase. Second, whether our classifier can dig out those potentially useful reviews need to be validated. To make our study more convincing, it is better to do a longitudinal study to find out those reviews marked as helpful in the future, and use those reviews to test our classifier again.

6. Reference


