Computers in Human Behavior 58 (2016) 75–81

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

A concept-level approach to the analysis of online review helpfulness

Aika Qazi [a, f, *], Karim Bux Shah Syed [b, g], Ram Gopal Raj [a], Erik Cambria [c], Muhammad Tahir [d], Daniyal Alghazzawi [e]

[a] Faculty of Computer Science and Information Technology, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia
[b] Faculty of Business and Accountancy, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia
[c] School of Computer Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore
[d] Faculty of Computing and Information Technology, University of Jeddah, Jeddah, Saudi Arabia
[e] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia
[f] Faculty of Computer Science and Information Technology, COMSATS Institute of Information Technology, Islamabad, Pakistan
[g] Institute of Business Administration, University of Sindh, 76080 Jamshoro, Pakistan

A R T I C L E I N F O

Article history:
Received 11 April 2015
Received in revised form
3 September 2015
Accepted 14 December 2015
Available online 1 January 2016

Keywords:
Online reviews
Review helpfulness
Electronic commerce
Suggestive reviews

A B S T R A C T

Helpfulness of online reviews serves multiple needs of different Web users. Several types of factors can drive reviews' helpfulness. This study focuses on uninvestigated factors by looking at not just the quantitative factors (such as the number of concepts), but also qualitative aspects of reviewers (including review types such as the regular, comparative and suggestive reviews and reviewer helpfulness) and builds a conceptual model for helpfulness prediction. The set of 1500 reviews were randomly collected from TripAdvisor.com across multiple hotels for analysis. A set of four hypotheses were used to test the proposed model. Our results suggest that the number of concepts contained in a review, the average number of concepts per sentence, and the review type contribute to the perceived helpfulness of online reviews. The regular reviews were not statistically significant predictors of helpfulness. As a result, review types and concepts have a varying degree of impact on review helpfulness. The findings of this study can provide new insights to e-commerce retailers in understanding the importance of helpfulness of reviews.

1. Introduction

The development of Web 2.0 has encouraged people to express their opinions about products/services. Opinions are central to most human activities and hence, are one of the key drivers of human behaviors (Hu & Liu, 2004). These opinions can help consumers in purchase decisions (Liu, 2010). There are varieties of opinions that discuss different aspects of a purchase of a product/service. Early research on online reviews has identified and studied two types of opinions, namely (1) regular and (2) comparative (Jindal & Liu, 2006b). Witnessing exponential proliferation of reviews in recent years, along with the diversity of the uses and functions these perform, this dual classification seems too narrow. More recently, suggestive have been identified as a third type of reviews (Qazi, Raj, Tahir, Waheed, et al., 2014). In linguistic, suggestives are defined as indirect speech acts. The speech acts used to direct someone to do something in the form of a suggestion are classified as suggestives. They can be considered polite in the sense that instead of telling someone to do something directly, they present it in the form of a suggestion, which the reader is not obliged to follow (Kumar, 2011). The appearance of multiple review types (regular, comparative and suggestive) significantly contributes in making variety of consumption choices and future guidelines that enables consumers as well as retailers to make better purchase decisions and business policies.

The reviews types are defined based on their linguistic construct (Liu, 2012) that expresses different sort of information. A *regular opinion* is often referred to simply an *opinion* in the literature (Jindal & Liu, 2006b). A *comparative opinion* expresses a relation of similarities or differences between two or more entities (Jindal & Liu, 2006a). A *suggestive opinion* is defined as directing someone to do something in a polite manner (Qazi, Raj, Tahir, Cambria, & Syed, 2014). The classification of these types of reviews assigned "A" to regular, "B" to comparative and "C" to suggestive opinions (Jindal &

* Corresponding author. Faculty of Computer Science and Information Technology, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia.
E-mail addresses: atikaqazium@gmail.com, atika@siswa.um.edu.my (A. Qazi).

http://dx.doi.org/10.1016/j.chb.2015.12.028
0747-5632/

Liu, 2006a; Qazi, Raj, Tahir, Waheed, et al., 2014). In the competitive business environment users experience difficulty in taking decisions if they only look at one aspect of a product (Ganapathibhotla & Liu, 2008; Liu, 2012). Clearly, different types of opinions carry variety of aspects, e.g. the notion of product comparisons is another aspect that is not only useful for product manufacturers, but also for potential buyers, thus helping in better decision making (Jindal & Liu, 2006b). Many studies suggest that online product reviews and related features have a significant impact on consumers' purchase decision and sales (Duan, Gu, & Whinston, 2008; Elwalda, Lü, & Ali, 2016; Forman, et al., 2008).

Among the many features associated with online product reviews, 'review helpfulness' is particularly important, as it represents the subjective evaluation of the review judged by others (Cao et al., 2011; Li, Huang, Tan, & Wei, 2013). Therefore, helpful reviews improve the value of business sites, and sites containing more helpful reviews are more likely to attract buyers and consumers seeking information. Major Websites, such as Amazon.com, Tripadvisor and Yelp.com, ask readers to rate the helpfulness of the reviews of products/services and make that information available. This implies that online retail sites with more helpful reviews offer greater potential value to customers. Such reviews are useful for better and well-informed decisions, and, hence, maximize users' satisfaction (Kohli, Devaraj, & Mahmood, 2004). However, helpfulness of online reviews is a multi-faceted concept that can be driven by several types of factors based upon quantitative and qualitative measures. In the early studies, the most common practice to measure the review helpfulness was based upon the quantitative factors of reviews such as the star rating or thumbs up/down and the review length (Otterbacher, 2009; Pang, Lee, & Vaithyanathan, 2002).

More recent studies have focused on qualitative measures in addition to quantitative ones (search goods, search experience, experience, reviewer impact, reviewer and cumulative helpfulness) to explore helpfulness (Huang, Chen, Yen, & Tran, 2015; Mudambi & Schuff, 2010). However, by looking into the multiple review types and associated vital aspects, helpfulness is quite a complex concept as one would equate quantitative measures of reviews to helpfulness, while others might consider qualitative instead. Therefore, this study was designed to extend existing research on online review helpfulness by viewing not just the quantitative factors (such as word count), but also qualitative aspects of reviews such as review types itself (including regular, comparative, suggestive reviews and cumulative helpfulness).

The study contributes to the conceptual development and understanding of the helpfulness components of reviews from a concept-level prospective. Built on the relevant online review literature, four hypotheses were proposed (H1, H1, H3 and H4) to study the proposed model for reviews' helpfulness. The dataset consisting of 1500 hotel reviews from Tripadvisor was employed to test these hypotheses. This study successfully validated the proposed model and found key factors to make an opinion helpful for readers. The results of the current research have contributed to relevant literature by providing further understanding of the morphological features (quantitative and qualitative) of reviews and their influence on helpfulness. Additionally, the findings of the paper have extended the results found in existing research (Mudambi & Schuff, 2010) by looking also at the review types (regular, comparative, and suggestive) to see whether each of those aspects influences online review helpfulness.

This paper is organized as follows: in Section 2, related work is presented; Section 3 presents the proposed model and related hypotheses; Section 4 presents the research methodology; Section 5 discusses evaluation results; Section 6 concludes the discussion; Section 7 presents conclusions and future work and Section 8 explains the implications of the study.

2. Literature review

The study of reviews is commonly termed opinion mining, defined as an interdisciplinary research field involving natural language processing, computational linguistics, and text mining (Thet, Na, & Khoo, 2010). Textual information is generally of two types: subjective and objective (Ganapathibhotla and Liu, 2008) and opinions are expressed by way of subjective expressions (Quigley, 2008).

Today opinion mining and sentiment analysis are mainly carried out at two levels: word-level and concept-level. Word-level analysis includes approaches such as keyword spotting (Poria et al., 2012), lexical affinity (Poria, Gelbukh, Cambria, Das, & Bandyopadhyay, 2012), and statistical methods (Vahdat, Oneto, Anguita, Funk, & Rauterberg, 2016). Concept-level analysis, instead, does not take words as basic elements for text analysis, but rather multi- word expressions (Cambria, Fu, Bisio, & Poria, 2015). An expression such a "cloud computing", for example, is a semantic atom in concept-level opinion mining, but two different words ("cloud" and "computing") in word-level analysis. Hence, concept-level analysis better preserves semantics associated with natural language (Cambria & White, 2014). Common approaches to concept-level sentiment analysis include taxonomy-based methods (Gangemi, Presutti, & Reforgiato Recupero, 2014) and common-sense-based approaches (Cambria & Hussain, 2015).

As mentioned earlier, sentiment analysis reviews are of different types: regular opinions, pertaining to a single object or entity, and comparative opinions, which discuss more than one object (Jindal & Liu, 2006a, 2006b). The regular opinion is mostly used to find good or bad views about a particular product whereas comparative opinions are significantly utilized for competitive intelligence (Jindal & Liu, 2006a). Existing works cover different aspects of regular opinions (Popescu & Etzioni, 2005) (Liu, Hu, & Cheng, 2005) (Cruz, Troyano, Enríquez, Ortega, & Vallejo, 2010) (Hariharan, Srimathi, Sivasubramanian, & Pavithra, 2010). The comparative sentence mining concept originates from Liu et al. in (Jindal & Liu, 2006a), and it is then considered further in (Ganapathibhotla & Liu, 2008; Hou & Li, 2008; Jindal & Liu, 2006; Li, Lin, Song, & Li, 2010; Xu, Liao, Li, & Song, 2011; Xu et al., 2011). The suggestive reviews as a third significant type of reviews have been examined recently by (Qazi, Raj, Tahir, Waheed, et al., 2014).

Classifying reviews is imperative because different types make different information-consumption choices (Jindal & Liu, 2006b). An example opinion sentence is "the service quality of hotel X is poor". An example comparative sentence is "the service quality of hotel X is not as good as that of hotel Y". Clearly, these two sentences give different information. Their language constructs are quite different too. Identifying comparative sentences is useful in practice because direct comparisons are perhaps one of the most convincing ways of evaluation, which may even be more important than opinions on each individual object (Jindal & Liu, 2006a; Liu, 2012). A suggestive review is characterized by the suggestion of a solution to a particular issue regarding an entity or a group of entities. For example "I suggest hotel X to better use the services and make your trip worth visiting". An important application area for such solution is business intelligence, as product manufacturers always wish to recognize consumers' opinions about several aspects of their services. These varieties of online reviews are widely used as convincing communication. The tourist's buying behavior is influenced by looking into different aspects of the reviews available through web 2.0 (Sparks, Perkins, & Buckley, 2013). The study by (Cui & Ryan, 2011), for example, found out that both urban and rural residents have favorable attitudes toward tourism. This indicates the urgency of more helpful online reviews to promote tourism business.

Helpfulness of reviews is determined by different set of features such as sentiments, user expertise, information type and information quality. The past research (Kim, Pantel, Chklovski, & Pennacchiotti, 2006) proposed an algorithm for automatically assessing helpfulness by using review length, its unigrams and sentiment words. To estimate the helpfulness voting ratio, Liu's (Liu, Huang, An, & Yu, 2008) used writing style, reviewer's expertise, and timeliness. The researchers (Danescu-Niculescu-Mizil, Kossinets, Kleinberg, & Lee, 2009) performed an analysis of several hypotheses and reported that helpfulness does not only depend on the content but on how the evaluation relates to other evaluations (of the same product).

By dividing the reviews into low, medium, high, duplicate and spam categories, (C. C. Chen & Tseng, 2011) proposed a review evaluation information quality to calculate helpfulness. A probabilistic distribution model for helpfulness binary voting (the review is either helpful or not) from text of online reviews was proposed by (Zhang & Tran, 2010). They used the algorithm "expectation-maximization (EM)" to search a distribution that maximizes the helpfulness distribution probability for a given training corpus. The study presented by (Hart-Davidson, McLeod, Klerkx, & Wojcik, 2010) determined helpfulness by using indicators for quality in online peer review. The researchers (Liu, Jin, Ji, Harding, & Fung, 2012) reported four different types of features for helpfulness measurement which express designers' interest in evaluating helpfulness. More recently, the helpfulness of reviews is also determined by quantitative factors (such as word count), and qualitative aspects (including reviewer's experience, reviewer's impact, reviewer's cumulative helpfulness) (Huang et al., 2015).

Although review helpfulness has become an important topic in business and information technology literature, little research has explored the effects of both quantitative and qualitative factors on review helpfulness. The current research is developed to bridge the gap in literature by shedding more light on this connection. In terms of quantitative factors, the current research uses word count by counting the number of concepts present in each review as a predictor of review helpfulness. Additionally, the quality of information is extremely crucial in online reviews, since high quality information provides reliable, current and concise information (Arazy & Kopak, 2011) (Alkhattabi, Neagu, & Cullen, 2011; Yaari, Baruchson-Arbib, & Bar-Ilan, 2011). Since there is currently no objective metrics to quantify qualifications, qualitative aspects of reviews including review types (comparative, and reviewer cumulative helpfulness) are also used as qualitative factors in this present study.

3. Research model and hypotheses

The proposed research model is presented in Fig. 1. The model is based on morphological properties of reviews and review types. The model leverages on the hypothesis that different types of review are characterized by a different number of concepts, which influence their helpfulness. According to past research, wordiness can increase information diagnosticity (Johnson & Payne, 1985). Similarly, the number of concepts per review (NCR) is key in assessing the semantic information carried by a review (Cambria, Gastaldo, Bisio, & Zunino, 2015) and, hence, it is an important feature for determining its helpfulness. Clearly, wordiness is helpful as it is generally viewed as directly proportional to the amount of information that a review delivers. However, instances in which wordiness accompanies excess concept repetition, diseconomy of concepts and unnecessary details may lead to poor scoring on review helpfulness.

The average length of a sentence determines the readability of writing as much as any other quality (Garner, 2001). Accordingly, readability formulas rely heavily on sentence length, which shows that better readability leads to better communication of the information, and hence increases the usefulness of a review. From the readership perspective, shorter and simpler sentences are usually preferred. However, excessive shortness and simplicity may also yield to inadequacy in communication and, hence, influence review helpfulness. Therefore, we also introduce the average number of concepts per sentence (ANCS) as a morphological factor that affects the helpfulness content of the review. We hypothesize that ANCS has an important relation with quality of text and readability; therefore, it would be one of the key components that persuade the review readers to vote on its helpfulness.

Review types provide competitive intelligence (CI) involving early identification of potential risks and opportunities by gathering and analyzing information in making strategic decisions (Liu, 2012). In addition, it is observed that consumers may provide detailed comparative opinions about a service or product for better decision-making. The suggestions, on the other hand, are polite in the sense that instead of telling someone to do something directly, they present it in the form of polite indirect speech (Kumar, 2011), which is usually written more concisely than comparisons. This leads us to the next hypothesis: review type moderates the effect of NCR and ANCS on review helpfulness. (See Table 1).

4. Research methodology

4.1. Data collection

We collected data using the online reviews available from TripAdvisor. We retrieved 1500 customer reviews along with Author, Content, Date, Number of Reader, Number of Helpful Judgment, Overall rating from freely available data source ("The database and information system laboratory," 2010). We parsed different sets of reviews for each hotel by removing HTML formatting and translating page contents to XML. This resulted in the separation of data into different records (reviews) and fields (review contents). Reviews were labeled into multiple classes (A, B and C) based on their morphological construct by using the card sorting method. The card sorting method involves manual labeling of content based on information and arranges it into different set of groups that make sense to users or participants (Warfel & Maurer, 2004). This classification was carried out to find the effect of different review types on online review helpfulness. Regular reviews were labeled A, comparative as B, and suggestive reviews as C. We excluded 164 reviews from the analysis because of no vote for usefulness. This led us to eliminate 10% (approx.) of the total, resulting in a data set of 1336 reviews.

4.2. Variables

We have extracted the ratio of number of votes in favor of usefulness to the total number of votes (Total Votes) cast on the helpfulness question usually worded as "was this review helpful to you?". The resulting ratio was used to mimic the relative helpfulness of the reviews. NCR was measured from review wordiness. We used number of concepts, i.e., multiword expressions, such as "living room", "hotel lobby", or "reserve restaurant table", extracted from SenticNet 3 (Cambria, Olsher, & Rajagopal, 2014) for each review as a measure of its length and calculated its relative ANCS. Due to the probable impact of relative amount of the total votes cast for reviews, we incorporated total votes in our model as a control variable.

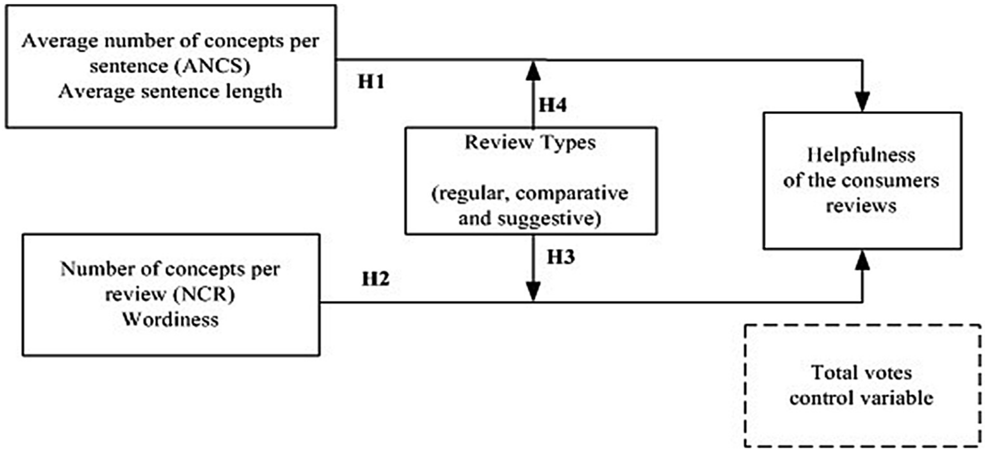

Fig. 1. Helpfulness of consumers review.

Table 1
Review helpfulness hypotheses.

H1	The average number of concepts per sentence influences review helpfulness.
H2	The number of concepts per review influences helpfulness.
H3	Review type moderates the effect of number of concepts per review on helpfulness, e.g., for comparative reviews, longer reviews have a positive effect on review helpfulness.
H4	Review type moderates the effect of the average number of concepts per sentence on helpfulness, e.g., for comparative reviews, longer sentences have a positive effect on review helpfulness.

4.3. Analysis method

The objective of this paper is to evaluate review helpfulness. We used the Tobit regression model because of the nature of the dependent variable, i.e. helpfulness. The ratio of favorable votes to the total votes on the question "was this review helpful to you?" is a float in the range (0, 1). Therefore, the dependent variable is continuous on a certain range. Furthermore, the dependent variable lacks the quality of being observable. The second reason to use Tobit regression was to overcome the problem of selection bias, which arises because it is not possible to determine the number of actual readers, which may be much more than the total vote cast. We can only infer from the collected data how many votes were casted and what proportion of these votes was in favor. Hence, the problem of selection bias is inherent in this type of sample. According to (Kennedy, 2003), with dependent variables possessing these characteristics, the Ordinary Least Square (OLS) estimates are neither unbiased nor consistent. Belonging to the family of limited variables models, the Tobit model addresses most of these computational problems. It employs maximum likelihood method and relies on likelihood ratio and Efron's Pseudo R-sq as measures of its goodness of fit.

Table 2 shows the descriptive statistics for our variables. On average, more than 80% of the voters who voted found the reviews helpful. Most of the reviews in our sample contained nearly 80 concepts, with 13 concepts per sentence. Each review attracted on average 11 votes. Finally, we investigated the impact of blending these morphological features with review types. We achieved it by introducing the interactions of review types with NCR and ANCS respectively. In order to explore the hypothesis that review type influences the helpfulness, we added dummy variable of review type, taking type A as a reference category. According to the above mentioned hypotheses and description, we have proposed the following equation for our model.

Table 2
Descriptive statistics for full sample.

Variable	Mean	SD	Observations
Helpfulness	0.8371	0.2357	1336
NCR	80.3254	59.9765	1336
ANCS	12.6764	5.0978	1336
Total votes	10.6527	9.7024	1336

$$Helpfulness = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \varepsilon$$

Where

X_1 = NCR
X_2 = ANCS
X_3 = TYPE B
X_4 = TYPE C
X_5 = Total number of helpful votes
X_6 = NCR * TYPE B
X_7 = NCR * TYPE C
X_8 = ANCS * TYPE B
X_9 = ANCS * TYPE C
ε = Random error

5. Results

5.1. Tobit analysis

We examined the effect of NCR and ANCS on helpfulness using Tobit regression. How NCR and ANCS affect helpfulness, given the type of review, was another objective of this study. To this end, we loaded the interaction terms of review type with NCR and ANCS respectively in our model. Table 3 summarizes the results for our model applied on the full sample set. Our model indicated a good fit

Table 3
Regression output for full sample.

Variable	Coefficient	Std. Error	z-Statistic	Sig
(Constant)	0.8619	0.0472	18.254	0.000***
NCR	−0.0005	0.0001	−3.484	0.000***
ANCS	0.0096	0.0038	2.547	0.011**
TYPE B	0.2807	0.0852	3.293	0.000***
TYPE C	2.0174	0.1672	12.066	0.000***
ANCS × B	−0.0125	0.0058	−2.158	0.023**
ANCS × C	−0.1099	0.0081	−13.639	0.000***
NCR × TYPE B	0.0005	0.0002	2.952	0.002***
NCR × TYPE C	−0.0006	0.0003	−1.924	0.045**
Total votes	−0.0721	0.0125	−2.147	0.070

*** significant at 99%, ** significant at 95%.

with likelihood ratio ($p = 0.000$) and Efron's pseudo R^2 value of 0.167.

For further analysis, we ran the Tobit regression separately for each of the three review types to test our hypothesis that review type affects the review length, and average sentence length on helpfulness. To do this, we extracted three subsamples, one for each review type. The results are included in Table 4, Table 5, and Table 6 for regular, comparative and suggestive types respectively.

From Table 3, we conclude that, generally, too long reviews are less helpful ($p < 0.01$), while relatively longer sentences have positive effect on helpfulness ($p < 0.05$). These results are supported by our hypotheses 1 and 3, that NCR and ANCS influence review helpfulness.

Further, we tested the interaction of review types with concept count and sentence length respectively, to test hypotheses 2 and 4. With reference to type A regular reviews, we tested the relative impact of review type on helpfulness. Results suggest that longer sentences are more helpful for regular reviews than both comparative and suggestive. This is evident from the negative coefficient of interaction terms for both comparative ($p < 0.05$) and suggestive ($p < 0.05$) in Table 3. Referring to Tables 4–6 we observe the varying sign of coefficient for each type, with comparative and suggestive significantly negative and regular significantly positive. We, therefore, find evidence in support of our hypothesis 2 that the impact of ANCS on helpfulness varies significantly with the underlying type of review. The type of review moderates the impact of ANCS on helpfulness.

Our results suggest that review type significantly influences the effect of review length over helpfulness. Interestingly, we conclude that in relation to regular, comparative reviews tend to be more helpful when wordier, while suggestive reviews more useful when less wordy. We found that, in general, review length affected helpfulness negatively (Table 3), i.e. too long reviews fail to attract the reader's attention and turn less helpful than shorter and concise reviews. However, when we tested the relationship of review length and helpfulness specifically for each type of review, results deviated significantly from the full sample set. This suggests that the nature of relationship between review length and helpfulness varies significantly given the review type. Although longer reviews appear to be less helpful in general, (Table 3), we find them more helpful if the review is comparative (Table 5). This is evident from the positive coefficient of concept count in Table 5.

Table 4
Regression output for regular reviews.

Variable	Coefficient	Std. error	z-Statistic.	Sig.
Constant	0.9859	0.0485	20.3466	0.000
ANCS	0.0085	0.0036	2.3743	0.017
NCR	−0.0004	0.0001	−3.1304	0.001
Total votes	−0.0161	0.0026	−6.0256	0.000

Table 5
Regression output for comparative reviews.

Variable	Coefficient	Std. error	z-Statistic.	Sig.
Constant	1.4401	0.112	12.7856	0.000
ANCS	−0.0363	0.006	−5.5941	0.000
NCR	0.0005	0.000	2.1184	0.034
Total votes	0.0006	0.002	0.2174	0.827

Table 6
Regression output for suggestive reviews.

Variable	Coefficient	Std. error	z-Statistic.	Sig.
Constant	3.0066	0.1524	19.7296	0.000
ANCS	−0.1029	0.0066	−15.6837	0.000
NCR	−0.0008	0.0002	−3.5248	0.000
Total votes	−0.0139	0.0024	−5.6915	0.000

6. Discussion

These results are not only interesting but intuitively plausible too. By definition, comparative reviews obtained comparing two entities. Longer comparative reviews, supposedly, provide extensive information on the entity under review not only in isolation but also in relative terms by comparing it with other comparable entities (Liu, 2012). This specific feature requires them to be objectively longer than other types. Thus, the longer the comparative review the more informative it is. From the reader perspective, the longer the description with comparison, the greater would be the helpfulness of review. Linking our results with consumer buying behavior would yield additional support for our results. Consumer buying decisions are seldom based on seeking information on a single entity in isolation. Rather, before making a buying decision, consumers want to collect information on all the available options in order to compare them, and choose the one that best suits their needs (Golicic, Fugate, & Davis, 2012; McKechnie, 1992). The more the information on available options at hand the lesser the degree of hesitation in buying decision and greater the perceived effectiveness of the decision itself.

Moreover, we suggest that readers treat reviews like a commodity and behave more like a consumer looking for maximizing the utility, gaining information in terms of description and comparison. Because comparative reviews satiate their need of both description about the object under review and comparison between competitive choices, review length of comparative is more likely to affect positively on helpfulness. For suggestive and regular reviews, length and helpfulness exhibit the same relationship as for full sample. The tendency for these reviews is to be more helpful as their wordiness decreases. Essentially, a discourse carrying a piece of suggestion or advice seems more attractive and appealing if it is polite in its tone, and terse in its context. In mundane matters, benignly expressed and concisely uttered suggestions are more likely to catch audience favor.

7. Conclusion and future work

This study contributes to examine both qualitative and quantitative measure for their joint effects on review helpfulness. Built on relevant online review literature, four hypotheses were proposed (H1, H2, H3 and H4). Tripadvisor data set was used to test hypotheses. The data set included 1336 hotel reviews. We have tackled the problem of identifying morphological features of different types of reviews that contribute to the helpfulness of online reviews. Two quantifiable morphological factors, namely number of concepts per review and average number of concepts per sentence, were tested for their possible impact on helpfulness. We found out that both these features help in explaining review helpfulness. Additionally, we concluded that the impact of wordiness and average sentence length differs significantly among types of reviews (regular, comparative, and suggestive), bolstering our hypothesis that review type moderates the effect of wordiness and sentence length on review helpfulness.

Our empirical findings support views embedded in linguistic literature for wordiness and average sentence length. We found that sentence length increases the diagnosticity of a regular review more than comparative and suggestive reviews and this relationship is inverse for wordiness. For example, writing a review that describes the service quality features of a particular hotel would usually be less wordy than the one that compares it too with another hotel. Type B review performs dual roles of description and comparison, while the other reviews (A and C) are more descriptive. Performing descriptive and comparative roles simultaneously would usually require type-B review author to use more concepts.

Regarding future research, there are several limitations of the present approach. Our sample of hotel reviews was sufficient to support our findings. However, our results are strictly generalizable only to hotel reviews. For future work, researchers are encouraged to sample from a different domain. For example this would allow an analysis of moderating effect of reviews types from different domains. Also we suggest researchers for analyzing rating patterns of malicious users and evaluate their potential for detecting shilling attacks. To eliminate and identifying shilling profiles happens to necessary. Attackers construct duplicate profiles to destroy the online recommender systems. The attack identifying algorithms can be planned to hold the problem. Finally, another limitation is related to the variables we used in the study. We did not include the ratings associated with online reviews. It is possible that the results of helpfulness could be different for the opinions having high or low ratings. Therefore, future research must incorporate such characteristics into their models.

8. Implications

The results of this study have implications for tourists, hotel managers and researchers. One practical implication is that managers and customers are able to see the most helpful evaluations of their businesses on travel blogs, websites and forums. Ideally, everyone desires to see online reviews that are perceived more helpful and useful, as such reviews add potential values to business (P.-Y. Chen, Dhanasobhon, & Smith, 2001; Zehrer, Crotts, & Magnini, 2011). However, less helpful reviews can also be useful in some other aspects based on morphological characteristics. Relatively longer sentences have positive effect on helpfulness rather than too long reviews. On the other hand, a long review with suggestive clues that is countered by a comparative clue should be considered less helpful. Hence, helpful opinions of this nature are imperative for the hotel managers and customers for better decision-making.

This study successfully validated the proposed model and found the key to make an opinion helpful for readers. Therefore, the proposed model might be used as an alternative theoretical model for evaluating e-business success in future studies. This study paves the way for the discovery of novel linguistic patterns (Poria, Cambria, Gelbukh, Bisio, & Hussain, 2015) for the identification of opinion helpfulness. Monitoring the helpful reviews is a challenging task for both customers and service providers alike. They both need to realize that, if they choose to grab reviews based on their morphological properties, they will find useful information relating to required service and might derive measures for further improvement.

Morphological features have proved to be an important aspect in understanding the importance of helpfulness. This is becoming a standard against which decisions are based to derive maximum success in decision-making. The consumer will focus on helpful votes prior to purchase based upon a variety of information encapsulated in reviews provided on different blogs, forums and websites such as TripAdvisor.

Helpfulness is often viewed as a simple “yes/no” choice, but our findings provide evidence that it is also dependent upon the review type and morphological markers. As customer reviews are widely used, our findings imply that it is important to recognize that regular, comparative and suggestive may make different information-consumption choices. The results of this study can be used to develop guidelines for creating more valuable online reviews. The review classification is particularly useful to hotel managers and tourists to get maximum benefit from each type and recognize the importance of each regular, comparative and suggestive review. The right choice according to user desire may lead them towards better decisions and ultimately enhance business intelligence.

Acknowledgements

This research is supported by University of Malaya under the research grant PPP for the project RP026-14AET.

References

Alkhattabi, M., Neagu, D., & Cullen, A. (2011). Assessing information quality of e-learning systems: a web mining approach. *Computers in Human Behavior, 27*(2), 862–873.

Arazy, O., & Kopak, R. (2011). On the measurability of information quality. *Journal of the American Society for Information Science and Technology, 62*(1), 89–99.

Cambria, E., Fu, J., Bisio, F., & Poria, S. (2015). In AAAI (Ed.), *AffectiveSpace 2: enabling affective intuition for concept-level sentiment analysis* (pp. 508–514). Austin.

Cambria, E., Gastaldo, P., Bisio, F., & Zunino, R. (2015). An ELM-based model for affective analogical reasoning. *Neurocomputing, 149*, 443–455. Part A(0).

Cambria, E., & Hussain, A. (2015). *Sentic computing: A common-sense-based framework for concept-level sentiment analysis*. Cham, Switzerland: Springer, ISBN 978-3-319-23654-4.

Cambria, E., Olsher, D., & Rajagopal, D. (2014). SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In *Paper presented at the twenty-eighth AAAI conference on artificial intelligence*.

Cambria, E., & White, B. (2014). Jumping NLP curves: a review of natural language processing research. *IEEE Computational Intelligence Magazine, 9*(2). pp. 48–57.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the “helpfulness” of online user reviews: a text mining approach. *Decision Support Systems, 50*(2), 511–521.

Chen, P.-Y., Dhanasobhon, S., & Smith, M. D. (2001). *All reviews are not created equal: The disaggregate impact of reviews and reviewers at amazon. com*.

Chen, C. C., & Tseng, Y.-D. (2011). Quality evaluation of product reviews using an information quality framework. *Decision Support Systems, 50*(4), 755–768.

Cruz, F. L., Troyano, J. A., Enríquez, F., Ortega, F. J., & Vallejo, C. G. (2010). A knowledge-rich approach to feature-based opinion extraction from product reviews. In *Proceedings of the 2nd international workshop on Search and mining user-generated contents* (pp. 13–20). ACM.

Cui, X., & Ryan, C. (2011). Perceptions of place, modernity and the impacts of tourism—Differences among rural and urban residents of Ankang, China: a likelihood ratio analysis. *Tourism Management, 32*(3), 604–615.

Danescu-Niculescu-Mizil, C., Kossinets, G., Kleinberg, J., & Lee, L. (2009). How opinions are received by online communities: a case study on amazon. com helpfulness votes. In *Proceedings of the 18th international conference on World wide web* (pp. 141–150). ACM.

Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233–242.

Elwalda, A., Lü, K., & Ali, M. (2016). Perceived derived attributes of online customer reviews. *Computers in Human Behavior, 56*, 306–319.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291–313.

Ganapathibhotla, M., & Liu, B. (2008). Mining opinions in comparative sentences. In *Proceedings of the 22nd International conference on computational linguistics* (pp. 241–248). Association for Computational Linguistics.

Gangemi, A., Presutti, V., & Reforgiato Recupero, D. (2014). Frame-based detection of

opinion holders and topics: a model and a tool. *Computational Intelligence Magazine, IEEE, 9*(1), 20–30.
Garner, B. A. (2001). *Legal writing in plain english: A text with exercises*. University of Chicago Press.
Golicic, S. L., Fugate, B. S., & Davis, D. F. (2012). Examining market information and brand equity through Resource-Advantage theory: a carrier perspective. *Journal of Business Logistics, 33*(1), 20–33.
Hariharan, S., Srimathi, R., Sivasubramanian, M., & Pavithra, S. (2010). Opinion mining and summarization of reviews in web forums. In *Proceedings of the third annual ACM Bangalore conference* (p. 24). ACM.
Hart-Davidson, W., McLeod, M., Klerkx, C., & Wojcik, M. (2010). A method for measuring helpfulness in online peer review. In *Proceedings of the 28th ACM international conference on design of communication* (pp. 115–121). ACM.
Hou, F., & Li, G. H. (2008). Mining chinese comparative sentences by semantic role labeling. In *Machine learning and cybernetics, 2008 international conference on (Vol. 5, pp. 2563-2568)*. IEEE.
Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 168–177). ACM.
Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior, 48*, 17–27.
Jindal, N., & Liu, B. (2006a). Identifying comparative sentences in text documents. In *Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 244–251). ACM.
Jindal, N., & Liu, B. (2006b). Mining comparative sentences and relations. In *Proceedings of the National Conference on Artificial Intelligence* (Vol. 21, p. 1331). Menlo Park, CA; Cambridge, MA; London: AAAI Press; MIT Press; 1999.
Johnson, E. J., & Payne, J. W. (1985). Effort and accuracy in choice. *Management science, 31*(4), 395–414.
Kennedy, P. (2003). *A guide to econometrics*. MIT Press.
Kim, S.-M., Pantel, P., Chklovski, T., & Pennacchiotti, M. (2006). Automatically assessing review helpfulness. In *Proceedings of the 2006 conference on empirical methods in natural language processing* (pp. 423–430). Association for Computational Linguistics.
Kohli, R., Devaraj, S., & Mahmood, M. A. (2004). Understanding determinants of online consumer satisfaction: a decision process perspective. *Journal of Management Information Systems, 21*(1), 115–136.
Kumar, R. (2011). Making machine translations polite: the problematic speech acts. In *Information Systems for Indian Languages* (pp. 185–190). Springer.
Li, M., Huang, L., Tan, C.-H., & Wei, K.-K. (2013). Helpfulness of online product reviews as seen by consumers: source and content features. *International Journal of Electronic Commerce, 17*(4), 101–136.
Li, S., Lin, C. Y., Song, Y. I., & Li, Z. (2010). Comparable entity mining from comparative questions. In *Proceedings of the 48th annual meeting of the association for computational linguistics* (pp. 650–658). Association for Computational Linguistics.
Liu, B. (2010). Sentiment analysis and subjectivity. In *Handbook of natural language processing* (Vol. 2, p. 568).
Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1–167.
Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer: analyzing and comparing opinions on the Web. In *Proceedings of the 14th international conference on World Wide Web* (pp. 342–351). ACM.
Liu, Y., Huang, X., An, A., & Yu, X. (2008). Modeling and predicting the helpfulness of online reviews. In *Data mining, 2008. ICDM'08. Eighth IEEE international conference on (pp. 443-452)*. IEEE.
McKechnie, S. (1992). Consumer buying behaviour in financial services: an overview. *International Journal of Bank Marketing, 10*(5), 5–39.
Mudambi, S. M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon. com. *MIS quarterly, 34*(1), 185–200.
Otterbacher, J. (2009). 'Helpfulness' in online communities: a measure of message quality. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 955–964). ACM.
Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on empirical methods in natural language processing* (Vol. 10)79-86): Association for Computational Linguistics.
Popescu, A. M., & Etzioni, O. (2005). Extracting product features and opinions from reviews. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 339–346). Association for Computational Linguistics.
Poria, S., Cambria, E., Gelbukh, A., Bisio, F., & Hussain, A. (2015). Sentiment data flow analysis by means of dynamic linguistic patterns. *IEEE Computational Intelligence Magazine, 10*(4). pp. 26–36.
Poria, S., Gelbukh, A., Cambria, E., Yang, P., Hussain, A., & Durrani, T. (2012). Merging SenticNet and WordNet-Affect emotion lists for sentiment analysis. In *Proceedings of ICSP* (pp. 1251–1255).
Poria, S., Gelbukh, A., Cambria, E., Das, A., & Bandyopadhyay, S. (2012). Enriching SenticNet polarity scores through semi-supervised fuzzy clustering. In *Proceedings of ICDM, Brussels* (pp. 709–716).
Qazi, A., Raj, R. G., Tahir, M., Cambria, E., & Syed, K. B. S. (2014). Enhancing business intelligence by means of suggestive reviews. *The Scientific World Journal*.
Qazi, A., Raj, R. G., Tahir, M., Waheed, M., Khan, S. U. R., & Abraham, A. (2014). A Preliminary investigation of user perception and behavioral intention for different review types: customers and designers perspective. *The Scientific World Journal*.
Quigley, M. (2008). *Encyclopedia of information ethics and security: IGI global*.
Sparks, B. A., Perkins, H. E., & Buckley, R. (2013). Online travel reviews as persuasive communication: the effects of content type, source, and certification logos on consumer behavior. *Tourism Management, 39*, 1–9.
The database and information system laboratory. (2010). *Retrieved 03-03-2014, 2014*.
Thet, T. T., Na, J. C., & Khoo, C. S. G. (2010). Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science, 36*(6), 823–848.
Vahdat, M., Oneto, L., Anguita, D., Funk, M., & Rauterberg, M. (2016). Can machine learning explain human learning? *Neurocomputing*.
Warfel, T., & Maurer, D. (2004). *Card sorting: A definitive guide. Boxes and arrows*.
Xu, K., Liao, S. S., Li, J., & Song, Y. (2011a). Mining comparative opinions from customer reviews for competitive intelligence. *Decision Support Systems, 50*(4), 743–754.
Xu, K. S. J., Wang, W., Ren, J., Xu, J. S. Y., Liu, L., & Liao, S. (2011b). Classifying consumer comparison opinions to uncover product strengths and weaknesses. *International Journal of Intelligent Information Technologies (IJIIT), 7*(1), 1–14.
Yaari, E., Baruchson-Arbib, S., & Bar-Ilan, J. (2011). Information quality assessment of community generated content: a user study of Wikipedia. *Journal of Information Science*, 0165551511416065.
Zehrer, A., Crotts, J. C., & Magnini, V. P. (2011). The perceived usefulness of blog postings: an extension of the expectancy-disconfirmation paradigm. *Tourism Management, 32*(1), 106–113.
Zhang, R., & Tran, T. T. (2010). A novel approach for recommending ranked user-generated reviews. *Advances in artificial intelligence*, 324–327.