Received January 28, 2019, accepted February 19, 2019, date of publication February 25, 2019, date of current version March 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2901472

Roles of Review Numerical and Textual Characteristics on Review Helpfulness Across Three Different Types of Reviews

YUSHENG ZHOU AND SHUIQING YANG
Department of Information Management and Engineering, Zhejiang University of Finance & Economics, Hangzhou 310018, China
Corresponding author: Shuiqing Yang (yangshuiqing@zufe.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 71810107003, Grant 71704153, and Grant 71701180, and in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY18G020013 and Grant LY18G020014.

ABSTRACT Understanding what factors make a helpful online review is critical to increase sales and drive revenue for online retailers. This paper examined the impacts of both reviews' numerical and textual characteristics on review helpfulness across three different review types including comparative, suggestive, and regular reviews. With an analysis of 30 338 product reviews collected from Amazon.com, the results indicated that the effects of numerical characteristics of reviews on review helpfulness are stronger for regular reviews than those for suggestive and comparative reviews. The impacts of text sentiment on review helpfulness are more significant for suggestive and comparative reviews when compared with regular reviews. Moreover, the text complexity of reviews has a significant invert U-shaped relationship with review helpfulness, and the relationships are stronger for regular reviews when compared with suggestive and comparative reviews. Furthermore, text sentiment has a negative effect on review helpfulness, and the effect is stronger for suggestive reviews than that for comparative and regular reviews. Finally, we employ a random forest method to predict review helpfulness based on its numerical and textual characteristics. This paper found that review length is the most helpful factor in predicting the helpfulness of online reviews. Our findings also indicated that the importance of numerical characteristics is greater than that of textual characteristics across three different review types. The theoretical and practical implications of the findings are presented.

INDEX TERMS Online reviews, review helpfulness, review type, electronic commerce, random forest.

I. INTRODUCTION

Review helpfulness is one of the most important concepts related to online customer reviews [1], [2], which have become growingly critical to customer purchase decisions. While the huge expansion of online reviews provided abundant product-related information for customers, the problem of information overload subsequently arises [3]. The overwhelming quantity of product reviews prevented customers to find the information they need, which greatly increases the risk and uncertainty related to consumer purchasing decisions. In order to eliminate the uncertainty and reduce the risk of decision-making, it is extremely necessary to find useful reviews, which are diagnostic and extremely important in customer purchase decisions [4].

Given the importance of product review helpfulness, a growing body of studies have been dedicated to explore what factors may affect the helpfulness of online reviews [2], [5]–[7]. In academia, the review helpfulness related studies have mainly focused on two perspectives: the number-special characteristics (e.g. review rating and review length) [8] and the text-special characteristics (e.g. review sentiment and review complexity) [9] of online reviews. The numerical characteristic of a review gives readers an impression before they can read the review carefully, while the textual characteristic of the review affects readers' review helpfulness evaluations when they read the review. Many previous studies have examined the impacts of the two types

The associate editor coordinating the review of this manuscript and approving it for publication was Hong-Mei Zhang.

of review characteristics on review helpfulness independently. However, they overlook the fact that the numerical and textual review characteristics are not exactly independent of each other but affect the customer's reviews helpfulness evaluations together. Motivated by the issue whether and how these review characteristics may affect consumers' evaluations of review helpfulness, the present study intends to extend existing research on review helpfulness by considering the effects of both its numerical characteristics and textual characteristics.

In addition, the present study also examines how different types of review opinion may affect the relationships between the review characteristics and review helpfulness. The review types are generally defined based on their linguistic construct [10], which expresses different opinions of reviewers with regard to product features [11]. There are generally three types of review opinion: comparative opinion, suggestive opinion, and regular opinion. A comparative opinion is a review that consumer tell the similarity or difference between two or more products [12], while a suggestive opinion, generally expresses an advice that the reviewer directs others to do or not to do something in a polite manner [13]. A regular opinion is usually referred to simply or abbreviated as an opinion [14]. Obviously, different types of product reviews can provide information on different aspects of the product to help customers or potential buyers make better decisions. For example, it is difficult for users to make decisions in a competitive business environment if they only look at one aspect of the product. Despite considerable effort has been devoted to understand the review helpfulness and consumers' purchase behaviors [15], the impacts of review opinions types on the relationships between review characteristics and helpfulness are still needed for a thorough investigation. The objective of our study is therefore to explore the impacts of the number-special and text-special characteristics on review helpfulness in the different review opinion types settings.

What is more, the present study adopted multi-analytic method to test the impacts of the numerical and textual review characteristics on review helpfulness across three different review types. Specifically, a negative binomial regression method was first employed to explore the influences of the numerical and textual characteristics on review helpfulness in three different datasets reflecting regular, comparative and suggestive reviews. Then, a random forest method was adopted to further predict the review helpfulness by ranking the important of these predictors. Specifically, the present study investigates:*(1) Do the number-special and text-special characteristics significantly shape review helpfulness, and if so, how? (2) How the impacts of the number-special and text-special characteristics on review helpfulness will be varied in the contexts of different review types? (e.g., comparative, suggestive, and regular reviews)*

This study contributes to the extant online review related literature in two ways. First, the present study explores the different impacts of the numerical and textual review characteristic on review helpfulness across three different review types including comparative, suggestive and regular reviews. Second, our study adopts a multi-analytic method including negative binomial regression and random forest method to investigate the helpfulness of online reviews. Employed a multi-analytic method would enhance the robustness of the findings of our research.

The structure of this paper is organized as follows. We presented the related work in the next Section. Section 3 develops the research hypotheses and put forward research model. Then it describes the research methodology in Section 4. It is subsequently followed by data analysis in Section 5, which discusses the results of hypothesis testing and random forest predicting. Section 6 concludes the discussion and conclusions.

II. LITERATURE REVIEW

A. REVIEW HELPFULNESS

Review helpfulness has been widely utilized as a tool to measure review's value in improving consumers' ability in assessing the quality of a product [5], [9], [16]. Mudambi and Schuff [5] have defined a helpful review as ''a peer-generated product evaluation that facilitates the consumer's purchase decision process''. Particularly, a useful review can create a greater value to potential consumers through offering appropriate information about the quality of product [16].

A growing body of study has devoted to investigating review helpfulness and what makes a helpful review [2], [3], [17]–[20]. With data collected from Amazon.com, Mudambi and Schuff [5] indicated that the perceived helpfulness of the review is affected by review extremity, review depth and product type. Review depth has a positive relationship with review helpfulness, and this relationship is moderated by product type (search or experience). Cao *et al.* [21] examined the effect of review rating, review length and reviews' posted time on review helpfulness. They found that reviews with extreme opinions are perceived as more helpful than those with neutral or mixed opinions. Huang *et al.* [22] stated that there is a threshold in the impact of review word count on review helpfulness. More recently, Eslami *et al.* [16] investigated review helpfulness based on reviews' specific characteristics (i.e., review length, review score, argument frame) and found that reviews associated with medium length, lower scores as well as non-positive argument frame are among the most helpful reviews. Siering *et al.* [23] found that both review content-related signals (i.e., review sentiment strength) and reviewer-related signals (i.e., reviewer ranking) affect the helpfulness of online review. Malik and Hussain [2] examined the impacts of review content and reviewer characteristics on review helpfulness. They found that syllables, the number of space, drives words, and aux verb in review text, as well as reviewer helpfulness per day and reviewer productivity score are significant predictors of review helpfulness. Ren and Hong [24] indicated that product type moderates the impact of three discrete emotions (sadness, fear, anger) on review helpfulness. They found that sadness embedded in a review negatively affects review helpfulness while fear embedded in a review positively influences review helpfulness. Moreover, anger embedded in a review has a more negative effect on review helpfulness for experience product than for search product. Malik and Iqbal [18] found that linguistic features including noun-singular, noun-general, preposition, personal-pronoun, and adverb are more influential predictors for helpfulness of review as a standalone model. Malik and Hussain [19] further examined the impacts of the features of review, reviewer and product category on review helpfulness. They found that these features are significant predictors for review helpfulness. Li *et al.* [20] found that reviewers' words have greater impacts on review helpfulness when compared with social relations. Although these studies provided significant insights into the effect of review and reviewer characteristics on the review helpfulness, little research has investigated the combination effects of numerical and textual characteristics on review helpfulness in the context of different types of reviews (e.g. regular, comparative and suggestive). The present study thus sheds light on this issue in online review literature.

In recent years, sentiment analysis has attracted increasing attentions from scientific communities and business circles [16], [25], [26]. Poria *et al.* [25] coupled ensemble classifier with a word-embedding model for sentiment analysis to improve the accuracy of model. Ma *et al.* [27] proposed a termed Sentic LSTM which outperforms state-of-the art methods in the tasks of targeted aspect sentiment. Additionally, the sentiment expressed in review is extremely crucial in online reviews, because it represents relevant information and influences consumers' purchase decisions [23], [28]. Cambria [26] suggested that "Emotions play an important role in successful and effective human-human communication". One of main factors in the area of sentiment analysis is business intelligence. In the context of online review, many researchers use the sentiment score of review to explain and predict the helpfulness of reviews [9], [16], [23]. On the other hand, previous studies also argue that text complexity that measures the degree of complexity in review content reflects the quality of review text from the cognitive perspective [29]. Therefore, text sentiment and text complexity are used as textual characteristics in this present study.

B. REVIEW TYPES AND REVIEW HELPFULNESS

Opinions are one of the main drivers of human behaviors [14]. Consumers can utilize the opinions to make better purchase decisions in that these opinions provide by consumers who bases on their own purchase experience and use experience [3]. Compared with the information provided by the merchant, these opinions are considered to be more credible and appealing by the consumer [30]. Early research on product reviews has studied and identified two types of opinions, namely comparative opinion and regular opinion [12]. This dual classification continued for a while until some scholars identified the suggestive opinion, a third type of review [14]. The classification of these three types of reviews can provide different levels of information to consumers and online retailers [10].

As we mentioned above, sentiment analysis reviews have different types: regular opinion, aiming at a single object [14], and comparative opinions, discussing more than one entity [12]. The regular opinion is significantly utilized to discover positive or negative points about a particular product, while the comparative opinion is usually used for competitive intelligence [31]. There are many existing works that examine the effect of regular opinions on review evaluations, which basically cover many aspects of the regular review [14]. Originally proposed by Jindal *et al.* [12], the comparative opinion has been examined by a number of online reviews related studies [32], [33]. Regular opinion and comparative opinion are also the two basic types of review established in previous studies. Recently, Qazi *et al.* [13] have examined a third significant type of review, namely suggestive review.

Different types of reviews are a key determinant of search costs, especially when consumers are confronted with thousands of product reviews. It is imperative to classify the reviews accurately, because different types of reviews convey different information to consumers, which lead them to make different consumption choices [34]. For example, an example of a regular review is "the memory of cell phone A is very good". Another example is "the memory of mobile phone A is better than that of mobile phone B". Obviously, these two examples provide different information and the latter is more referential. Therefore, it is important to distinguish between comparative opinions in review analysis, since direct comparisons are more persuasive than opinions on a single individual product. The suggestive review is characterized by recommendations for solutions to specific problems of an object or group of objects [14]. For instance, "Do not buy this phone, bad touch screen, bad system, bad resolution, it is absolutely terrible".

III. RESEARCH MODEL AND HYPOTHESES

Based on existing studies, the present study proposed a research model which reflects the effects of review characteristics and review types on review helpfulness. Specifically, as shown in Fig. 1, the reviews' numerical characteristics including rating and length, as well as textual characteristics including sentiment and complexity will affect review helpfulness. In addition, three types of reviews including regular, suggestive and comparative reviews were identified. Then the present study examines whether the impacts of the reviews' numerical characteristics and textual characteristics on review helpfulness were different across these different review types.

A. NUMERICAL CHARACTERISTICS OF ONLINE REVIEW

Numerical characteristics of customer reviews refer to the word count of review, average sentence length, review rating, and other numeric-related aspects of information [35].

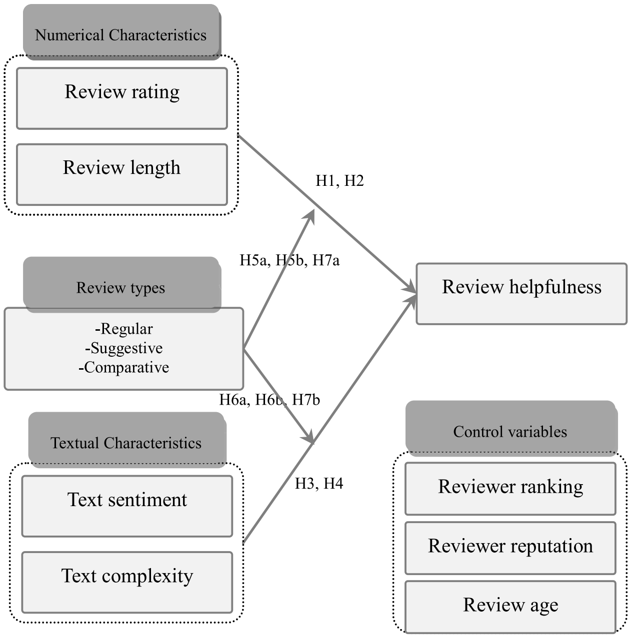

FIGURE 1. Research model.

They could give customers a visual impression before customers read review content. However, not all numerical characteristics are relevant to the context of this study. Review rating can be measured using the quantity of stars of each review in Amazon.com. For text-based information, the word counts of a review seem to be a suitable measure.

Previous studies suggested that reviews with low rating have a stronger impact on review helpfulness than ones with high rating [9], [36], [37]. This view is mainly based on the following three theories. Firstly, based on the Prospect Theory, people are more sensitive to loss than to gain [38]. Negative reviews mean that users would lose profit after purchasing products. In the interest of risk aversion and avoidance, review readers would pay more attention to such reviews and tend to think that these reviews are more informative. Secondly, based on the Attribution Theory, customer will infer the motivation of reviewer when they read the review [39], [40]. Generally speaking, customers tend to attribute negative reviews to internal factors (product) and consider them credible, while attributing positive reviews to external factors (reviewers) and consider them unauthentic [1]. Thirdly, based on the Information Diagnostic Theory, consumers generally believe that positive reviews are less diagnostic than negative ones [41]. Hence reviews with negative rating are thought to be more helpful. Therefore, we propose the following hypothesis:

H1: Review rating will negatively affect product review helpfulness.

Reviews containing more words and concepts can be perceived higher information diagnosticity and therefore, are perceived more helpful [9]. When consumers have no enough confidence to make decisions on purchasing a product, they will spend more time on evaluating the product. People are irresolute and hesitant in making decisions when information is less diagnostic [5]. The added length of information can increase the diagnosticity of reviews, thus increasing the customer's confidence in the purchase decision. Moreover, reviews with a great deal of words usually include more details about product features. Therefore, the following hypothesis was developed:

H2: Review length will positively affect product review helpfulness.

B. TEXTUAL CHARACTERISTICS OF ONLINE REVIEW

Textual characteristics of online reviews refer to the content, text complexity, text sentiment, and any other text-related aspects of information. The textual characteristics enable customers to have a cognitive understanding of the review after reading it. For text sentiment of online reviews, the previous studies have suggested it has as asymmetric impact on the helpfulness of a review. In particular, negative reviews have high customer perceived helpfulness than positive reviews [9], [36]. One significant interpretation is that positive reviews are less diagnostic than negative ones [42]. Therefore, it is expected that reviews with positive sentiments are expected to be less helpful than those with negative sentiments. The following hypothesis was therefore proposed:

H3: Review sentiment will negatively affect product review helpfulness.

Understandability is defined as the level of comprehension that an online review requires so that consumers can understand and make informed decisions after they have read the review [29]. In general, the understandability of review text is directly related to its complexity, which lies in how hard it is to read and understand a piece of text that is relevant to a product's features and functions. Previous study has suggested that a simpler piece of review text will be considered more helpful than difficult one [6]. On the one hand, the cognitive abilities of consumers, especially the cognitive ability for review text, are so limited that they just have a normal professional level in evaluating products [1]. Theoretically, cognitive matching occurs when the information contained in the review text matches the client's own information processing strategy. On the other hand, the complexity of the text plays an important role in assessing the rationality of the review. Therefore, we hypothesize:

H4: Review complexity will negatively affect product review helpfulness.

C. REVIEW TYPES AND REVIEW HELPFULNESS

The online product review rating system allows customers to talk about their experience of purchasing and their suggestions for their purchased products in a narrative manner. As such, reviews that contain various contents will form different styles. Review type was then introduced to capture the opinion expressed by various reviews on the online shopping platform [13]. It is expected that the impact of review rating on review helpfulness may be stronger for suggestive reviews when compared with comparative reviews. As we mentioned earlier, suggestive reviews are mostly used to express reviewers' advice about one product, whereas comparative reviews are usually utilized for discussing more than one product [14]. Since the review rating is only for a single object or entity, it is thus expected that the impact of review rating on review helpfulness will be weaker for comparative reviews than for suggestive reviews.

Review length refers to the overall word count of a review [43]. As mentioned earlier, review length can improve the diagnosticity of information and thus, positively affect review helpfulness [21]. However, review length may not be equally important for all review types, and may depend on whether the reader is reading a suggestive or comparative review. With reference to the numerical characteristics of review rating, it is expected that while longer review is more helpful, the incremental worth of review length for suggestive reviews may be greater than the incremental value of review length for comparative reviews. Therefore, we hypothesize:

H5: The effect of review rating (A) and review length (B) on review helpfulness will be stronger for suggestive reviews, when compared with comparative reviews.

Sentiment embedded in the review text can provide more acquiescent and text-specific emotions of the reviewers, beyond the review rating [44]. Previous study has suggested that negative reviews are more diagnostic and therefore more helpful than positive ones [1]. In the context of the present study, it is expected that the effects of text sentiment on review helpfulness to be different between suggestive reviews and comparative reviews. For comparative reviews, the social presence provided by reviews may be very important. In line with the social comparison theory, everyone has the motivation to compare themselves with others. In a retail environment, consumers often look for social cues from other consumers [5], while comparative reviews can be seen as making a declaration about the reviewer's preference and taste. In contrast, a suggestive review is utilized to direct someone to do something (in our case, to buy or not to buy) in a straightforward or euphemistic manner. In other words, the sentiments embedded in suggestive reviews are more persuasive than that in comparative reviews. Therefore, we argue that the impact of text sentiment on review helpfulness may be stronger for suggestive reviews when compared with comparative reviews.

Text complexity of a review is an important predictor of its helpfulness [6]. Simple reviews are more likely to be superficial and lack a comprehensive assessment of product features. Complicated reviews, in contrast, are more likely to contain more information about product features and in-depth analysis of the product. However, reading complicated reviews may increase consumer's search costs through decreased information diagnosticity [9]. Therefore, it is expected that review complexity will have a negative impact on the helpfulness of a review, but the effects of text complexity on review helpfulness may be weaker for comparative reviews than for suggestive reviews. This is because those suggestive reviews can increase information diagnosicity by providing clear advice. Consequently, we hypothesize that:

H6: The effects of text sentiment (A) and text complexity (B) on review helpfulness will be stronger for suggestive reviews, when compared with comparative reviews.

Numerical characteristics and textual characteristics resemble quantitative and qualitative features of online reviews, respectively. However, the effect of these two characteristics on different types of reviews may not always be consistent [44]. In the context of the present study, when consumers read the suggestive and comparative reviews, they may form different evaluations on the reviews when compared with the regular reviews. To some extent, both numerical characteristics and textual characteristics affect consumers' review helpfulness evaluations by influencing consumers' perceptions, whereas both suggestive and comparative reviews can promote consumers' perception of product reviews [14]. Thus, we hypothesize that:

H7: The effects of numerical characteristics (A) and textual characteristics (B) on review helpfulness will be stronger for suggestive and comparative reviews, when compared with regular reviews.

IV. RESEARCH METHODOLOGY

A. DATA COLLECTION

The dataset of online reviews used for analysis in the present study was collected from Amazon.com. The reason for selecting Amazon.com as our test bed is that it is one of the most popular online retail websites with thousands of reviews, which provides an abundant setting for search on online reviews. Based on Nelson's research, the present study included both search and experience products [45], [46]. For search goods, consumers can have a more accurate and comprehensive understanding of the attributes and quality of the products based on the product information provided by the online retailer. They can browse online customer reviews just to verify whether the product information provided by the online retailer is true. For experience goods, consumers cannot have a comprehensive understanding of the product solely based on the product information provided by the online retailer. They need to combine the online customer reviews with product information to gain a more thorough understanding of the product to determine whether the product meets their preferences. Considering the target products must have a relatively large number of online consumer comments compared with similar products, cell phone and book were selected as the search and experience products in our study, respectively. In addition, these selections were also supported and adopted in [5], [6], and [47]. Following Salehan and Dan [9] sampling strategy, we collected reviews of 20 different products that had at least 100 reviews randomly. Then, a total of 31,357 reviews were collected. We eliminated 1,019 analysis reviews that did not have complete reviewer's personal information. Finally, we got 30,338 reviews, which include 13,551 for search products and 16,787 for experience products.

B. VARIABLES

Based on the prior studies, the present study operationalizes the variables of proposed model by using the dataset collected from Amazon.com. The dependent variable is review helpfulness, measured by the number of consumers who consider the review helpful. In other words, review helpfulness is measured by the overall helpful votes that response to the question "Was this review helpful to you". Besides, the explanatory variables are text complexity, text sentiment, review rating, review length, review type. Review rating was measured as the number of stars of the review. Review length was measured by the overall count of words of the review. Text complexity was calculated by the average sentence length of each review. Text sentiment of a review was automatically performed by the LIWC program [48]. Following Ludwig *et al.* [49], we derived the text sentiment score of a review by the following formula:

$$TextSentiment_i = \frac{PW_i - NW_i}{SW_i}$$

where the overall intensity of text sentiment in review i is represented by $\mathrm{TextSentiment}_i$; the count of positive words in review i is represented by PW_i; the count of negative words in review i is represented by NW_i; the count of all words in review i is represented by SW_i.

The review type was divided into three categories by the card sorting method, which involved artificial labeling of text content based on message and classified it into different groups that made sense to participants [50]. The whole classification process was completed in optimalworkshop.com, where we can use the provided OptimalSort tabs to create a card sort. Therefore, we could find the impact of different types of reviews on online review helpfulness by carrying out this classification [14].

We included reviewer ranking, reviewer reputation and review age as control variables. Reviewer ranking is a weighted calculation based on the overall quality and the number of reviewer's entire reviews, which is displayed in the reviewer's personal page. Reviewer reputation was measured by the overall number of helpful votes got by the reviewer [51]. Review age was measured by the number of days between comment collection date and comment release date. For example, the comment release date is September 8, 2017, and the comment collection date is September 27, 2017, then review age is 19 days.

Table 1 lists the descriptive statistics of the variables in the complete dataset. Table 2 displays a comparison of the descriptive statistics for the regular, suggestive and comparative reviews subsamples.

The average review is obviously positive, with an average review rating of 4.3075. In addition, as Table 2 shown, the average number of useful votes for both suggestive reviews and comparative reviews is larger than regular reviews, which is consistent with the common sense and judgment in practical matters.

TABLE 1. Descriptive statistics for the full sample.

Variable	Mean	SD	N
Review Rating	4.307535	1.235159	30338
Review Length	46.18287	76.84183	30338
Text Sentiment	8.253679	12.53072	30338
Text Complexity	14.51448	15.90124	30338
Reviewer Ranking	1.73E+07	1.76E+07	30338
Reviewer Reputation	55.25911	178.5808	30338
Review Age	1247.413	1395.079	30338
Helpfulness	1.182675	26.7526	30338

TABLE 2. Descriptive statistics and comparison of means for subsamples.

	Regular (N=18285)	Suggestive (N=3731)	Comparative (8322)
	Mean (SD)	Mean (SD)	Mean (SD)
Rating	4.34(1.21)	3.90(1.47)	4.40(1.11)
Length	24.26(24.87)	117.28(167.41)	62.45 (63.39)
Sentiment	10.01(14.63)	3.56 (5.77)	6.49 (8.19)
Complexity	12.01(12.56)	19.64 (22.80)	17.70 (17.50)
Ranking	1.71E+07(1.73 E+07)	1.77E+07(1.85 E+07)	1.76E+07(1.79 E+07)
Reputation	53.95(164.96)	66.18(180.69)	53.21(204.38)
Age	966.03 (1019.37)	1751.94(1834.78)	1639.45(1691.59)
Helpfulness	0.41(5.84)	4.47(61.13)	1.38(29.19)

Table 3 presents the correlation coefficient matrix of each variable. As shown in Table 3, the correlation coefficients of each variable are relatively low; the highest correlation coefficient is only 0.4871, indicating the multicollinearity was unlike a serious problem in our dataset.

C. ANALYSIS METHOD

Because the dependent variable (helpfulness) in this study is a count variable, it is not suitable to be analyzed by general Multiple Regression Model. Given that the variance of the dependent variable (26.75) is overwhelmingly greater than its mean (1.18), there may be over-dispersion problems that will lead to a lower standard error of the model parameter estimation. In this case, the Poisson regression model analysis method is also not suitable, because the Poisson regression model generally requires that the variance of dependent variables equals its mean. In contrast, negative binomial regression model is more appropriate, which can effectively correct

TABLE 3. Correlations of variables.

	1	2	3	4	5	6	7	8
Helpfulness	1							
Rating	-0.03	1						
Length	0.08	-0.06	1					
Complexity	0.02	-0.06	0.19	1				
Sentiment	-0.02	0.27	-0.20	-0.21	1			
Reputation	0.15	-0.07	0.05	-0.02	0.01	1		
Ln (Ranking)	-0.04	-0.03	-0.05	-0.01	0.01	-0.49	1	
Ln (Age)	0.02	0.17	0.22	0.14	-0.13	-0.08	0.16	1

the over-dispersion problem [52]. Therefore, referring to the method of Zhou and Guo [3], this study adopts Negative binomial regression model analysis method, and the estimation method adopts a robust standard error. Under large samples, robust standard error is often used regardless of heterogeneity among error terms. Even with heterogeneity problems, the use of robust standard errors can avoid incorrect interval estimates and statistical tests.

In the present study, in terms of review numerical characteristics, review rating was hypothesized to have a negative impact on review helpfulness (H1), and review length will have a positive impact on review helpfulness (H2). In terms of review textual characteristics, text sentiment was hypothesized to have a negative impact on review helpfulness (H3), and text complexity has a negative impact on the helpfulness of a review. In order to test the nonlinear effect of text complexity on review helpfulness, the present study included both a linear term (Complexity) and a quadratic term (Complexity^2) of text complexity in the research model.

$$\begin{aligned}\text{Helpfulness}\\ &= exp[\beta_0+\beta_1(\text{Rating}) + \beta_2(\text{Length}) + \beta_3(\text{Sentiment})\\ &\quad +\beta_4(\text{Complexity})+\beta_5(\text{Complexity}^2)+\beta_6(\text{Ranking})\\ &\quad +\beta_7(\text{Reputation}) + \beta_8(\text{Age})+\rho + \varepsilon]\end{aligned}$$

where

ρ = Product fixed effect

ε = Random error

V. RESULTS

A. HYPOTHESES TESTING

The present study examined the effect of numerical and textual characteristics on review helpfulness by using Negative binomial regression method. Table 4 summarizes the results for the model testing on the full dataset. In addition, the present study also examined the impacts of numerical and textual characteristics on review helpfulness in the regular, comparative and suggestive data sets, respectively. To do this, we extracted three subsamples, each of which corresponds to a type of review. The Negative binomial regression results for the datasets of regular, comparative and suggestive types were presented in Table 5.

TABLE 4. Negative binominal regression results for review helpfulness.

Variable	Coefficient	Std. error	z-Statistics	Sig.
Constant	0.811	0.138	5.890	0.000
Rating	-0.196	0.040	-4.910	0.000
Length	0.010	0.001	14.430	0.000
Sentiment	-0.010	0.003	-3.840	0.000
Complexity	0.007	0.004	1.460	0.144
Complexity^2	-3.90E-05	1.64E-05	-2.370	0.018
Reputation	0.002	3.85E-04	5.800	0.000
Ln (Ranking)	-0.157	0.031	-5.030	0.000
Ln (Age)	0.813	0.095	8.590	0.000

TABLE 5. Group analysis of helpfulness based on review type.

Variable	Regular (TY-A)	Suggestive (TY-B)	Comparative (TY-C)	T-value		
				TY-A VS TY-B	TY-A VS TY-C	TY-B VS TY-C
Rating	-0.210***	-0.183***	-0.046	40.98	346.15	167.20
Length	0.018***	0.006***	0.009***	434.95	421.69	188.55
Sentiment	-0.003	-0.042**	-0.011*	342.30	152.22	169.58
Complexity	0.002	0.001	0.010	12.75	94.61	53.00
Complexity^2	-1.06E-04*	-1.02E-05	-8.17E-05	128.52	30.57	51.88
Reputation	-3.90E-04	0.003***	0.004	578.88	469.54	34.94
Ln(Ranking)	-0.220***	-0.220	-0.042	0.55	284.48	155.43
Ln (Age)	0.974***	0.764***	0.529**	124.97	302.70	72.75

Note: T=(PC1-PC2)/[Spooled*SQRT(1/N1+1/N2)]; Spooled= SQRT[[(N1-1)/(N1+N2-2)]*SE$_1^2$+[(N2-1)/(N1+N2-2)]*SE$_2^2$]; SE = standard error of regression coefficient in the Negative binominal regression model; PC=regression coefficient in the Negative binominal regression model.
***p<0.001, **p<0.01, *p<0.05.

From Table 4, we found that review rating ($\rho = -0.196$, $p < 0.001$) and text sentiment ($\rho = -0.010$, $p < 0.001$) are negatively and significantly related to review helpfulness, which supports H1 and H3, respectively. In addition, the results show that review length ($\rho = 0.010$, $p < 0.001$)

has a significant positive impact on review helpfulness, indicating that the more words the review has, the more helpful the review is. Hence, H2 is supported. Finally, the results of Model 1 show that the square of the text complexity ($\rho = -3.90E\text{-}05$, $p < 0.01$) has a significant negative impact on review helpfulness, which means the text complexity of the review has an invert U-shaped relationship with review helpfulness. In other word, reviews with extreme complexity are less helpful than reviews with moderate complexity. Therefore, H4 is not supported.

We further conducted a comparison testing to examine the differences in the strength of regression coefficients among the three types of reviews. As presented in Table 5, most of regression coefficients were found to be significantly different among the three types of reviews. For suggestive reviews ($\rho = -0.183$, $p < 0.001$), review rating has a significant negative relationship with review helpfulness, while comparative reviews do not have a significant relationship with review helpfulness ($p > 0.05$). In particular, the effect of rating on review helpfulness is stronger for suggestive reviews than for comparative reviews. Hence, our hypothesis 5A is supported. However, the effect of review length on review helpfulness, in contrast, is weaker for suggestive reviews than for comparative reviews. Therefore, Hypothesis 5B is not supported.

Hypothesis 6 postulates that the effect of text sentiment and text complexity on review helpfulness will be stronger for suggestive reviews than comparative reviews. The results in Table 4 suggest that text sentiment has a negative impact on review helpfulness for both suggestive ($\rho = -0.042$, $p < 0.01$) and comparative ($\rho = -0.011$, $p < 0.05$) reviews. Moreover, as shown in Table 5, the effect of text sentiment on the helpfulness of online review is weaker for comparative reviews when compared with suggestive reviews. Thus, hypothesis 6A is supported. In contrast, the correlation between text complexity and review helpfulness was non-significant for suggestive ($\rho = -1.02E\text{-}05$, $p > 0.05$) and comparative ($\rho = -8.17E - 05$, $p > 0.05$) reviews. These results show that the ways in which text sentiment and text complexity are related to review helpfulness are similar for suggestive and comparative reviews. Hence, hypothesis 6B is not supported.

Hypothesis 7 assumes that the effects of numerical characteristics and textual characteristics on review helpfulness will be stronger for suggestive and comparative reviews when compared with regular reviews. However, as shown in Table 5, the results indicate that the effects of numerical characteristics on review helpfulness are weaker for suggestive and comparative reviews than for regular reviews. Thus, hypothesis 7A is rejected. For textual characteristics, the effect of review sentiment on helpfulness is significantly weaker for regular reviews than for suggestive reviews and comparative reviews, while the effect of text complexity on review helpfulness is stronger for regular reviews than for suggestive and comparative reviews. Therefore, Hypothesis 7B is also rejected.

TABLE 6. Robustness test results for review helpfulness.

Variable	(1)	(2)
Constant	0.811***	-1.036***
Rating	-0.196***	-0.201***
Length	0.010***	0.009***
Sentiment	-0.010***	-0.009***
Complexity	0.007	0.007
Complexity^2	-3.90E-05*	-3.86E-05*
Reputation	0.002***	0.002***
Ln (Ranking)	-0.157***	-0.172***
Ln (Age)	0.813***	0.236*
Order		-2.26E-04***

Note: ***p<0.001, **p<0.01, *p<0.05.

B. ROBUSTNESS CHECK

In online product rating systems, consumers are allowed to discuss their own opinions in an asynchronous way. Therefore, a sequence forms as product reviews arrive sequentially [3]. Specifically, the Matthew effect could kick in when review sequence is regard as a heuristic cue [53]. As such, helpfulness estimates may be sensitive to review sequence [3]. To verify if the estimated effect of numerical and textual characteristics is robust, we include the review order in our model to control for the possible sequential bias. Following the procedure conducted by Zhou and Guo [3], the present study ordered the reviews for a specific product by review date. For a given product, let d_r represent the day on which review r arrives, and then some reviews have the same value of d. For each d', we formed $S_{d'} \equiv \{r : d_r = d'\}$, which denotes the set of reviews for which $d_r = d'$. Then we operationalized the variable*review order*as order $(d') \equiv \sum_{d<d'} N(S_d) + 1$, where $N(S_d)$ is the cardinality of set S_d. As showed in Table 6, our main findings still hold after controlling the possible sequential bias effect. In addition, review order is proved to negatively affect the helpfulness of online reviews, which is consistent with the findings of Zhou and Guo [3].

The present study also tested the effect of review characteristics on review helpfulness in different product type conditions. The reviews were categorized into two categories based on product type: search product and experience product. The model was tested for the two product categories separately. As shown in Table 6, the complexity of reviews has a strong invert U-shaped relationship with review helpfulness for search product, but not for experience product. In addition, reviewer reputation has a significantly positive relationship with review helpfulness for experience product, while this relationship is non-significant for search product.

TABLE 7. Group analysis of helpfulness based on product type.

Variable	Search product	Experience product
Rating	-0.098***	-0.504***
Length	0.008***	0.013***
Sentiment	-0.009***	-0.019***
Complexity	0.007***	0.013
Complexity^2	-3.39E-05***	-7.32E-05
Reputation	3.67E-04	0.004***
Ln (Ranking)	-0.155***	-0.123**
Ln (Age)	1.584***	0.576 ***

Note: ***p<0.001, **p<0.01, *p<0.05.

C. PREDICTING THE MOST HELPFUL REVIEWS BY RANDOM FORESTS

To ensure the theoretical and potential practical contributions of this study, we employed a Random Forest that considers all review characteristics simultaneously to forecast its helpfulness. Moreover, online retailer can use the proposed Random Forest as a tool to rank the reviews based on their predicted helpfulness.

A random forest is a classifier consisting of a large number of trees, which is random in two aspects: (1) ''each tree is based on a random subset of the observations'', and (2) ''each split within each tree is created based on a random subset of candidate variables'' [54]. Random forests are very popular in the field of machine learning since they can handle a much great number of variables with relatively small amounts of observations. More importantly, random forests provide an evaluation of variable importance [55], [56]. A well-known variable importance metric in random forests is average impurity reduction and its analog, Gini importance.

We conceptualized review helpfulness counts in the original data as a binary variable, i.e., the helpfulness variable of the reviews which did not obtain helpful votes were set as ''0'' and the others were set as ''1''. During all the experiments, 80% of the data set was used as training set and the rest was for testing set. We build the predictive model for helpfulness of reviews and calculate the relative important ratio of numerical and textual characteristics based on average impurity reduction metric. R programming language is utilized for all model prediction and variable importance measurement.

Table 7 summaries the prediction performance of the trained Random Forests for the training and testing sets. The performance on the independent test sets with accuracy rates is all above the minimum accepted range of 80% [57] except the accuracy rate of suggestive set, which demonstrates that our trained Random Forests have a good performance to accurately predict review helpfulness based on the review characteristics.

TABLE 8. Prediction accuracy levels.

	Prediction Accuracy Rate on Training Set	Prediction Accuracy Rate on Testing Set
Regular	88.77%	85.69%
Comparative	86.58%	80.67%
Suggestive	81.34%	74.41%

TABLE 9. The relative importance of factors.

Predictor	Relative importance ratio		
	Regular	Suggestive	Comparative
Review length	100.00%	100.00%	100.00%
Review rating	80.97%	11.50%	29.66%
Review complexity	74.87%	7.76%	17.18%
Review sentiment	50.13%	2.85%	13.54%

Finally, we checked the relative importance of each numerical and textual characteristic. As shown in Table 8, among our four characteristics, across all three datasets, review length is the most important factor that affects review helpfulness. This findings are consistent with our Negative binominal regression results which is also supported by several prior studies in the field of online reviews [9], [16]. In addition, across three datasets, the importance of numerical characteristics is greater than textual characteristics and this result is stable in the three review types contexts.

VI. DISSCUSSION AND CONCLUSIONS

A. FINDINGS

In the present study, we try to analyze the impacts of both the numerical characteristics and textual characteristics of reviews on review helpfulness and identify the review type under which the effects of review characteristics on review helpfulness will be strengthened or weakened. Grounded in existing online review literature, a theoretical framework was developed and tested with 30,338 product reviews collected from Amazon.com. The results of our study indicate that, in terms of numerical characteristics of reviews, the rating of a review has a significant and negative impact on the review helpfulness, which is consistent with [8] and [44]. However, this effect is significant for regular and suggestive reviews, but not for comparative reviews. In addition, the impact of review length on review helpfulness is significant and positive for all the three types of online reviews, which suggests that review length is a robust indicator for supporting and guiding consumer purchasing decisions. Longer review has protentional to provide more helpful information, thus consumers' evaluations of review helpfulness on longer review would be higher. In addition, the length of reviews increases the diagnosticity of product reviews [5], in that it is easier for consumers to gather information on product quality prior to purchase.

In terms of textual characteristics of reviews, text sentiment has significant effects on review helpfulness. Specifically, reviews with negative sentiment are more helpful than reviews with positive sentiment for suggestive and comparative reviews, but not for regular reviews. More importantly, we found that review text complexity has an invert U-shaped relationship with review helpfulness. In other words, reviews with moderately difficult sentences are associated with higher levels of helpfulness than reviews with extremely simple or complex sentences. However, this effect was not found in the datasets of suggestive and comparative reviews.

Furthermore, our findings indicate that the effects of numerical characteristics and textual characteristics on review helpfulness are significantly different in different types of reviews. On the one hand, the effect of numerical characteristics on review helpfulness is stronger for regular reviews when compare with suggestive and comparative reviews. A plausible interpretation is that regular reviews provide less textual information than suggestive and comparative reviews, and thus the effect of numerical characteristics on review helpfulness is dominant in the regular reviews. On the other hand, text sentiment, one of textual characteristics, has a stronger effect on review helpfulness for suggestive reviews when compare with regular and comparative reviews. As we discussed earlier, the sentiments embedded in suggestive reviews which are utilized to help consumer to make a decision are straighter and more persuasive than that from other types of reviews. Additionally, text complexity is proved to have a stronger effect on review helpfulness for regular reviews than that for suggestive and comparative reviews. One plausible explanation is that the regular review could not provide sufficient information if review text is too simple. In contrast, if review text content is extremely complex, it would be difficult for reader to understand. However, both suggestive and comparative reviews are able to express unequivocal goals to readers with clear meaning. Thus, the effect of text complexity in these two types of review on review helpfulness is weaker.

In terms of the different products, the linear and nonlinear effects of text complexity on review helpfulness were both significant for search product, but nor for the experience product. This suggests that review readers are more likely to care about the review readability for the search product. In fact, the search product with standard parameter and capacity features will impress readers relying on their readability and clarity. In addition, the results show that the reviewer reputation significantly affects review helpfulness for the experience product, but not for the search product. This suggests that review readers who evaluate an experience product's review as helpful will rely on the reviewer's reputation and experience. The reason may be that it is more difficult for readers to judge the quality of an experience product review solely on the review features itself.

Finally, our results revealed that our proposed Random Forest model has a great performance in predicting the helpfulness of online review. Besides, we found that review length is the most important predictor in predicting review helpfulness, which is consistent with our regression results and the findings of Eslami *et al.* [16]. Furthermore, our results also suggested that the importance of numerical characteristics is greater than that of textual characteristics. One possible reason is that the numerical characteristics are more intuitive and allow the reader to make a fast assessment for the helpfulness of a review.

B. THEORETICAL CONTRIBUTIONS

The present study contributes to the literature from several aspects. First, the present study explores the combination effects of numerical and textual characteristics across three different review types. In consistent with [14], the present study found that different review types have different impacts on review helpfulness. For example, numerical characteristics including review rating and length are significant predictors of review helpfulness in regular reviews. Text sentiment was found to be significant predictors of review helpfulness in the suggestive and comparative reviews. Specifically, in terms of the coefficient and significant level, the impact of text sentiment on review helpfulness was stronger for suggestive reviews than that for comparative reviews. This suggests that text sentiment plays an important role in determining the helpfulness of the suggestion related reviews. In addition, the present study also reveals that the complexity of reviews has a strong invert U-shaped relationship between review helpfulness in the regular review context. To sum up, our results provide additional support to the impact of review characteristics on review helpfulness in different review types contexts.

Second, we adopted a multi-analytic method including linear regression and random forest method to explain and predict the helpfulness of a review. The results of the present study address a gap in current researches by identifying the characteristics of the most helpful reviews among regular, comparative and suggestive reviews. Especially, our study compared the importance of the impacts of numerical and textual characteristics on review helpfulness. The present study provided empirical evidence that the importance of numerical characteristics on review helpfulness is greater than that of textual characteristics. Our results also reveal that review length is the most important factor in affecting review helpfulness as potentially longer reviews contain more cues. Our multi-analytic investigation further enhanced credibility to our findings.

C. PRACTICAL CONTRIBUTIONS

The current study also has important practical implications for online retailers and rating system managers. First, online reviews are found to have important impacts on consumers purchase decision making and can benefit online retailers with enhanced revenue and product sales [3]. However, not all product reviews have the same utility on decision making and sales enhancing. The present study found that the reviews' numerical characteristics and textual characteristics both affect product review helpfulness. An important impaction of the finding is that online retailers and rating system managers should devote more effort to manage the reviews' numerical and textual elements. For instance, the reviews' length and complexity should be controlled within a reasonable range.

Second, the results of the present study show that the impacts of numerical characteristics including review rating and review length on review helpfulness are significantly stronger for regular reviews than for suggestive and comparative reviews. The implication for online retailers is that they should pay more attention to the length and rating of the regular reviews, especially longer negative regular reviews. In addition, text sentiment was found to be a significant predictor of review helpfulness for suggestive and comparative reviews. An important impaction for online retailers and rating system managers is that they can provide some influential and sentimental review note for guiding consumers' review writing.

Finally, in terms of search product reviews, the effects of text complexity on review helpfulness were both linearly and nonlinearly significant. The implication for online retailers is that they can take measures to restrict the maximum and minimum words number to manage text complexity and consequently the review helpfulness. In addition, reviewer reputation has positively influence on review helpfulness for the experience product, but not for the search product. This suggests that reviewer reputation can be regarded as an important predictor for distinguishing a helpful product review. The implication for online retailers is straightforward: they can design or renew their reviewer reputation systems to better capture the reviewers' reputation. For instance, online retailers can calculate their reviewer reputation by including a longer time span.

D. LIMITATIONS AND FUTURE RESEARCH

Like all other empirical researches, the present study suffers from some limitations. First, the data used in the present study was collected from a representative E-commerce platform, Amazon.com. As different kinds of product rating sites existed, such as, Ebay.com, Tmall.com and JD.com, future studies are encouraged to address these additional rating sites to test the generalizability of our findings.

Second, a major theme of the present study is to explore the effects of reviews' numerical and textual characteristics on review helpfulness. Although essential, the two sets of elements are insufficient to draw a complete picture of a helpful online review. The literature of behavioral science suggests that review order and social location are also potentially important determinates of review helpfulness [3], [58]. Therefore, future studies are encouraged to investigate how review order and social location in combining with reviews' numerical and textual characteristics affect review helpfulness.

Finally, another protentional limitation is the self-selection bias in our sample, in which we only include review posters. The opinions from the non-posting consumers who brought a particular product from current E-commerce sites did not analyzed in our study. This is the common drawback for studies on related topic that collected secondary data from online product review sites. Therefore, further empirical researches are encouraged to explore the issue of review helpfulness using other complementary research methods, such as survey or eye tracking experiments.

REFERENCES

[1] H. Hong, D. Xu, G. A. Wang, and W. Fan, "Understanding the determinants of online review helpfulness: A meta-analytic investigation," *Decis. Support Syst.*, vol. 102, pp. 1–11, Oct. 2017.

[2] M. Malik and A. Hussain, "An analysis of review content and reviewer variables that contribute to review helpfulness," *Inf. Process. Manage.*, vol. 54, no. 1, pp. 88–104, 2018.

[3] S. Zhou and B. Guo, "The order effect on online review helpfulness: A social influence perspective," *Decis. Support Syst.*, vol. 93, pp. 77–87, Jan. 2017.

[4] D. Yin, S. Bond, and H. Zhang, "Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews," *MIS Quart.*, vol. 38, no. 2, pp. 539–560, 2014.

[5] S. M. Mudambi and D. Schuff, "What makes a helpful review? A study of customer reviews on Amazon.com," *MIS Quart.*, vol. 34, no. 1, pp. 185–200, 2012.

[6] A. Y. K. Chua and S. Banerjee, "Analyzing review efficacy on Amazon.com: Does the rich grow richer?" *Comput. Hum. Behav.*, vol. 75, pp. 501–509, Oct. 2017.

[7] S. Karimi and F. Wang, "Online review helpfulness: Impact of reviewer profile image," *Decis. Support Syst.*, vol. 96, pp. 39–48, Apr. 2017.

[8] A. H. Huang, K. Chen, D. C. Yen, and T. P. Tran, "A study of factors that contribute to online review helpfulness," *Comput. Hum. Behav.*, vol. 48, pp. 17–27, Jul. 2015.

[9] M. Salehan and D. J. Kim, "Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics," *Decis. Support Syst.*, vol. 81, pp. 30–40, Jan. 2016.

[10] A. Qazi, R. G. Raj, G. Hardaker, and C. Standing, "A systematic literature review on opinion types and sentiment analysis techniques: Tasks and challenges," *Internet Res.*, vol. 27, no. 3, pp. 608–630, 2017.

[11] J. K. Rout, A. Dalmia, K.-K. R. Choo, S. Bakshi, and S. K. Jena, "Revisiting semi-supervised learning for online deceptive review detection," *IEEE Access*, vol. 5, no. 1, pp. 1319–1327, 2017.

[12] N. Jindal and B. Liu, "Identifying comparative sentences in text documents," in *Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2006, pp. 244–251.

[13] A. Qazi, R. G. Raj, M. Tahir, M. Waheed, S. U. R. Khan, and A. Abraham, "A preliminary investigation of user perception and behavioral intention for different review types: Customers and designers perspective," *Sci. World J.*, vol. 2014, Feb. 2014, Art. no. 872929.

[14] A. Qazi, K. B. S. Syed, R. G. Raj, E. Cambria, M. Tahir, and D. Alghazzawi, "A concept-level approach to the analysis of online review helpfulness," *Comput. Hum. Behav.*, vol. 58, pp. 75–81, May 2016.

[15] C. Forman, A. Ghose, and B. Wiesenfeld, "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," *Inf. Syst. Res.*, vol. 19, no. 3, pp. 291–313, 2008.

[16] S. P. Eslami, M. Ghasemaghaei, and K. Hassanein, "Which online reviews do consumers find most helpful? A multi-method investigation," *Decis. Support Syst.*, vol. 113, nos. 32–42, Sep. 2018.

[17] Y. Wang, J. Wang, and T. Yao, "What makes a helpful online review? A meta-analysis of review characteristics," *Electron. Commerce Res.*, pp. 1–28, Jun. 2018. doi: 10.1007/s10660-018-9310-2.

[18] M. Malik and K. Iqbal, "Review helpfulness as a function of linguistic indicators," *Int. J. Comput. Sci. Netw. Secur.*, vol. 18, pp. 234–240, Jan. 2018.

[19] M. Malik and A. Hussain, "Exploring the influential reviewer, review and product determinants for review helpfulness," *Artif. Intell. Rev.*, pp. 1–21, Oct. 2018. doi: 10.1007/s10462-018-9662-y.
[20] S.-T. Li, T.-T. Pham, and H.-C. Chuang, "Do reviewers' words affect predicting their helpfulness ratings? Locating helpful reviewers by linguistics styles," *Inf. Manage.*, vol. 56, no. 1, pp. 28–38, 2019.
[21] Q. Cao, W. Duan, and Q. Gan, "Exploring determinants of voting for the 'helpfulness' of online user reviews: A text mining approach," *Decis. Support Syst.*, vol. 50, no. 2, pp. 511–521, 2011.
[22] A. H. Huang, K. Chen, D. C. Yen, and T. P. Tran, "A study of factors that contribute to online review helpfulness," *Comput. Hum. Behav.*, vol. 48, pp. 17–27, Jul. 2015.
[23] M. Siering, J. Muntermann, and B. Rajagopalan, "Explaining and predicting online review helpfulness: The role of content and reviewer-related signals," *Decis. Support Syst.*, vol. 108, pp. 1–12, Apr. 2018.
[24] G. Ren and T. Hong, "Examining the relationship between specific negative emotions and the perceived helpfulness of online reviews," *Inf. Process Manage.*, to be published. doi: 10.1016/j.ipm.2018.04.003.
[25] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowl.-Based Syst.*, vol. 108, pp. 42–49, Sep. 2016.
[26] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, Mar./Apr. 2016.
[27] Y. Ma, H. Peng, and E. Cambria, "Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 5876–5883.
[28] M. W. Uhl, "Explaining US consumer behavior with news sentiment," *ACM Trans. Manage. Inf. Syst.*, vol. 2, no. 2, pp. 1–18, 2011.
[29] M. Smith and R. Taffler, "Readability and understandability: Different measures of the textual complexity of accounting narrative," *Accounting Auditing Accountability J.*, vol. 5, no. 4, pp. 183–208, 1992.
[30] K. Y. Goh, C.-S. Heng, and Z. Lin, "Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content," *Inf. Syst. Res.*, vol. 24, no. 1, pp. 88–107, 2013.
[31] X. Yang, G. Yang, and J. Wu, "Integrating rich and heterogeneous information to design a ranking system for multiple products," *Decis. Support Syst.*, vol. 84, pp. 117–133, Apr. 2016.
[32] K. S. J. Xu, W. Wang, J. Ren, J. S. Y. Xu, L. Liu, and S. Liao, "Classifying consumer comparison opinions to uncover product strengths and weaknesses," *Int. J. Intell. Inf. Technol.*, vol. 7, no. 1, pp. 1–14, 2011.
[33] K. Xu, S. S. Liao, J. Li, and Y. Song, "Mining comparative opinions from customer reviews for competitive intelligence," *Decis. Support Syst.*, vol. 50, no. 4, pp. 743–754, 2011.
[34] D. H. Zhu, Z. Q. Ye, and Y. P. Chang, "Understanding the textual content of online customer reviews in B2C websites: A cross-cultural comparison between the U.S. and China," *Comput. Hum. Behav.*, vol. 76, pp. 483–493, Nov. 2017.
[35] D. Yin, S. D. Bond, and H. Zhang, "Keep your cool or let it out: nonlinear effects of expressed arousal on perceptions of consumer reviews," *J. Marketing Res.*, vol. 54, no. 3, pp. 447–463, 2017.
[36] R. Ullah, N. Amblee, W. Kim, and H. Lee, "From valence to emotions: Exploring the distribution of emotions in online product reviews," *Decis. Support Syst.*, vol. 81, pp. 41–53, Jan. 2016.
[37] M. S. I. Malik and A. Hussain, "Helpfulness of product reviews as a function of discrete positive and negative emotions," *Comput. Hum. Behav.*, vol. 73, pp. 290–302, Aug. 2017.
[38] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econ., J. Econ. Soc.*, vol. 47, no. 2, pp. 263–291, 1979.
[39] Z. Chen and N. H. Lurie, "Temporal contiguity and negativity bias in the impact of online word of mouth," *J. Marketing Res.*, vol. 50, no. 4, pp. 463–476, 2013.
[40] S. Sen and D. Lerman, "Why are you telling me this? An examination into negative consumer reviews on the Web," *J. Interact. Marketing*, vol. 21, no. 4, pp. 76–94, 2007.
[41] M. Lee, M. Jeong, and J. Lee, "Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website: A text mining approach," *Int. J. Contemp. Hospitality Manage.*, vol. 29, no. 2, pp. 762–783, 2017.
[42] Z.-P. Fan, Y.-J. Che, and Z.-Y. Chen, "Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis," *J. Bus. Res.*, vol. 74, pp. 90–100, May 2017.
[43] Y. Pan and J. Q. Zhang, "Born unequal: A study of the helpfulness of user-generated product reviews," *J. Retailing*, vol. 87, no. 4, pp. 598–612, 2011.
[44] N. Hu, N. S. Koh, and S. K. Reddy, "Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales," *Decis. Support Syst.*, vol. 57, no. 1, pp. 42–53, 2014.
[45] P. Nelson, "Advertising as information," *J. Political Econ.*, vol. 82, no. 4, pp. 729–754, 1974.
[46] P. Nelson, "Information and consumer behavior," *J. Political Econ.*, vol. 78, no. 2, pp. 311–329, 1970.
[47] L. M. Willemsen, P. C. Neijens, F. Bronner, and J. A. de Ridder, "Highly recommended!' The content characteristics and perceived usefulness of online consumer reviews," *J. Comput. Mediated Commun.*, vol. 17, no. 1, pp. 19–38, 2011.
[48] K. S. Ziemer and G. Korkmaz, "Using text to predict psychological and physical health: A comparison of human raters and computerized text analysis," *Comput. Hum. Behav.*, vol. 76, pp. 122–127, Nov. 2017.
[49] S. Ludwig, K. De Ruyter, M. Friedman, E. C. Brüggen, M. Wetzels, and G. Pfann, "More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates," *J. Marketing*, vol. 77, no. 1, pp. 87–103, 2013.
[50] D. Spencer and T. Warfel, "Card sorting: A definitive guide," *Boxes Arrows*, Apr. 2004. [Online]. Available: http://boxesandarrows.com/card-sorting-a-definitive-guide/
[51] S. Banerjee, S. Bhattacharyya, and I. Bose, "Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business," *Decis. Support Syst.*, vol. 96, pp. 17–26, Apr. 2017.
[52] Z. Lin, M. W. Peng, H. Yang, and S. L. Sun, "How do networks and learning drive M&As? An institutional comparison between China and the United States," *Strategic Manage. J.*, vol. 30, no. 10, pp. 1113–1132, 2010.
[53] Y. Wan, "The matthew effect in social commerce," *Electron. Markets*, vol. 25, no. 4, pp. 313–324, 2015.
[54] U. Grömping, "Variable importance assessment in regression: Linear regression versus random forest," *Amer. Statist.*, vol. 63, no. 4, pp. 308–319, 2012.
[55] H. Ishwaran, "Variable importance in binary regression trees and forests," *Electron. J. Stat.*, vol. 1, pp. 519–537, Nov. 2007. doi: 10.1214/07-EJS039.
[56] C. Strobl, A.-L. Boulesteix, A. Zeileis, and T. Hothorn, "Bias in random forest variable importance measures: Illustrations, sources and a solution," *Bioinformatics*, vol. 8, no. 1, pp. 1–25, 2007.
[57] M. M. Ghiasi, A. Bahadori, S. Zendehboudi, A. Jamili, and S. Rezaei-Gomari, "Novel methods predict equilibrium vapor methanol content during gas hydrate inhibition," *J. Natural Gas Sci. Eng.*, vol. 15, pp. 69–75, Nov. 2013.
[58] R. P. Mohammadiani, S. Mohammadi, and Z. Malik, "Understanding the relationship strengths in users' activities, review helpfulness and influence," *Comput. Hum. Behav.*, vol. 75, pp. 117–129, Oct. 2017.

YUSHENG ZHOU is currently pursuing the master's degree with the School of Information, Zhejiang University of Finance & Economics. His main research interests include mobile commerce and mobile social networks.

SHUIQING YANG received the Ph.D. degree from the Huazhong University of Science and Technology, in 2012. He is currently an Associate Professor with the School of Information, Zhejiang University of Finance & Economics. His research has been published in *Decision Support Systems*, *Information and Management*, *Computers in Human Behavior*, *Industrial Management & Data Systems*, the *International Journal of Mobile Communications*, the *International Journal of Human–Computer Interaction*, and several other journals. His researches focus on electronic and mobile business, and technology adoption.