A Novel Approach to Predict the Helpfulness of Online Reviews

Morteza Namvar
The University of Queensland
m.namvar@business.uq.edu.au

Abstract

Online reviews help consumers reduce uncertainty and risks faced in purchase decision making by providing information about products and services. However, the overwhelming amount of data continually being produced in online review platforms introduce a challenge for customers to read and judge the reviews. This research addresses the problem of misleading and overloaded information by developing a novel approach to predict the helpfulness of online reviews. The proposed approach in this study, first, clusters reviews using reviewer-related, and temporal factors. It then uses review-related factors to predict online review helpfulness in each cluster. Using a sample of Amazon.com reviews, the empirical findings offer strong support to the proposed approach and show its superior predictions of review helpfulness compared to earlier approaches. The outcomes of this study help customers in online shopping and assist online retailers in reducing information overload to improve their customers' experience.

1. Introduction

In today's competitive business environment, establishing a strong online presence and reputation is of crucial importance in attracting new customers. Today's online shoppers are increasingly relying on online resources to benefit in purchasing decisions. They expect to have access to reviews, and they shape their pre-purchase behaviors by the product ratings that they observe [1].

Recent studies have shown that customers consult online reviews before they make their online purchase decision. In a study by Vlachos [2], 43 percent of participants indicated that they read online reviews before making a final purchase decision. In another study, Santos [3] reported that nearly half of online customers actively read and post reviews after experiencing service products.

Hence, consumers tend to rely on online reviews, as a form of word of mouth [3], that allow them to obtain sufficient information to reduce their level of perceived uncertainty in online shopping [5], and online reviews have received considerable attention, as they have been useful in the purchase experience of online customers [6]. Businesses, therefore, have found that providing a platform in which online customers can write reviews regarding the quality of the products or services is one way to establish trust and motivate online customers in purchasing products or services [7].

The problem of information overload and conflicting opinions, however, can confuse consumers with identifying and considering attributes relevant to their decision [8]. Also, since positive reviews can augur significant financial benefits, they offer incentives for organizations to engage in online malpractices such as opinion spamming [11]. Hence, it is questionable whether online reviews are helpful for other readers [7].

Mudambi and Schuff [9] define online review helpfulness as consumers' perceived value of online reviews while shopping online. They propose to use the ratio of helpful votes to total votes received as the measure of perceived review helpfulness. Most of the reviews newly posted have not received helpfulness voting by the readers as they are newly written [10].

Scholars have, therefore, investigated approaches to predict the helpfulness of online reviews. They have mainly studied online review helpfulness on the basis of review-, and reviewer-related factors, whereas temporal elements have been largely ignored. Theses studies have also proposed holistic predictive models for all possible review and reviewer segments.

This study develops temporal factors related to the reviews and uses them in clustering online reviews. It then develops different models in each cluster to predict the perceived helpfulness of customer online reviews.

For this purpose, it uses a dataset of Amazon.com customer reviews. By developing review clusters based on temporal-, and reviewer-related factors, this study distinguishes between reviews on various products, written on different product lifetime and by different reviewers. Further analysis shows that review clustering assists in classifying helpful and unhelpful reviews, and the proposed approach outperforms the existing works for classifying helpful reviews.

HĭCSS

The next section reviews the literature on the helpfulness of online reviews. Then, I present the research method and process. I further elaborate on the results of the study. The paper conclusion reveals the emerging insights into online reviews helpfulness, followed by opportunities for further research.

2. Background

Several studies have investigated the helpfulness of online reviews and have identified the deterministic factors in their helpfulness (e.g. [6], [11]–[15]). For example, Korfiatis [7] analyzed 37,221 reviews gathered from Amazon.com to determine the dependency between the proportions of valuable votes provided to the reviews and review content bombastic components. In another study, Karimi and Wang [16] analyzed 2,178 reviews about mobile games to investigate if the identity of the reviewer has impacts on customers' evaluations towards the helpfulness of online reviews. Similarly, Zhou and Guo [17] studied 70,610 hotel reviews gathered from Yelp.com to investigate the impacts of the order of review on its helpfulness.

These studies have identified different factors as important for the helpfulness of an online review. For example, Huang et al. [18] have found that the feeling of ease in reading led consumers to judge the reviewers as more helpful. In another study, Hernández-Ortega [19] analyzed whether the social-psychological distance affects the readers' responses to the reviews. Hong et al. [18] stated that expertise and reputation of the reviewer, depth of review, the age of review and disclosure of reviewer's identity have an impact on the helpfulness of an online review. They, however, claimed that review rating and readability of an online review does not affect the helpfulness of a review. Zhou & Guo [17] have identified length and sentiments of reviews as essential factors on the helpfulness of an online review. Furthermore, Karimi and Wang [16] have identified the use of visual cues (profile picture), the real name of a reviewer, the nationality of reviewers and length of reviews deterministic factors on the helpfulness of an online review.

This research identifies three categories of determinants of review helpfulness; namely 1) reviewer-related variables 2) temporal variables, and 3) review-related variables (see Table 1). The following sections elaborate on these categories by presenting the views of different researchers.

2.1. Reviewer-related variables

The reviewer's identity has an impact on readers' perception of online reviews. Reviewer-related factors include reviewer expertise and reviewer reputation [14]. In many e-commerce platforms, visitors provide feedback on the reviews which affect the reviewers' reputation [20]. Many hospitality and tourism websites have reputation evaluation system to check the reviewers' past reviews and make their decision of following or not following them; however, in many other websites, such as online retailers, this evaluation system does not exist.

Table 1 Determinants of factors influencing the helpfulness of online reviews

Category	Measure	Definition	Reference
Reviewer-related variables	Reviewer activity	Total number of reviews by a reviewer	[21], [16], [14]
	Reviewer helpfulness	Total number of helpfulness votes received about the reviews posted by a reviewer	[14]
Temporal variables	Recency	Total number of days since the review was published	[23], [22]
	Frequency	The number of reviews written on the product divided by the number of days between the first and the last review on the product	[20]
	Order	Total number of previous reviews posted about the product before the review	[17], [13]
Review-related factors	Sentiment	Attitude, thought or judgment expressed in the review	[24], [14], [25]
	Score (star rating)	The average score provided by different reviewers about the product	[20], [9], [22]
	Length	Total number of words in the review	[11]

This study defines the activity of reviewers as the total number of reviews published by them. It describes the helpfulness of reviewers as the total number of helpfulness votes received about the reviews posted by a reviewer.

2.2. Temporal variables

Online retailers, such as Amazon.com, tend to sort the reviews based on their helpfulness votes; hence, customers are more likely to read older reviews [6], [13]. Although this approach sounds helpful to the customer in avoiding information overload, old reviews receive more votes as compared to new ones due to their more extended visibility on the websites. Furthermore, it is hard to rank new reviews which have not received any votes [20]. In this line, Zhou and Guo [17] have investigated hotel reviews from Yelp.com and found that the order of review influence the helpfulness vote that they receive.

This study, therefore, considers three temporal variables that can have an impact on the helpfulness of online reviews, namely frequency, order and recency of reviews. Frequency of reviews on a product is the number of reviews written on the product divided by the number of days between the first and the last review on the product. This study measures recency based on the number of days between the day a selected review was posted and the day when the previous review about the product was posted. It measures order based on the total number of reviews written on the product before a selected review.

2.3. Review-related variables

Several researchers have investigated review-related features, and their impacts on the way reviews are written and consequently, the helpfulness of the online reviews (e.g. [20]). Score (star rating), review sentiment and review length are three factors related to the way review is written on the product, and have an impact on the helpfulness of online reviews.

Score (star rating) is an overall rating of a product by a reviewer, on a five-point scale reflecting the reviewer's attitudes toward a product. While a very low rating reflects an extremely negative experience, a very high rating represents a highly positive attitude toward a product [9].

Hong et al. [14] define review sentiment as the attitude, thought, or judgment expressed in online reviews. If the sentiment expressed in the review signifies the reviewer's product evaluation, the review should be perceived as more helpful by consumers as it clarifies whether they should consider the product should or not [25]. Consequently, the strength of the review sentiment influences the vote received by online reviews. Also, review depth, the total number of words in a review, effect the helpfulness of online reviews [26].

3. Research method

This research adopts a structured knowledge discovery process [27, p. 120] to predict the helpfulness of online reviews. Figure 1 shows the research process consisting of four main steps: data pre-processing, review clustering, helpfulness prediction, and model evaluation. The following sections elaborate on each of these steps.

3.1. Data pre-processing

This research uses a dataset of online reviews and follows four steps for data pre-processing; namely data cleaning, data integration, data reduction, and data transformation. It integrates review dataset with other available datasets. Then, it filters out missing values and other possible redundant records in the data reduction step. Applying techniques for transforming unstructured text data into a structured format [28], it operationalizes the study variables.

3.2. Cluster analysis

The aim of clustering or cluster analysis is to find clusters that encompass observations that are similar to one another and dissimilar to those observations in other clusters [27, p. 443]. To develop clusters that have small inter-point distances in relation to the distance to observations in other clusters, this research 1) identifies ideal variables for cluster analysis, 2) chooses the appropriate clustering techniques, and 3) determines the ideal number of clusters. The following sections elaborate on these steps and explain how this research applies them to cluster analysis and model building.
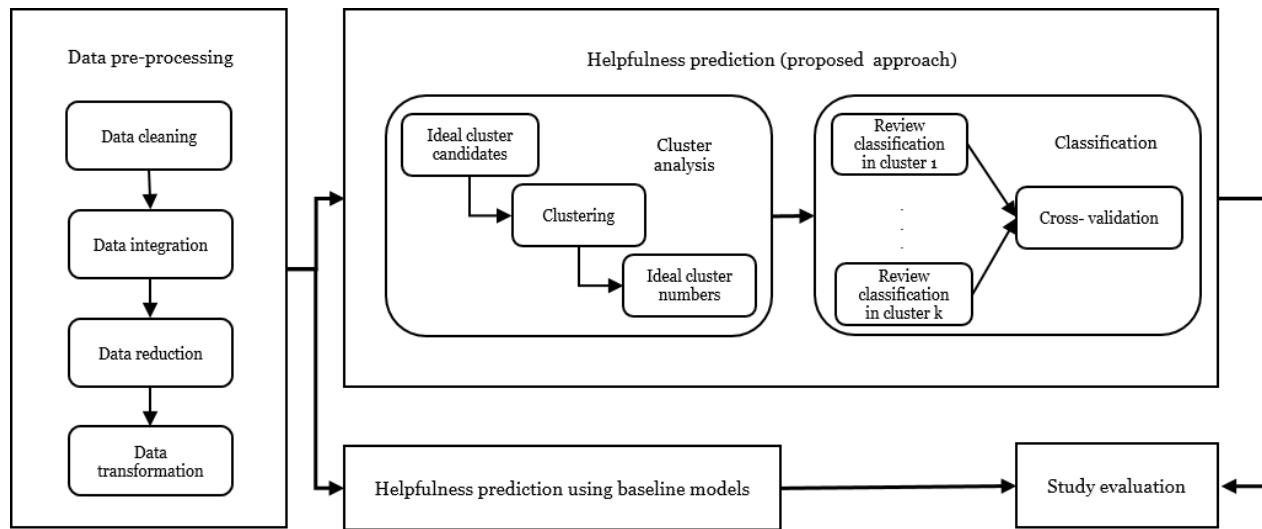
Figure 1 Research process

3.2.1. Ideal cluster candidates: The effectiveness of cluster analysis highly depends on the dimensionality of a dataset. If cluster analysis uses highly correlated features, the performance of clusters will be decreased [29]. While extracting many features might include irrelevant information or noise, extracting too few features result in a loss of information [30]. This research, therefore, uses principal component analysis [31] to transform the data into uncorrelated principal components.

3.2.2. Clustering: The objective function of most of the clustering algorithms is the sum of squared distances of each observation to its cluster center (e.g., K-means algorithm). The assignment of observation to a cluster with the closest cluster center will repeat until they minimize the objective function. This study examines several clustering algorithms, including Self-Organizing Maps [32], and K-means [33], to choose the best performing one for the dataset in hand.

3.2.3. Ideal cluster numbers: Clustering algorithm partitions the dataset into K clusters. For most of them, we need to specify the number of clusters, K, to conduct the clustering [27, p. 453]. Choosing the number of clusters is challenging as clustering algorithms prefer to increase the number of clusters to optimize the objective function [32]. Therefore, using this way of evaluating clusters will end up choosing as many clusters as observations are available [27, p. 453]. Scholars have used several approaches to determine a suitable number of clusters, including elbow method, gap statistic, and jump method. This study chooses the elbow method, as this approach is theoretically motivated and applies to a wide range of problems [34]. In the elbow method, a researcher

chooses the number of clusters in a way that adding another cluster doesn't provide much better interpretation of clusters. It looks at the percentage of variance explained as a function of the number of clusters.

3.3. Helpfulness prediction and evaluation

This research trains a dataset of online reviews using a classification algorithm to label online reviews. For each review, the classifier learns from a pre-processed set of online reviews with the assigned label. It uses cross-validation technique along with the neural network to avoid overfitting and achieve more accurate measures for classification performance [27, p. 370].

4. Empirical results

I used a dataset from Amozon.com, which includes 568,454 reviews from 256,059 users about 74,258 products. I integrated review dataset with product dataset, which is also gathered from Amazon.com, to provide more depth and context for the analysis. After I filtered out missing values, I had reviews of 15,582 products. I restricted the analysis to products with more than 30 reviews, due to the methodological requirements of this study (see the research method section). I also kept only review records which had been evaluated at least by 20 readers. The reviews are written on various products in different categories including grocery and gourmet food, health and household, pet supplies, and TV and movies.

After data filtering, the dataset reached to 4675 reviews on 1611 products written by 2687 different users, voted by 215,542 readers. From these, 85% (n = 4000) were helpful reviews, and 15% (n = 675) were unhelpful reviews.

In the next step, I applied six text processing techniques to transform the unstructured online reviews to structured format; namely stemming, editing stop words, N-Grams [35], Singular Value Decomposition (SVD) [36] and sentiment analysis. I evaluated both "Snowball" and "Porter" methods of stemming and found that with the available dataset, they produced identical results. I, however, selected the "Snowball" method as it demonstrated superior performance in other studies (e.g. [37]).

The initial text exploration showed that many of the words used in the reviews mentioned the products' name, and filtering these out could improve the analysis as I can focus on the remaining words in the review that were more likely to be directly related to the evaluation of the product. I also removed generic stop words and generated a stemmed word list. I then converted this to a custom stopwords dictionary and modeled with the sample review set. I extended the concept of stop words further and generated a wordlist from the entire set of reviews. When I sorted words by frequency, I found that many words did not add value and were related to the product rather than the review. These words were mainly a noun. Hence I assumed that verbs and adjectives are more useful in evaluating review helpfulness [38]. Therefore, I created another custom stop word list based on the words that had a frequency higher than 500, as I assumed they don't contribute to the quality of a review.

Next, I used n-Gram generation [35] to identify words that appear together. For example, "bad" would be considered positive; however, "not bad" changes the meaning. I modeled n-Grams of 2 and three words with control against our sample set. Then, I used Singular Value Decomposition (SVD) [36] to reduce the number of word attributes created by stemming. There are two main algorithms for SVD; Keep Percentage and Fixed Dimension. Keep percentage retains components that have a variance less than the specified threshold, and a fixed dimension retains the number of components specified [38]. I ran analysis across both with a range of attributes. The results suggested that the Keep Percentage with a threshold of 25% performs better than other options. Finally, I conducted sentiment analysis and associated a sentiment value, from extremely negative (-1) to extremely positive (+1), to each review.

4.1. Variable operationalization

After the data preparation step, I operationalized nine variable, which I describe next.

Review helpfulness (binary) is a binomial measure for helpfulness. It is the label of each review, and I use it for training the classification algorithm and evaluating the predictive or classification power of the proposed models and comparing them with baseline models. In this measure, following the instruction from Salehan and Kim's model [13], a review is labeled as helpful if the number of readers who voted the review as helpful is more than 60% of the total number of votes.

Reviewer activity is the total number of reviews written by the reviewer. *Reviewer helpfulness (%)* is the total number of readers who have mentioned that review(s) written by a specific reviewer have been useful divided by the total number of people who have mentioned whether the review(s) written by the reviewer has been useful.

Review recency is "1" divided by the number of days between the day a selected review was posted and the day when the previous review about the product was posted.

Frequency is the number of reviews written on the product divided by the number of days between the first and the last review on the product.

Review order (%) is the total number of reviews written on the product before a selected review divided by the total number of reviews written on the product.

Review sentiment is from extremely negative (–1) to extremely positive (+1), to each review, and quantifies the emotional direction of a review. *Score (star rating)* is an overall rating of a product in a review, on a five-point scale reflecting the reviewer's attitudes toward a product. *Review length* is the number of words in each review.

Table 2 presents the summary statistics of the measures of this study.

Table 2 Descriptive statistics of independent (predictor) varibales

Variables	Mean	SD
Reviewer activity	22.5	19.13
Reviewer helpfulness (%)	2.5	4.13
Recency	1.36	1.42
Frequency	252	117
Order (%)	0.45	0.32
Sentiment	0.10	0.11
Score (star rating)	3.138	1.40
Length	57.59	55

Table 3 Variable correlations

	Variable	1	2	3	4	5
1	Order	1.00				
2	Recency	-0.06	1.00			
3	Frequency	-0.03	**0.61**	1.00		
4	Reviewer activity	0.02	0.01	-0.03	1.00	
5	Reviewer helpfulness	-0.04	0.07	0.12	-0.14	1.00

4.2. Ideal cluster candidates

I use Helpfulness (binary) and Helpfulness (%), as dependent variables in various steps of the rest of analysis. 8 other variables presented in Table 2 are independent variables. 5 of independent variables are candidate ones for clustering analysis (see Table 3). I used 3 others, review-related variable, in classification.

To identify the ideal candidates for cluster analysis, I used the reduced dataset, which contains 4675 online reviews, to analyze the variable correlations and ensure that multicollinearity between 5 candidate variables does not drive the results.

I observed correlation when considering frequency and recency: when there are more frequent reviews on a product, the recency of reviews on the product increases (see Table 3). Apart from this correlation, there is not any significant correlation between the independent variables. I, therefore, removed frequency measure from cluster analysis and reduced the number of fo variables to 4.

Next, I investigated the impacts of 4 identified variables from the previous step in the helpfulness of online reviews (see Table 4). Considering the reviewer-related variables, I observed that reviews provided by experienced reviewers are more helpful than reviews contributed by less experienced reviewers. Furthermore, a reviewer activity has a significant influence on review helpfulness.

Table 4 Regression estimates explaining review helpfulness

Variable	Coefficient	Std. error	p-value
Order	0.0005	0.00007	0.000
Recency	-0.0006	0.00023	0.002
Reviewer activity	0.0004	0.00007	0.000
Reviewer helpfulness	0.0012	0.00007	0.000
Intercept	0.5149	0.01084	0.000

Table 4 provides strong evidence that the review order and recency play key roles in the interpretation of reviews. Specifically, this analysis finds that later reviews are likely to be more helpful than what the online platform shows and recent reviews are likely to be less helpful than what the system shows. In line with the previous works on online reviews, the analysis confirms the impact of the reviewer experience and activity on the helpfulness of online reviews.

All 4 variables in Table 4 show significant impacts on the helpfulness of online reviews. Hence, I use these variables for the clustering step, in which I turn next.

4.3. Review clustering and ideal cluster numbers

The k-means clustering algorithm is one of the first and widely applied clustering algorithms. K-means algorithm randomly chooses one observation for each cluster and uses it as the centroid for the initial cluster [27, p. 451]. Then, in an iterative process, it assigns each observation first to the nearest cluster and, second, it adjusts the cluster center to represent all observations in the cluster.

I used k-means algorithm to develop review clusters (see 3.2.2). This method uses the value of k, as determined by the user, to make a k number of clusters. First, I used an initial value for k by using the square root of the total number of records divided by two [27, p. 451]. In an attempt to find the optimal number of clusters, I used the elbow method to adjust the value for k. I evaluated the clusters using the Davies-Bouldin Index [39], which assesses intra-cluster similarity and inter-cluster differences. This method measures the average distance between the center of a cluster and the objects it contains. K-means desirer lower value of this index as lower values indicate a more precise grouping of records and higher differentiation between individual clusters. Since I did not find significant improvement by increasing the number of clusters (see Table 5),

I chose 3 clusters to achieve lower cluster numbers which assisted in easier interpretation of clusters and their profiles. After clustering using, K=3, the analysts resulted in 3213 reviews in cluster

1, 1064 reviews in cluster 2, and 398 reviews in cluster 3.

Table 5 Clustering performance (Davies-Bouldin Index)

Number of clusters	3	4	5	6	7
Average within centroid distance	0.059	0.074	0.065	0.058	0.052

I normalized all the 4 variables from 0 to 1 to improve the interpretability of clusters. Figure 2 shows the profile of clusters. Each column in this figure shows the average value of the associated variable.
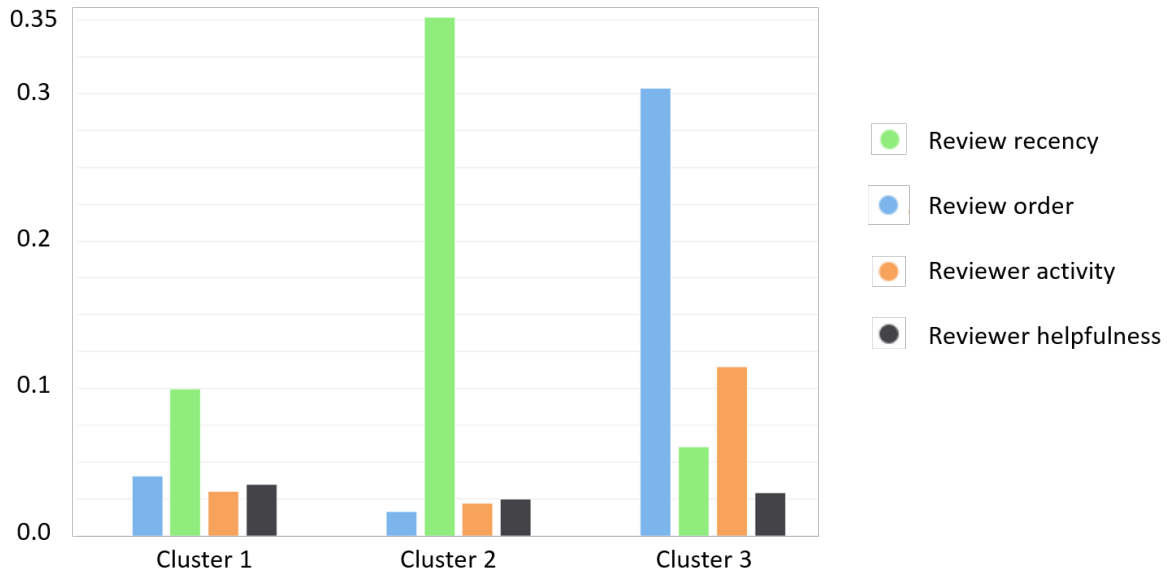
Figure 2 indicates that reviews in cluster 1 have moderate level of recency. They are not too early or too late. Also, their order is fairly moderate. Overal we can name cluster 1 as the moderate cluster without many outlier reviews.

In contrast, cluster 2 includes many recent reviews, and the average recency is very high (see recency measure in Section 4.1). Also, the reviews in cluster 2 are among the early reviews so that they have higher age.

The reviews in cluster 3 have a high order number, meaning that they have not been posted early. These reviews are not recent, meaning that the time distance between them and the last review on product has been high. Also, mainly active reviewers have written these reviews.

The next section uses four different classification models, one for each cluster.

Figure 2 Profile of clusters

4.4. Helpfulness prediction and evaluation

In this analysis, I used the binomial helpfulness measure (see Table 1) as a class label or target variable. I trained four fully connected multilayer perceptron [27, p. 398] with the backpropagation algorithm to classify online reviews. For each review, the classifier learns from a pre-processed set of online reviews with an assigned label ("helpful" or "unhelpful"). The inputs of this multilayer perceptron are review-related variables (see Table 1). There is one output neuron in all the developed models which indicate if the review is helpful or not. One hidden layer [27, p. 400] of 25–40 neurons performed better than other possible configurations of the neural network.

I computed the classic evaluation metric, predictive accuracy, which is the total number of true predictions divided by the total number of predictions. Table 6 compares the accuracy of the proposed model with three seminal models and one recent model for predicting the helpfulness of online reviews. These models are 1) Foreman et al.'s model [40] 2) Mudambi and Schuff's model [9] 3) Salehan and Kim's model [13] 4) Siering et al.'s model [20].

The results of cross-validation indicate that the proposed model outperforms the existing models for predicting the helpfulness of online reviews.

Table 6 Models comparison

	Accuracy	Recall	Variance
Foreman et al. (2008)	57.46%	-	-
Mudambi and Schuff (2010)	68.48%	-	-
Salehan and Kim (2016)	61.20%	-	-
Siering et al. (2018)	81.09%	-	-
Model (cluster) 1	88.73%	95.81%	+/- 1.4%
Model (cluster) 2	82.24%	93.37%	+/- 2.8%
Model (cluster) 3	85.94%	92.38%	+/- 5.1%

Focusing on predictive accuracy by showing the proportion of true results among the total number of cases examined, Table 6 shows that the proposed models in cluster 1 and 3 significantly outperform the other models, being able to classify with an accuracy higher than 85% (note that the authors of baseline models have not reported recall and variance). The developed models in cluster 2 also outperform three seminal models with an accuracy of higher than 82%.

Recall of the proposed model show that the model does not perform well in classifying the positive class (helpful) at the expense of poor performance in classifying unhelpful reviews. This evaluation is of crucial importance, especially for unbalanced datasets such as the one used in this study, as the proportion of different classes are significantly different (Namvar et al. 2011).

5. Conclusion and future work

Each review posted online by a customer is a form of advertising for businesses. Company's name and product are exposed to readers, increasing their awareness of who the company is and what they do. Many businesses have realized that getting online reviews from (happy) customers has many benefits in their reputation marketing – and that list of benefits keeps growing. This study proposed a novel approach for predicting the helpfulness of online reviews in different review clusters.

This study identified the factors that can potentially influence the helpfulness of online reviews. It categorized the identified factors into three groups; namely, review related-factors, temporal factors, and reviewer-related factors. It then chose the ideal candidate attributes to cluster the collected reviews; namely, reviewer helpfulness and activity, and review order and recency. It then developed k-means clustering and developed predictive models in each cluster. The results of review classification using neural networks along with cross-validation show that the proposed research approach improves the accuracy of predicting the helpfulness of online reviews compared to the existing methods.

The research reported in this paper sheds light on the understanding of online review helpfulness and the design of a better helpfulness voting mechanism for online review platforms. It advances our understanding of the inter-relationship between reviews and reviewers. It also has implications for consumers to leverage online product reviews to infer actual product quality. Currently, most of the online retailers sort reviews based on the number of helpfulness votes. The method proposed in this study help the online retailers to sort reviews when there are not enough customers votes for the reviews.

The analysis of insights provides benefits to online retailers planning to implement online reviews to improve their customer experience. The discrepancy in the review helpfulness measure, however, is an essential cause for the mix findings in the literature on online reviews, and the operationalization of the review helpfulness measure may explain the inconsistent relationships between determinants and online review helpfulness.

Therefore, future studies in this area still need to be conducted to develop new helpfulness metrics to enhance the feedback from the reviewers. Future works can investigate the new metrics that may emerge for specific products or reviewers. Also, as this study examined the proposed approach using a dataset from an online retailer, future works can examine this approach in other e-commerce platforms which collect reviews regarding services rather than products. In adition, as this study only used a dataset form Amazon.com, the proposed approach can be applied to datasets from more online review platforms.

6. References

[1] Y.-C. Ho, J. Wu, and Y. Tan, "Disconfirmation effect on online rating behavior: A structural model," *Inf. Syst. Res.*, vol. 28, no. 3, pp. 626–642, 2017.

[2] G. Vlachos, "Online travel statistics," in *Info Graphics Mania*, 2012.

[3] S. Santos, "Social media and tourism industry statistics," *Retrieved Dec.*, vol. 22, p. 2012, 2012.

[4] J. Wu, S. Fan, and J. L. Zhao, "Community engagement and online word of mouth: An empirical investigation," *Inf. Manage.*, 2017.

[5] Q. Ye, R. Law, B. Gu, and W. Chen, "The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings," *Comput. Hum. Behav.*, vol. 27, no. 2, pp. 634–639, 2011.

[6] 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, pp. 42–53, Jan. 2014.

[7] N. Korfiatis, E. GarcíA-Bariocanal, and S. Sánchez-Alonso, "Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content," *Electron. Commer. Res. Appl.*, vol. 11, no. 3, pp. 205–217, 2012.

[8] Q. B. Liu and E. Karahanna, "The dark side of reviews: The swaying effects of online product reviews on attribute preference construction," *MIS Q.*, vol. 41, no. 2, pp. 427–448, 2017.

[9] S. M. Mudambi and D. Schuff, "Research note: What makes a helpful online review? A study of customer reviews on Amazon. com," *MIS Q.*, pp. 185–200, 2010.

[10] 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.

[11] A. Y. K. Chua and S. Banerjee, "Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality," *Comput. Hum. Behav.*, vol. 54, pp. 547–554, Jan. 2016.

[12] T. L. Ngo-Ye and A. P. Sinha, "The influence of reviewer engagement characteristics on online review helpfulness: A text regression model," *Decis. Support Syst.*, vol. 61, pp. 47–58, 2014.

[13] 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.

[14] 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, 2017.

[15] R. Safi and Y. Yu, "Online product review as an indicator of users' degree of innovativeness and product adoption time: a longitudinal analysis of text reviews," *Eur. J. Inf. Syst.*, vol. 26, no. 4, pp. 414–431, Jul. 2017.

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

[17] S. Zhou and B. Guo, "The interactive effect of review rating and text sentiment on review helpfulness," in *International Conference on Electronic Commerce and Web Technologies*, 2015, pp. 100–111.

[18] Y. Huang, C. Li, J. Wu, and Z. Lin, "Online customer reviews and consumer evaluation: The role of review font," *Inf. Manage.*, 2018.

[19] B. Hernández-Ortega, "Don't believe strangers: Online consumer reviews and the role of social psychological distance," *Inf. Manage.*, vol. 55, no. 1, pp. 31–50, 2018.

[20] M. Siering, J. Muntermann, and B. Rajagopalan, "Explaining and predicting online review helpfulness: The role of content and reviewer-related signals," *Decis. Support Syst.*, 2018.

[21] Y.-H. Cheng and H.-Y. Ho, "Social influence's impact on reader perceptions of online reviews," *J. Bus. Res.*, vol. 68, no. 4, pp. 883–887, 2015.

[22] N. Hu, P. A. Pavlou, and J. Zhang, "On self-selection biases in online product reviews," *MIS Q.*, vol. 41, no. 2, pp. 449–471, 2017.

[23] Q. Gan, Q. Cao, and D. Jones, "Helpfulness of online user reviews: More is less," 2012.

[24] M. W. Uhl, "Explaining US consumer behavior with news sentiment," *ACM Trans. Manag. Inf. Syst. TMIS*, vol. 2, no. 2, p. 9, 2011.

[25] R. M. Schindler and B. Bickart, "Perceived helpfulness of online consumer reviews: The role of message content and style," *J. Consum. Behav.*, vol. 11, no. 3, pp. 234–243, 2012.

[26] A. Y. Chua and S. Banerjee, "Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth," *J. Assoc. Inf. Sci. Technol.*, vol. 66, no. 2, pp. 354–362, 2015.

[27] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*, 3rd ed. Elsevier, 2011.

[28] G. J. Spilich, G. T. Vesonder, H. L. Chiesi, and J. F. Voss, "Text processing of domain-related information for individuals with high and low domain knowledge," *J. Verbal Learn. Verbal Behav.*, vol. 18, no. 3, pp. 275–290, 1979.

[29] C. A. Sugar and G. M. James, "Finding the number of clusters in a dataset: An information-theoretic approach," *J. Am. Stat. Assoc.*, vol. 98, no. 463, pp. 750–763, 2003.

[30] R. Cangelosi and A. Goriely, "Component retention in principal component analysis with application to cDNA microarray data," *Biol. Direct*, vol. 2, no. 1, p. 2, 2007.

[31] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 2, no. 4, pp. 433–459, 2010.

[32] M. Namvar, S. Khakabimamaghani, and M. R. Gholamian, "An approach to optimised customer segmentation and profiling using RFM, LTV, and demographic features," *Int. J. Electron. Cust. Relatsh. Manag.*, vol. 5, pp. 220–235, 2011.

[33] S. KhakAbi, M. R. Gholamian, and M. Namvar, "Data Mining Applications in Customer Churn Management," 2010, pp. 220–225.

[34] G. J. King *et al.*, "A standardized method for assessment of elbow function," *J. Shoulder Elbow Surg.*, vol. 8, no. 4, pp. 351–354, 1999.

[35] Y. Guo, H. Wang, and J. Van Genabith, "Dependency-based n-gram models for general purpose sentence realisation," *Nat. Lang. Eng.*, vol. 17, no. 4, pp. 455–483, 2011.

[36] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[37] A. Wiese, V. Ho, and E. Hill, "A comparison of stemmers on source code identifiers for software search," in *Software Maintenance (ICSM), 2011 27th IEEE International Conference on*, 2011, pp. 496–499.

[38] M. Barter, Leticia Hotez, and Brian King, "Predictive Analytics," Deakin University, Melbourne Australia, 2, Sep. 2017.

[39] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1650–1654, 2002.

[40] B. Dervin and L. Foreman-Wernet, "Sense-Making Methodology as an Approach to Understanding and Designing for Campaign Audiences," in *Public communication campaigns*, R. E. Rice and C. K. Atkin, Eds. Sage, 2012, pp. 147–162.