Review helpfulness as a function of Linguistic Indicators

Hamad[†] **MSI Malik**[††] **Khalid Iqbal**[†††]

Department of Computer Science
Comsats Institute of Information
Technology Attock, Pakistan

Department of Computer Science
Comsats Institute of Information
Technology Attock, Pakistan

Department of Computer Science
Comsats Institute of Information
Technology Attock, Pakistan

Abstract
Online reviews are playing an important role in customer's decision making procedure for buying any product online. As buying products online is becoming customer's first choice while shopping. It is very helpful to make purchase decision for any product by reading online reviews related to that particular product. However, such a large volume of online reviews that is being generated can be considered as a big data challenge for both entities i.e. e-commerce websites and customers. These online reviews are usually ranked on the basis of helpful votes. This article examined the important factors that contribute to the helpfulness of online reviews and built a helpfulness predictive model for online reviews. Five novel linguistic characteristics are proposed and popular machine learning algorithms are applied to construct an effective predictive model for review helpfulness. LCM and visibility features are also used as baseline. We have performed experimental analysis on two popular Amazon review datasets and results reveals that hybrid set of features deliver the best predictive performance. We also found that the proposed Linguistic features are better predictors for review helpfulness as a standalone model. The findings of our study can provide new wisdom to e-commerce vendors for effective ranking of online reviews on the basis of their helpfulness. This research would also help customers in making better decisions before purchasing any product.

Keywords:
Online reviews, Helpfulness, Random forest, Noun, Amazon, Linguistic.

1. Introduction

The advent of Web 2.0 technology brought about a new way of distributing personal knowledge, suggestions and skills online[1]. User generated content (UGC) is now a progressively more useful resource for many users on the internet. As it's an era of web 2.0, different firms have started doing business buy social media platforms [1]. In the e-commerce context, Web 2.0 allows customers to share their purchase and usage experiences in the form of online product reviews (e.g. Amazon product reviews) [2]. These websites allow people to express their personal feelings, emotions, attitudes and feelings regarding not only products and services but also political and economic issues of the real world because they have to compete with the world market for maintaining their space in the current era. To achieve best response and better solution for attracting the users, appropriate business strategies should be made [3].

User generated online reviews have turn out to be today's word-of-mouth for individuals who are using e-commerce. In electronic marketplaces, the internet not only allows consumers to buy products online but also encourages them to inform others about their experience for making purchase decisions[2]. Mostly, reviews which consist of earlier purchase experiences of ordinary users are found to be more helpful than the information generated by professional vendors [4]. Hence, these online reviews create trust for other prospective users [5-7].

Online reviews are mostly found formless because big data have both pessimistic and optimistic impacts on customers. The consumers are getting the real-time experiences of other users which help them in making intelligent decisions to purchase an item [8]. But at the same time, the large amount of online generated reviews about any product creates information overloaded problem [9]. In a few cases, it is not possible for any person to take overview of all the reviews and then make any decision [4]. For example, an average ranked article on social web site can have more than several numbers of online reviews. Similarity, the number of reviews can be in the thousands for a new gadget. In such situations, it is impossible for consumers to read all the reviews before making purchase decisions, especially for products that have been reviewed by hundreds and sometimes thousands of customers with their inconsistent opinions [5].

Specifically, a review sentiment could be helpful, not-helpful or mixed. Some customers consider favorable and unfavorable reviews helpful because such polarized entries simplify the process of confirming or eliminating purchase options [10]. Yet others find mixed reviews helpful because such ambiguous entries highlight both advantages and advantages of the items under evaluation [11]. Customer's opinion about review helpfulness lies on two main product types i.e. search and experience on any e-commerce website [9]. Search products such as DSLR cameras are those whose qualities can be found easily even before purchasing it by going through online reviews [12]. In contrast experience, products such as songs are those whose quality is difficult to predict before listing to

Manuscript received January 5, 2018
Manuscript revised January 20, 2018

it [12]. Customers looking for beneficial products may need to rely more on others' online generated reviews, where they share their purchase experiences about any product [13]. Online social media facilitates the consumers for an opportunity to voice their opinions and learn from their peers about products and services of interest.

In previous research people have used multiple features for computing the helpfulness of online reviews. Sentiment analysis is performed using consumer reviews for extracting helpfulness [14]. Some researchers used metadata features to predict the helpful votes of online generated reviews [15]. Sentiment analysis or opinion mining are also used for predicting the helpfulness of user generated reviews [16]. Many online customers share their experiences about products or services by generating online reviews for that entity. Several experiments proved that sales of any particular product is affected by user generated reviews [17]. Some recent investigations highlighted that reviewer and review characteristics such as information quantity, semantic factors, reviewer location and identity opened new dimensions in the line of research [18]. Studying online reviews can help manufacturers better understand consumer responses to their products [6]. The assessment and predictions are usually based on the number of helpful votes in conjunction with other content characteristics of reviews that have received at least some helpful votes. However, the definitions and the number of the features of reviews are not consistent to helpfulness in prior literature [4]. Majority of websites who enable users to write helpful reviews for any product such as Amazon.com and Yelp.com provide a "Most Helpful Review First" option in sorting and presenting customer reviews. However, the helpfulness voting is not an answer. Not all online reviews received helpful votes, instead a large number of online reviews on many popular websites do not attract helpful votes. The accuracy of the prediction relies on the accurate and precise assessment of the characteristics that have received helpful votes.

The main objective of this study is to examine the impact of novel linguistic features on online review helpfulness and to highlight their influences for review helpfulness prediction. Two Amazon product review datasets are used to evaluate the contribution of novel features. We used four popular machine learning methods to build an effective helpfulness prediction model using novel linguistic features. Random forest based helpfulness prediction model outperforms using linguistic features. Theoretically, the findings of the article added contributions to the prior research by highlighting influential linguistic variables and their contributions to helpfulness of online reviews. In addition, the study disclosed the strong relationships between linguistic features and review helpfulness.

1. Five novel linguistic features are proposed. Two Amazon review datasets and four popular machine learning are utilized for experimentation.
2. An effective review helpfulness prediction model using random forest is constructed.
3. The findings indicate that Singular noun, general noun, proposition and personal pronoun are the most effective features for review helpfulness.

The rest of the article is organized as follows; related work is presented in the next section. Section 3 presents the proposed methodology in detail followed by the section 4, which presents the experimental analysis of proposed solution. Section 5 provides concluding remarks.

2. Related work

A regression model for predicting the utility of product reviews is presented by [19] . They used lexical similarity, syntactic terms based on Part-Of-Speech (POS) and lexical subjectivity as features. Some researchers [11] formulated a linear regression model for determining factors that contribute to prediction of review helpfulness. Their work was replicated by [20] and achieved just 15% explanatory power. A multilayer perceptron neural network based helpfulness model is presented by [21] that make use of product, review metadata, reviewer and review characteristics as features. User engagement related features are also used to predict review helpfulness [22]. However, they did not consider other important set of features such as readability, subjectivity and metadata [10] that are empirically proven to be better predictors of review helpfulness. Researchers also explored the interplay between review helpfulness, rating score, and the qualitative characteristics of the review text (using readability and other features) [12].

A non-linear regression model based on radial basis function for predicting helpfulness of movie reviews is presented by [7]. Other research works in the literature that used regression model to predict the review helpfulness are presented by [12, 19, 23]. Authors [24] have also developed a helpfulness prediction model for travel product websites. Some researchers have studied real life scenarios and found different results for why enterprises use viral marketing and how we can apply different models on them [25]. Similar research studies have also conducted for helpfulness prediction of product reviews [26]. Later, two ranking mechanisms for ranking product reviews: a consumer-oriented ranking mechanism that ranks the reviews according to their expected helpfulness and a manufacturer-oriented ranking mechanism that ranks them according to their expected effect on sales are proposed [23]. With the help of viral marketing datasets, researchers found an influential node in the web content [27]. RFM model is used to measure the standards of customers for enterprises in which customers who recently bought (Recency), who paid for many times (Frequency), and who spent more cash (Monetary) typically represent the best targets for new contributions [28].

Reliability is formulated as the use of positive and negative reviews, better suggestions for buying any product and claims of expertise [29-30]. Another study concluded that reviews comprising both negative and positive comments could be designated as unbiased. Authors studied the extent to which comprehensibility, specificity and reliability of reviews are related to review helpfulness [10]. Recently, a helpfulness prediction model is proposed that uses discrete positive and negative emotions of online reviews and produced very effective results [31]. They used deep neural network based classifier for binary helpfulness prediction. According to them, hybrid set of features with positive emotions produce the best predictive performance. In their recent work, authors conducted another research using Amazon reviews to study the influences of reviewer, psychological, summary language, and text complexity variables on helpfulness of product reviews [32]. Stochastic gradient boosting ML model is the most effective method and hybrid proposed determinants have shown best performance.

Recently, Hu and Chen presented a study to analyze the influence of review visibility, interaction between hotel stars and review ratings on hotel review helpfulness using Model tree (M5P). They concluded that interaction effect exists between hotel stars and review ratings. Furthermore, review visibility has a strong impact on review helpfulness [33]. Later, Berlo communication model based index system is designed by [34], to analyze the impact of multi-typed factors on review helpfulness. A recent study is proposed by [35] to investigate the consistency of reviewer's pattern of rating over time and predictability. Authors summarized that reviewers' rating behavior is consistent over time and across products. Moreover, reviews having higher absolute bias in rating in the past receive more helpful votes in future. Other studies that utilizes regression models explore significant textual and non-textual features include [12, 36-39].

More recently, Linguistic category features (LCM model) is introduced by [30]. Authors used five LCM features to predict the binary helpfulness of product reviews. The results revealed that random forest is the best classifier and hybrid set of features produced best predictive performance [30]. We will prove that the proposed linguistic indicators are the influential predictors for review helpfulness. The findings of this study can provide new insights to e-commerce retailers for better organization and ranking of online reviews and help customers in making better product choices [40].

3. Proposed Methodology

In this research work, we addressed the problem of helpfulness prediction of user generated online reviews as a regression problem. We have proposed five linguistic characteristics. In addition, the state of the art baseline features are also considered for comparisons. The detail of datasets and number of machine learning models are also discussed.

3.1. Dataset Description

We used two real-life review datasets for conducting multiple experiments to evaluate the effectiveness of proposed features. Data cleaning process is applied on both datasets to reduce the redundancy of reviews. The steps are: 1) Duplicate reviews are identified and removed 2) Reviews having blank text are also removed. First dataset (DS1) is a publicly available multi-domain sentiment analysis dataset [41]. This dataset has 109356 customer reviews across twenty two different product categories. The second dataset (DS2), a more recent review dataset is obtained by crawling reviews from amazon.com. This dataset has 2062 reviews which are collected from six different product categories. These reviews are collected from top-10 rated products. The details of both datasets are described in Table 1.

3.2. Variables

In this study, we have introduced five novel linguistic features to build an effective predictive model for helpfulness of online reviews. In addition, LCM [30] and visibility features are also considered for comparison with proposed features. The details are provided as follows.

Table 1: Dataset description

Dataset	#Reviews	Product types
DS1	109356	Apparel, automotive, baby, beauty, camera_&_photo, cell_phones_&_service, computer_&_video_games, electronics, gourmet_food, grocery, health_&_personal_care, jewelry_&_watches, kitchen_&_housewares, magazines, musical_instruments, office_products, outdoor_living, software, sports_&_outdoors, tools_&_hardware, toys_&_games, video
DS2	2062	Camera, Cell phone, Laser printer, Mp3, music, video game

3.2.1. Proposed Linguistic features

This study has proposed five novel linguistic features to build an effective helpfulness predictive model. The features are: 1) Noun-Singular (NS), 2) Noun-General

(NG), 3) Preposition (P), 4) Personal-Pronoun (PRP) and 5) Adverb (ADV). The feature computation process for proposed features is as follows: First, we used the NLTK Parts-Of-Speech (POS) tagger [20] to parse and tag each word of the all the reviews. Second, we used python language to traverse each word and assigning it linguistic tags in each review. Third, we then count the occurrence of each tag in each review to compute the final value of proposed features.

3.2.1.1. Noun-Singular

A noun refers to person, city, product or opinion, e.g. woman, Scotland, book, informative. Nouns mostly appear after determiners and adjectives. We extracted singular noun by counting all words with Noun-Singular tag using NLTK POS tagger[30].

3.2.1.2. Noun-General

This feature is extracted by counting the all words with noun tag in each review. We used noun count function for computing this feature by using NLTK POS tagger.

3.2.1.3. Preposition

A word that combines with a noun to complete a phrase is known as preposition. For computing this feature, we count the number of words with 'IN' tag in each review using NLTK POS tagger [30]e.g. for, up to, with etc.

3.2.1.4. Personal-Pronoun

A word that is associated primarily with a particular person is known as personal pronoun .This linguistic feature is extracted by finding all words with 'PRP' tag in a sentence of review using NLTK POS tagger [30]. For example; I, he, she.

3.2.1.5. Adverb

Adverb in nature refers to emotional, affective or mental state of a person. It is used in the sentences that are further than specific behavior. For computing this linguistic feature, we count all the words with 'RB' POS tag in each review. E.g. In a sentence of a review i.e. "this camera was working efficiently (RB)". Here we found that efficiently is the adverb referring to the past state of the camera.

3.2.2. LCM Features

LCM features [42] are used as baseline for comparing the effectiveness of proposed features. There are five linguistic features proposed in previous research works [30]. We compared the performance of proposed linguistic features with LCM so that influential features could be highlighted. The features in LCM model are: 1) ADJ (adjective), 2) SV (State verb), 3) SAV (state action verb), 4) IAV (iterative action verb) and 5) DAV (descriptive action verb). The five features are computed using NLTK POS tagger [30].

3.2.3. Visibility features

In this study, visibility features are also considered as baseline to compare the performance of proposed features in order to develop an effective helpfulness predictive model. The number of visibility features utilized here are: 1) Rating of the review and 2) Review age. The mathematical formulation of the review age is computed in Eq.1.

Review_age = Current_date - date of review posted (1)

3.3 Machine Learning Models and Evaluation Metrics

In this study, we utilized four machine learning models to build the helpfulness prediction models for online reviews. The models are: 1) Linear Regression 2) Multivariate Adaptive Regression (MAR) 3) Classification and Regression Tree (CART) and 4) Random Forest (RandF). In addition, three evaluation metrics are also utilized to evaluate the performance of machine learning models. The metrics are: mean square error (MSE), mean absolute deviation (MAD), and mean absolute precision error (MAPE).

4. Experimental results

In this research, we performed various experiments to predict the helpfulness of user generated reviews. We conducted three types of experiments, such as helpfulness prediction analysis, feature-wise analysis and feature importance analysis. To conduct these experiments, we utilized multiple types of features i.e. proposed linguistic features, LCM features and visibility features[30]. These experiments are performed to evaluate the effectiveness of the proposed features for review helpfulness prediction.

4.1. Predicting review helpfulness

In the first set of experiments, we built predictive models for review helpfulness and analyzed their accuracies. The models used the combination of proposed linguistic features, LCM features and visibility features. The results are presented in Table 2 in which best results are underlined.

We used four machine learning techniques i.e. Random Forest [30], CART decision tree, Multivariate Adaptive Regression and Linear Regression to develop four helpfulness predictive models. The performances of these models are then compared and best performance is highlighted. Both datasets (DS1, DS2) are utilized for experimentations. To evaluate the predictive accuracies of models using datasets (i.e. DS1, DS2), we used 10-fold cross-validation method for performance comparisons. Three error-based measures such as Mean square error

(MSE), Mean Absolute Deviation (MAD) and Mean Absolute Precision Error (MAPE) are also used. It is evident from Table 2 that the RandF machine learning model demonstrated the best results. The DS1 dataset has demonstrated MSE (0.06388), MAD (0.18335) and MAPE (0.28839) respectively whereas DS2 dataset has shown better performance in the form of MSE (0.0395), MAD (0.15362) and MAPE (0.72525) as compared to DS1 dataset.

Table 2: Helpfulness prediction performance using both datasets

Models	DS1 (dataset)			DS2 (dataset)		
	MSE	MAD	MAPE	MSE	MAD	MAPE
LR	0.094	0.245	0.410	0.085	0.253	1.207
MARS	0.086	0.226	0.374	0.054	0.186	0.881
CART	0.080	0.213	0.348	0.0627	0.205	0.908
RANDF	**0.063**	**0.183**	**0.288**	**0.039**	**0.153**	**0.725**

Random Forest has performed with more effective results and it is proved to be a best machine learning model among four for designing the effective helpfulness predictive model. The overall results are quite promising and demonstrated the utility of the proposed novel features for review helpfulness prediction. We conducted further set of experiments using the Rand Forest method as it is found to outperform other methods on both the datasets. We also found that by using linear regression model, we got the minimum performance as compared to other models. CART DT and MARS Regression models have demonstrated approximately same results.

4.2 Feature-wise Analysis

We have used Random Forest machine learning model for feature-wise comparison analysis. We found best results by using proposed features as a standalone model. We got MSE (0.08029), MAD (0.20923) and MAPE (0.3443) by using DS1 and MSE (0.05723), MAD (0.18928) and MAPE (0.85050) by using DS2 dataset with proposed features as a standalone model. The performance of purposed features is also compared with baseline features. For baseline features, we considered LCM features and visibility features. We found better results by using proposed features as compared to LCM and visibility features. In addition, DS2 dataset demonstrated better results as compared to DS1 dataset. This proved the utility of proposed feature as a standalone model. By combining all features (Proposed, LCM, Visibility), we got optimal performance as shown in Table 3. Similarly, DS2 dataset presented minimum MSE, MAD and MAPE as compared to DS1 dataset.

Table 3: Feature- wise analysis

Features	DS#1 (109356 records)			DS#2 (2062 records)		
	MSE	MAD	MAPE	MSE	MAD	MAPE
Visibility	0.097	0.240	0.396	0.060	0.186	0.908
LCM	0.088	0.220	0.365	0.064	0.203	0.961
Proposed	**0.080**	**0.209**	**0.344**	**0.057**	**0.189**	**0.850**
All features	**0.063**	**0.183**	**0.288**	**0.039**	**0.153**	**0.725**

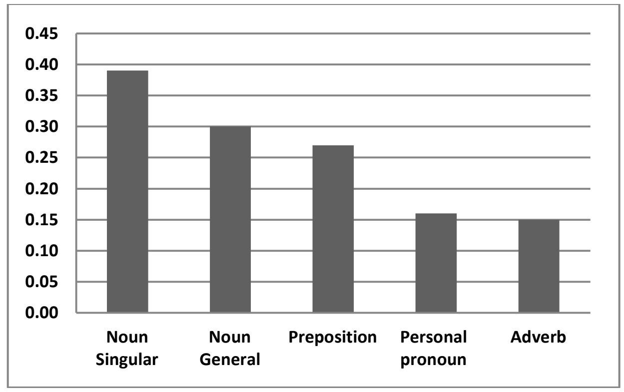

Fig. 1: Feature Importance using DS1

4.3 Feature Importance

In this section, we utilized Rand Forest ML model for computing importance of proposed features using both datasets. We obtained Noun-Singular as the best indicator among five linguistic characteristics by using DS1 dataset as shown in Fig. 1. This indicates that reviews which contain more singular nouns will attract more helpful votes. Similarly, Noun-General is the second best indicator for review helpfulness prediction problem. Preposition stands at the third position by using DS1 dataset. However, Adverb indicator stands at the last position and have marginal effect on review helpfulness. In other way, Noun-Singular and Noun-General are the most influential indicators when DS1 dataset is considered.

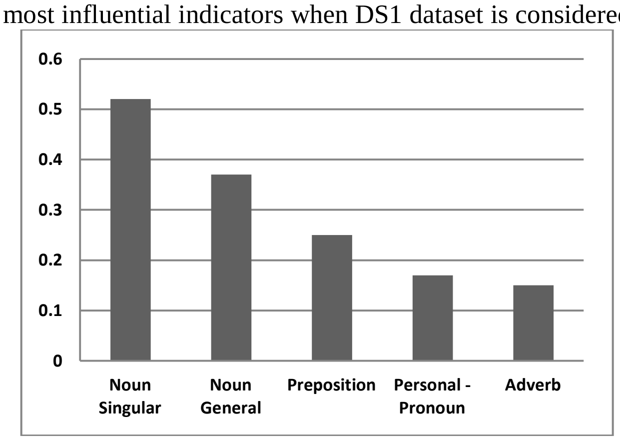

Fig. 2: Feature Importance using DS2

We obtained similar ranking of proposed indicators by using DS2 dataset. However, the influence of each indicator is higher when DS2 dataset is used as compared to DS1 dataset. We found that reviews which contain more adverbs have not significant impact on review helpfulness. However, reviews which contain more singular nouns receive more helpful votes. Therefore, noun-singular is the most effective indicator. DS2 dataset presents higher importance of each indicator as compared to DS1 dataset.

5. Conclusion

Review helpfulness prediction and its relationship with other features is an area of interest to researchers from many field. In this article, we addressed the problem that predicts the helpfulness of online reviews. In this regard, various determinants are computed to examine their influences on review helpfulness. In this study, we proposed five novel linguistic features to examine their effectiveness for review helpfulness prediction. We also utilized LCM features and Visibility features as baseline. For experiments, we used two popular datasets i.e. (multi-domains sentiment analysis dataset (DS1) and crawled Amazon review dataset (DS2). Four machines learning methods i.e. Random forest, Cart Decision Tree, Multivariate Adaptive Regression (MARS) and Linear Regression are utilized. With the help of these methods, we made four predictive models for review helpfulness prediction. After conducting multiple experiments, we concluded that proposed features have given very effective results as compared to baseline features i.e. LCM features and Visibility features. The results drawn by combining all features (i.e. proposed features and Baseline features) are very promising. The outcome of our research work is that the proposed features for predicting helpfulness are very influential.

For future directions, numerous interesting extensions could be explored. The proposed features can be used in other research domains. For example, sentiment analysis, documentation analysis, positivity and negativity prediction etc. Future work can also explore use of some novel features for making our research problem more effective. By using these novel features, we can also improve the efficiency of current models. This research work can also be improved by using new hybrid evolutionary approaches.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] A. E. Crowley and W. D. Hoyer, "An integrative framework for understanding two-sided persuasion," Journal of Consumer research, vol. 20, pp. 561-574, 1994.
[2] B. Bickart and R. M. Schindler, "Internet forums as influential sources of consumer information," Journal of interactive marketing, vol. 15, pp. 31-40, 2001.
[3] A. Ghose and P. G. Ipeirotis, "Designing ranking systems for consumer reviews: The impact of review subjectivity on product sales and review quality," in Proceedings of the 16th annual workshop on information technology and systems, 2006, pp. 303-310.
[4] C. Forman, et al., "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," Information Systems Research, vol. 19, pp. 291-313, 2008.
[5] D.-H. Park and S. Kim, "The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews," Electronic Commerce Research and Applications, vol. 7, pp. 399-410, 2008.
[6] A. Y. Chua and S. Banerjee, "Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality," Computers in Human Behavior, vol. 54, pp. 547-554, 2016.
[7] Y.-M. Li, et al., "Identifying influential reviewers for word-of-mouth marketing," Electronic Commerce Research and Applications, vol. 9, pp. 294-304, 2010.
[8] M. Dredze, et al., "Frustratingly Hard Domain Adaptation for Dependency Parsing," in EMNLP-CoNLL, 2007, pp. 1051-1055.
[9] F. Zhu and X. Zhang, "Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics," Journal of marketing, vol. 74, pp. 133-148, 2010.
[10] L. M. Willemsen, et al., "“Highly recommended!” The content characteristics and perceived usefulness of online consumer reviews," Journal of Computer-Mediated Communication, vol. 17, pp. 19-38, 2011.
[11] S. M. Mudambi and D. Schuff, "What makes a helpful review? A study of customer reviews on Amazon. com," 2010.
[12] N. Korfiatis, et al., "Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content," Electronic Commerce Research and Applications, vol. 11, pp. 205-217, 2012.
[13] A. J. Flanagin and M. J. Metzger, "Trusting expert-versus user-generated ratings online: The role of information volume, valence, and consumer characteristics," Computers in Human Behavior, vol. 29, pp. 1626-1634, 2013.
[14] M. Pontiki, et al., "SemEval-2016 task 5: Aspect based sentiment analysis," in ProWorkshop on Semantic Evaluation (SemEval-2016), 2016, pp. 19-30.
[15] A. Khan, et al., "Longitudinal trends in global obesity research and collaboration: a review using bibliometric metadata," Obesity Reviews, vol. 17, pp. 377-385, 2016.
[16] E. Cambria, "Affective computing and sentiment analysis," IEEE Intelligent Systems, vol. 31, pp. 102-107, 2016.
[17] W. Duan, et al., "Do online reviews matter?—An empirical investigation of panel data," Decision support systems, vol. 45, pp. 1007-1016, 2008.
[18] G. Ganu, et al., "Improving the quality of predictions using textual information in online user reviews," Information Systems, vol. 38, pp. 1-15, 2013.
[19] Z. Zhang and B. Varadarajan, "Utility scoring of product reviews," in Proceedings of the 15th ACM international conference on Information and knowledge management, 2006, pp. 51-57.

[20] A. H. Huang, et al., "A study of factors that contribute to online review helpfulness," Computers in Human Behavior, vol. 48, pp. 17-27, 2015.
[21] S. Lee and J. Y. Choeh, "Predicting the helpfulness of online reviews using multilayer perceptron neural networks," Expert Systems with Applications, vol. 41, pp. 3041-3046, 2014.
[22] W. C. Warren, et al., "The genome of a songbird," Nature, vol. 464, pp. 757-762, 2010.
[23] A. Ghose and P. Ipeirotis, "The EconoMining project at NYU: Studying the economic value of user-generated content on the internet," Journal of Revenue and Pricing Management, vol. 8, pp. 241-246, 2009.
[24] B. Fang, et al., "Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics," Tourism Management, vol. 52, pp. 498-506, 2016.
[25] A. Dobele, et al., "Controlled infection! Spreading the brand message through viral marketing," Business Horizons, vol. 48, pp. 143-149, 2005.
[26] M. Kenteris, et al., "A mobile tourism recommender system," in Computers and Communications (ISCC), 2010 IEEE Symposium on, 2010, pp. 840-845.
[27] M. Li, et al., "Helpfulness of online product reviews as seen by consumers: Source and content features," International Journal of Electronic Commerce, vol. 17, pp. 101-136, 2013.
[28] M. Allahbakhsh, et al., "Robust evaluation of products and reviewers in social rating systems," World Wide Web, vol. 18, pp. 73-109, 2015.
[29] J. Liu, et al., "Low-quality product review detection in opinion summarization," in Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), 2007.
[30] S. Krishnamoorthy, "Linguistic features for review helpfulness prediction," Expert Systems with Applications, vol. 42, pp. 3751-3759, 2015.
[31] M. Malik and A. Hussain, "Helpfulness of product reviews as a function of discrete positive and negative emotions," Computers in Human Behavior, vol. 73, pp. 290-302, 2017.
[32] M. Malik and A. Hussain, "An analysis of review content and reviewer variables that contribute to review helpfulness," Information Processing & Management, vol. 54, pp. 88-104, 2018.
[33] Y.-H. Hu and K. Chen, "Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings," International Journal of Information Management, vol. 36, pp. 929-944, 2016.
[34] C. Xiang and Y. Sun, "Research on the evaluation index system of online reviews helpfulness," in Service Systems and Service Management (ICSSSM), 2016 13th International Conference on, 2016, pp. 1-6.
[35] B. Gao, et al., "Follow the herd or be myself? An analysis of consistency in behavior of reviewers and helpfulness of their reviews," Decision Support Systems, 2016.
[36] Q. Cao, et al., "Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach," Decision Support Systems, vol. 50, pp. 511-521, 2011.
[37] A. Y. Chua and S. Banerjee, "Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth," Journal of the Association for Information Science and Technology, vol. 66, pp. 354-362, 2015.
[38] Y. Pan and J. Q. Zhang, "Born unequal: a study of the helpfulness of user-generated product reviews," Journal of Retailing, vol. 87, pp. 598-612, 2011.
[39] A. Ghose and P. G. Ipeirotis, "Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics," IEEE Transactions on Knowledge and Data Engineering, vol. 23, pp. 1498-1512, 2011.
[40] A. Qazi, et al., "A concept-level approach to the analysis of online review helpfulness," Computers in Human Behavior, vol. 58, pp. 75-81, 2016.
[41] J. P. Ferraro, et al., "The effects of natural language processing on cross-institutional portability of influenza case detection for disease surveillance," Applied clinical informatics, vol. 8, pp. 560-580, 2017.
[42] L. Coenen, et al., "The linguistic category model (LCM)," Manual. Free University Amsterdam, pp. 1151434261594-8567, 2006.