Toward Understanding Review Usefulness: A Case Study on Yelp Restaurants

Lei Li[1, 2], Ke Zhang[2], Qingqing Zhou[1], Chengzhi Zhang[1]
[1]Department of Information Management, Nanjing University of Science and Technology
[2]School of Information Science, University of Pittsburgh

Abstract
The quality of online reviews has become a critical element for opinion-sharing-enabled platforms. Crowd-sourced feedback, such as votes, on previously shared reviews can provide signals about the quality of the reviews. Prior studies examined some shallow features to explain the usefulness of votes to reviews, regardless of the detailed information described in the review. By using 1052 restaurant reviews from Yelp, we extensively explore four types of features that are related to review content details, business neighborhoods, user profiles, and business profiles, to better understand the usefulness of voting on reviews. Our main findings indicated that decisions made on whether a review is useful or not might lead to a biased result based only on review voting, since the review may receive votes due to other factors, regardless if the review discusses valuable aspects information of the business.
Keywords: online reviews; review usefulness; useful votes; quality of review
doi: 10.9776/16536

Contact: lileiwelldone@gmail.com, kez11@pitt.edu, breeze7zhou@163.com, zhangcz@njust.edu.cn

1 Introduction

A product or service on an online review website can attract a large amount of reviews from customers. As a result, users are challenged to distinguish useful reviews from those that are either frivolous or biased. To solve this problem, many online review websites enable the generation of social signals, such as useful-review votes. Recent literatures suggest that votes on the helpfulness of reviews play a significant role in influencing users' purchase decisions (Ghose & Ipeirotis, 2011). Hence, the factors that shape review usefulness can provide a rich area for research.

Previous studies on the usefulness of online reviews mainly launched from the following two aspects. One part of studies examined some basic content features of reviews, including length (Mudambi & Schuff, 2010) or positive/negative sentiment of review (Cheung et al., 2009). The other part focused on the user's reputation, such as rating (Forman et al., 2008) or number of published reviews (Hu et al., 2008) from the users. However, both of these have ignored the detailed aspects of information mentioned by the reviewers, such as food quality or overall service. Previous studies indicated that customers really care about detailed information on various aspects of a product or service (Liu & Seneff, 2009). As a result, detailed information mentioned in reviews plays an important role for users to make decisions.

In this paper, we focus on restaurant reviews from Yelp. Our goal is to understand the usefulness of user voting on online reviews. As an extension from prior studies, we extract four types of features and apply the linear regression model to examine the relationship between the number of useful votes and each feature. Our findings can further help online review websites to manage online review quality and reduce bias in review recommendations.

2 Methodology

2.1 Dataset

By Yelp API, we queried the webpage of 550 restaurants in Pittsburgh via HTTP to get the restaurant profiles and the full list of reviews which include17, 654 reviews on November 2014. For the author of each review, we visited the author's profile pages to get the user's information.

The duration of time since a review was published is significantly correlated with its number of useful votes (Pearson Coefficient r=.09, p=.00). In order to reduce the time bias, we examined reviews that were made in a specific period. At the same time, we also need sufficient reviews and useful votes during the selected time period. We finally chose a set of reviews during a four-month period from 2/11/2012 to 6/09/2012, which leaves us 1, 052 reviews with the same publish duration, and the average number of useful votes per review is 1.53. We further checked that there was no significant correlation between review publish duration and number of useful votes (r=. 02, p=. 44).

2.2 Experiment

We explored four types of features that may attract users to give usefulness votes for restaurant reviews and used the linear regression model to test which factors play a critical role that cause a review to get more useful votes. The four types of features are described as follows.

Content-detail-based: Previous research suggested that longer reviews often include more details about the products (Nah et al., 2005). Similar to previous studies, we considered **the length of the review**. Other than the sentiment expressed in the review, we elaborately examined if the user mentions an opinion on different aspects of the restaurant, including its **food, location, service, price, waiting time, surroundings, and the methods of payment**. Two coders manually labelled the 1, 052 restaurant reviews and took each aspect as a binary feature. The mean pair-wise Cohen's kappa of the two coders was 0.89, showing the consistency of the labeling process (Viera & Garrett, 2005). Furthermore, we calculated an aggregate score by counting the number of aspects mentioned in a review, and defined this feature as the **total number of mentions**.

Neighborhood-based: The chance that a review gets a "useful" vote relies on its visibility to the public, which is highly related to the urban business environment that surrounds the restaurant. Prior studies indicated that neighborhood where the business located has a strong impact on the popularity of the business (Karamshuk et al., 2013). Similarly, we considered all neighboring restaurants are within distance d=1 mile from the target restaurant R. We denoted the number of neighboring restaurants around restaurant R as the **density**. We considered **competitiveness** as the number of restaurants with the same category as R. Intuitively, given a target area with multiple restaurants providing similar types of food, customers might be more likely to focus on high-quality reviews in different restaurants and compare between them to help making the final decision. The total number of reviews observed in all nearby restaurants indicates the businesses' overall **popularity** in that area.

User-based: Reviews contributed by influential users can have a significant impact on the sales of products or services (Forman et al., 2008). The reputation of a Yelp user is the aggregation of his or her social status and content-sharing activities. Here, we extracted user-based features from user profiles, including **# friends, # users' reviews, # review updates, # fans, # local photos, # history useful votes, and # compliments**.

Business-based: The existing research focuses on how online reviews affect the popularity of the product (Zhu & Zhang, 2010; Park et al., 2007). We expanded on existing research to understand the relationship between the useful voting on the online reviews and the restaurant's quality and popularity. Intuitively, the better and more popular the restaurant, the higher the probability that people will go there, and the more chance its reviews will be visible to the public. Therefore, we consider several business-based features as measurements of restaurant quality and popularity. The first one is the restaurant's overall **star rating**, which is an average score of all users' rating. Another one is the total number of **reviews of this restaurant**. We also considered the types of services provided by the restaurant, such as delivery, take-out, or Wi-Fi. We then take the total **number of services provided** as another feature.

3 Results

In linear regression analysis, the importance of each feature is determined by the coefficient. In Table 1, the significant features with high coefficient values that affect the number of the useful votes were concentrated in user-based features, including #friends, #fans, and #history useful votes. However, detailed aspects of the restaurant, such as food, location, service, and #aspects mentioned, have low coefficient values and show no significant relationship to the usefulness number of the review. This further explains that the user's reputation is the most important reason for the usefulness of review votes, and that the content of the review is underweighted by other factors to explain the number of useful review votes. This indicates that a review with more useful votes may not actually be able to provide more detailed and valuable information on the restaurant.

4 Conclusion

In this paper, we taken restaurant reviews in Yelp as a study case and investigated four types of factors (review content details, business neighborhood, user profiles, and business profiles) to better understand the usefulness of voting on reviews. We found that user-based features play a more important role than detailed aspects about features of restaurants on the number of useful votes.

We found a bias for useful votes, such that even if a review gets more useful votes, we cannot simply declare that the review provides enough comprehensive information.

Features		Coeff	t	p
Content-detail-based	Length of the review	.189***	11.033	.000
	Food	.000	-.006	.995
	Location	.002	.051	.959
	Service	-.011	-.328	.743
	Price	-.003	-.091	.928
	Waiting time	-.005	-.142	.887
	Surrounding	-.001	-.030	.976
	Method of payment	.003	.074	.941
	# Aspects mentioned	.031	.180	.857
Neighborhood-based	Density	.063	1.663	.097
	Competitiveness	.002	.297	.766
	Area popularity	-.074	-1.744	.081
User-based	# Friends	.432***	8.160	.000
	# Users' Reviews	-.010	-.207	.836
	# Review updates	-.088	-1.690	.091
	# Fans	-.722***	-3.959	.000
	# Local photos	-.155*	-2.421	.016
	# History usefulness votes	.530**	3.053	.002
	# Compliments	.332	1.179	.239
Business-based	Star rating	-.030	-1.691	.091
	# Restaurant' Reviews	-.029*	-2.508	.012
	# Services provided	-.015	-1.145	.252

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

Table 1. Regression Results

5 References

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering, 23*(10), 1498-1512. doi:10.1109/TKDE.2010.188.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com. *MIS quarterly*, *34*(1), 185-200.

Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. *International Journal of Electronic Commerce*, 13(4), 9-38. doi:10.2753/JEC1086-4415130402

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

Hu, N., Liu, L., & Zhang, J. J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management*, 9(3), 201-214. doi :0.1007/s10799-008-0041-2.

Liu, J., & Seneff, S. (2009, August). Review sentiment scoring via a parse-and-paraphrase paradigm. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1* (pp. 161-169). Association for Computational Linguistics.

Nah, F. F. H., Lee, H. H., & Chen, L. (2005). Information Search Patterns in E-commerce Product Comparison Services. *SIGHCI 2005 Proceedings*, 3.

Karamshuk, D., Noulas, A., Scellato, S., Nicosia, V., & Mascolo, C. (2013, August). Geo-spotting: Mining online location-based services for optimal retail store placement. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 793-801). ACM. doi: 10.1145/2487575.2487616

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

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2), 133-148. doi: http://dx.doi.org/10.1509/jmkg.74.2.133.

Park, D. H., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, 11(4), 125-148. doi: 10.2753/JEC1086-4415110405.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Fam Med*, *37*(5), 360-363.