

The Impact of Emotions and Information Entropy on Review Helpfulness based on Ebay Review Data

Xin Zeng¹

¹School of Economics and Management, XiDian University, Xi'an, China

²School of Economics and Management, XiDian University, Xi'an, China

Abstract: Many e-commerce sites use vote machine to help consumers find helpful reviews quickly and make decisions. However, there is the Matthew effect in the review voting. This study identifies the influence of information entropy and emotions on the review helpfulness, and explores the adjustment of emotions in the influence of information entropy on review helpfulness based on the Ebay data. Research shows that the impact of emotions on review helpfulness is positive, and information entropy has different effects on different products, which has positive impact on experiential product reviews and negative impact on search products. The influence of information entropy on the review helpfulness is positively regulated by the emotions. The conclusions of this paper can help e-commerce websites improve the ranking mechanism of online reviews and enhance the user experience.

Keywords: Review helpfulness; information entropy; emotions.

1. INTRODUCTION

Online reviews are user-generated content that consumers describe based on their experiences. Online comments help consumers eliminate certain information asymmetry in online shopping, which also affects the sales of products on e-commerce platforms. However, with the popularity of online shopping, the number of online comments began to surge, and it was difficult for consumers to effectively identify useful comments from a large number of comments. Recently, scholars have paid much attention to the usefulness of criticism, and discussed the influential factors of the usefulness of criticism from various aspects. due to the encouragement measures launched by the e-commerce platform and the comment incentive activities of some merchants, the platform is full of a large number of invalid long comments, and it is not accurate to use the length to express the amount of comment information. For the emotional analysis of comments, many literatures divide the feelings of comments into positive or negative ones. However, each comment may contain multiple emotions, and the total amount of emotions contained in comments with the same emotional polarity may vary greatly. The total amount of emotion received by consumers may vary.

2. LITERATURE REVIEW

Mudambi [1] began to study earlier comments usefulness, defines evaluation usefulness as a help to promote the degree of consumer purchase decision making process, using the star rating and comment depth to determine usefulness, to reflect the opinion of the product with star rating, and length of comments depth to comment, said on behalf of the comment on the amount of information, the research results show that the length of product ratings and comments for comments usefulness has significant positive impact.

Subsequently, on the basis of Mudambi's research, many scholars conducted further studies on the usefulness of reviews from different perspectives to explore the factors affecting the usefulness of reviews. From the perspective of comment text information and features, scholars are no longer limited to use the length of comments to express the information content of comments, but refine the text content of comments to extract important information that affects consumer decisions and the usefulness of comments. Weathers [2] divided the comment text information into several categories, including product information, experience information and brand comparison information, indicating that the product information, experience information and brand comparison information in the comment have a positive impact on the usefulness of the comment. It also proves that the more information contained in the comment, the greater the amount of information has a greater impact on the usefulness of the comment. Guo [3] et al. focus on the similarity of language style and professional knowledge in comment texts from the perspective of information search. These two dimensions of similarity play different roles in shaping people's perception of the usefulness of comments.

From the perspective of commenting on emotions, scholars mainly judge the polarity of emotions through star rating or digging from the comment text, so as to study the influence of emotions on the usefulness of comments. However, there are many controversies about the research results of commenting on the polarity of emotion and exploring the usefulness of commenting. According to the negative bias theory, some studies suggest that negative evaluations are thought to be more helpful than positive ones because they deviate from accepted norms of remaining positive. Lee [4] studied online hotel reviews and found that when potential customers read online hotel reviews during their future stay, negative reviews were more helpful than positive ones. Yin [5] found that negative emotions affect users' efforts perceived from online comments, thus affecting the usefulness of comments.

3. RESEARCH HYPOTHESES

3.1 The influence of information entropy on review helpfulness

According to information theory, information entropy is the average amount of information contained in the received information after excluding redundancy. According to the information entropy, we can get the average information amount of each comment. While online comments are natural language texts, the values of parameters are qualitative rather than quantitative, and the values of parameters are words or sentences rather than Numbers [6]. Therefore, we first observe the probability distribution of words in the text. For example, a

word that often appears in multiple comments has a high probability and a low entropy. The entropy of each comment can be obtained from the information entropy of the word [7]. The higher the entropy of the comment, the more information it contains, and the difference between the information it contains and other comments.

However, the information entropy of reviews may not be the same for all products, and there may be differences among different products. Nelson [8] proposed three types of products: search products, experiential products and trust products. Among the studies of online reviews, the first two have been more extensive. Different types of products pay different attention to comment information. Kumar and Benbasat [9] found that the mere presence of comments established a social network, whereas online and open comments could be viewed as offline interpersonal interactions. For experiential products, consumers' comments on experiential products may pay more attention to the information of the reviewers' feelings, which is more favorable for consumers to compare and judge. The higher the entropy of review information of experiential commodities is, the more systematic and chaotic the review is. The more information the review contains, the more personal the information it describes. For search products, consumers need more objective information, such as product attributes or characteristics. The higher the entropy of information, the more disorderly the comments will be, and the information obtained from the comments will not meet the needs of consumers. Therefore, we make the following assumptions:

H1: information entropy has different effects on the usefulness of reviews for different categories of products. For experiential products, information entropy has a positive effect on the usefulness of reviews, while it has a negative effect on search products.

3.2 The influence of emotion on review helpfulness

Until now, psychologists have not reached a consensus on the concept of emotion, but the importance of emotion has been recognized in many research fields. In the field of psychology, many studies have proved that emotions play a particularly important role in people's behavior. Emotions have different effects on cognition, decision-making and attention. Research on emotions in marketing also shows that emotions play a powerful central role in changing consumers' beliefs, attitudes, decisions and behaviors. The emotions expressed by people can influence the behavior of consumers by triggering mutual emotional reactions.

In traditional word-of-mouth communication, emotion is an important factor influencing consumers. With the development of the Internet, people no longer just spread information and emotions face to face, but use the Internet to spread the information and emotions they want to express. Although this mode of communication is not as strong as the feelings in face to face communication, the network communication is more rapid and extensive. A number of studies have also demonstrated the impact of emotions transmitted online on users. Berger and Stieglitz [10] used psychological field experiments and twitter messages to show that emotional content significantly affects users' behavior of sharing online content. Online comments, as a kind of user-generated content, are written by consumers about their experiences or feelings in the whole shopping process. The comments often contain the

personal feelings of many reviewers and a large number of emotional words. Research on emotional language shows that, in general, readers can easily perceive and distinguish the emotions embedded in the author's writing, and emotional words can be processed more quickly and effectively by readers than non-emotional words, or even automatically. Therefore, consumers will be affected by the emotions in the comments to some extent. Therefore, we propose the following hypothesis:

H2: the emotions contained in the comments have a positive impact on the usefulness of the comments. The more emotion a comment contains, the more useful it becomes.

4. STUDY DATA AND VARIABLES

In this study, a python crawler was used to collect product reviews from the EBAY e-commerce platform. According to the existing research, different product reviews will be different, and the usefulness of reviews will be affected differently. This study chose the Ebay website comments more two products, one kind is one product - movie DVDS, is a kind of search products - iphone, a total of 32162 comments, respectively, including the film DVD 19646 comments and iphone 12516 comments, removed without votes for blank text reviews and comments, go unless English text comments, end up with 6437 valid samples, including movie DVD, 4342, 2095 iphone. Each sample contains product information and review information, including product name, price, average score, and number of reviews received for the product.

Dependent variable comments usefulness is each comments get useful votes accounted for the proportion of the total number of votes, the scope of [0, 1]. Through crawl to the data, we know every comment count and useless vote vote of useful and the total number of voting which together for a quick review. According to the existing research, control variables including product information and comments score, time lag, because of the different products, will affect the comment on the usefulness of, so we will product information as control variables. Product information included in the price of the product, the average rating of product, the product gained by the total number of comments. The price is getting the products unit price when you pick up the data, Mean score is getting products of star rating when you pick up the data, value range of [0, 5], the total number of comments is to collect data at the end of each product contains the total number of comments. Review score is the score in each comment text numerical, value range of [0, 5], time difference is that every comments comments and crawl data time.

5. RESEARCH RESULTS AND DISCUSSION

We use Tobit model to solve the selection bias problem of such samples. Combined with the previous assumptions, the regression model of this study is as follows:

$$\begin{aligned}
 Help &= \beta_1 Info + \beta_2 Sent + \beta_3 AverR + \beta_4 Rating + \beta_5 Price \\
 &+ \beta_6 Time + \beta_7 Count + e
 \end{aligned}$$

In this paper, two models are used for testing. Model 1 represents the model only containing control variables, and model 2 represents the model containing control variables and independent variables. The model regression analysis results of movie disc products are shown in table 1, and the model regression results of iphone products are shown in table 2.

From the regression results of model 1 in table 1, we can find that all the control variables have a certain impact on the review usefulness of movie disc products, and the average score of movie disc products ($= -0.397, p < 0.01$) has the greatest impact on the review usefulness, which means that the products with higher average score get more useful reviews. However, from the regression results of model 1 in table 2, it is found that the average score has no significant influence on the usefulness of iphone product reviews, indicating that the difference in the average score does not affect the usefulness of iphone product reviews. Among them, the major influence is the price of products. This is because the price elasticity of iphone products is relatively large, and the price of new phones is often higher, but there will be fewer useful comments without accumulation.

From the model 2 in table 1 and table 2, we can see that the information entropy of reviews in different products has different effects on the review usefulness. The information entropy has a positive influence on the review usefulness of experience products, but a negative influence on the review usefulness of search products, which supports hypothesis 1. For movie discs, the coefficient of information entropy ($= 0.068, p < 0.05$) is greater than 0 and the regression coefficient is significant, indicating that the information entropy in reviews has a positive impact on the usefulness of reviews. Comments with high information entropy, namely comments with large amount of information, are more useful. However, in the regression results of iphone, the coefficient of information entropy ($= -0.135, p < 0.01$) is negative and more significant, which indicates that the influence of information entropy on the usefulness of comments is negative. The greater the information entropy is, the smaller the usefulness of comments will be.

According to the results of model 2 in table 1 and table 2, the total amount of emotions has a positive impact on the usefulness of comments, which is consistent with our hypothesis 2. Among them, the effect of the total amount of emotion ($= 0.025, p < 0.01$) contained in the film review was smaller than that in the iphone review ($= 0.053, p < 0.01$). It can be seen that the emotional content in reviews can improve the usefulness of reviews, but the emotional content in reviews of experiential products need not be too rich, and adding emotional content in reviews of search products can more effectively improve the usefulness of reviews.

Table 1. Tobit regression result 1

变量	模型1	模型2
Info		0.0683** (0.0346)
Sent		0.0250*** (0.00933)
Price	0.0784*** (0.0202)	0.0778*** (0.0201)
AverR	0.397*** (0.0544)	0.409*** (0.0544)
Rating	0.194*** (0.0133)	0.197*** (0.0134)
Count	-0.210*** (0.0143)	-0.210*** (0.0143)
Time	-0.198*** (0.0259)	-0.199*** (0.0260)
Constant	0.521 (0.343)	0.466 (0.342)
Observations	4,342	4,342
Pseudo R2	0.103	0.105
log likelihood	-3392	-3386

Table 2. Tobit regression result 2

变量	模型1	模型2
Info		-0.135*** (0.0444)
Sent		0.0534** (0.0212)
Price	-0.288*** (0.0406)	-0.285*** (0.0403)
AverR	0.00435 (0.149)	0.0306 (0.149)
Rating	0.188*** (0.0182)	0.178*** (0.0183)
Count	0.130*** (0.0264)	0.130*** (0.0263)
Time	-0.117*** (0.0357)	-0.119*** (0.0355)
Constant	2.773*** (0.747)	2.687*** (0.742)
Observations	2,095	2,095
Pseudo R2	0.0557	0.0606
log likelihood	-1631	-1623

Note: *** p<0.01, ** p<0.05, * p<0.1

6. CONCLUSION

In this paper, from the Angle of information entropy and emotional, explores how the information in the comment and emotion affects comments usefulness, and by establishing the model and empirical test, confirmed that the following results: (1) different products for different processing, to experience the kind of product, need more personality information, can display self feeling, don't have to be similar to others. In contrast, for search products, the information in the comments should be consistent and in an orderly state, which is conducive to the viewer to receive information. (2) the total amount of emotion contained in comments has a positive impact on the usefulness of comments. The more emotional the comment, the more helpful it will be to the viewer. When you add emotion to a comment, it makes people more likely to believe it's true. Therefore, when merchants encourage consumers to write comments, they should encourage consumers to include their feelings and feelings.

REFERENCES

- [1] Mudambi S M, Schuff D. What Makes a Helpful Online Review? a Study of Customer Reviews on Amazon.com[M]. Society for Information Management and the Management Information Systems Research Center, 2010, 34 (1):185-200.
- [2] Weathers, D., Swain, S. D., and Grover, V. Can Online Product Reviews Be More Helpful? Examining Characteristics of Information Content by Product Type [J], Decision Support Systems 2015, (79):12-23.
- [3] Guo, B., and Zhou, S. S. What Makes Population Perception of Review Helpfulness: An Information Processing Perspective [J], Electronic Commerce Research [J], 2017, 17 (4): 85-608.
- [4] Lee, M., Jeong, M., and Lee, J. Roles of Negative Emotions in Customers' Perceived Helpfulness of Hotel Reviews on a User-Generated Review Website a Text Mining Approach[J], International Journal of Contemporary Hospitality Management 2017,29(2),762-783.
- [5] Yin, D. Z., Bond, S. D., and Zhang, H. Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews [J], Mis Quarterly, 2014, 38(2): 539-560.
- [6] LEE J H, SUN H J, PARK J H. The role of entropy of review text sentiments on online WOM and movie box office sales [M]. Elsevier Science Publishers B. V. 2017.
- [7] SINGH J P, IRANI S, RANA N P, ET AL. predicting the “helpfulness” of online consumer reviews [J]. Journal of Business Research, 2017, 70:346-355.
- [8] NELSON P. Advertising as Information [J]. Journal of Political Economy, 1974, 82(4):729--754.
- [9] KUMAR N, BENBASAT I. Research Note: The Influence of Recommendations and Consumer Reviews on Evaluations of Websites [J]. Information Systems Research, 2006, 17(4):425-439.
- [10] Stieglitz S, Dang-Xuan L. Emotions and Information Diffusion in Social Media—

Sentiment of Microblogs and Sharing Behavior [J]. *Journal of Management Information Systems*, 2013, 29(4):217-248.