

Predicting Early Review Ratings for Product Marketing In E-Commerce Websites

KAVITI SARAL KUMAR#1, MALLIPUDI REVATHI#2

1Student, Computer Science and Engineering, NOVA College of Engineering and Technology, Vegavaram(V), Jangareddigudem(M) West Godavari(D) Andhra Pradesh

2Guide, Associate Professor, Computer Science and Engineering NOVA College of Engineering and Technology, Vegavaram(V), Jangareddigudem(M) West Godavari(D) Andhra Pradesh

ABSTRACT: The level of buying items by the client has been expanded definitely through web. Clients even have the office of sharing their musings about the specific item on web as surveys, web journals, remarks and so on. Numerous clients read survey data given on web to take choices for purchasing items. A few clients may give the audits for building up the closeout of the item or to diminish the deal. This may confound the clients who depend on the audits to purchase an item. In this way, there is a need to locate the legitimate audits and expel phony surveys that are included by noxious or extortion client. The proposed framework thinks of the answer for this issue. Driving occasions has been utilized to discover the time interim between the audits. The proposed framework mines the dynamic time frames, for example, driving sessions to precisely find the progressive extortion. These driving sessions can be helpful for recognizing the nearby peculiarity rather than worldwide abnormality of item surveys. After this to break down the rating, audits and progression of the item we analyze three certainties, they are appraising based actualities, survey based realities and chain of importance certainties. Likewise, we propose a streamlining based collection technique to incorporate every one of the certainties for misrepresentation location. The assessments of this advancement are done on manufactured dataset that are gathered. The ordered and condensed item audit data causes web clients to comprehend survey substance effectively in a brief timeframe.

KEYWORDS: E-commerce; reviews; feature identification; opinion mining^[1].

1.INTRODUCTION

The emergence of e-commerce websites has enabled users to publish or share purchase experiences by posting product reviews, which usually contain useful opinions, comments and feedback towards a product. As such, a majority of customers will read online reviews before making an informed purchase decision. It has been reported about 71% of global online shoppers read online reviews before

purchasing a product. Product reviews, especially the early reviews (i.e., the reviews posted in the early stage of a product), have a high impact on subsequent product sales. We call the users who posted the early reviews early reviewers. Although early reviewers contribute only a small proportion of reviews, their opinions can determine the success or failure of new products and services. It is important for companies to identify early reviewers since their feedbacks can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products. For this reason, early reviewers become the emphasis to monitor and attract at the early promotion stage of a company. The pivotal role of early reviews has attracted extensive attention from marketing practitioners to

induce consumer purchase intentions. For example, Amazon, one of the largest e-commerce company in the world, has advocated the Early Reviewer Program, which helps to acquire early reviews on products that have few or no reviews. With this program, Amazon shoppers can learn more about products and make smarter buying decisions. As another related program, Amazon Vine² invites the most trusted reviewers on Amazon to post opinions about new and pre-release items to help their fellow customers make informed purchase decisions. Based on the above discussions, we can see that early reviewers are extremely important for product marketing. Thus, in this paper, we take the initiative to study the behaviour characteristics of early reviewers through their posted reviews on representative e-commerce platforms, e.g., Amazon and Yelp. We aim to conduct effective analysis and make accurate prediction on early reviewers. This problem is strongly related to the adoption of innovations. In a generalized view, review posting process can be considered as an adoption of innovations³, which is a theory that seeks to explain how, why, and at what rate new ideas and technology spread. The analysis and detection of early adopters in the diffusion of innovations have attracted much attention from the research community.

II. LITERATURE SURVEY

[1] Ghose and Ipeirotis (2006) proposed 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. They used econometric analysis with text mining to make their ranking work

[2] Wu et al. (2013) carried out an analysis on both seller and customer reviews. Before purchasing any item, customers go through various things, such as Volume 5, Issue 2, March-April -2019 | <http://ijsrcseit.com> customer reviews, seller reviews, and price comparison with other marketplaces. The authors used all these parameters to determine the willingness to pay of customers using a conceptual model.

[3] Li et al. (2013) analyzed content-based and source-based review features that directly influence product review helpfulness. It was also found that customer-written reviews that were less abstract in content and highly comprehensible result in higher helpfulness

[4] Lee and Shin (2014) investigated whether the quality of reviews affects the evaluations of the reviewers and the e-commerce website itself. They conducted pilot tests prior to the main experiment. The participants were asked questions such as (a) how frequently they use online shopping malls, and (b) if they had ever used the target product. They investigated (a) how the reader's acceptance depends on the quality of online product reviews and (b) when such effects are more or less likely to occur. Their findings indicated that participants' intention to purchase the product increases with positive high-quality reviews as opposed to low-quality ones.

[5] Huang et al. (2015) examined message length together with aspects of review patterns and reviewer characteristics for their joint effects on review helpfulness. They found that the message length in terms of word count has a threshold in its effects on review helpfulness. Beyond this threshold, its effect diminishes significantly or becomes near nonexistent.

[6] Allahbakhsh et al. (2015) proposed a set of algorithms for robust computation of product rating scores and reviewer trust ranks. They harvested user feedback from social rating systems. Social rating systems collect and aggregate opinions (experience of using a service, purchasing a product, or hiring a person that is shared with other community members, in order to help them judge an item or a person that they have no direct experience with) to build a rating score or level of trust worthiness for items and people.

[7] Chua and Banerjee (2016) found a relation between helpfulness and review sentiment, helpfulness and product type, and helpfulness and information quality. Review sentiment was classified in three categories: favorable, unfavorable, and mixed. The products were categorized as search products and experience products. The information quality has three major dimensions: comprehensibility, specificity, and reliability.

[8] Qazi et al. (2016) explained why some reviews are more helpful compared to others. As the helpfulness of online reviews helps the online web user to select the best product, they read several reviews of that product and finally conclude whether the review was helpful or not.

In this section, the details of the proposed system are going to be present. In fig.2. The flow chart is describing the overview of our proposed system. Firstly we are going to collected all the reviews of the consumer from those reviews the aspects are to be identified and opinions are collected and then data preprocessing is done to remove all the noisy words from the collected opinions. After data gets classified by using data classification, the most ranking products are to be collected according to term frequency and opinions collected. Simultaneously are going to get the best rated product.

Let us consider the set of consumer reviews^[5] for a desired product are $R = \{r_1, r_2, r_3, \dots, r_{|R|}\}$ for all $r \subseteq R$ and by

considering multiple aspects of the product the overall rating can be given Let us consider the reviews are O_{\min} and O_{\max} this rating is a numerical score that indicates the overall opinion of the product in a particular review r , i.e., $O_r \subseteq [O_{\min}, O_{\max}]$. Whereas O_{\min} and O_{\max} are the minimum and maximum ratings respectively. Generally the ratings are from 1 to 5 and for some websites it will be from 1 to 10. In the next subsections we are going to introduce the algorithms which are used in the proposed system.

Aggregate ranking algorithm

In this algorithm we combine the three techniques.

- (a) Frequency-based method
- (b) Correlation-based method, and
- (c) Hybrid method

a. Frequency based method

Frequency-based method is the method which is used in our aggregate ranking algorithm, in which it gives the features according to term frequency of the product. This method takes only the frequency of the term and which will impact on the customer opinions on the particular product, it helps in rating the product. There are some usual features of the product will appear frequently those are consider as the important features.

b. Correlation- based method

Correlation-based method, which measures the correlation between the reviews on particular products and the final rankings. It ranks the aspects based on the number of cases when such two kinds of opinions are consistent. Correlation-based method ranks

the aspects by simply counting the consistent cases between reviews on particular products and the final rankings. It ignores to model the uncertainty in the generation of overall ratings, and thus cannot achieve satisfactory performance.

c. Hybrid method

Hybrid method , that captures both aspect frequency and the correlation The hybrid method simply aggregates the results from the frequency-based and correlation-based methods, and cannot boost the performance effectively.

Advantages

By aggregating these things we can achieve the high accuracy and efficiency and we can classify the items in efficient manner. We are going to give the highest ranking product directly without reading all the reviews.

III.CONCLUSION

In this paper, we have proposed a framework to predict the best product in the e-commerce website by taking all the important aspects and opinions given by various customers. The framework mainly contains five components, i.e., product feature identification, opinion collecting, opinion mining^[7], classification, and Product Rating. We utilize the *Pros* and *Cons* opinions for improve the feature identification and opinion classification on text reviews. We then developed an aggregate ranking algorithm to summarize the importance of various features of a product from numerous users' reviews. The algorithm simultaneously inspects the aspect frequency and the influence of customer opinions are given to each feature over collected opinions. More over in this paper we implemented the comparison of products belonging to only one website, In future we will enhance this work to implement in comparing the product in different websites in order to get the best product with good quality, cost and more to satisfy the user requirements.

REFERENCES

- [1] X.Ding, B. Liu and P.S.Yu. :A Holistic Lexicon-based Approach to Opinion Mining. in Proc. of WSDM, pp. 231-240. USA. 2008
- [2] A. Ghose and P. G. Ipeirotis.: Estimating the Helpfulness andEconomic Impact of Product Reviews: Mining Text and Reviews Characteristics. in IEEE Trans. on Knowledge and Data Engineering,vol. 23, pp. 1498-1512. 2010.

- [3] V. Gupta and G. S. Lehal.: A Survey of Text Summarization Extractive Techniques. in Journal of Emerging Technologies in Web Intelligence, vol. 2, pp. 258-268. 2010.
- [4] W. Jin and H. H. Ho.: A novel lexicalized HMM-based learning framework for web opinion mining. in Proc. of ICML, pp. 465-472.
- [5] M. Hu and B. Liu.: Mining and Summarizing Customer Reviews. in Proc. of SIGKDD, pp. 168-177. Seattle, WA, USA, 2004. Montreal, Quebec, Canada, 2009.
- [6] F. Li, C. Han, M. Huang, X. Zhu, Y.-J. Xia, S. Zhang, and H. Yu: Structure-Aware Review Mining and Summarization. in Proc. of COLING, pp. 653-661. Beijing, China, 2010.
- [7] B. Pang and L. Lee.: Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, pp. 1-135. Now publisher. 2008.
- [8] A. M. Popescu and O. Etzioni.: Extracting Product Features and Opinions from Reviews. in Proc. of HLT/EMNLP, pp. 339-346, Vancouver, Canada. 2005.

Author's Profile:



KAVITI SARAL KUMAR Studying M.Tech (2nd year Pursuing) Computer Science and Engineering.



MALLIPUDI REVATHI Associate Professor
Qualification: M.Tech