

CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON E-COMMERCE WEBSITES

BARRE SUMATHI*, M TILAK**

PG SCHOLAR*, ASST.PROFESSOR**

E-MAIL: SUMATHIVADADI@GMAIL.COM*, TILAK_MP@YAHOO.COM**

SKBR PG COLLEGE, AMALAPURAM, E.G.DIST, ANDHRA PRADESH – 533201

ABSTRACT:

Online reviews have become an important source of information for users before making an informed purchase decision. Early reviews of a product tend to have a high impact on the subsequent product sales. In this paper, we take the initiative to study the behavior characteristics of early reviewers through their posted reviews on two real-world large e-commerce platforms, i.e., Amazon and Yelp. In specific, we divide product lifetime into three consecutive stages, namely early, majority and laggards. A user who has posted a review in the early stage is considered as an early reviewer. We quantitatively characterize early reviewers based on their rating behaviors, the helpfulness scores received from others and the correlation of their reviews with product popularity. We have found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our analysis of product reviews also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. By

viewing review posting process as a multiplayer competition game, we propose a novel margin-based embedding model for early reviewer prediction. Extensive experiments on two different e-commerce datasets have shown that our proposed approach outperforms a number of competitive baselines

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 [1]. It has been reported about 71% of global online shoppers read online reviews before purchasing a product [2]. 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 [3]. 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 [4], [5]. 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 [6]. For example, Amazon, one of the largest e-commerce company in the world, has advocated the Early Reviewer Program¹, 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 prerelease 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 behavior 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 [8]. The analysis and detection of early adopters in the diffusion of innovations have attracted much attention from the research community. Three fundamental elements of a diffusion process have been studied: attributes of an innovation, communication channels, and social network structures [8]. However, most of these studies are theoretical analysis at the macro level and there is a lack of quantitative investigations. With the rapid growth of online social platforms and the availability of a high volume of social networking data, studies of the diffusion of innovations have been widely conducted on social networks [9]–[12]. However, in many application domains, social networking links or communication channel are unobserved. Hence, existing methods relying on social network structures or communication channels are not suitable in our current problem of predicting early reviewers from online reviews. To model the behaviors of early

reviewers, we develop a principled way to characterize the adoption process in two real-world large review datasets, i.e., Amazon and Yelp. More specially, given a product, the reviewers are sorted according to their timestamps for publishing their reviews. Following [8], we divide the product lifetime into three consecutive stages, namely early, majority and laggards. A user who has posted a review in the early stage is considered as an early reviewer. In our work here, we mainly focus on two tasks, the first task is to analyze the overall characteristics of early reviewers compared with the majority and laggard reviewers. We characterize their rating behaviors and the helpfulness scores received from others and the correlation of their reviews with product popularity. The second task is to learn a prediction model which predicts early reviewers given a product. To analyze the characteristics of early reviewers, we take two important metrics associated with their reviews, i.e., their review ratings and helpfulness scores assigned by others. We have found that (1) an early reviewer tends to assign a higher average rating score to products; and (2) an early reviewer tends to post more helpful reviews. Our above findings can find relevance in the classic principles of personality variables theory from social science, which mainly studies how innovation is spread over time among the

participants [8]: (1) earlier adopters have a more favorable attitude toward changes than later adopters; and (2) earlier adopters have a higher degree of opinion leadership than later adopters. We can relate our findings with the personality variables theory as follows: higher average rating scores can be considered as the favorable attitude towards the products, and higher helpfulness votes of early reviews given by others can be viewed as a proxy measure of the opinion leadership. Our analysis also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. We further explain this finding with the herd behavior widely studied in economics and sociology [13]–[15]. Herd behavior refers to the fact that individuals are strongly influenced by the decisions of others. To predict early reviewers, we propose a novel approach by viewing review posting process as a multiplayer competition game. Only the most competitive users can become the early reviewers w.r.t. to a product. The competition process can be further decomposed into multiple pairwise comparisons between two players. In a two-player competition, the winner will beat the loser with an earlier timestamp. Inspired by the recent progress in distributed representation learning [16], [17], we propose to use a margin-based embedding model by first mapping both

users and products into the same embedding space, and then determining the order of a pair of users given a product based on their respective distance to the product representation. Previous studies have highly emphasized the phenomenon that individuals are strongly influenced by the decisions of others, which can be explained by herd behavior [6], [13]–[15], [18]–[20]. The influence of early reviews on subsequent purchase can be understood as a special case of herding effect. Early reviews contain important product evaluations from previous adopters, which are valuable reference resources for subsequent purchase decisions. As shown in [19], when consumers use the product evaluations of others to estimate product quality on the Internet, herd behavior occurs in the online shopping process [19]. Different from existing studies on herd behavior, we focus on quantitatively analyzing the overall characteristics of early reviewers using large-scale real-world datasets. In addition, we formalize the early reviewer prediction task as a competition problem and propose a novel embedding based ranking approach to this task.

EXISTING SYSTEM:

Previous studies have highly emphasized the phenomenon that individuals are strongly influenced by the

decisions of others, which can be explained by herd behavior. The influence of early reviews on subsequent purchase can be understood as a special case of herding effect. Early reviews contain important product evaluations from previous adopters, which are valuable reference resources for subsequent purchase decisions. As shown in, when consumers use the product evaluations of others to estimate product quality on the Internet, herd behavior occurs in the online shopping process. Different from existing studies on herd behavior, we focus on quantitatively analyzing the overall characteristics of early reviewers using large-scale real-world datasets. In addition, we formalize the early reviewer prediction task as a competition problem and propose a novel embedding based ranking approach to this task. To our knowledge, the task of early reviewer prediction itself has received very little attention in the literature. Our contributions are summarized as follows:

We present a first study to characterize early reviewers on an e-commerce website using two real-world large datasets. We quantitatively analyze the characteristics of early reviewers and their impact on product popularity. Our empirical analysis provides support to a series of theoretical conclusions from the sociology and economics. We view review

posting process as a multiplayer competition game and develop an embedding-based ranking model for the prediction of early reviewers. Our model can deal with the cold-start problem by incorporating side information of products. Extensive experiments on two real-world large datasets, i.e., Amazon and Yelp have demonstrated the effectiveness of our approach for the prediction of early reviewers.

PROPOSED SYSTEM:

To predict early reviewers, we propose a novel approach by viewing review posting process as a multiplayer competition game. Only the most competitive users can become the early reviewer's w.r.t. to a product. The competition process can be further decomposed into multiple pairwise comparisons between two players. In a two-player competition, the winner will beat the loser with an earlier timestamp. Inspired by the recent progress in distributed representation learning, we propose to use a margin-based embedding model by first mapping both users and products into the same embedding space, and then determining the order of a pair of users given a product based on their respective distance to the product representation.

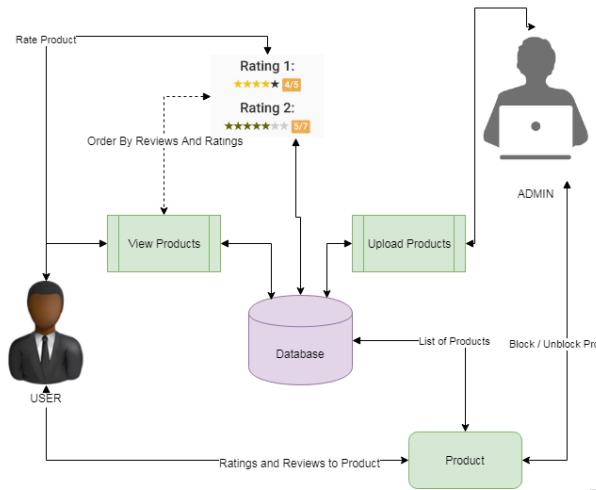
ALGORITHM:

Naïve Bayes Algorithm:

In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features. Naive Bayes has been studied extensively since the 1950s. It was introduced under a different name into the text retrieval community in the early 1960s, and remains a popular (baseline) method for text categorization, the problem of judging documents as belonging to one category or the other (such as spam or legitimate, sports or politics, etc.) with word frequencies as the features. With appropriate pre-processing, it is competitive in this domain with more advanced methods including support vector machines. It also finds application in automatic medical diagnosis.[3] Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers. In the statistics and computer science literature, naive Bayes models are known under a variety of names, including simple Bayes

and independence Bayes. All these names reference the use of Bayes' theorem in the classifier's decision rule, but naive Bayes is not (necessarily) a Bayesian method

ARCHITECTURE:



MODULES:

1. UPLOAD PRODUCTS

Uploading the products is done by admin. Authorized person is uploading the new arrivals to system that are listed to users. Product can be uploaded with its attributes such as brand, color, and all other details of warranty. The uploaded products are able to block or unblock by users.

2. PRODUCT REVIEW BASED ORDER

The suggestion to user's view of products is listed based on the review by user and rating to particular item. Naïve bayes

algorithm is used in this project to develop the whether the sentiment of given review is positive or negative. Based on the output of algorithm suggestion to users is given. The algorithm is applied and lists the products in user side based on the positive and negative.

3. RATINGS AND REVIEWS

Ratings and reviews are main concept of the project in order to find effective product marketing. The main aim of the project is to get the user reviews based on how they purchased or whether they purchased or not. The major find out of the project is when they give the ratings and how effective it is. And this will helpful for the users who are willing to buy the same kind of product.

4. DATA ANALYSIS

The main part of the project is to analysis the ratings and reviews that are given by the user. The products can be analysis based on the numbers which are given by user. The user data analysis of the data can be done by charts format. The graphs may vary like pie chart, bar chart or some other charts.

CONCLUSION:

In this paper, we have studied the novel task of early reviewer characterization and prediction on two real-world online review datasets. Our empirical analysis strengthens a series of theoretical conclusions from sociology and economics. We found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our experiments also indicate that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity at a later stage. We have adopted a competition-based viewpoint to model the review posting process, and developed a margin based embedding ranking model (MERM) for predicting early reviewers in a cold-start setting.

REFERENCES

- [1] J. McAuley and A. Yang, "Addressing complex and subjective product-related queries with customer reviews," in WWW, 2016, pp. 625–635.
- [2] N. V. Nielsen, "E-commerce: Evolution or revolution in the fastmoving consumer goods world," nngroup. com, 2014.
- [3] W. D. J. Salganik M J, Dodds P S, "Experimental study of inequality and unpredictability in an artificial cultural market," in ASONAM, 2016, pp. 529–532.
- [4] R. Peres, E. Muller, and V. Mahajan, "Innovation diffusion and new product growth models: A critical review and research directions," International Journal of Research in Marketing, vol. 27, no. 2, pp. 91 – 106, 2010.
- [5] L. A. Fourt and J. W. Woodlock, "Early prediction of market success for new grocery products." Journal of Marketing, vol. 25, no. 2, pp. 31 – 38, 1960.
- [6] B. W. O, "Reference group influence on product and brand purchase decisions," Journal of Consumer Research, vol. 9, pp. 183–194, 1982.
- [7] J. J. McAuley, C. Targett, Q. Shi, and A. van den Hengel, "Imagebased recommendations on styles and substitutes," in SIGIR, 2015, pp. 43–52.
- [8] E. M.Rogers, Diffusion of Innovations. New York: The Rise of HighTechnology Culture, 1983.
- [9] K. Sarkar and H. Sundaram, "How do we find early adopters who will guide a resource constrained network towards a desired distribution of behaviors?" in CoRR, 2013, p. 1303.
- [10] D. Imamori and K. Tajima, "Predicting popularity of twitter accounts through the discovery of link-propagating early adopters," in CoRR, 2015, p. 1512.

- [11] X. Rong and Q. Mei, “Diffusion of innovations revisited: from social network to innovation network,” in CIKM, 2013, pp. 499– 508.
- [12] I. Mele, F. Bonchi, and A. Gionis, “The early-adopter graph and its application to web-page recommendation,” in CIKM, 2012, pp. 1682–1686.
- [13] Y.-F. Chen, “Herd behavior in purchasing books online,” Computers in Human Behavior, vol. 24(5), pp. 1977–1992, 2008.
- [14] Banerjee, “A simple model of herd behaviour,” Quarterly Journal of Economics, vol. 107, pp. 797–817, 1992.
- [15] A. S. E, “Studies of independence and conformity: I. a minority of one against a unanimous majority,” Psychological monographs: General and applied, vol. 70(9), p. 1, 1956