

Improving Product Sales on E-Commerce Websites Based on Reviewers Opinions

S. Kayalvili¹, Muthu Priyadharshini. A²

¹M.E, Associate Professor (Sr. Gr), Department of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India

²PG Scholar, Department of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India

ABSTRACT

Online reviews are the important source of information for users before selecting a product for making a decision. Reviews of product particularly early reviews have impact on the product sales. Study the behavior characteristics of early reviewers through their posted early reviews. At first divide the product lifetime into three stages- Early, majority and laggards. A person who posts reviews in early stage is considered as early reviewers. The Early reviewers are the first one who responds to the product at the beginning stage. Rating behaviors of early reviewers are predicted based on k-means with Page Rank.

Keywords: Online reviews, Early reviewers, K-means with pageRank

I. INTRODUCTION

A.DATA MINING

Data mining is the process of extracting hidden information from large databases that are predictable, is a new technology with great potential to help companies to focus on the important information in their data warehouses. Data mining tools predict future trends and behaviors, allowing businesses to form knowledge-driven decisions. The automated tools offer analyses of past events. The support provided by retrospective tools is typical of decision support systems. Data mining tools can answer business questions to improve marketing that traditionally were time consuming to resolve. Finding predictive information that experts may find difficult to resolve because it lies outside their expectations.

B.THE FOUNDATION OF DATA MINING

The Foundation of Data Mining techniques are the result of a long process of research and product

development. AR& D group conducted a recent study on data warehouse projects found that 19% of respondents are beyond the 50 gigabyte level, while 59% expect to be there by second quarter of 1996. Computational engines can now be met in a cost-effective manner with parallel multiprocessor computer technology. Data mining algorithms use techniques that have existed for more than 5 years, but have only recently been implemented as mature, reliable, understandable, predictable tools that perform better than old statistical methods.

C.SOCIAL FEASIBILITY

The characteristic of study is to check the possibility of acceptance of the system by the user. This includes the process of training the user with the appropriate tool to use the system efficiently. The user must not feel that the system is threatening, instead must accept it as a necessity. Their level of confidence must be raised so that they are able to make some

constructive condemnation, which is welcomed, as the final user of the system.

D. OBJECTIVES

Input Design is the process of converting a user-perception description of the input into a computer-based system. The design created by the user is important to avoid errors in the data input process and show the exact path to the management for getting accurate information from the computerized system that has been trained by the user.

The error rectification is achieved by creating user-friendly screens for the data entry to handle large volume of data. The aim of designing user friendly input screen is to make data entry easier and to be free from errors. The designing process involves data entry screen such that all the data manipulates can be performed. The easy to use screen provides record viewing facilities.

II. RELATED WORKS

A. PREDICTING POPULARITY OF TWITTER ACCOUNTS THROUGH THE DISCOVERY OF LINK-PROPAGATING EARLY ADOPTERS

The method adopted in this paper is used for ranking recently created Twitter accounts according to their prospective popularity. Early detection of new promising accounts is useful for trend prediction, viral marketing, user recommendation, and so on. New accounts are, however, difficult to evaluate because they have not established their reputations, and they cannot apply existing link-based or other popularity-based account evaluation methods. Users who often find new good information sources earlier than others. Our method then regards new accounts followed by good early adopters as promising, even if they do not have many followers now. How many times the follow links from the account have been copied by its followers. If its followers have copied

many of its follow links in the past, the account must be an early adopter, who find good information sources earlier than its followers. A method of inferring links which are created by copying which links. One advantage of this method is that the method only uses information that can be easily obtained only by crawling neighbors of the target accounts in the current Twitter graph. Evaluation of this method is conducted by an experiment on Twitter data. Chose then-new accounts from an old snapshot of Twitter, compute their ranking by our method, and compare it with the number of followers.

B. PREDICTING MATCHUPS AND PREFERENCES IN CONTEXT

Present a general probabilistic framework for predicting the outcome of pairwise matchups and pairwise preferences both of which have widespread applications ranging from matchmaking in computer games to recommendation in e-commerce. Unlike existing models for these tasks, our model not only learns representations of the items in a more expressive latent vector space, but also models how context modifies matchup and preference outcomes. The context “weather” may alter the winning probability in a tennis match, or the fact that the user is on a mobile device may alter his preferences among restaurants. More generally, the model is capable of handling any symmetric game comparison problem that can be described by vectored player item and game context features. A comprehensive evaluation of its predictive performance with real datasets from both domains to show its ability to predict preference and game outcomes more accurately than existing models.

C. DISTRIBUTED REPRESENTATIONS FOR CONTENT-BASED AND PERSONALIZED TAG RECOMMENDATION

Consider the problem of learning distributed representations for documents from their content and

associated tags, and of distributed representations of users from documents and tags provided by users. The documents, words, and tags are represented as low-dimensional vectors and are jointly learned with a multi-layered neural language model. Propose a two stage method where in the first stage which consists of two layers, we exploit the corpus wide topic-level information contained in tags to model one layer of neural language model and use document level words sequence information to model other layer of the proposed architecture. In the second stage, obtained document and tags representations to learn user representations. Utilize these jointly trained vector representations for personalized tag recommendation tasks.

D. INFERRING NETWORKS OF SUBSTITUTABLE AND COMPLEMENTARY PRODUCTS

In a modern recommender system, it is important to understand how products relate to each other. While a user is looking for mobile phones, it might make sense to recommend other phones, but once they buy a phone, instead want to recommend batteries, cases, or chargers. We formulate this as a supervised link prediction task, where we learn the semantics of substitutes and complements from data associated with products. The primary source of data we use is the text of product reviews, though our method also makes use of features such as ratings, specifications, prices, and brands. Methodologically, we build topic models that are trained to automatically discover topics from text that are successful at predicting and explaining such relationships. Experimentally, we evaluate our system on the Amazon product catalog, a large dataset consisting of 9 million products, 237 million links, and 144 million reviews.

E. IMPROVING MUSIC RECOMMENDATION USING DISTRIBUTED REPRESENTATION

A music recommendation approach based on distributed representation is presented. The proposed

approach firstly learns the distributed representations of music pieces and acquires users' preferences from listening records. Then, it recommends appropriate music pieces whose distributed representations are in accordance with target users preferences.

F. EXPERIMENTAL STUDY OF INEQUALITY AND UNPREDICTABILITY IN AN ARTIFICIAL CULTURAL MARKET

Hit songs, books, and movies are many times more successful than average, suggesting that the best alternatives are qualitatively different from the rest yet experts routinely fail to predict which products will succeed. We investigated this paradox experimentally, by creating an artificial "music market" in which 14,341 participants downloaded previously unknown songs either with or without knowledge of previous participants' choices. Increasing the strength of social influence increased both inequality and unpredictability of success.

G. IMAGE BASED RECOMMENDATIONS ON STYLES AND SUBSTITUTES

Humans inevitably develop a sense of the relationships between objects, some of which are based on their appearance. Some pairs of objects might be seen as being alternatives to each other, while others may be seen as being complementary. This information guides many of the choices that people make, from buying clothes to their interactions with each other. We seek here to model this human sense of the relationships between objects based on their appearance. Problem defined on graphs of related images, and provides a large-scale dataset for the training and evaluation of the same.

H. LEARNING HIERARCHICAL REPRESENTATION MODEL FOR NEXT BASKET RECOMMENDATION

Next basket recommendation is a crucial task in market basket analysis. Given a user's purchase history, usually a sequence of transaction data, one attempt to build a recommender that can predict the next few items that the user most probably would like. Ideally, a good recommender should be able to explore the sequential behavior, as well as account for users' general taste for recommendation. Moreover, these two factors may interact with each other to influence users' next purchase.

I. MODELING TRANSITIVITY INMATCHUP AND COMPARISON DATA

A method for learning potentially intransitive preference relations from pair wise comparison and matchup data. Unlike standard preference-learning models that represent the properties of each item/player as a single number, our method infers a multi-dimensional representation for the different aspects of each item/player's strength.

III. PROPOSED SYSTEM

To overcome the problems in the previous versions the text based reviews are introduced. The reviews are divided into three stages Early, majority and laggards. The person who posting the review in the first stage is said to be Early reviews. This Early reviews are the important for the product success and failure. By analyzing the early reviewers by using their reviews we can predict the product future. K-means with PageRank is used to analyzing the early reviewers.

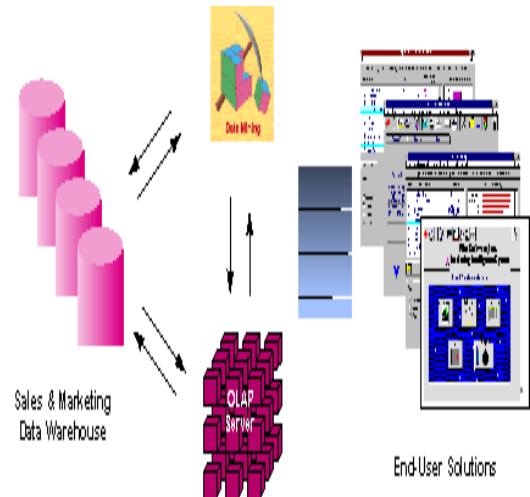


Fig 1 AN ARCHITECTURE OF E-COMMERCE

IV. IDENTIFICATION OF EARLY REVIEWERS

- Administration and Registration
- View reviews and Pre-processing
- Feature extraction
- Classification

A. ADMINISTRATION AND REGISTRATION

Admin can view the details about the users Admin module get all the user rating and review the finalised result which is early and other reviews.

Web system is connected with local host. The user enters the personal details for registering in the web system.

B. VIEW REVIEWS AND PREPROCESSING

The admin can view all the users review. The user can also view the reviews.

The pre-processing step is used for separating the words to find meaning for the words. Tokenization, stemming are processed in pre-processing steps.

English words like 'write' can be inflected with a morphological suffix to produce 'writes, writer, written'. They share the same stem 'write'. Often it is beneficial to map all inflected forms into the stem. This is a complex process, since there can be many exceptional cases

The most commonly used stemmer is the Porter Stemmer. There are many others. They are not perfect and they do make mistakes. Many are not designed for special domains like biological terms. Some other languages are particularly hard.

The words are stored in product database. It contains product details for example company name, price, date of manufacturing, quantity, brand and image.

C. FEATURE EXTRACTION

Based on the features extracted from the early review and other reviews we should analyze the word counts. It's based on word count and stamping should be trained in database.

D. CLASSIFICATION

The decision makes rule which provides the results for early or other reviews.

Based on the sentence we trained into the database it provides the result for classification. The extracted data from database is used to provide positive and negative opinions for the products.

Graph is drawn based on reviews. From overall product ratings interested and uninterested product ratings are identified. Based on the positive and negative count, the product final review is stored in the database list. So the Online companies can choose which is the best product to buy for sale that improves business development.

V. CONCLUSION

The novel tasks of early reviewer characterization and prediction on two real-world online review datasets have been studied. Our empirical analysis strengthens a series of theoretical conclusions from sociology and economics. An early reviewer tends to assign a higher average rating score; and 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.

The review content is not considered. We will explore effective ways in incorporating review content into our early reviewer prediction model. Currently, we focus on the analysis and prediction of early reviewers, while there remains an important issue to address, i.e., how to improve product marketing with the identified early reviewers. We will investigate this task with real e-commerce cases in collaboration with e-commerce companies in the future.

VI. REFERENCES

- [1]. S. Chen and T. Joachims, "Modeling intransitivity in matchup and comparison data," in WSDM, 2016, pp. 227-236.
- [2]. S. L. G. X. Dongjing Wang, ShuiGuang Deng, "Improving music recommendation using distributed representation," in WWW, 2016, pp. 125-126.
- [3]. D. Imamori and K. Tajima, "Predicting popularity of twitter accounts through the discovery of link-propagating early adopters," in CoRR, 2015, p. 1512.
- [4]. B. Liu, Sentiment Analysis - Mining Opinions, Sentiments, and Emotions. Cambridge University Press, 2015.
- [5]. J. J. McAuley, C. Targett, Q. Shi, and A. van den Hengel, "Image-based recommendations on styles and substitutes," in SIGIR, 2015, pp. 43-52.
- [6]. J. McAuley and A. Yang, "Addressing complex and subjective product-related queries with customer reviews," in WWW, 2016, pp. 625-635.

- [7]. N. V. Nielsen, "E-commerce: Evolution or revolution in the fast- moving consumer goods world," nngroup. com, 2014.
- [8]. X. Rong and Q. Mei, "Diffusion of innovations revisited: from social network to innovation network," in CIKM, 2013, pp. 499- 508.
- [9]. W. D. J. Salganik M J, Dodds P S, "Experimental study of in- equality and unpredictability in an artificial cultural market," in ASONAM, 2016, pp. 529-532.
- [10]. 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.
- [11]. A. A. SaurabhKataria, "Distributed representations for content- based and personalized tag recommendation," in ICDM Workshop- s, 2015, pp. 1388-1395.
- [12]. S. Usami, "Individual differences multidimensional bradey-terry model using reversible jump markov chain montecarloalgorith- m," *Behaviormetrika*, vol. 37(2), p. 135155, 2010.
- [13]. P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng, "Learning hi- erarchical representation model for next basket recommendation," in SIGIR, 2015, pp. 403-412.

Cite this article as :

S. Kayalvili, Muthu Priyadharshini. A, "Improving Product Sales on E-Commerce Websites Based on Reviewers Opinions", *International Journal of Scientific Research in Science and Technology (IJSRST)*, Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 7 Issue 3, pp. 361-366, May-June 2020. Available at
doi : <https://doi.org/10.32628/IJSRST207366>
Journal URL : <http://ijsrst.com/IJSRST207366>