



International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

2019 IJIEMR. Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 7th Mar 2018. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-03](http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-03)

Title: **PREDICTING USER OPINION IN E-COMMERCE WEBSITE USING AGGREGATE RANKING ALGORITHM**

Volume 08, Issue 03, Pages: 96–103.

Paper Authors

MR.K.HARIKRISHNA, S. MUZAMMIL HUSSAIN

Vignan's Lara Institute of Technology & Science



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code

PREDICTING USER OPINION IN E-COMMERCE WEBSITE USING AGGREGATE RANKING ALGORITHM

MR.K.HARIKRISHNA¹, S. MUZAMMIL HUSSAIN²

Assistant Professor¹, Department of M.C.A, Vignan's Lara Institute of Technology & Science

M.C.A Student², Department of M.C.A, Vignan's Lara Institute of Technology & Science

ABSTRACT: Online business is the easiest way of shopping. In online business, users can buy the products by viewing the feedbacks or reviews of the other users who are used the products earlier. Based on those opinions the product can get rank. But the user has to read a lot of reviews for a particular product in order to get the best product. It was the time taking process. In this paper we are supposed to propose a system that we can directly collect the reviews of the products from online and by comparing those reviews we can get the best product based on the good opinions given by earlier users of that product.

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

INTRODUCTION

Online business, in other words it will be called as E-Commerce or Electronic business. In Online business all the transactions are going to be done through internet only. Users can easily get the desired products. The services, payments and guidance for the usage of the product will completely be done by using web based technologies. The online business is same as the ordinary business, But the only difference between is in E-business all the transactions are done by web based technology only. In online we can get all the services like banking, movie tickets, hotel booking, air tickets, E-booking, trading, etc. In online we can get any type of product. There are several websites for Online business example Amazon, Flipkart, Paytm, Snapdeal, etc. Each website is having a numerous type of product varieties. For example, Amazon website is one of the

greatest website in E-commerce, firstly it started an online book store with a wide variety of books later it became a store for all the products. Now, Amazon will sell more than 200 million products in USA under 35 categories. In clothing it has 5 million items in Electronics it is having more than 24 million products. Today in India the average sale of products only from the Amazon is about 18 million products[2]. For such online marketing, huge numbers of reviews are given by the users for the products they purchased from the site. Based on those reviews the other users can be able to know what is the good product. Such a user comments are having a high knowledge on the product. Every consumer wants the good quality product for them, so they must go with the reviews posted by others. These reviews are very much important for both the consumer and the firms. As the

Page 97

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 $r \in R$ 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 considered 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. PROPOSED SYSTEM

The Existing system contains, feature identification[3] include both the supervised and unsupervised methods. In the supervised method, by using the labeled reviews extract

the collection of words. Existing aspect identification include the supervised and unsupervised methods. Most existing supervised methods are based on the successive learning mechanism such as Hidden Markov Models and Conditional Random Fields. HMM model to extract aspects and opinion expressions. All these Existing methods require sufficient labeled samples for training. The next approach is unsupervised approach. In this product aspects are nouns and noun phrases. First all the nouns and phrases are going to be collected in this approach. The term frequency is going to be calculated and the most frequent terms are taken as the features of the products[8]. To overcome the Disadvantages in the existing system i.e., in supervised approach. Supervised method is time-consuming and labor intensive to label samples. It has no content related to mining aspect importance and ranking aspects according to their importance. In the paper proposed system, they propose the ranking framework based on the important aspects of the products by using the probabilistic aspect ranking algorithm. In that, summarize the documents based on the rankings given by the user for the particular product and get the overall opinions of the product. In this, we just get the opinions and feedbacks of the users about some products. We will enhance the work by implementing the recommendation systems. By using the rating and reviews of the particular products we can get the exact results of particular product or items using recommendation system. In this we combine the three

methods for estimating the best products among various products. They are

- (a) Frequency-based method,
- (b) Correlation-based method and
- (c) Hybrid method as discussed in the above section.

In the proposed we are also comparing the products similarities by a model-based algorithm i.e., an item based collaborative filtering model. This is used in giving recommendations to the user about the good opinioned products which are having high ranking. In the algorithm, we calculate the similarities among the various products in the dataset. The similarity values among the various products are taken by seeing all the users who have rated both the items. As shown in the fig.3. The similarities between different products are totally depending upon the rankings given by the users.

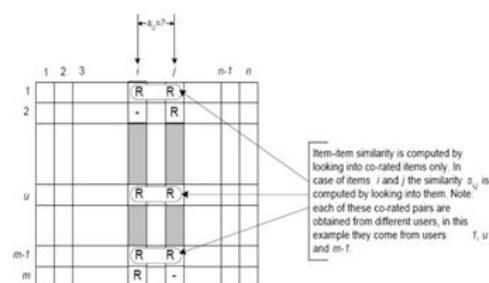


Fig.3. Item-Item Similarity Computing

Item to Item similarity is done by taking the reviews from all the users of different products.

IV. DATA SETS AND EXPERIMENTAL RESULTS

In our proposed system, the review datasets of the products are directly taken[6] in online from the amazon website www.amazon.com by using the import.io

tool After feeding the online data set we have to download a .csv file of our desired product, We can collect the reviews of numerous users for the product and now these reviews can be viewed by product wise. With the help of online review data sets, the user's opinions are getting collected and we can get the users feedback for the products. After applying the data preprocessing technique and the stop words are removed .Now the wordlist is going to be viewed comparing this word list with a bag of words containing Good opinions and Bad opinions after this opinion mining[4] we can get the high opinioned products .Those products can be viewed in the browser.

A.Data sets for mobile

Tab.1. Data Set for Mobile

Product	Consumer Name	Review Date	Color	Comments
Meizu m2 Note	M.Ganesan	23sep2015	White	9of19 found review helpful
Meizu m2 Note	Sonu Biradar	20sep2015	Grey	216of241 found review helpful
Meizu m2 Note	Aux	20sep2015	White	532of608 found review helpful
Meizu m2 Note	VenkataTapseva Manthene	27sep2015	Grey	22of25 found review helpful

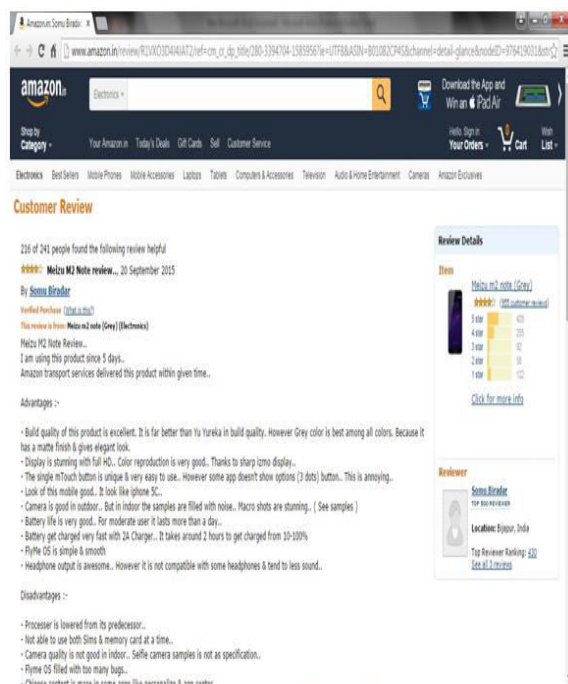
The above data set is the sample data set for the product mobiles containing the reviews of four consumers for two different models. This is the data set of reviews downloaded from the amazon website in the form of .csv file. For this data set we have to implement the Data preprocessing and the data classification techniques in order to get the

good opinioned product directly. The four customer reviews for the product mobile are shown in fig.4. Collected from the amazon website for the mobiles of different varieties. And various number of other users also gave their feed back to the product in the form of rating, Here the users can able to give the textual reviews in the paragraphs by writing .These reviews contains the pros and cons of the product, with the help of these reviews others can take the decision which is the best product and the firm can improve the product by the feedback given by other users. In the particular paragraph review we first split the textual reviews into the sentences and these sentences are splitted into nouns and phrases then these are classified into the good opinions and bad opinions by comparing[9] with the bag of words.

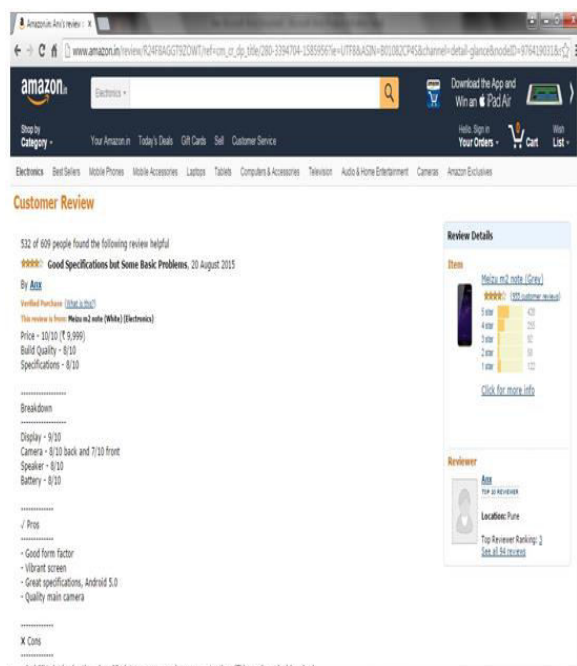


Review given by Ganesan on the meizu m2 note white colour in the amazon website

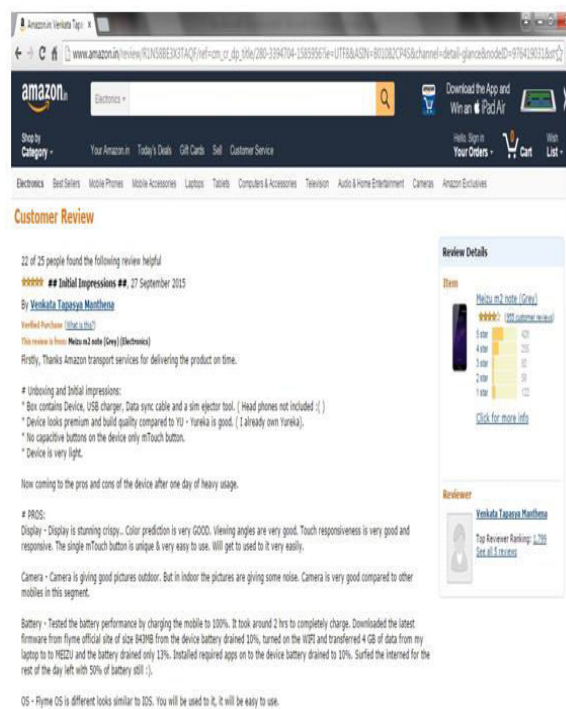
which was taken in our dataset.



Mr.Sonu Biradar's review on the Meizu m2 note(Grey) Review given by Sonu Biradar on the meizu m2 note grey colour in the amazon website which was taken in our dataset.



Mr.Anx's review on the Meizu m2 note(white) Review given by Anx on the meizu m2 note white colour in the amazon website which was taken in our dataset.



Mr. VenkataTapaseya Manthenea's review on the Meizu m2 note(Grey) Review given by Venkata Tapaseya Manthenea on the meizu m2 note grey colour in the amazon website which was taken in our dataset.

FIG.4. Reviews of Consumers

The Comparison results of existing system to the proposed system are shown in the fig.5. the exisisting system is having low efficiency in time whereas the proposed system is having the higher efficiency of time i.e., the time taken by the proposed system to give the best product is low when comparing to the existing system.

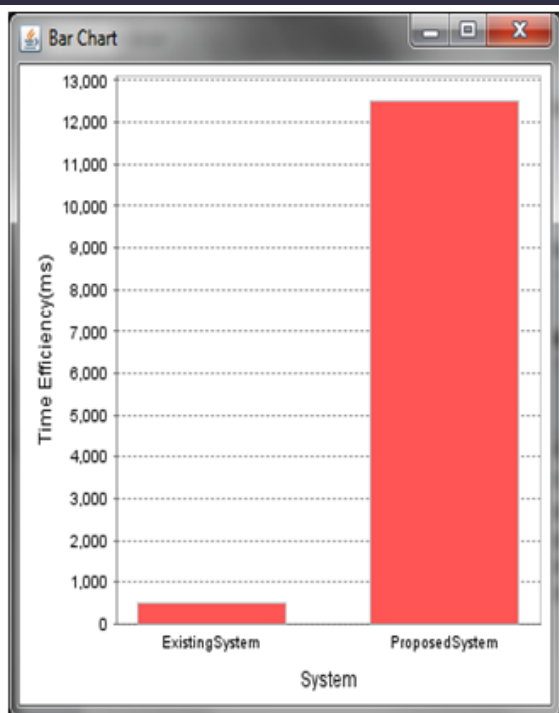


Fig.5. Time Efficiency Comparison

In this time comparison graph, we proved that the time efficiency of our proposed system is better than the existing system.

V.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.

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 learningframework 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. Nowpublisher. 2008.
- [8] A. M. Popescu and O. Etzioni.: Extracting Product Features and Opinions



International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org

from Reviews. in Proc. of HLT/EMNLP, pp. 339-346, Vancouver, Canada. 2005.

[9] B. Liu, M. Hu, and J. Cheng.: Opinion Observer: Analyzing and Comparing Opinions on the Web. in Proc. of WWW, pp. 342-351. Chiba, Japan. 2005.