

Online Product Ranking Based on Important Features Identified in Consumer Reviews

S.Nandhini¹, K.Selvam²

¹Research Scholar, Department of Computer Applications, Dr. M.G.R. Educational and Research Institute University, Chennai

²Professor, Department of Computer Applications, Dr. M.G.R. Educational and Research Institute University, Chennai
Email- nandhumca23@gmail.com

Abstract- The purpose of the work are used to identify the important features of the consumer reviews based on the product they purchased and used through online shopping. Consumer comments can play a significant role in the world of online shopping. Every comments of individuals are important as well as it enables others to know about the details of particular product and make them to take correct decision before they purchase the product. Reviews are beneficial for customer at the time of purchasing as well as firm to develop their concern. This paper consists the product ranking depends upon the features identified in consumer reviews. In a given customer reviews, important features are identified using shallow dependency parser, and classified them in to positive or negatives via sentiment classification using NLP, finally apply the ranking algorithm to determine the particular product ratings.

Keywords : Online shopping, consumer reviews, shallow dependency parser, sentiment classifier, product ranking.

I. INTRODUCTION

Online reviews can play a vital role for consumers to purchase products as well as firms to develop their marketing, customer relationship management and product development. Most of the retail websites needs consumer reviews based on features of products. Many retail websites like Bing shopping 1 can offered 5 millions of products through online, Amazon.com offer 36 million products. Shopper.com records 5 million products. CNet.com consists nearly 7 million product reviews whereas millions of reviews contains in pricegrabber.com. Most of the consumer says signal connection of the iphone4 was not good, but even though iphone4 still in high ratings. Comparing to the signal connection the design and speed must be important, but they are not frequently commented. Hence its prove truly important features cannot be identified by frequency based solutions. The product consists many features, but eventually some features are important than others. It provides greater difficulty for consumer to take decision and the firm for setting the strategies to develop their products. For e.g. When we see the features of camera like lenses, and their quality of the picture are more

important than other features like wrist strap and a/v cable. However, the reviews are disorganized and provide difficulties in information navigation and knowledge acquisition and also reduces the overall true ratings about the product both in worst case and good cases. Hence, it is impractical for people to identify the important features of products from large number of reviews through manually. Therefore, an approach to find out important features is highly demanded. The purpose of product ranking framework identified the important features of products based on online consumer reviews automatically and also aim to improve the usability of numerous reviews. Our proposed work specifies, based on the consumer reviews of a product . First identify features of the product using shallow dependency parser and determine their opinions into positive and negative via sentiment classifier using NLP, and develop probabilistic ranking algorithm for providing ratings to the particular product. Hence product ranking framework elaborates more discussions and analysis of the product features and also perform extensive evaluations on more products in more diverse domains .

II. METHODOLOGY

The three methodology are used in the proposed work.

- 1.Features or Aspect Identification of the product.
2. Classify the positive and negative comments of the product via sentiment classifier using NLP.
3. Probabilistic ranking algorithm used for product ranking .

Customer reviews In a product review website users commonly share their thoughts about a product used. That review generally contains both good and bad about the things. If new user search about a product review it displays in a same form. Hence its difficult to find out the original issue with the product. For ex a mobile have good camera resolution but worst battery features, means finding the reality is difficult in the existing system.

Features identification are used to identify the particular and important features of the product. Identity management (IM) are also used here to find out the features of product. The product feature is the important

factor, based on this a consumer can buy product are convey their opinions related to the product, hence consumer opinions are important for the firm to develop the business in various fields like marketing, product development, customer relationship management.

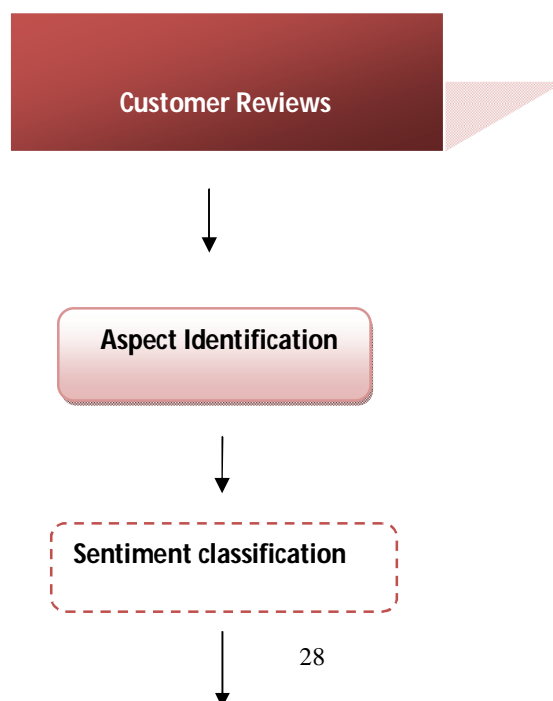
Sentiment classifier are used to classify the positive and negative comments on the user reviews, depends upon the user comments each words should be splitted and identified how many positive and negative words. Based upon the words ratings particular product will be ranked. In the general review extracted from the aspect identification have to classify the review whether which is shared for the good rating or bad rating. Ex the word headphone and earphone targets about the same product but in different words. Thus the review might be contain good rating without exclamation words are differentiated. corpus collection are used here to find out the negative and positive emotions in consumer review like Happy emoticons: “:-)”, “:.)”, “=:)”, “:D”. Sad emoticons: “:-(”, “:(”, “=(”, and also count how many times the particular words comes when we analyzing numerous reviews. Eg. The picture quality of camera is good. The camera lenses are too good. From that we examine the good comes in two times. NLP techniques are used find out the consumer reviews from their own language and make them into understandable format. Cross domain classification are used to specify the positive and negative comments from different domains. Eg. The book is well researched and interesting. Energy saving grill for using to make burger, it comes under kitchen appliances.

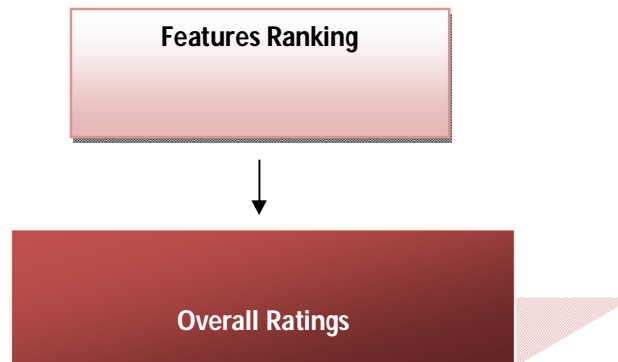
Product ranking are used to perform ratings to the particular product. It consists features, positive and negative comments of the user reviews and analyse them, within that apply the ranking algorithm and finally assume ratings. With the help of both aspect identification and the sentiment word classification the product overall and individual parts review is classified and ranked separately. This framework may differ from one website to other website.

III. FLOW AND ARCHITECTURE OF THE PROCESS

The user opens the retailer sites and enter their details for registration, once the user enter successfully they are able to search variety of products in which they want to buy. Before choosing the particular products, view all the features belong to the product, and also read the reviews already posted by numerous consumers, analyze the product ratings present for the product and also posted their opinions and comments to the particular product. Collect the set of consumer reviews for the particular product, if each individual consumer reviews consists m important features, identifying all features, classify them into positive and negatives. The particular features are ranked, based on the weightage the overall ratings assumed to the product. The Architecture diagram specifies that in a consumer reviews. The important features of the product identified, then analyze the positives and negative reviews of the product via sentiment classifier, based upon the overall opinions and the positive and negative ratings each product classified and ranked separately.

ARCHITECTURE DIAGRAM





IV. RANKING ALGORITHM

Input : collect the overall reviews of consumers , each review associated with overall ratings.

Output : perform ratings to the m features identified in the reviews.

Step 1 : classify the m features into positive and negatives

Step 2: Each features consists ratings, hence based on the ratings as well as the weightages.

Step 3: compute important features rating, based on that calculate overall ratings

Algorithm for the software

Step 1 : The admin can enter the retail website and post the products with its relevant features and also capable of him to modified and adding the required features of the product.

Step 2 : The user can enter the details in the user registration page before viewing the products.

Step 3 : The user information stored in the database and shows the message user enter the site successfully.

Step 4 : Then, the user can viewing the required product with its features and also possible for them to see the previous customer reviews and the product ratings for the particular product.

Step 5 : Finally the user can enter the review page to post their reviews related to the product.

Step 6 : The review page contains the three things, that to be filled by the user, it consists product name and the particular property of the product they want to write review, enter their own reviews and close the website.

Step 7 : Based on the consumer reviews, the important features are identified, and the features are classified into positive and negatives using sentiment classifier, finally

based on the weightage as well as the overall ratings of positive and negatives ranked will be allotted to the particular product.

V. ADVANTAGES OF THE PROPOSED WORK

1.This framework can provide true ratings of the product based on the analyzing of various consumer reviews.

2. The true ratings are identified based on the overall opinions of positive and negative comments .

3. The overall opinion generated based on weighted aggregation, whereas weights generally measure the degree of importance of the features of product.

4. It ensures more discussions, analyses and various comments.

5. The proposed work performs correct evaluations on more products in various domains.

6. It demonstrates the effectiveness of the product aspect ranking approach.

7. To achieve significant performance improvements.

8. It determines the overall opinion of a review document based on the number of positive and negative terms

VI. CONCLUSION AND FUTURE ENHANCEMENT

The framework contains three main components, i.e., Feature identification, aspect sentiment classification, and aspect ranking. This framework helps us to analyse features, negative and positive comments, ratings to the particular product automatically and easily. It provide a beneficial way for both the consumer and business firms.In future the framework can be implemented for a particular product with a help of multiple website comments. Hence the overall product aspect ratio in the open market can be analyzed. Our classifier is able to

determine positive, negative and neutral sentiments of documents based on the user mind. The classifier is based on the multinomial Naive Bayes classifier that uses N-gram and POS-tags as features. We will improve our future work with other learning methods. We also want to cluster the related object features to provide more concise review summary.

REFERENCES

- [1] J. C. Bezdek and R. J. Hathaway.: Convergence of alternating optimization.in *Journal of Neural, Parallel & Scientific Computations* vol. 11, pp. 351-368. USA. 2003.
- [2] C. C. Chang and C. J. Lin.: Libsvm: a Library for Support Vector Machines. <http://www.csie.ntu.edu.tw/~simjlin/libsvm/>, 2004.
- [3] G. Carenini, R. T. Ng, and E. Zwart.: Multi-document Summarization of Evaluative Text. in *Proc. of ACL*, pp. 3-7. Sydney, Australia. 2006.
- [4] China Unicom 100 Customers iPhone User Feedback Report, 2009.
- [5] ComScore Reports [http://www.comscore.com/Press Events/Press Releases](http://www.comscore.com/Press%20Events/Press%20Releases), 2011.
- [6] 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.
- [7] G. Erkan and D. R. Radev.: LexRank: Graph-based Lexical Centrality as Salience in Text Summarization. in *Journal of Artificial Intelligence Research*, vol. 22, pp. 457-479. 2004.
- [8] O. Etzioni, M. Cafarella, D. Downey, A. Popescu, T. Shaked, S. Soderland, D. Weld, and A. Yates.: Unsupervised Named-entity Extraction from the Web: An Experimental Study. in *Journal of Artificial Intelligence*, vol. 165, pp. 91-134. 2005.
- [9] A. Ghose and P. G. Ipeirotis.: Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Review Characteristics. in *IEEE Trans. on Knowledge and Data Engineering*, vol. 23, pp. 1498-1512. 2010.
- [10] 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.