A SURVEY ON PRODUCT ASPECT RANKING BASED ON ONLINE CUSTOMER REVIEWS

Vaibhav Anil Songire[1] and R. V. Patil[2]
[1]ME II (CSE), SSVPS'S B.S.Deore, College of Engg, Dhule
[2]Asst. Prof. (CSE), SSVPS'S B.S.Deore, College of Engg, Dhule

ABSTRACT -- Now day's e-commerce is rapidly growing which gives facility for customers to purchase products online. Varieties of brands and lots of products have been offered online and numbers of customer reviews are available on internet. These reviews are important for the customers as well as for the sellers. Most of the reviews are disorganized so it generates difficulty for using valuable information. So the product aspect ranking system, which identifies important aspects of products from online customer reviews and improve the usability of different reviews. The important product aspects are recognized using two observations: - 1) the important aspects are mostly commented by a large number of users 2) customers opinions on the important aspects are greatly effect on the overall opinions of the product. This paper provides the survey on various techniques of aspect identification and sentiment classification.

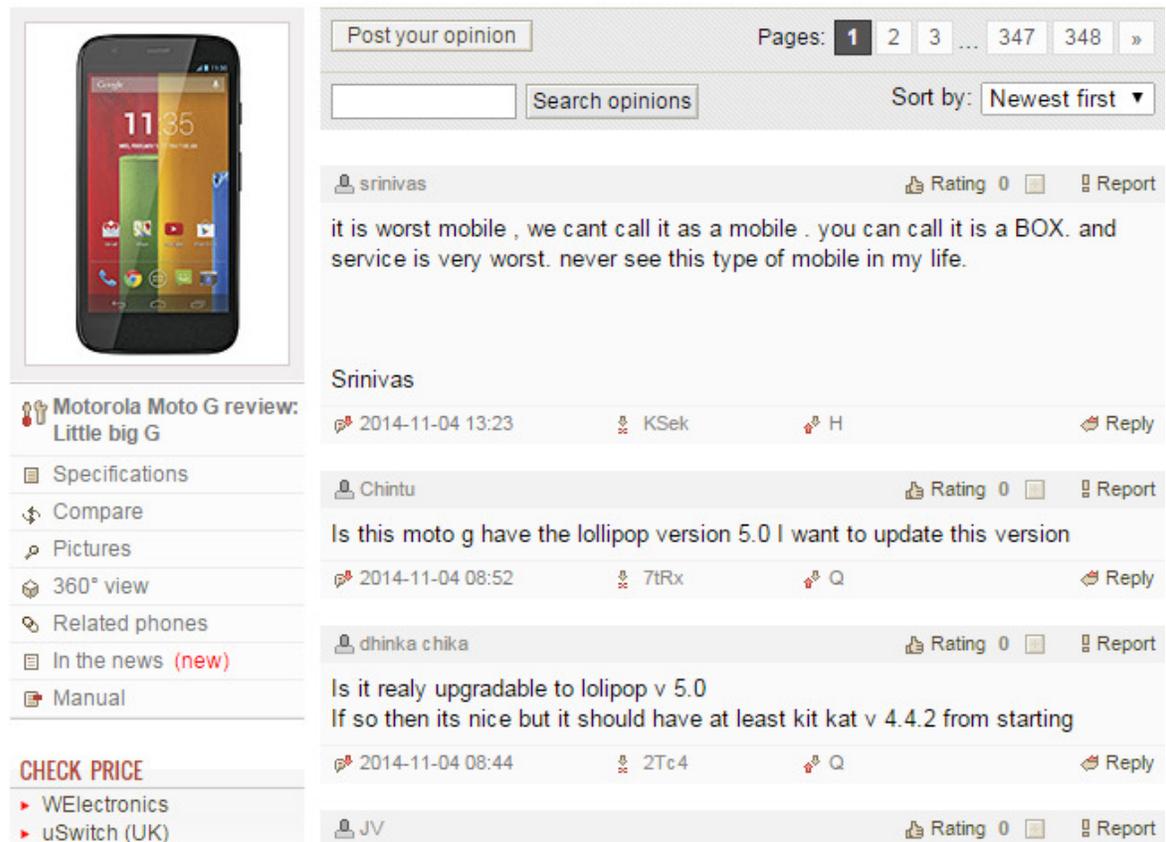**Keywords --** Customer reviews, product aspects, aspect identification, sentiment classification, aspect ranking

I. INTRODUCTION

In the recent years peoples trend towards online shopping increases day by day. There is a rapid expansion in e-commerce. There are many online retail shopping sites available and they indexed millions of products for selling. For example, Amazon.com has collection of more than 36 million products. Most retail websites encourage customers to write reviews to express their opinions on different *aspects* of the product. Here, an *aspect*, also called *feature*, means an attribute of a particular product. An example of review is "*The sound quality of Moto G is amazing.*" shows positive opinion on the aspect "*sound quality*" of product *Moto G*.

Other than retail websites, many forum websites also available. They provide a platform for customers to post reviews on millions of products. For example, CNet.com contains more than seven million product reviews, whereas Pricegrabber.com contains millions of reviews on more than 32 million products in 20 categories over 11,000 sellers [3]. Such numerous customer reviews contain rich and valuable knowledge and have become an important source for both customers and sellers. Customers commonly search quality information from online reviews for purchasing a product, while many sellers use online reviews as important feedbacks in their product development, marketing, and customer relationship management strategies [2].

Generally, a product may have number of aspects. For example, *Moto G* has more than hundred aspects such as "*memory*", "*screen size*", "*sound quality*", "*camera*." Some aspects are most important than the others, and have more impact on the customers decision making as well as seller's product development schemes. For example, aspects of *Moto G*, e.g., "*memory*" and "*camera*," are considered important by most of the customers, and are most important than the others such as "*color*" and "*button*s". Hence, identification of important product aspects plays an important role to improve the usability of numerous reviews and it is beneficial to both customers and sellers [1].

Figure 1. Example of Customer Reviews of Product Moto G [4]

Customers can easily make purchasing decision by giving attention to the important aspects, while sellers can focus on the enhancement of product quality. However, manual identification of important aspects from huge number of reviews is impractical. Therefore, an approach product aspect ranking is proposed to automatically identify the important aspects from online customer reviews.

II. LITERATURE SURVEY

This section briefly survey previous work on product aspect ranking system, starting with the product aspect identification with sentiment classification. Existing product aspect identification technique can be classified into two main approaches:- supervised and unsupervised.

Supervised learning technique learns an extraction model which is also called as aspect extractor, that aspect extractor is then used to identify aspects in new reviews. For this task Hidden Markov Models, Conditional Random Fields, Maximum Entropy and Naive Bayes Classifier approaches have been used. Wong and Lam used a supervised learning technique to train an aspect extractor. They learned aspect extractor using Hidden Markov Model and Conditional Random Field. All supervised techniques are reasonably effective, but preparation of training examples is time consuming.

In contrast, unsupervised approaches automatically extract product aspects from customer reviews without using training examples. Hu and Liu's works focuses on association rule mining based on the Apriori algorithm to mine frequent item sets as explicit product aspects. In association rule mining, the algorithm does not consider the position of the words in the sentence. In order to

remove incorrect frequent aspects, two types of pruning principles were used: compactness and redundancy pruning. The technique is effective which does not require the use of training examples or predefined sets of domain-independent extraction patterns. However, it suffers from two main shortcomings. First, frequent aspects discovered by the mining algorithm might not be product aspects. The compactness and redundancy pruning rules are not able to eliminate these false aspects. Second, even if a frequent aspect is a product aspect, customers may not be expressing any subjective opinion about it in their reviews [9].

Wu et al also used the unsupervised method. They used the phrase dependency parser to extract noun and noun phrases and then they used a language model to filter out the unwanted aspects. This language model was used to predict the related score of candidate aspects and was built on product reviews. Candidate having low score were filtered out. However, this language model might be biased to frequent terms in the reviews and cannot predict the aspect score exactly as a result cannot filter out noise very efficiently. Then, Popescu and Etzioni developed the *OPINE* system, which extracts aspects based on the *KnowItAll* web information extraction system [8].

After identification of the important aspects next step is sentiment classification which is used to determine the orientation of sentiment on each aspect. Aspect sentiment classification can be done by using two approaches unsupervised approach and supervised learning approach. Lexicon based approach is typically unsupervised. Lexicon consists of list of sentiment words, which may be positive or negative. This method usually employs a bootstrap strategy to generate high quality lexicon. Hu and Liu have used this lexicon based method. They obtained the sentimental lexicon by using synonym/antonym relation describe in WordNet to bootstrap the seed word set [12].

Hu's method is improved by Ding et al by addressing two issues: opinion of sentiment word would be content sensitive and conflict in review. They derived the lexicon by using some constraints [13].

Second approach is supervised learning approach which classifies opinions on aspects by using sentiment classifier. Sentiment classifier is learned from training corpus which is used to classify the new aspects opinions. Many learning models are applicable for this purpose. Bopong and Lee used three machine learning techniques SVM, Naive Bayes and Maximum Entropy for determining whether the review is positive or negative [10].

There is no previous work study in the product aspect ranking. The product aspect ranking is to predict the ratings on individual aspects. Wang *et al* developed a latent aspect rating analysis model, which aims to determine reviewer's latent opinions on each aspect and the relative emphasis on different aspects. This work concentrates on aspect-level opinion estimation and reviewer rating performance analysis, not on aspect ranking. Snyder and Barzilay expressed a multiple aspect ranking problem. However, the ranking is actually to predict the ratings on individual aspects [1].

III. CONCLUSION

This survey paper presents an overview on the product aspect ranking system. Product aspect ranking system contains three main steps i.e. product aspect identification, sentiment classification and aspect ranking. We surveyed that aspect identification and sentiment classification have some supervised and unsupervised techniques. Also the important aspects of a product would be the aspects that are frequently commented by customers and customer's opinions on these important aspects are greatly effective on the overall opinions of the product.

REFERENCES

[1] Zheng-Jun Zha, Jianxing Yu, Jinhui Tang, Meng Wang, Tat-Seng Chua, "Product Aspect Ranking and Its Applications" IEEE Transactions On Knowledge And Data Engineering, vol. 26, no. 5, pp. 1211-1224, May 2014

[2] Shahuraj Patil, Jyoti Raghatwan , "An Efficient Product Aspect Ranking and its Application: A Review", International Journal of Science and Research, vol. 3 , no.12, pp. 2048-2051, December 2014

[3] Mr. S. P. Ghode, Prof. S. S. Bere, "Exploiting and Ranking Dominating Product Features through Communal Sentiments", International Journal on Recent and Innovation Trends in Computing and Communication, vol. 3, no. 6, pp. 3894 – 3900, June 2015

[4] Mr. A. S. Kamale, Mr. S. P. Ghode, Prof. P. B. Dhainje, Mr. A. V. Moholkar, "A Survey on Feature-Sentiment Classification Techniques", International Journal on Recent and Innovation Trends in Computing and Communication , vol. 2, no.12, pp. 3972 – 3978, December 2014

[5] Charushila Patil, Prof. P. M. Chawan, Priyamvada Chauhan, Sonali Wankhede, "A Survey on Product Aspect Ranking", International Journal of Innovative Research in Science, Engineering and Technology, vol. 4, no. 12, pp. 12196-12204, December 2015

[6] Aarati Mahadik, Asha Bharambe, "Aspect Based Opinion Mining and Ranking: Survey", International Journal of Current Engineering and Technology, vol. 5, no. 6, pp. 3589-3592, December 2015

[7] R. Suganya, " Identifying and Ranking Product Aspects based on Consumer reviews," International Journal for Research in Applied Science & Engineering Technology, vol. 3, no. 1, pp. 290-293, January 2015

[8] A. M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in Proc. HLT/EMNLP, Vancouver, BC, Canada, pp. 339–346, 2005

[9] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proc. SIGKDD, Seattle, WA, USA, pp. 168–177, 2004

[10] B. Liu, "Sentiment analysis and subjectivity," in Handbook of Natural Language Processing, New York, NY, USA: Marcel Dekker, Inc., 2009

[11] B. Liu, "Sentiment Analysis and Opinion Mining". Mogarn & Claypool Publishers, San Rafael, CA, USA, 2012.

[12] B. Ohana and B. Tierney, "Sentiment classification of reviews using SentiWordNet," in Proc. IT&T Conf., Dublin, Ireland, 2009

[13] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in Proc. WSDM, New York, NY, USA, pp. 231–240, 2008