

ASPECT BASED PRODUCT EVALUATION USING SENTIMENT ANALYSIS

G.Kavitha¹, Dr.M.Pushparani², Dr.K.Kavitha³

¹M.Phil Scholar, ²Professor & Head, ³Assistant Professor, Computer Science,

Mother Teresa Women's University, (India)

ABSTRACT

With the rapid growth of online users most of websites, forums and blogs allow customers to post their view about a variety of products or services they consumed. This customer reviews can help the organization to improve their product quality. And also help other potential customers to understand the opinion of customer and help them to make better decisions. These reviews are large in number and difficult to analyze. Manual analysis of such large number of reviews is practically impossible. Sentiment analysis is a discipline which deals with analysis of such reviews. Aspect based or Aspect level opinion mining is one of the opinion mining tasks. In this paper, a feature-opinion (Aspect) based method is proposed to perform sentiment classification on a review using the product features mentioned in the reviews and the relevant opinions to the corresponding features.

Keywords: Feature Based Opinion Mining, Information Retrieval, Opinion Mining, Product Evaluation, Sentiment Analysis

I. INTRODUCTION

A vast number of websites, blogs and forums allow customers to post their opinions about a mixture of product or services. These customer opinions are important source of other customers and product manufacturers which help them to make better decisions. As the customer reviews expands, it is difficult for users to obtain a comprehensive view of opinions of previous customers about various aspects of product through manual analysis. But proper analysis and summarization about a particular product review is essential. This has inspired research in opinion mining and sentiment analysis to develop methods for automatically detecting opinions, emotions and other evaluations from texts. One of the most relevant applications of opinion mining is aspect based summarization [1]. Broadly speaking, given a collection of opinion posts, this task is aimed at obtaining relevant aspects(such as product features),along with associated sentiment information expressed by customers(usually an opinion word and/or a polarity score)[2].

Using opinion mining a review can be evaluated at three different levels- at document level, sentence level and feature level. When review is evaluated at document level, whole review is classified into either positive or negative depending upon the opinion expressed in that review. When review is evaluated at sentence level, then each sentence in a review is classified into either positive or negative. Whereas feature level or feature based opinion mining gives summary which feature of product is liked or disliked by reviewer[3].

Aspects are topics on which opinions are expressed. In the field of sentiment analysis, other names for aspect are: features, product features or opinion targets. Aspects are important because without knowing them, the opinions expressed in a sentence or reviews are limited use [5]. For example, in the review sentence "after using iPhone6, I found the battery to be perfect for two days", "battery" is the aspect for which an opinion is expressed. Product feature detection is critical to sentiment analysis.

II. RELATED WORK

Previous work on extracting product features from online customer reviews has mainly relied on natural language processing (NLP) [1]. Part-of-speech (POS) tagging, shallow parsing techniques, and dependency grammar have been widely applied to identify both noun phrases that act as potential features and opinion words that affect them through syntactical dependencies. Using the double-propagation strategy [5] allows the incremental identification of features and opinion words from a predefined initial set (usually a lexicon of opinion words). Generally, NLP-based approaches present good precision but low recall figures because they depend on the definition of extraction patterns, which are dependent on both the particular language and the reviews application domain.

Opinion mining is a relatively recent discipline that studies the extraction of opinions using Artificial Intelligence and/or Natural Language Processing techniques. More informally, it's about extracting the opinions or sentiments when given a piece of text. This provides a great source of unstructured information especially opinions that may be useful to others, like companies and their competitors and other consumers. For example, someone who wants to buy a camera, can look for the comments and reviews from someone who just bought a camera and commented on it or written about their experience or about camera manufacturer. He can get feedback from customer and can make the decision. Also a manufacturing company can improve their products or adjust the marketing strategies. Opinion Mining needs to take into account how much influence any single opinion is worth. This could depend on a variety of factors, such as how much trust we have in a person's opinion, and even what sort of person they are. It may differ from person to person like an expert person and any non-expert person. There may be spammers. Also we need to take into account frequent vs. infrequent posters [6].

The paper focuses on the problems of double propagation. Double propagation assumes that features are nouns/noun phrases and opinion words are adjectives. Opinion words are usually associated with features in some ways. Thus, opinion words can be recognized by identified features, and features can be identified by known opinion words. The extracted opinion words and features are utilized to identify new opinion words and new features, which are used again to extract more opinion words and features. This propagation or bootstrapping process ends when no more opinion words or features can be found. The advantage of the method is that it requires no additional resources except an initial opinion lexical analyzer [7]. The paper presents a method for identifying an opinion with its holder and topic. Opinion holders are like people, organizations and countries, i.e. the entity who has given opinion for some product. An opinion topic is an object an opinion is about. In product reviews, for example, opinion topics are often the product itself or its specific features, such as design and quality as "I like the design of iPod video", "The sound quality is amazing". Opinion topics can be social issues, government's acts, new events, or someone's opinions. The overall steps include identifying

opinions, labeling semantic roles related to the opinions, find Holders and topics of opinions among the identified semantic roles and storing <opinion, holder, topic> triples into a Database [8].

III. SENTIMENT ANALYSIS

When the buyer purchases any product, then he/she purchase it for some purpose and has some condition in his/her mind about the product. For example if buyer wishes to purchase any Laptop, then he/she may have specification in his/her mind that the laptop should be low price but with good battery life, it should be medium size and handy, may not concentrate on the color and quality. So, when the customer purchases any product, then he/she is having his/her own preferences about features of the product. In Fig.1, reviews of opinions given by people, the people use different words/adjectives for giving opinions about the features of any product. If in reviews, one person says “Best” about any particular feature of one product and another person says “too good” about the same feature of that product, and then there exists difference of opinions. Suppose for that product and for same feature, the customer preference is “too good” Then the customer will try to find the number of reviews in which the same feature of that product is given “too good”. Finding this statistical data is very difficult for the customer.

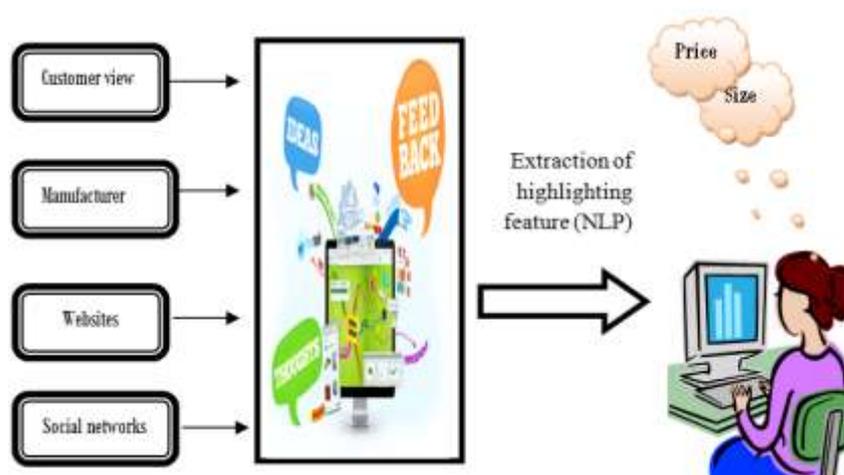


Fig.1 Product Evaluation

As given in fig.1, Collect the ideas, suggestions and reviews from various sources and extract the needed reviews from all the sources with the help of natural language processing tools.

3.1 Feature Based Opinion Mining

As stated Earlier, opinion mining can be done at document level, sentence level and feature level. This paper focus on feature level/ feature based opinion mining. The major tasks of feature based opinion mining are - (1) to discover the products features in review, (2) to decide opinion expressed by the reviewer (positive, negative or neutral),(3) summarize discovered information. Some researchers have proposed ontology based feature based opinion mining, domain-specific opinion extraction and some other automated techniques for feature based opinion mining. The following Fig.2 Explained the steps for feature based opinion mining and summarization.

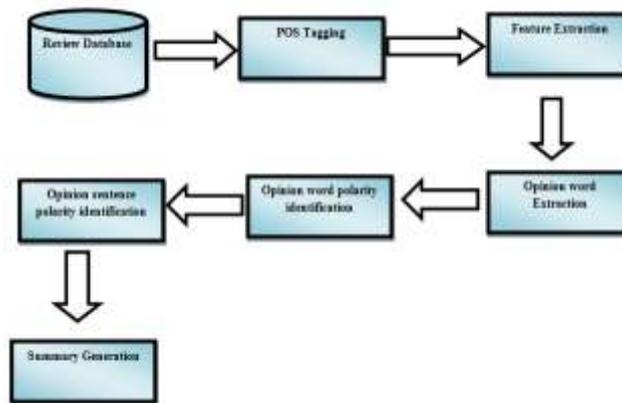


Fig 2: Steps for Feature Based Opinion Mining & Summarization

3.1.1 Review Database

Reviews are retrieved from different sources and then we can store those retrieved reviews database. Each Sources i.e., Websites, blogs, social networks has its own structure. We can extract the reviews from the Web crawlers which is also known as spiders or robots, are programs that automatically download Web pages. A crawler can accumulate customer reviews by accessing various websites, that can be done either online (as it is downloaded) or off-line (after it is stored). After completion of this task pre-processing is done where unnecessary text is isolated and then reviews are stored into database.

3.1.2 Part-of-Speech Tagging (POS Tagging)

The main objective of sentiment analysis is to identify the product features and opinion words (opinion words means words which state opinion). And then find polarity of each opinion word. Normally, opinion words are adjectives and product features are nouns. Consider following example

“The size of my mobile is excellent”

In above sentence, size (product feature) are noun and Excellent (opinion word) is adjective. with the help of POS Tagger, each word in review is assigned with separate tag. The tag will be separately assigned for noun, adjective, adverb, verb, etc., By the end of POS tagging, nouns are retrieve as product features and adjectives as opinion words.

3.1.3. Feature Extraction

In this process, extraction of product feature is done in each sentence. Commonly Product features mentioned in the customer reviews can be implicit or explicit. Features which are mentioned in a sentence directly are called as explicit features and features which are not mentioned directly are called implicit features. For example,

“Battery Life of a Laptop is less”

In this sentence reviewer has mentioned battery life directly so it is explicit feature. It is easy to extract such features. Now consider following sentence,

“This Laptop needs to charge many times in a day”

In this sentence reviewer is talking about battery of Laptop but it is not mentioned directly in the sentence. So here battery is implicit feature. It is complicated to understand and extract such features from sentence.

3.1.4. Opinion Word Extraction

Opinion words are identified in opinion word extraction process. If a sentence contains one or more product features and one or more opinion words, then the sentence is called an opinion sentence. As mentioned earlier opinion words are usually adjectives.

3.1.5 Opinion Word Polarity Identification

In opinion word polarity identification, semantic orientation of each opinion word is recognized. Semantic orientation used to identify whether opinion word is expressing positive opinion, negative opinion or neutral opinion.

3.1.6 Opinion Sentence Polarity Identification

Opinion sentence polarity identification indicates the orientation of an opinion sentence. Consider following sentence,

“This is not good Laptop”

Above sentence contains opinion word ‘good’ which expresses positive opinion. But sentence expresses negative opinion as the word ‘not’ is presented. Hence, there is a need to find the polarity of the opinion sentence along with the finding the polarity of opinion word. In order to proceed to the opinion sentence polarity identification, a list of negation words such as ‘no’, ‘not’, ‘but’ etc. can be prepared and negation rules can be formed.

3.1.7 Summary Generation

Finally Summary generation is generated after opinion sentence orientation identification process. This summary is aspect based product. With the help of information discovered in previous steps summary can be generated. Summary can be generated in the form of graphs or tables. A table or graph will give summary of all the reviews related to a product.

V.EXPERIMENTAL RESULT

Table 1: Extraction of Customer Reviews

Product	Total number of customer reviews	Positive reviews	Negative reviews
Product A	2000	1419	581
Product B	286	283	3
Product C	3642	2980	662

Here, the dataset used in the experiment is downloaded from the Amazon product home page; three kinds of products tested in the system are “Product A”, “Product B”, and “Product C”, as shown in TABLE 1. In this experiment, we test five individual features for these three products.

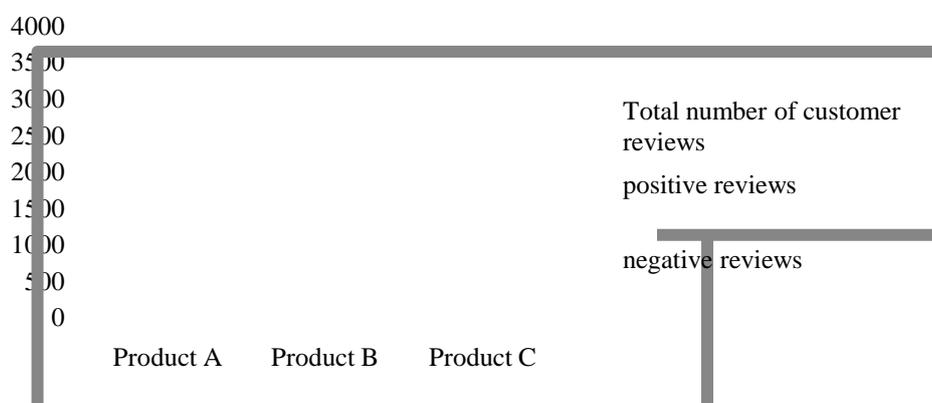
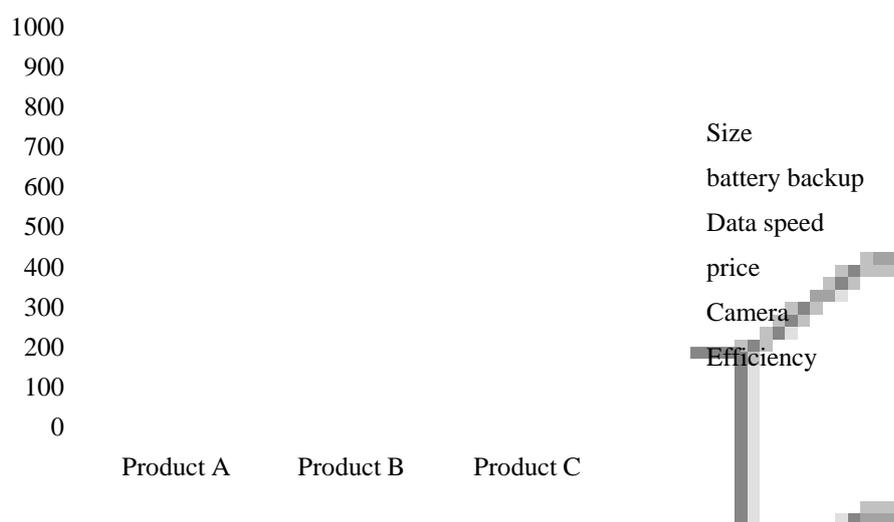


Fig.3 Extraction of Customer Reviews

Table 2: Aspect Based Product Extraction

Products	Size	Battery backup	Data Speed	Price	Camera	Efficiency
Product A	132	366	212	282	199	228
Product B	22	139	3	60	7	52
Product C	364	789	273	927	391	236

**Fig.4 Product Feature Extraction**

Here, POS tagged sentences are processed to extract features and the values given to those features. These features and values are given rates and stored in opinion database. In Fig.4, battery backup got maximum votes for product A and product B. Product C got maximum positive reviews for price. From that we can easily find the product feature. After getting these values we can get the percentage of positivity and negativity given to product in a particular corpus.

IV. CONCLUSION

Opinion mining is a new field of study. Feature based opinion mining and summarization is Challenging field for researchers. It is useful for individuals as well as for organization. This is important because, in this economical world, every customer try to compare multiple products before purchasing. Also the business party needs customer opinion about their products to be in the competition and to place improvements in their products. This is a recent trend in research also. While the development of the opinion mining tools described shows very much work in progress and initial results are hopeful, but still it requires a lot of refinements. Since this is a study of sentiments of a person, so it requires a lot of precision. Whenever any person talk about something, then the context in which he is talking and how the sentence is formed may change the parsing method to catch the exact opinion said by that person. If try to concentrate on one pattern of sentence then there may lose any other pattern of sentence from our parsing method. This is a major challenge in front of the opinion mining methods.

REFERENCES

- [1] G. Carenini, R. Ng, and A. Pauls, "Multi-Document Summarization of Evaluative Text," Proc. European Chapter Assoc. for Computational Linguistics, ACL, 2006, pp. 305–312.
- [2] Lisette García-Moya, Henry Anaya-Sánchez, and Rafael Berlanga-Llavori, "Retrieving Product Features and Opinions from Customer Reviews" Universitat Jaume, Published by the IEEE Computer Society, 1541-1672/13/\$31.00 © 2013 IEEE
- [3] Padmapani P. Tribhuvan , S.G. Bhirud , Amrapali P. Tribhuvan , "A Peer Review of Feature Based Opinion Mining and Summarization", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (1) , 2014, 247-250
- [4] Ayoub bagheri, Mohamed Saraee, Franciska de jong, "care more about customers : unsupervised domain-independent aspect detection for sentiment analysis of customer reviews" 52(2013)201-2013
- [5] J. Yu et al., "Aspect Ranking: Identifying Important Product Aspects from Online Consumer Reviews," Proc. 49th Ann. Meeting of the Assoc. for Computational Linguistics: Human Language Technologies—Volume 1, ACL, 2011, pp. 1496–1505.
- [6] Vrushali Yogesh Karkare, Sunil R. Gupta, "Product Evaluation using Mining and Rating Opinions of Product Features"
- [7] Lei Z., Bing L., Suk H. L., Eamonn O'Brien-Strain. Extracting and Ranking Product Feature in Opinion Documents. In the Proceedings of COLING '10 The 23rd International Conference on Computational Linguistics. Pages 1462-1470.
- [8] Soo-Min K., Eduard H. Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text. In Proceedings of the COLING/ACL, an ACL Workshop on Sentiment and Subjectivity in Text.