International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 10 (2015) pp. 25629-25636
© Research India Publications
http://www.ripublication.com

Opinion Mining And Market Analysis

Mr. Vinayak Hegde[1], Ms. Karthika P [2], Ms. Madhu M G [3]

Amrita Vishwa Vidyapeetham, Mysore Campus,
#114 Bogadi 2nd Stage Mysore. 570026, Karnataka
India, vinayakhegde92@gmail.com, Mob No: 9886484778

Abstract

It is obvious that customers are important stakeholders in an organization, and their satisfaction and review is priority to management. Online customer review constitute new channel of information acquisition. Many firms after consumers the possibility of positive and negative polarities on their goods. Online customer reviews become a strategic investment in business strategies. So online customer product reviews has been considered as an effective way, to analyze the product reach to customers. The rapid growth of web in last decades helps this out. Customer reviews are increasingly available online, ranking this as a bib asset to know their market level is important to organization. The impact of customer satisfaction the positive and negative of a product can be analyzed by organization through online reviews posted on websites, blogs, and social networking sites. Our project mainly concentrates on gathering reviews from the sites based on certain ontology extracting the words and classifying them into percentage of positive and negative of a product. This helps organization to automate their market analysis quickly by reducing manual efforts on evaluating the reviews according to features of product. The outcome of system is set of reviews organizing by their degree of positive and negative on a product.

Keywords: Opinion mining, Sentiment analysis, Naive Bayes algorithm.

I. INTRODUCTION

In the present age, it has become common practice for people to communicate or express their opinions and feedbacks on social media. An increase in online activities like blogging, social networking, review posting has resulted in incremental lot of user-generated content. Most of these online reviews are in the form of text. This in turn has led to increased research in automatic classification, summarization, and opinion mining from web-based data. Opinion mining applications are the basic infrastructure for decision making. They help making sense of thousands of interventions. Traditionally, surveys are used to collect feedback in a structured manner. However, this kind of data collection is excessive, as it deserves an investment in design and data collection; it is difficult, as people are not interested in answering surveys; but fails to detect the most important problems. Opinion mining is helpful to identify problems and to make better decision to improve their market value. Not only organization but also potential buyers want to know what the value for the product they buy is.

For any organization customers are the biggest asset, when organization release new product to market and want to know their market value. Now a days customer express view through internet forms, blogs, social network sites and carry out search on web. Retrieving market value simply lies on surveying but it is not sufficient in today's web. We require smart search for analyzing and mining people opinion on product attributes which indicates positive and negative sentiments. As major companies are increasingly coming to realize these consumer reviews can produce enormous influence in shaping their market range. Brand loyalty, their own purchases, improving customer satisfaction. The field of customer opinion mining and sentiment analysis is well suited to various types of intelligence application, indeed business since to be one of the main factor in corporate field. Not only corporate field but also helps in mining review about the movies, politicians' position and other areas. Which are not visible to human eyes. Output is the last stage in which result can be altered image or report that is based on image analysis.

Select the appropriate name on the style menu. The style will adjust your fonts and line spacing.

Opinion can be expressed as

- **Direct Opinion:** "This is **good** Phone but Battery life is **poor** and performance is **better**".
- **Comparative Opinion:** "Nokia Lumia is **better than** Nokia in its performance"

II. LITERATURE SURVEY

Numerous techniques are reported in the literature to perform the opinion mining. The overview of existing experimentations reported in the literature is discussed. As mining the opinion is now a day's taking a explosive growth numerous techniques are been discussed.

Mining the opinion not only on the product but also for political situation, public issues are also being taken into account. Finding the sentiment based on user or consumers view with respect to both potential customers and organization benefit is the main thing focused upon. Using the review posted on social media, forums and Internet blogs many research works is happening to find the positive, negative and neutral result .With the help of data mining algorithm the research has been taken. Many challenges and the application related to the field and what all will be the limitation are also been mentioned. Many methods like finding opinion based on stored dictionary words, using SentiWordNet are discussed. Feature based sentiment analysis for a product with respect its features is also discussed in research work carried by [3] Amruta Sankhe and Prachi Gharpu.

III. PROPOSED METHODOLOGY

The proposed methodology helps in deciding the polarity of data in which opinions / reviews / arguments can be classified as positive or negative which is facilitated by collection of positive or negative examples already fed.Naïve Bayes algorithm is implemented to estimate the probability of a data to be negative or positive. The probability is calculated by studying positive and negative example.
This theorem provides conditional probability of occurrence of event E1 when E2 has already occurred, the vice versa can also be calculated by following mathematical statement.

$$P(E1 / E2) = \frac{P(E2 / E1)\, P(E1)}{P(E2)}$$

For proposition E1 and evidence E2,
P (E1), the prior, is the initial degree of belief in E1.
P (E1 | E2), the posterior, is the degree of belief having accounted for E2.
P (E2 | E1) / P (E2) represents the support E2 provides for E1.

In this algorithm, the probabilities of the labels, according to the words are found. It means that how many words from the opinion belong to which label [positive or Negative] is found.

a) If the probability of positive label is greater than the negative, then the opinion is positive.

Example: “This is **good Phone** but Battery life is **poor** and **performance is better**”.

b) If the probability of negative is greater than the positive, then the opinion is negative.

Example: “This phone is **not good** and battery life is also **bad** but **price is low**”.

Algorithm 1: Algorithm to classify positive and negative polarities

Step 1: Input the Document.
Step 2: Initialize positive_count =0 and negative_count = 0
Step 3: search for the positive and negative polarities in the document.
Step 4: Count the number of positive and negative words
If(positive_count == 7 && negative_count == 3)
than

shows the positive opinion is more than negative
else
shows the negative opinion more than positive
End
Step 5: Identify theprobabilities of the labels [positive and negative]
If (probability_positive> probability_ negative)
than
shows opinion is positive
else if (probability_ negative > probability_positive)
than
shows opinion is negative
End
End
Step 6: Stop

IV.IMPLEMENTATION STEPS

The three main modules that we are dealing with are-

- Load the document
- Classification
- Presentation of Result

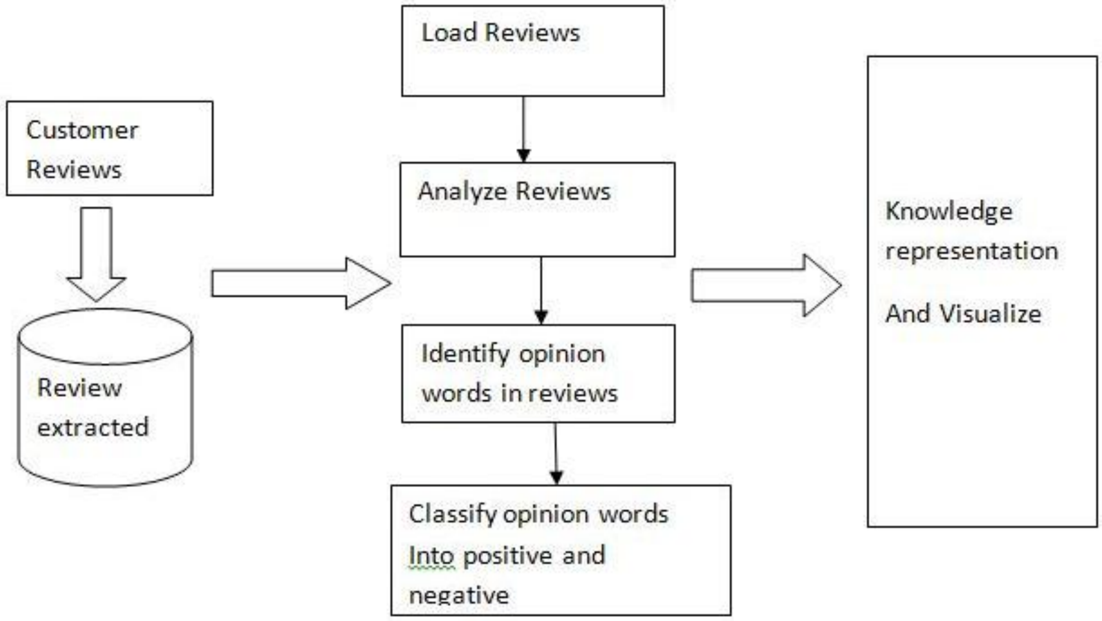

Figure1: Architecture Diagram

<u>Load the Document</u>

This is the basic and the first step for any text classification. Here we are collecting the reviews from different sites and uploading it as document. Once the document has been uploaded, the next is the extraction of words to find out positive and negative words in a text.

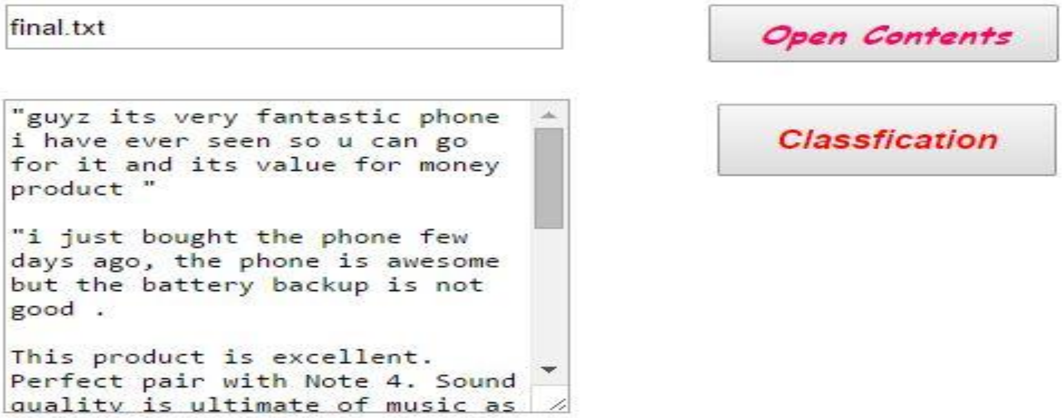

Figure 2: Document loading

<u>Classification:</u>

In this phase, content polarity is identified. Once the document has been submitted for identification, the classifier breaks down the sentence and accordingly gives it its sentiment. Usually two classes are used for classification: positive, negative. After the first step word extraction those extracted words are compared with the set of stored positive and negative words. The text is manually classified with corresponding sentiment as either positive or negative.

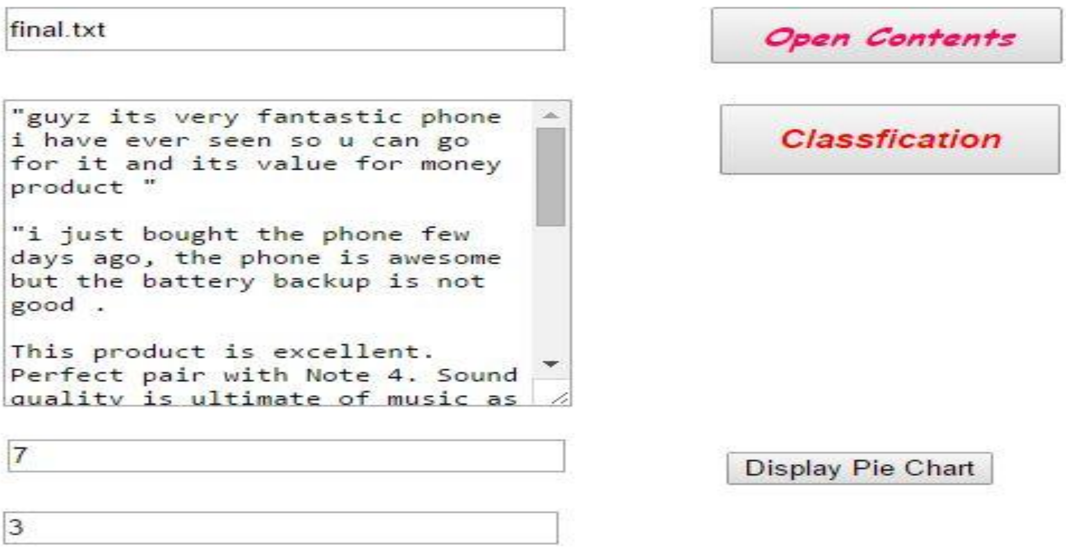

Figure 3: Classificationof positive and negative polarity

Presentation of Result:

At this stage, classified result will be represented using graphs. It shows the percentage of positive and negative words in the document. It helps us to make decision on whether sentence implies a positive or a negative impact on a product to make better decision.

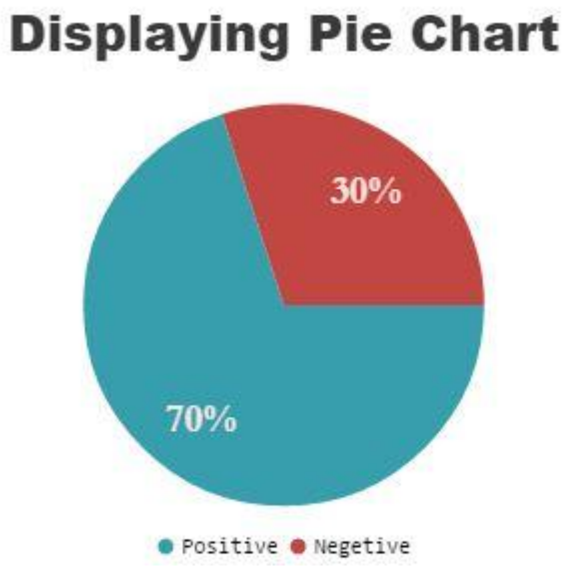

Figure 4: showing the percentage of positive and negative polarity

V.EXPERIMENTAL ANALYSIS

In this paper we have collected reviews from different websites (flip kart, amazon etc...) for Micromax phone. It has helped us in market analysis, which is used to determine whether the opinion towards the product is positive or negative by considering the percentage of positive and negative polarity in the result. As per the result shown in this paper, the product micromax has more positive percentage than negative. So positive opinions are more when compared to negative regarding this product. We can apply this method to any product so as to make correct business decisions.

APPLICATIONS

Opinions are so important that whenever one needs to make a decision, one wants to know what others opinion is. This will be helpful for both individuals and organizations. So the process of opinion mining has a wide scope for practical applications.

- **Individual consumers:** If an individual wants to purchase a product, it is useful to see a summary of opinions of other users so that he/she can make a better decision. He/she will get to know what is the positive and negatives of a product before investing money on that product.

- **Organizations and businesses:** Opinion mining is equally, if not even more, important to businesses and Organizations. For example, it is critical for a product manufacturer to know how consumers perceive its products and what the reach in the market to their product is. Those of its competitors. This information is not only useful for marketing and product benchmarking but also useful for product design and product developments.

CONCLUSION

Opinion mining has become a fascinating research area due to the availability of a huge volume of user-generated content in review sites, forums and blogs. Opinion mining has applications in a variety of fields ranging from market research to decision making to advertising. With the help of opinion mining, companies can estimate the extent of product acceptance and can devise strategies to improve their product. Individuals can also use opinion mining tools to make decisions on their buying by comparing competitive products not just based on specifications but also based on user experience and public opinions. Our focus is to mine the reviews and classifying it into positive and negative polarities.

MORE CHALLENGES IN OPINION MINING

We see a few but important challenges for text analytics tasks like opinion mining. It is very difficult to distinguish between objective and subjective information, generally opinion words also occur in objective sentences, so it is very tough to handle these challenges. Many times we see customers posting the reviews in the blogs or forums with a lot of spelling mistakes which our dictionary cannot catch them and resulting in less accuracy of desired output. Most times of the times we see many spam blogs and spam reviews posted by the users. If we consider these reviews for performing Opinion Mining, we may get deviated from our desired results. So, lot of work has to be done in this field for identifying spam blogs, considering spelling mistakes and for other challenges.

REFERENCES

[1] Mining of Sentence Level Opinion Using Supervised Term Weighted Approach of Naïve Bayesian Algorithm, Trivedi Khushboo N, Swati K. Vekariya, Prof.Shailendra Mis ra, Int.J.Computer Technology & Applications.

[2] A Peer Review of Feature Based Opinion Mining and Summarization, PadmapaniP.Tribhuvan , S.G.Bhirud , Amrapali P. Tribhuvan.

[3] “Amruta Sankhe and Prachi Gharpu [Computer Department, Mumbai University]”Feature Based SentimentAnalysis for OnlineReviews in Car Domain.

[4] Arti Buche, Dr. M. B. Chandak, Akshay Zadgaonkar, ”OPINION MINING AND ANALYSIS: A SURVEY”, International Journal on Natural Language Computing (IJNLC) Vol. 2, No.3, June 2013.
[5] Beltrán Borja Fiz Pontiveros , “Opinion Mining from a Large Corpora Of Natural Language Reviews”, Master in Artificial Intelligence (UPC-URV-UB).
[6] Songbo Tan, “Adapting Naive Bayes to Domain Adaptation for Sentiment Analysis”, Key Laboratory of Network, Institute of Computing Technology, China.
[7] Prof. Shailendra Mishra, “Mining of Sentence Level Opinion Using Supervised Term Weighted Approach of Naïve Bayesian Algorithm”, Assistant Professor, Department of Computer Engineering, Parul Institute Of Technology, Vadodara, Gujarat.
[8] Padmapani P. Tribhuvan, “A Peer Review of Feature Based Opinion Mining and Summarization”, Deogiri Institute of Engineering and Management Studies, Auarangabad, India.
[10] Rohini K. Srihari, ”OpinionMiner: A Novel Machine Learning System for Web Opinion Mining and Extraction”, Department of Computer Science & Engineering University of New York at Buffalo.
[11] Amruta Sankhe, Prachi Gharpure, ” Feature Based Sentiment Analysis for Online Reviews in Car Domain”, Computer Department, Mumbai University, SPIT, India, International Journal of Current Engineering and Technology, Vol.4, No.2 (Feb 2014).
[12] Thomas Y. Lee, ” Needs-Based Analysis of Online Customer Reviews”, University of Pennsylvania, The Wharton School.
[13] “Fast and accurate sentiment classification using an enhanced Naive Bayes model.”, Vivek Narayanan, Department of Electronics Engineering, Indian Institute of Technology (BHU), Varanasi, India.