A CASE STUDY ON CUSTOMER REVIEW BASED SENTIMENT ANALYSIS

NeetuNarwal *

Abstract – Today large number of companies are shifting their businesses online due to growing trend among customer to shop online. There arises a need for effective visual analysis of online customer opinions. It has a significant impact on building successful business strategies and at the same time helping the customers in decision making process. In this paper we studied the aspect of using these summarized feedbacks to provide customer sentiment regarding a product. For the purpose of study we have focused on customer reviews related to Google Pixel 2 newly launched mobile phone released in Oct 2017. These sentiment analyses are of great value to the customer who is planning to purchase the product from the market. We have gathered two months reviews of customers who have purchased the mobile and have provided feedback related to the product. The result shows unfavorable response towards Google Pixel 2 mobile phone.

Keywords—Electronics Commerce, Sentiment Analysis, Social Media .

I. INTRODUCTION

According to definition Opinion mining is a type of natural language processing for tracking the mind of the public about a particular product [1]. Opinion mining, which is also called sentiment analysis, involves developing a system to gather opinion and categorize them to provide analysis about a product or topic. Automated opinion mining often uses machine learning, a type of artificial intelligence (AI), to mine text for sentiment.

Opinion mining is of great value to both the customer and the organization. It can help the business organization to evaluate the success of new product, its popularity and shortcoming if

* **Asst.Prof., Maharaja Surajmal Institute, Affiliate College of GGSIP University, New Delhi**

any exist. It also provides useful information about the customer demographic which can help organization formulate business plans to target the area.

There are certain challenges in the area of opinion mining. The first challenge is that same word when used in different situation has different meaning. A second challenge is that people are sometime sarcastic in their remarks. It is very difficult for the tool to understand the sarcasm in the opinion. A third challenge is that people often use both positive and negative comments at the same time. This becomes difficult to judge the remark as positive or negative.

Since, the social media comments are more informal hence people often combine different opinion in their single post. For example: "the phone bombed even though the battery backup is amazing". This remarks is easy for a human to understand, but more difficult for a computer to parse. Sometimes even other people have difficulty understanding what someone thought based on a short piece of text because it lacks context.

According to Economic Time article [2] the number of consumers who purchase online is expected to cross 100 million by 2017 end with e-retail market likely jumping 65% on year in 2018, an ASSOCHAM-Resurgent India study. Mobile commerce is likely to jump to 45-50% in 2017 against the current 30-35% of e-commerce sales. The study added that 50% sales come from mobile with majority being first time users. It also observes a surge in the number of people shopping on mobile across India with tier II and III cities displaying increased dominance.

According to Statistica.com as on Jan 2017, India's had an estimate of 462.1 million Internet users [3]. About 442.7 million people in India accessed the internet through their mobile phones in 2017, which corresponds to about 33% percent of the country's population [4]. Both figures are forecast to increase in the coming years, with projections to amount to 524.5 million and around 37.4 percent respectively in 2021. Mobile internet usage in India varies according to people's living areas.

In this paper, we explored the possibility to use sentiment analysis of the customer reviews to better understand and rate the product. In this study we have taken the review comments for newly launched mobile phone Google Pixel 2 from https://www.gsmarena.com/google_pixel_2-review-1680.php website [3].

 We have collected customer reviews for the period of three months from 6 Oct 2017 till 15 Nov 2017. Total 1850 reviews are collected for study. As the number of review increases it become difficult to summarize them manually. These reviews are gathered, cleaned and processed for

finding sentiment analysis of the customer reviews. We designed few R programming module to perform data cleaning, processing and sentiment analysis.

This paper is organized as follows, Section II presents the current research work done in this area, Section III proposes the methodology adopted in this study, Section IV presents the results and discussion. Section V present the conclusion and future directions.

II. RELATED RESEARCH WORK

Sentiment Analysis of Social Media data has been widely used by researchers to study people behavior towards certain event. Song H. et. al. [5] in their paper used customer blog to extract opinion about a product. Yadav, M. P., et al [6] in their paper inferred customer behavior towards E-commerce companies using popular clustering technique using customer feedback. Today Consumers prefer to choose one among millions from online store instead of one among few in the superstore nearby. It shows that consumers prefer e-commerce site to engage in trade. Singh P.K. et al [7] in their paper gathered millions of reviews on social networking sites on the basis of their geographical location and provided the consumer (business/manufacturer) summarized data about the expressed sentiments in an intuitive and easy to understand graphs, charts and other visualization.

Yu C et. al. [8] discussed about the changing behavior of customers today they prefer to purchase online as compared to window shopping as online shopping provides user with millions of options for the same product. As, companies are able to attract most of the customers because ecommerce is not just buying and selling over internet but it also act as to get advantage on big giants of market. For this purpose data mining also know as knowledge discovery can be used by these E-Commerce companies for increasing their sales and customer satisfaction.

Hu and Liu [9] summarized customer reviews and prepared a list of positive words and a list of negative words, respectively. The positive list contains 2006 words and the negative list has 4783 words. Both lists also include some misspelled words that are frequently present in social media content. Their bag of words are used for Sentiment categorization. It is essentially a classification problem, where features that contain opinions or sentiment information should be identified before the classification.

Pang and Lee [10] in their work suggested to remove objective sentences by extracting subjective ones for feature selection. They proposed a text-categorization technique that is able to identify subjective content using minimum cut.

Gann et al. [11] selected 6,799 tokens based on Twitter data, where each token is assigned a sentiment score, namely TSI(Total Sentiment Index), featuring itself as a positive token or a negative token. Specifically, a TSI for a certain token is computed as:

$$TSI = \frac{p - \frac{tp}{tn}*n}{p + \frac{tp}{tn}*n} \qquad (1)$$

Where, p is the number of times a token appears in positive tweets and n is the number of times a token appears in negative tweets. $\frac{tp}{tn}$ is the ratio of total number of positive tweets over total number of negative tweets.

In our previous work [12] we have used customer feedback to provide sentiment analysis about the product which can help new consumer to gain insight about the product before making the purchase.

III. METHODOLOGY USED

In this research the following methodology has been adopted and it is shown in the Figure1. It comprises of four phases.
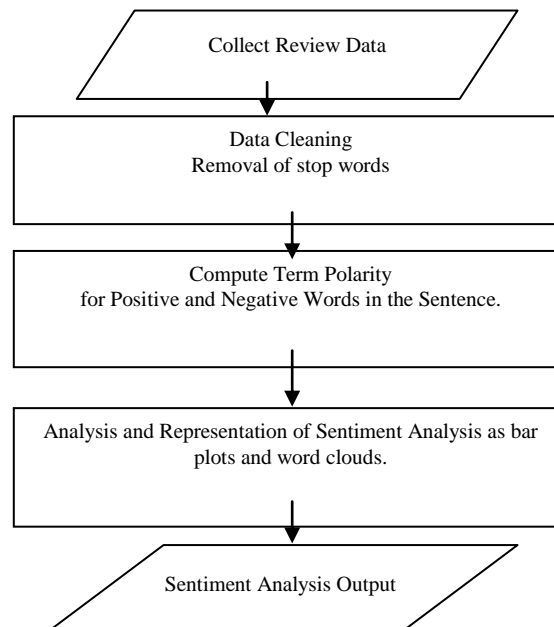

Figure 1: Methodology used in the research

Step 1: In our study the customer reviews for Google Pixel2 from https://www.gsmarena.com/google_pixel_2-review-1680.php website [3]. We have collected customer reviews for the period of two months from 6 Oct 2017 to 15 Nov 2017.

The primary data comprises of total 1850 customer reviews which is used further to perform sentiment analysis.

Step 2: The dataset of 1850 reviews comprises of 21120 words. In the second step dataset is cleaned before further processing. Using library functions stop words like a, the, is, are etc. and non-meaningful words are removed before these words are processed for sentiment analysis. We have designed a module in R language that utilizes different library functions available in tm package to clean the dataset. The output is 18640 words obtained after performing cleaning process. These words are used further to obtain sentiment of the Google Pixel2 mobile phone.

Step 3: In this step 18640 words are processed for sentiment analysis. We have designed a module in R language using library package syuzhet and get_nrc_sentiment function. It is used to calculate the presence of eight different emotions and their corresponding valence in a word corpus.

Step 4: In the final presentation step the term frequency of each sentiment group are plotted as bar plots using R library function and word cloud of review comments is plotted. In this module we have used some of the R language packages viz. ggplot2, scales, reshape2, dplyr to facilitate understanding of emotions and plotting the same.

IV RESULTS AND DISCUSSIONS

The customer reviews are collected for two months period 6 Oct 2017 till 15 Nov 2017 for newly launched Google Pixel2 mobile phone. All the reviews after cleaning have been analyzed from following perspective:

- Word cloud generation
- Sentiment analysis

The Word Cloud provides a graphical representation of most talked about words in the customer review comments. Sentiment Analysis performs analysis on the basis of ten different domains from the customer reviews. The observations shared by people who have bought the product is valuable and their sentiment gives correct indication regarding future of the products. Keeping in view the customer reviews related to mobile phone, its feature and their experiences are analyzed for sentiment analysis.

Sentiment analysis of Customer reviews for Google Pixel2 mobile phone has been plotted in Figure 2. It shows 39% of the people have positive emotions, joy and trust towards the Google Pixel2 Mobile Phone. Only 61% have sentiments in terms of negative emotion, sadness, fear and surprise.
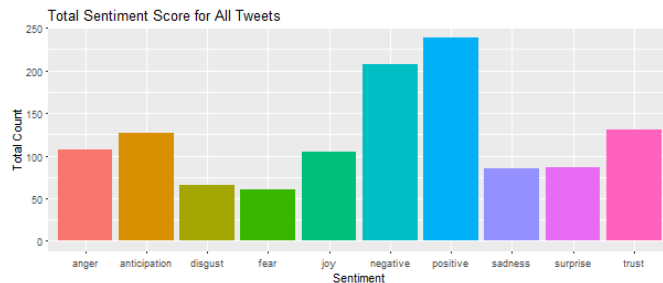

Figure 2: Sentiment Analysis of Customer Reviews for Google Pixel 2

On analyzing the word cloud for Customer Review for Google Pixel 2 as shown Figure 3 reflects that most talked terms amongst people are battery, high hope, card slot, price, doesn't even, screen etc.

Figure 3: Word Cloud of Customer Reviews for Google Pixel 2

V CONCLUSIONS AND FUTURE DIRECTIONS

It can be concluded that E-Commerce organization can use the customer reviews for their future planning and strategizing. This study focuses on the question that whether the data from

customer reviews can be utilized in better way to predict sentiment analysis of customer towards any product. In the present study we analyzed over 1850 reviews of customers for the newly launched Google Pixel 2 mobile phone. It has been concluded that social media is a powerful and reliable source of public opinion as far as trust and faith of people toward any product.

REFERENCES

[1]   http://searchbusinessanalytics.techtarget.com/definition/opinion-mining-sentiment-mining

[2]   https://economictimes.indiatimes.com/industry/services/retail/online-retail-consumers-to-cross-100-million-by-2017-assocham-resurgent-india-study/articleshow/56417797.cms

[3]   https://www.statista.com/statistics/309866/india-digital-population/

[4]   http://www.indiaonlinepages.com/population/india-current-population.html

[5]   Song, H., Chu, J., Hu, Y., & Liu, X. (2013, December). Semantic Analysis and Implicit Target Extraction of Comments from E-Commerce Websites. In *Software Engineering (WCSE), 2013Fourth World Congress on* (pp. 331-335). IEEE.

[6]   Yadav, M. P., Feeroz, M., & Yadav, V. K. (2012, July). Mining the customer behavior using web usage mining in e-commerce. In *Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on* (pp. 1-5). IEEE.

[7]   Kumar Singh, P., Sachdeva, A., Mahajan, D., Pande, N., & Sharma, A. (2014, September). An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites. In *Confluence The Next Generation Information Technology Summit (Confluence), 2014 5th International Conference-* (pp. 329-335). IEEE.

[8]   Yu, C., & Ying, X. (2009, December). Application of Data Mining Technology in E-Commerce. In *Computer Science-Technology and Applications, 2009. IFCSTA'09. International Forum on* (Vol. 1, pp. 291-293). IEEE.

[9]   Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD International conference on Knowledge discovery and data mining. ACM, New York, NY, USA. pp 168–177.

[10]   Pang B, Lee L (2004) A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42Nd Annual Meeting on Association for Computational Linguistics, ACL '04. Association for Computational Linguistics, Stroudsburg, PA, USA.

[11]   Gann W-JK, Day J, Zhou S (2014) Twitter analytics for insider trading fraud detection system. In: Proceedings of the second ASE international conference on Big Data. ASE, 2014.

[12]   Neetu Narwal, A Study of Product Sentiment Analysis based on Reviewers Comments, Jamshedpur Research Review, ISSN: 2320-2750, Year 2017 Vol. V Issue XXIV, pp 83-87.