

# SENTIMENT ANALYSIS ON TWITTER SOCIAL MEDIA PRODUCT REVIEWS

Ms.Vimali.J.S

Research Scholar, Sathyabama Institute of Science and Technology,  
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600 119, Tamilnadu, India

S. Murugan

Professor, Department of Computer Science and Engineering,  
Sathyabama Institute of Science and Technology, Jeppiaar Nagar,  
Rajiv Gandhi Salai, Chennai – 600 119, Tamilnadu, India

**Abstract** - Well timed and profitable analytics over “Social Networks” is now a lead for gain in many industries to increase brand exposure and broaden customer reach by analyzing the reviews and recommending products to consumers. The objective is to extract the unstructured conversation or posts from Twitter and present a methodology for sentiment analysis on Twitter that reviews the products by considering the consumer’s sentiments Naïve Bayesian Classifier. Samsung smartphones were taken into analysis and the experiment reveals six possible different sentiments: joy, surprise, anger, disgust, fear and sadness in the analysing process of the tweets. From the experimental results, it is observed that 80% of the customers has given positive feedback on the product and 20% has given negative feedback and the level of accuracy of the proposed approach is 90%.

**Keywords:** Social Networks; Twitter; Sentiment Analysis; Naïve Bayesian Classifier.

## 1. Introduction

Sentiment analysis is the progression of extracting the public’s attitude and emotions from English language. It is an enduring research domain of text-mining, which mainly concentrates on predictions. An enhanced consumer behaviour prediction can be achieved by combining machine learning algorithms with data mining techniques. In real-time applications and businesses, the sentiment analysis plays a vital role for behavior (consumer) predictions. Sentiment analysis focuses to get the feelings of a person with respect to an area of interest or the complete lexical polarity in a text document. Attitude can be a person’s judgment or assessment, emotional state of the author while posting the message, or the intentional emotional transformation i.e., the review outcome of the customer. Since the data for the reviews are fetched from social media, the data will be in “Huge Volume” and can be either structured or unstructured data (“Variety”). “Velocity” deals with the intensity at which data floods from source like social media sites. The flow of data is massive and continuous. The data that was residing in the form of Database can be either relational or hierarchical or embedded as object in database servers, storage in Web servers can be from Social Media, Email, Blog Contents, Photos, Video, Online chat, scientific data etc., are viewed only as data and they are organized in a structured manner. In today’s scenario, the data in any of the above form is termed as “Big Data”. The data residing in the servers of any organization, data from the Web especially from social media was quite data previously – arranged and filed. Big Data also contains data cached in clouds, the URLs that a Web user bookmarked and several other data such as Key-value stores, Document stores, Column-oriented Data, Graph-based models [25]. The organizations or Websites may perhaps not have the data in structured format. All these data can be used to get dissimilar outcomes with different methods of analysis. For the purpose of analysis, it is not mandatory to make use of all the data available at the repository. Different analysis requires various categories of big data for arriving at appropriate decisions, predictions and behavior.

## 2. Related Work

Social media analytics is the art of analyzing the collected information from social media servers and blogs. The primary role of social media analytics is to extract consumer sentiment for a consumption based marketing and service activities. Minqing Hu and Bing Liu [19], The feature selection method used by authors to mine the product features was “Part of Speech,” which concentrated on mining the opinion and features of products that reviewers had written. The main focus on the data retrieved from the social media, is to recognize the business objectives for raise in turnover or sales. The main objective of social network analytics endeavor is to govern the corporate goal (maximize profit/ high valued consumer/ number of clients) based on the analysis performed on the data collected. Other goals include increasing sales, lowering customer service costs, gathering input on goods and services, and enhancing public perception of a specific product or division [10-18]. Bogdan Batrinca et al., in [24] also discussed about the Classic objectives stated above. New products are launched in a high frequency and the customers are provided with a wide range of options for each and every product to choose

from before they decided to buy. In this era of digital world, most of the customers share their opinion of the products they own/purchased through different blogs, social media, etc. With the help of these reviews, customers share the parameters of their view regarding the purchased products to anyone, which leads to a great opportunity for the manufacturers to have an opinion of their product's sale in the future along with the drawbacks which they need to rectify. Marco Balduini et al.[21], Aggarwal Charu C et al[22] focused the opportunity for marketing intelligence which regardless of gender, race, nationality, religion; today's generation share each and everything via email, blogs or commenting on Websites, but social media is the widely used medium. Opinions can be about anything such as on an election poll prediction, an event, a product, a job or anything. These opinions are circulated through the Web and provides an interactive platform for consumers to influence someone or get influenced. The gap between the manufacturer or designer and the consumers is reduced and the marketers can understand the sentiments of their consumers and their perception in real time. But analyzing the data has thrown a real time challenge to marketers, as the data in social media are found in fragments in a completely unorganized and sprinkled manner. Wu He et al.,[1] focused on social media combative analytics structure with sentiment benchmarks that will be used to garner corporate-specific promotional intelligence. Sentiment analysis is a method of text mining which overcomes these challenges by extracting the data and analyzing it from the Web without any delay in time. With the help of this technique, marketers have an ample of opportunity to understand the reality of their business products.

1. Find the features of the product on which consumers have expressed their opinions (called opinion features) and grade them based on how often they give the impression in the reviews. [2,3,4,5].
2. Determine how many positive or negative customer reviews there are for each feature. The functionality is accompanied by specific reviews that share these viewpoints. This makes it easier for prospective buyers to comb through the reviews.

For example assume that the reviews of a specific smart phone, Smartphone\_1 are summaraized as follows:

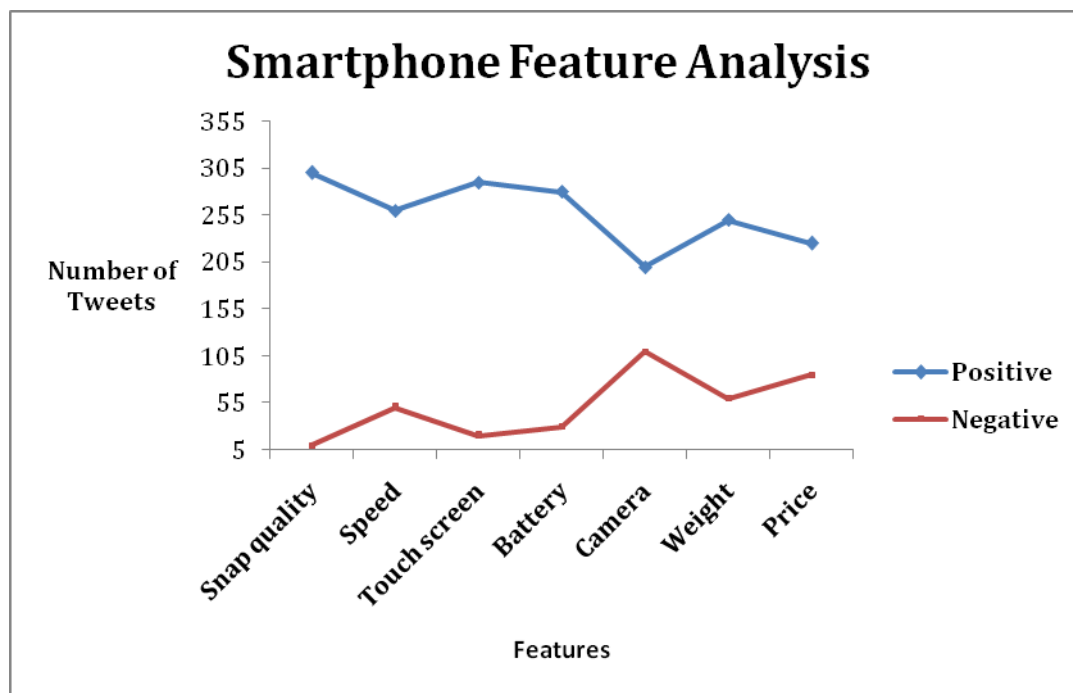


Fig. 1. Sample Smartphone Analysis

Walaa Medhat et al[3] discussed, the key techniques and samples of the ways for classifying different sentiments and feature selection for an improved sentiment analysis. The main objective is to use sentimental analysis as a predictive analysis tool. Manufactures has understood the only way to connect their user to the Web and so they are trying to unfold the challenges required to extract the data and use for their own benefits. Also the enhancement in technology has opened up a way for the vendors to analyze the huge amount of data without any delay in time or unobstructed by some external sources. With this powerful tool, the marketing research cost and human error has been reduced reasonably (i.e.,customer's preferences can be monitored through social media data) and helped in an organized analysis of data in this competitive world.

Kerstin Denecke et al., [9] provided an overview of how entity extraction and sentiment analysis findings can be used to improve classification efficiency and effectiveness. Based on the extraction of instructive and sentimental material, the proposed method will differentiate between sentimental and informative medical articles. In [21], BOTTARI, a sophisticated use of semantic technology that allows use of the rich and collective information gleaned from the continuous analysis of social media streams. The results indicate that when implementing social media marketing campaigns, businesses should distinguish and harness the specific influence of numerous media outlets. [6].

### 3. Methodology

The proposed methodology for the sentiment analysis is defined in Fig.2. Initially, the data extraction process from Twitter was carried out. The process of stemming is

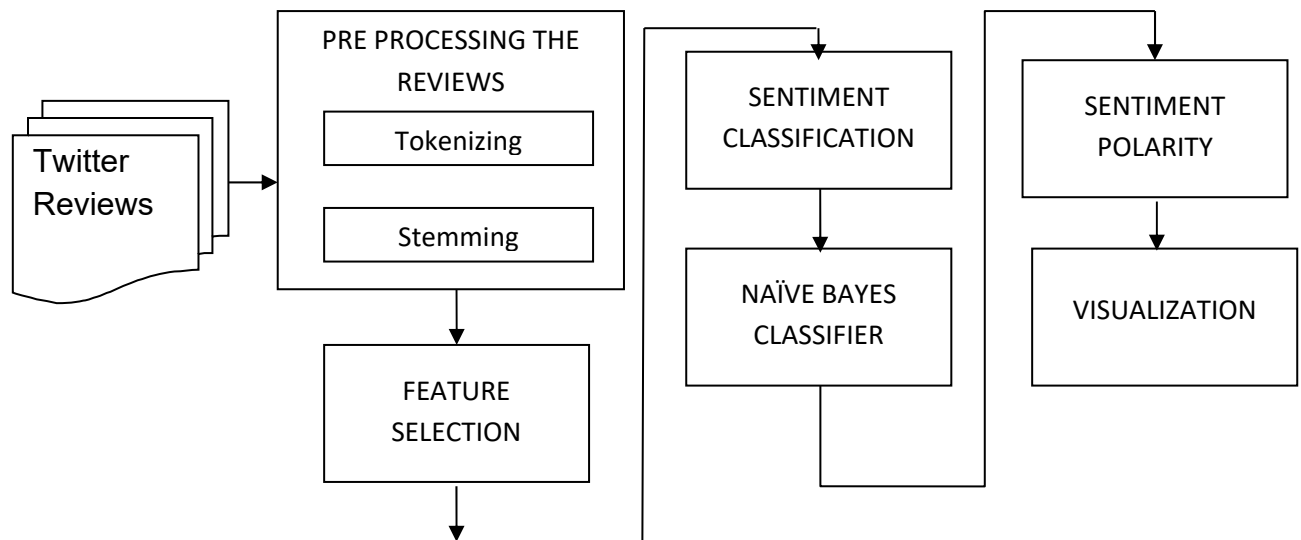


Fig. 2. Proposed Methodology for Sentiment Analysis

achieved by removing suffixes from words to get the common origin. The method of choosing a subset of the terms found in the training/testing set is known as feature selection. Text classification is carried out by using these subset features. The data obtained from feature selection process is further processed for sentiment classification to extract subjective information from the reviews obtained. In this step supervised learning method called Naïve Bayes Classifier is applied for the text using the probabilistic approach. Polarity for the reviews are assigned for obtaining the aspect belongs to positive, negative or neutral for finding the sentiments of every tweet. Finally, based on the obtained polarities, various sentiments are visualized.

### 4. Implementation

#### 4.1 Preprocessing the Tweets

The content of the tweets are in complete raw form. Understanding this textual comment is quite impossible. People tend different sentiments in different statements which make this work quite

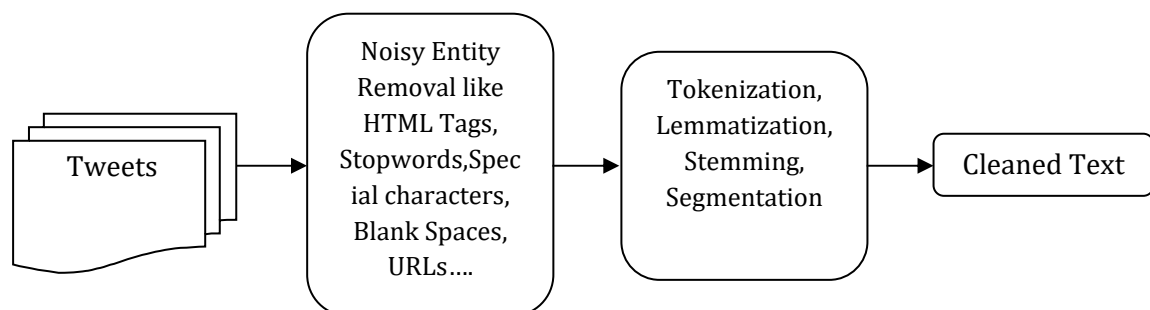


Fig. 3. Preprocessing Steps

difficult. The data extracted are in a complete disorganized fragmented manner with opinions written in their own style of vocabulary and expressions. Many statements, even use slang and some pseudo words in their reviews. So it can be said that it is impossible to find the sentiments of emotion showed in them. However the existance of advanced technologies of artificial intelligence, machine learning (innovative text analytics) and sentiment analysis the data is extracted and analyzed. Before the data is forwarded for analysis, the data of disinterest should be removed. These data are not considered for the sentiment analysis process. Let's clean the corpus first prior to implement some intuitive algorithms to extract more perceptions from the tweets we've gathered. To clean up, we must first comprehend the structure of the tweet list. It contains a lot of meta data including @ people, URLs, and # hashtags, as well as unnecessary spaces, numbers, and punctuation marks. A number of such tweets are retweets by same users, for the provided app the necessary data like the reviewer's name, id, date at which they have expressed their reviews etc., alone should be considered. Also data which is not relevant to the analysis, such as reviews which are not in the context of the analysis are also removed. Fig.3 explains the preprocessing of tweets in detail. The input tweets are read in the form of text. Special symbols “#” and “@” associated with the textual reviews are removed from the reviews. Then the punctuations, numbers, blank spaces, tabs and NA values are removed and the cleaned text is stored in new list for further analysis.

#### 4.1.1 Sample Tweets

```
> head(samsung_tweets)

[[1]]

[1] "lilbeastgirl16: Dear @samsung @SamsungMobile stop installing apps we dont need! We dont like it, we our phone memory! #samsungphones #samsungissues"

[[2]]

[1] "mark2ak: @DarlenJGMC @justbware #Iphones dont have shit on #samsungphones period,, overall"

[[3]]

[1] "bot_innovation: RT @TypesyApp: How the Samsung Galaxy S7 Came To Be #SamsungS7 #innovation #smartphones #SamsungPhones #Technology https://t.co/zXBlo8AOaC"

[[4]]

[1] "TypesyApp: How the Samsung Galaxy S7 Came To Be #SamsungS7 #innovation #smartphones #SamsungPhones #Technology https://t.co/zXBlo8AOaC"

[[5]]

[1] "erefect: How the Samsung Galaxy S7 Came To Be #SamsungS7 #innovation #smartphones #SamsungPhones #Technology https://t.co/sijL1qWe4I"

[[6]]

[1] "taedreamin: RT @jeffqian345: Very strange that Samsung chose to announce the Galaxy Note 7 instead of the Galaxy Note 6 #weird #whyskipit #samsungphones"
```

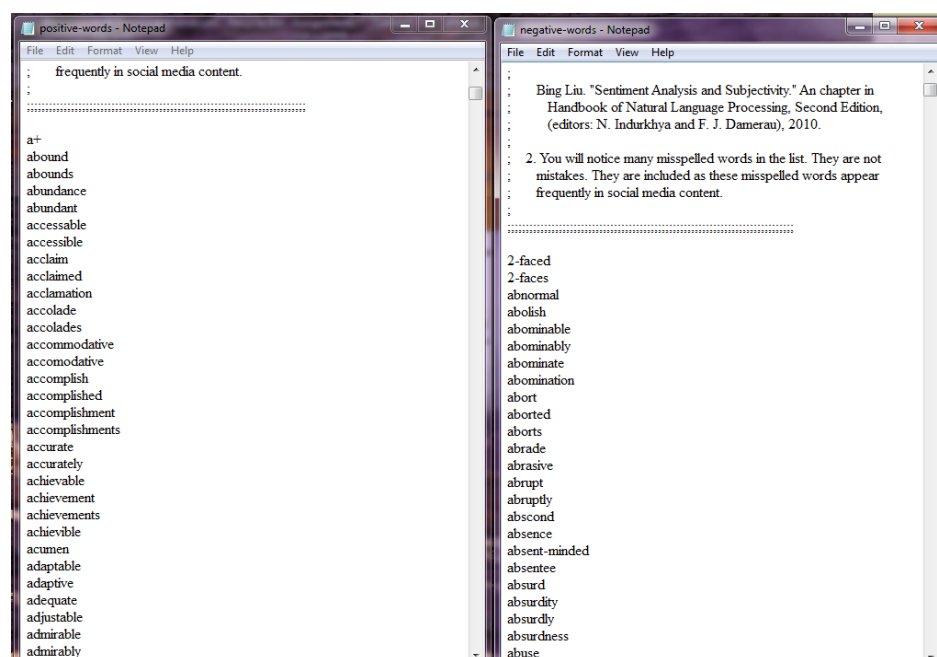
Fig. 4. Sample Samsung PhoneTweets

Fig 4. Shows the sample of tweets extracted from Twitter using hashtag #samsungmobiles. Sample sentences will be split based on matching list of positive and negative words. Each sentence can be positive or negative. The score will be difference between the sum of positive sentence and negative sentence, the sample scores shown in Fig 5..

#### Pseudocode to compute the Score (POSITIVE INTEGER and NEGATIVE INTEGER)

```
Let,    Input  : Sample Tweets, Predefined List of Positive and Negative word
        Output : Score
        Calculate the Score (POSITIVE INTEGER and NEGATIVE INTEGER)
Let,    Min_Score= max(negative word)
        Max_Score = max(positive word)
        No_Score = equal(positive word and negative word)
```

Text	Score
1 dear stop installing apps we dont need we dont like it we our phone memory	-2
2 dont have shit on period overall	-1
3 how the samsung galaxy s came to be	0
4 very strange that samsung chose to announce the galaxy note instead of the galaxy note	-1
5 samsung extends lead over apple in smartphone market	2



### 4.3 Tweaking Naive Bayes Classifier

The naive Bayes classifier has a big benefit in terms of training process. Furthermore, since the structure is in the case of products, it can be transformed to a number using logarithms, resulting in substantial computational benefits. Another general technique for measuring probability distributions of data is maximum entropy. When nothing is understood, the overarching concept of maximum entropy states that the distribution has to be as homogeneous as possible, i.e., have max entropy. Labelled data is used to generate a collection of model constraints that describe the distribution's class-specific objectives.

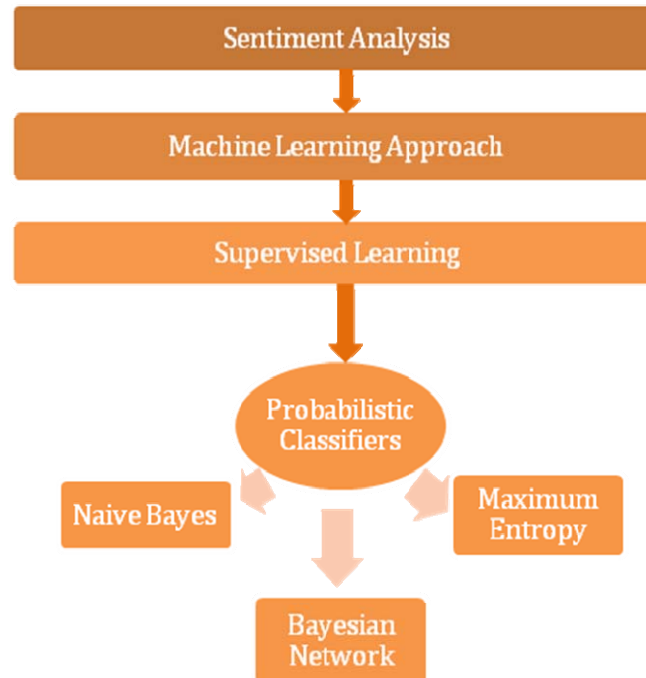


Fig. 7. Probabilistic Classifiers for Sentiment Classification

As shown in Fig.7, Set of machine learning techniques are available for performing sentiment analysis on various platforms like movie reviews, product reviews, etc., In supervised learning approach, under probalistic classifiers as shown in figure 2, there are three classifiers available. 1. Naïve Bayes 2. Maximum Entropy 3. Bayesian Network. A Bayesian Network (BN) is a statistical approach that depicts the probable associations between a group of variables. Sentiment Analysis [8,10,18] is a part of machine learning that is in its early days, the process of mining information from existing data has been employed by enterprises for a while but Sentiment Analysis has very little information lying around. Sentiment Analysis could be a powerhouse in deciding the success of various enterprises (“Alert Vendor if someone complains about our product”) and it is a relatively new concept that there are a lot of improvements to be made. Due to the dependence of the general public on social sites like Facebook, Twitter, there is a huge volume of data that can be processed to get knowledge from. In the proposed approach, A Naive Bayes classifier algorithm[26], is applied for performing the sentiment analysis. The Naive Bayes classifier is developed as a tool to analyze the sentiments that are posted on Twitter.

It is called naive since it assumes that the natural language can be processed as individual words, but this is not the case. The word ‘Tower’ is more likely to be followed by ‘of England’ than ‘of potatoes’. This shows that the language isn’t built on single words, but a series of words that convey a message according to the situation it is used upon. Sentiment Analysis is different from classification as it is subjective. Objective language is easier to process as it follows a set of grammatical rules that can be taught to an algorithm. Subjective language is prone to idioms, phrases, sarcasm and slangs (“I like the game.... Not.”).

Processing the sentence as it is in objective language may lead to it, because of the word "good," this sentence is known as a positive sentence. use of the word “like”. But the word “not” negates the meaning of the sentence. This is the reason why data should be considered as bigrams instead of unigrams. If the data set is sufficiently large, identification and classification of data on bigram will be efficient. It is much better if the data are classified beforehand. (The word “Modi” is found more frequently in documents classified as ‘politics’ rather than those classified as ‘Sports’). Considering the words as bigrams is achieved by considering the words as

pairs. Instead of tokenizing the sentence “ The food was not good” as ‘The’, ‘food’, ‘was’, ‘not’, ‘good’, it can be tokenized as pairs such as ‘The food’, ‘food was’, ‘was not’, ‘not good’. Now the algorithm can identify the difference between ‘was good’ and ‘not good’.

A number of multinomial records assumptions have been successfully corrected using feature normalizations like TF(Term Frequency)-IDF(Inverse Document Frequency) [7]. In context of information extraction, there are a number of TF-IDF features. The generic form of TF-IDF [27] used in this study is as follows:

$$\omega_m = \log \left[ 1 + \frac{\omega_m^u}{s(\omega^u)} \right] \log \left[ \frac{doc}{doc_m} \right] \quad (1)$$

in Equation (1),  $\omega_m^u$  is the number of unsmoothed word,  $doc_m$  the amount of experimental corpus the string  $m$  happens in it and  $doc$  the count of experimental corpus.

Unique size normalisation and word frequency log transform are performed by the first component in TF-IDF algorithm. The term "length normalisation" refers to the method of accounting for variations in document length. Number of words and cosine length normalizations are also common options for length normalisation, but specific length normalisation has been more consistent across datasets. By dampening elevated counts to enhanced suit the Multinomial distribution, the term frequency log transform exactness for the "burstiness" cause on term calculations. The IDF weighting of words is done by the next factor in TF-IDF function. This gives insightful terms more weight, allowing the weighted counts to better distinguish the groups.

Use the document's words (or terms/phrases) to label it on the appropriate entities in a text classification issue. According to [7], the following classifier is obtained using the “maximum a posteriori (MAP)” decision law: is given in Equation (2)

$$e_{map} = \arg \max (P(e|doc)) = \arg \max \left( P(e) \prod_{1 \leq m \leq n_{doc}} P(s_m|e) \right) \quad (2)$$

Where  $s_m$  are the symbols (phrases/terms) this implies that, in order to determine which group it fit into as a novel document, we ought to measure the product of likelihood of every term in the corpus provided a specific entity (possibility) represents the probability of the specific entity (preceding). Once computing the maximum posteriori for all entities of the given set  $e$ , select the one with the highest probability. Estimating the outcome of these probabilities would result in float point underflow because computers can manage figures with precise decimal point precision. The result will be a very tiny proportion that will not suit in memory, so it'll be accumulated to zero, making the study ineffective. To evade this, as an alternative of increase in the product of probabilities, maximize the total of all their logarithms as in equation (3):

$$e_{map} = \arg \max \left( \log P(e) + \sum_{1 \leq m \leq n_{doc}} \log P(s_m|e) \right) \quad (3)$$

Choose the class mostly with greatest log value rather than the one with the highest likelihood. Since the logarithm method is monotonic, MAP's opinion stays consistent.

**Pseudocode to compute the Emotions (JOY, DISGUST, ANGER, FEAR, SADNESS and SURPRISE)**

```

Let,   Input  : TWEET
        Output : BEST_FIT
        Calculate the EMOTION (JOY, DISGUST, ANGER, FEAR, SADNESS and SURPRISE) Value (in
        Decimal)
Let,   Min_Value=7 (for each emotion)
        For each TWEET
        IF Value = Max(Value) and Value>=Min_Value then BEST_FIT = “EMOTION”
        ELSE BEST_FIT = “NA”
    
```



>head(SamsungTweetsClassEmo,15)						
	ANGER	DISGUST	FFAR	JOY	SADNESS	BEST_FIT
[1,]	"1.46871776464786"	"3.09234031207392"	"2.78695866252273"	"1.02547755260094"	"1.7277074477352"	"2.06783599555953" NA
[2,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[3,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[4,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[5,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"7.34083555412328"	"1.7277074477352"	"2.78695866252273" "joy"
[6,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[7,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[8,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"13.6561935556456"	"1.7277074477352"	"7.34083555412327" "joy"
[9,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[10,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"7.34083555412328"	"1.7277074477352"	"2.78695866252273" "joy"
[11,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"7.34083555412328"	"1.7277074477352"	"2.78695866252273" "joy"
[12,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"7.34083555412328"	"2.78695866252273" "sadness"
[13,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"1.7277074477352"	"2.78695866252273" NA
[14,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"1.02547755260094"	"7.34083555412328"	"7.34083555412327" "sadness"
[15,]	"1.46871776464786"	"3.09234031207392"	"2.06783599555953"	"7.34083555412328"	"1.7277074477352"	"2.78695866252273" "joy"

Fig. 8. Emotions of Samsung Smart Phone Tweets

Fig.8 shows the results obtained through Bayesian analysis for the Samsung Smart Phone tweets. The sentiment bundle was created with a qualified collection of sentiment terms in mind (nearly 1,500 words). The polarity value of the following six emotions is seen in the induced results: Joy, Sadness, Anger, Disgust, Surprise and Fear. NA is returned while it is unable to measure the final emotion as one of the above mentioned options.

#### Pseudocode to compute the positive and negative livelihoods

Let,     Input : TWEET  
           Output : BEST\_FIT  
           Compute POS, NEG  
           Calculate the Ratio POS/NEG  
           IF Ratio>1 then BEST\_FIT -> "POSITIVE"  
           ELSE IF Ratio<1 then BEST\_FIT -> "NEGATIVE"  
           ELSE BEST\_FIT -> "NEUTRAL"

>head(SamsungTweetsClassPol,15)				
	POS	NEG	POS/NEG	BEST_FIT
[1,]	"8.78232285939751"	"0.445453222112551"	"19.7154772340574"	"positive"
[2,]	"1.03127774142571"	"9.47547003995745"	"0.108836578774127"	"negative"
[3,]	"1.03127774142571"	"0.445453222112551"	"2.31512017476245"	"positive"
[4,]	"17.2265151579293"	"18.5054868578024"	"0.930886892644283"	"negative"
[5,]	"8.78232285939751"	"0.445453222112551"	"19.7154772340574"	"positive"
[6,]	"1.03127774142571"	"18.5054868578024"	"0.0557282145209216"	"negative"
[7,]	"9.47547003995745"	"0.445453222112551"	"21.2715265477714"	"positive"
[8,]	"49.6169899909363"	"0.445453222112551"	"111.385410471674"	"positive"
[9,]	"1.03127774142571"	"0.445453222112551"	"2.31512017476245"	"positive"
[10,]	"17.2265151579293"	"9.47547003995745"	"1.81801167491282"	"neutral"
[11,]	"17.2265151579293"	"9.47547003995745"	"1.81801167491282"	"neutral"
[12,]	"8.78232285939751"	"9.47547003995745"	"0.926848253686942"	"negative"
[13,]	"25.670707456461"	"9.47547003995745"	"2.70917509613869"	"positive"
[14,]	"16.5333679773693"	"17.8123396772424"	"0.928197433742679"	"negative"
[15,]	"17.9196623384892"	"0.445453222112551"	"40.2279329207804"	"positive"

Fig. 9. Sentiment polarity of Samsung Smart Phone Tweets



The Fig.9 shows the aggregated values of emotions, where joy can be considered as positive emotion. Sadness, disgust, anger and fear can be categorised as negative emotion. Surprise categorised as neutral because it can be positive emotion or neagative emotion based on the state of customer.

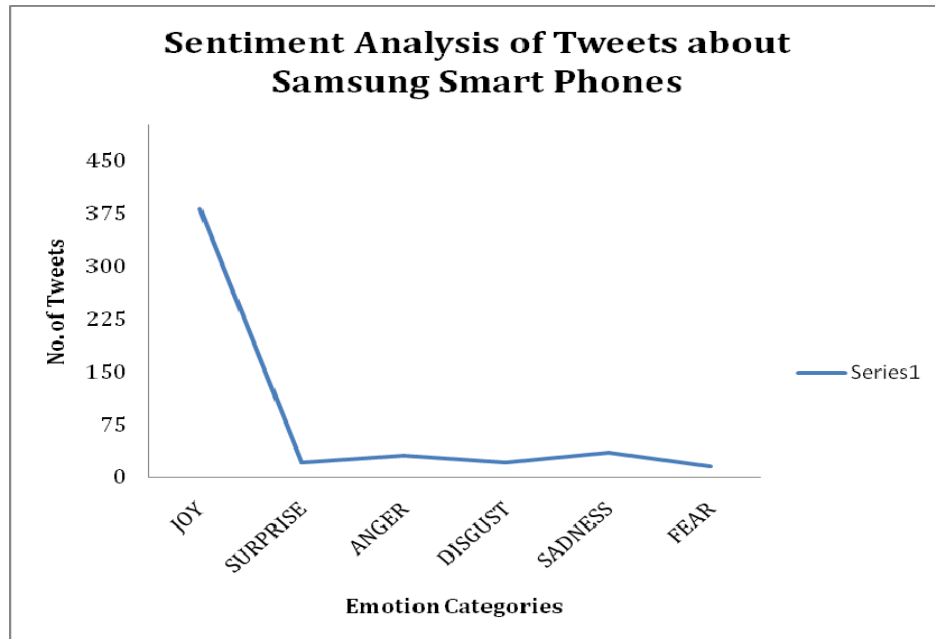


Fig. 10. Visualization of Emotions

Fig. 10 shows the visualization of emotions which were assigned with polarity values previously. Out of five-hundred tweets maximum of tweets expressed emotion 'joy', all other emotions were also expressed by the customers. The level of accuracy of the proposed approach is 90%.

## 5 Conclusion

A machine learning based approach has been developed for finding the customer's sentiments by analysing the reviews posted by them in social medias. In order to clean the reviews, text preprocessing which involves stemming and tokenizing were carried out. To select features in the text, POS technique was used and by using Naïve Bayes Classifier, the sentiments such as Joy, Sadness, Anger, Disgust, Surprise and Fear were classified. Sentiment polarities were assigned to find the overall sentiment towards each customer's reviews. Finally the six different sentiments were visualized. The visualized result shows each review was classified with accurate sentiment none of the review came under Not Applicable Category. Manufacturers can use the results to improve the value of the upcoming products to fullfil the requirements of clients and to increase the sale percentage.

## References

- [1] Adam Hale Shapiro, Moritz Sudhof, Daniel J. Wilson, Measuring news sentiment, Journal of Econometrics, 2020, ISSN 0304-4076.
- [2] Ahmad, S., Asghar, M.Z., Alotaibi, F.M. and Awan, I., 2019. Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. Human-centric Computing and Information Sciences, 9(1), p.24.
- [3] Barbieri, N., Manco, G., Ritacco, E., Carnuccio, M. and Bevacqua, A., 2013. Probabilistic topic models for sequence data. Machine learning, 93(1), pp.5-29.
- [4] García-Pablos, A., Cuadros, M. and Rigau, G., 2018. W2VLDA: almost unsupervised system for aspect based sentiment analysis. Expert Systems with Applications, 91, pp.127-137.
- [5] Guerreiro, J. and Rita, P., 2020. How to predict explicit recommendations in online reviews using text mining and sentiment analysis. Journal of Hospitality and Tourism Management, 43, pp.269-272.
- [6] Huang, F., Zhang, S., Zhang, J. and Yu, G., 2017. Multimodal learning for topic sentiment analysis in microblogging. Neurocomputing, 253, pp.144-153.
- [7] Jha, V., Savitha, R., Shenoy, P.D., Venugopal, K.R. and Sangaiah, A.K., 2018. A novel sentiment aware dictionary for multi-domain sentiment classification. Computers & Electrical Engineering, 69, pp.585-597.
- [8] Lamprier, S., Gisselbrecht, T. and Gallinari, P., 2019. Contextual bandits with hidden contexts: a focused data capture from social media streams. Data Mining and Knowledge Discovery, 33(6), pp.1853-1893.
- [9] Li, W., Guo, K., Shi, Y., Zhu, L. and Zheng, Y., 2018. DWWP: Domain-specific new words detection and word propagation system for sentiment analysis in the tourism domain. Knowledge-Based Systems, 146, pp.203-214.
- [10] Li, Z., Fan, Y., Jiang, B., Lei, T. and Liu, W., 2019. A survey on sentiment analysis and opinion mining for social multimedia. Multimedia Tools and Applications, 78(6), pp.6939-6967.
- [11] Liao, S., Wang, J., Yu, R., Sato, K. and Cheng, Z., 2017. CNN for situations understanding based on sentiment analysis of twitter data. Procedia computer science, 111, pp.376-381.

- [12] Ma, R., Wang, K., Qiu, T., Sangaiah, A.K., Lin, D. and Liaqat, H.B., 2019. Feature-based compositing memory networks for aspect-based sentiment classification in social internet of things. *Future Generation Computer Systems*, 92, pp.879-888.
- [13] Preoțiuc-Pietro, D., Liu, Y., Hopkins, D. and Ungar, L., 2017, July. Beyond binary labels: political ideology prediction of twitter users. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 729-740).
- [14] Puustjärvi, J. and Puustjärvi, L., 2017. Increasing the expression power of persons' profiles in semantic social networks. *Procedia Computer Science*, 111, pp.8-16.
- [15] Shaalan, Y., Zhang, X., Chan, J. and Salehi, M., Detecting singleton spams in reviews via learning deep anomalous temporal aspect-sentiment patterns. *Data Mining and Knowledge Discovery*, pp.1-55.
- [16] Simpson, E. and Gurevych, I., 2020. Scalable Bayesian preference learning for crowds. *Machine Learning*, pp.1-30.
- [17] Singh, N.K., Tomar, D.S. and Sangaiah, A.K., 2020. Sentiment analysis: a review and comparative analysis over social media. *Journal of Ambient Intelligence and Humanized Computing*, 11(1), pp.97-117.
- [18] Vadicamo, L., Carrara, F., Cimino, A., Cresci, S., Dell'Orletta, F., Falchi, F. and Tesconi, M., 2017. Cross-media learning for image sentiment analysis in the wild. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* (pp. 308-317).
- [19] Wang, W., Tang, B., Fan, X., Mao, H., Yang, H. and Zhu, M., 2017. Efficient visibility analysis for massive observers. *Procedia computer science*, 111, pp.120-128.
- [20] Xiao, W., Zhao, H., Pan, H., Song, Y., Zheng, V.W. and Yang, Q., 2020. Social explorative attention based recommendation for content distribution platforms. *Data Mining and Knowledge Discovery*, pp.1-35.
- [21] Xiao, Z., Li, X., Wang, L., Yang, Q., Du, J. and Sangaiah, A.K., 2018. Using convolution control block for Chinese sentiment analysis. *Journal of Parallel and Distributed Computing*, 116, pp.18-26.
- [22] Yang, J., She, D., Sun, M., Cheng, M.M., Rosin, P.L. and Wang, L., 2018. Visual sentiment prediction based on automatic discovery of affective regions. *IEEE Transactions on Multimedia*, 20(9), pp.2513-2525.
- [23] Zeng, D., Wang, S., Shen, Y. and Shi, C., 2017. A GA-based feature selection and parameter optimization for support tucker machine. *Procedia computer science*, 111, pp.17-23.
- [24] Zhang, S., Wei, Z., Wang, Y. and Liao, T., 2018. Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary. *Future Generation Computer Systems*, 81, pp.395-403.
- [25] Zhou, F., Jiao, J.R., Yang, X.J. and Lei, B., 2017. Augmenting feature model through customer preference mining by hybrid sentiment analysis. *Expert Systems with Applications*, 89, pp.306-317..

#### Authors Profile



Ms.J.S.Vimali, completed B.Tech and M.Tech in Information Technology. Working as Assistant Professor in Sathyabama Institute of Science and Technology since 2006. I'm pursuing my Ph.D in the domain of Social Networks in Sathyabama Institute of Science and Technology. My areas of interest are Sentiment Analysis, Data Analytics, Deep Learning Supply chain management.