Int.J.HIT.TRANSC:ECCN.Vol.7: Issue 1A (2021) Page 62-71 ISSN: 0973-6875

Available Online at www.hithaldia.in/locate/ECCN


ORIGINAL CONTRIBUTION

# An Analysis and Discussion of Human Sentiment Based on Social Network Information

Sangita Bose[1], Siddhartha Chatterjee[2], Bidesh Chakraborty[3], Pratik Halder[4], Saikat Samanta[5]

[1]Department of Computer Science and Information Technology, NIELIT Centre, Barasat Govt. College, Kolkata - 700124, West Bengal, India

[2]Department of Computer Science and Engineering, Gargi Memorial Institute of Technology, Kolkata - 700144, West Bengal, India

[3]Department of Computer Science and Engineering, Haldia Institute of Technology,

Haldia - 721657, West Bengal, India.

[4]Department of Computer Science and Engineering, Gargi Memorial Institute of Technology, Kolkata - 700144, West Bengal, India

[5]Department of Computer Science and Engineering, Bengal College of Engineering, Durgapur – 713212, West Bengal, India

**ABSTRACT**

Sentiment Analysis shows, for example, the emotions or feelings present behind any social media message blog post, Facebook status or tweet. This information is essential to understand the feelings of the author which play a critical role in influencing the readers. Any individual's social impact depends on many factors, the most important of which is sentiment analysis as it reveals human emotions. This is one of the rapidly increasing fields of research that explores the influence of social media by writers focused on how the emotions of the reader are affected. This paper looks at the various aspects of study of feelings and their effect on the power of social media. Sentiment analysis, also named as opinion mining, is the camp of research that scanned the opinion of people, their feelings, attitudes, habits, and emotions towards institutions such as products, services, organizations, individuals, problems, events, topics, and their highlights. This offers enough space for problems.

**Keywords:** Twitter, sentiment analysis, sentiment classification, opinion mining, emoticons.

## 1. INTRODUCTION

How others still say a significant piece of news for us during the decision. Opinions and reviews can now be found almost everywhere-blogs, social media such as Facebook and Twitter, online forums, e-commerce pages, Play store applications etc. Business organizations and individuals are getting so much support from these opinions, but users are overwhelmed by the tremendous amount of these opinionated text data. The way opinions can be analyzed and summarized in these large, opinionated text data is very interesting domain for researchers. This modern research field is commonly referred to as Sentiment Analysis, or Opinion Mining. Sentiment analysis is the automated outpouring by moving Natural Language Processing of behaviours, perceptions and emotions from the origins of text, speech, and databases. There are three types of opinion that include an interpretation of emotions, i.e. "positive" or "neutral." While linguistics and natural language processing (NLP) have a long history, prior to the year 2000 little work was done on the thoughts and feelings of people. The field has since grown into a very active area of study. There are several explanations for that. First, it

*Corresponding Address: sangitabose21@gmail.com

has a wide range of implementations, nearly in any area. The market underlying sentiment analysis has also flourished due to the emergence of commercial applications. This gives the researchers great inspiration in the field of science. It also provides many daunting research problems which have never before been studied [1, 2].

## 2. APPLICATIONS OF SENTIMENT ANALYSIS

Analysis of emotions may be carried out for various reasons in different fields. This segment treats some of the rising ones [3].

***Ecommerce-*** Sentiment analysis is often used in e-commerce operations. Websites request that their users apply their experience or provide input on shopping and the quality of the items. By awarding ratings or scores they provide a description of the product and various features of the product. Opinions and suggestions on the whole product as well as relevant product features & graphical overview of the overall product can be easily seen by the consumer and their preferences are presented to users. Popular retailer websites such as amazon.com provide feedback from editors and even from customers with rating details.

***Voice of the Market (VOM)*** **–** Market's most critical feature and voice is the study and determination of what consumers think about rival goods or services. Competitive advantage can be achieved by reliable and timely consumer voice knowledge and the creation of new products. Any company can get consumer opinion through Sentiment Analysis in real time, which helps them plan new marketing campaigns, develop product functionality and predict product failure chances.

***Brand Reputation Management -*** Image Strategy is preoccupied with maintaining the business image. Opinions from customers or some other party can harm or improve your credibility. Brand Reputation Management (BRM) focuses not on a customer but on product and the business. Now-a-days, one-to-many conversations are held at a high pace online. Analysis of the feelings helps to decide how online culture perceives the brand, product or service of the company.

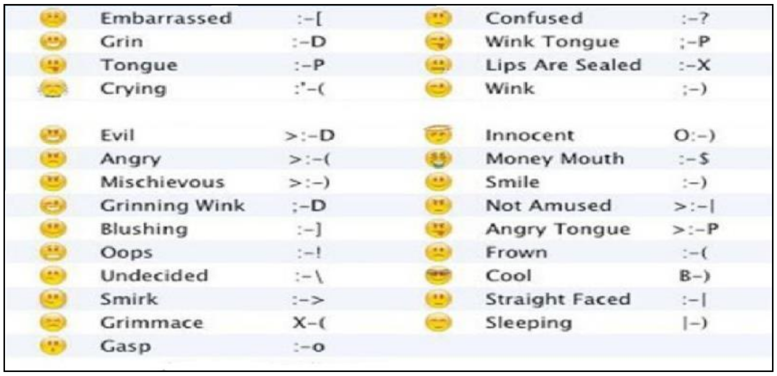
***Government-*** Governments can determine their strengths and weaknesses by using sentiment analysis to evaluate public perceptions. If that's the case, for example, how do you expect the truth to come out? The MP itself prosecuting 2 g fraud is highly corrupt. But this is an indication of negative policy opinion.

## 3. EMOTICONS

Emoticons of emotion are used for useful and accurate measures, and may therefore be used either to automatically produce a training corpus or to serve as proof to improve the classification of feelings.

Emoticons are articulated into the written language as non-verbal components that played the role of facial expressions in speech. Their function is primarily pragmatic: emoticons have two forms of visual communication, i.e. positive or negative senses. There is a relationship between the emotion orientation of the emoticons and messages according to this consideration. The emoticons were divided into two major groups, i.e. positive and negative. Positive emoticon instances are :-),;), =): D, while negative examples are:-: ('(= (; (. Any details for polarity classification emoticons are certainly an valuable source. In addition, emoticons on social media are mainly used for positive and negative messages. When social media is one of the main forms of communication, it makes them an desirable data source.

By analyzing emoji's, the precision of recognizing emotions can increase and improve. They offer a vital piece of information and this is important for businesses to better understand the feelings of their customers. The way people interact is constantly evolving thanks to technology; in order to understand what they want, it is necessary for businesses to adapt the way they respond to these changes, and this often means taking into account emoticons [4].



| | | | |
|---|---|---|---|
| Embarrassed | :-[ | Confused | :-? |
| Grin | :-D | Wink Tongue | ;-P |
| Tongue | :-P | Lips Are Sealed | :-X |
| Crying | :'-( | Wink | ;-) |
| Evil | >:-D | Innocent | O:-) |
| Angry | >:-( | Money Mouth | :-$ |
| Mischievous | >:-) | Smile | :-) |
| Grinning Wink | ;-D | Not Amused | >:-\| |
| Blushing | :-] | Angry Tongue | >:-P |
| Oops | :-! | Frown | :-( |
| Undecided | :-\ | Cool | B-) |
| Smirk | :-> | Straight Faced | :-\| |
| Grimmace | X-( | Sleeping | \|-) |
| Gasp | :-o | | |

Figure 1: Emoticon image and its meaning

## 4. IMPORTANCE

Sentiment Analysis shows, for example, the emotions or feelings present behind any social media message-a blog post, Facebook status, or even a tweet. This information is essential to examine the sentiments of the author that play a critical role in influencing the readers on these platforms. A research study shows that an update to a single Facebook status can affect the feelings of multiple Facebook status updates. This is because of a single user's control over his Facebook peers. The study showed that a positive status update raises the number of positive messages by an average of 1.75, and a negative status update reduces the number of negative messages by an average of 1.29 [5].

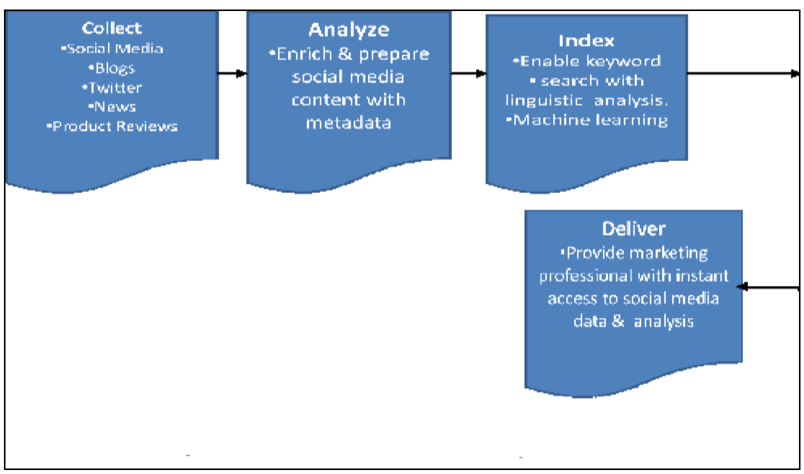



Figure 2: Sentiment Analysis of Social Networking DATA

## 5. CHALLENGES FOR SENTIMENT ANALYSIS

Analysis of sentiment classifies text as positive, negative or objective, so it can be considered a function of interpretation of the language. Classification of text has many classes because there are many subjects but study of emotions has just three classes. Compared to conventional text classification, however, there are several variables that make analysis of sentiments difficult. Below are a few of the factors [6].

***Co-reference resolution***– The problem of co-reference resolution is an understanding of what a pronoun or noun phrase is referring to. For example, "We watched the film and went for dinner; it was terrible." What does "it" mean?

Co-reference resolution can be helpful in understanding the topic / aspect-based feelings. Co-reference resolution could increase opinion mining accuracy.

***Temporal Relations -*** Review times can be critical for evaluating sentiments. The reviewer might think that Windows Vista is nice in 2008, but now, due to the new Windows 7, he may have a negative opinion in 2009. Thus evaluating this type of opinion, which is modified over time, can boost the efficiency of the system of sentiment analysis. It helps us to understand how a particular product is changed over time, or whether people change their minds about a product.

***Ironic sentences-*** Text may include sentences of sarcasm and irony. What a marvelous car, for starters! By the second day, it stopped working. In such a case, positive terms may have a sense negative. It can be difficult to recognize cynical or humorous sentences which can contribute to incorrect mining of opinions.

***Grouping synonyms-*** Many times, text includes different words that have the same meaning. Therefore, this word should be defined and grouped together for precise classification. Identifying such words is a challenging job, because people often use different words to identify the same characteristics. The voice and tone, for example, both refer to the same feature in phone analysis.

***Thwarted Expectations***-Some text contains sentences starting from different context and ends up having different meaning. The cast wasn't perfect, for instance, the actors performed poorly but I liked it. The last sentence in the above summary makes the whole summary positive. Because of more negative terms in the analysis, if the term frequency considered the above statements would classify as negative.

***Review Spam Detection***- Many people write fake reviews on the product review platform, called review spam, to promote their goods by giving undeserving positive views, or defaming the goods of their rivals by giving false negative views. The role of detection of spam opinion has significant impacts on industrial communities. Unless the service given an opinion includes a significant number of spams, it can impact the experience of the customer. Therefore, if the consumer is generated by the opinion given, he will never again use the program.

## 6. FEATURES

For our classification experiments, we use a number of functions. Comparison purpose here we use Unigrams and Bigrams. We do have features commonly used in sentiment analysis, namely features that mimic details from a dictionary and features related to POS emotion. Eventually, we have apps that cover few of the more basic segments for micro blogging languages [7].

***N-gram features-*** First we delete stop words to define a collection of useful n-grams. Then, by putting the word we detect rudimentary negation, not a phrase that goes forward or follows a negative expression. This has proven valuable in earlier investigations. Finally, both unigrams and bigrams are graded and scored according to their gain in information in the results of the study, measured using Chi-squared. We use the top 1000 n-grams in a word-type bag for our experiments. The apps are both unigrams and bigrams. Accuracy increased for Naive Bayes (81.3Ent (from 80.5 to 82.7) as compared with unigram applications. However, SVM declined (from 82.2a decline for Naive Bayes and SVM but Max-Ent improved.

***Lexicon features-*** Words listed in the MPQA lexicon on subjectivity are characterized by their prior polarity: positive, negative, or neutral. We have established three characteristic which depends from the present lexicon on the certain words.

***Part-of-speech features-*** With every tweet, we have apps for counting verbs, adverbs, adjectives, nouns and any other spoken section. Section of the voice we use (POS) tags as apps because depending on its use, the same term can have different meanings. For instance "over" may have a negative verb connotation. "Out" can also be used as a cricket noun which has no positive or negative connotations. We noticed the POS tags had not been helpful.

***Micro-blogging features-***We create differential apps representing the attendance of positive, negative, and neutral emotions and abbreviations as well as intensifier presence (e.g., all-caps and character repetitions). We also utilize the internet Lingo Dictionary and online slingshot dictionaries of the emotions and commentary.

## 7. EXPERIMENT AND RESULTS

Within this section we discuss the findings and observations for two classification tasks:

1. Positive vs. Negative.
2. Positive vs. Negative vs. Neutral.

We present three models for each of the classification tasks, as well as findings for two variations of these models:

1. Unigram model (our baseline).
2. Tree kernel model.
3. 100 senti features model.
4. Kernel plus senti features.
5. Unigram plus senti features.

Measuring feelings on your own, depending on the scale of the interaction, can be quite a time commitment. You will read each to record the mood of the mentions, Assess the tone, and give a positive, negative or neutral score. There are a few free tools available that monitor and measure feeling, and provide predictive sentiment analysis on content channels for social media marketing. If a person were to tweet about their shopping experience with Sears, they would decide the sentiment based on the words they use to describe. Significant sales at Sears! This would be classified as positive, while customer service for Sears would be the poorest reported as negative. A report is developed using sentiment analysis to determine the overall classification of sentiments shown in Figure 3. [8].

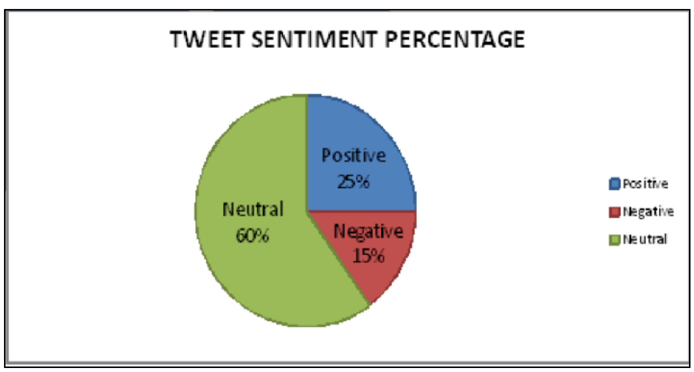



Figure 3: Report using Sentiment Analysis of Reviews

## 8. METHODOLOGY

When considering methodologies for social network intelligence we need a clearer understanding of the word "analysis of sentiment." Text analysis and sentiment analysis methods are also implemented for consumer views on social media analytics. Sentiment analysis and opinion mining have similar definitions, as both apply to automated evaluative text analysis and predictive decision monitoring. Analysis of sentiment is a type of natural language processing (NLP) and machine learning and its main application is the analysis of unstructured text and the extraction of key concepts from that text. Analysis of emotions can be applied in a number of commercial and non- ways. For example, review-related websites produce a large amount of feedback and are important for business intelligence purposes to summarize such data. Another application of this technique refers to enabling applications for other systems, such as recommendation systems, flames identification (overly heated or antagonistic language) and inappropriate material for ad placement, question answering among others.

Sentiment analysis is an option accessible to organizations to enhance their decision-making process by the application of a data-driven approach. Yet this mining strategy still faces several deployment challenges. Vinodhini et al offer a description of some of these issues. They note that in another circumstance an opinion word which is considered optimistic can be viewed in a negative way. A second difficulty is that people do not necessarily voice new identically and in the same paragraph, they can mix a positive and a negative opinion. [9].

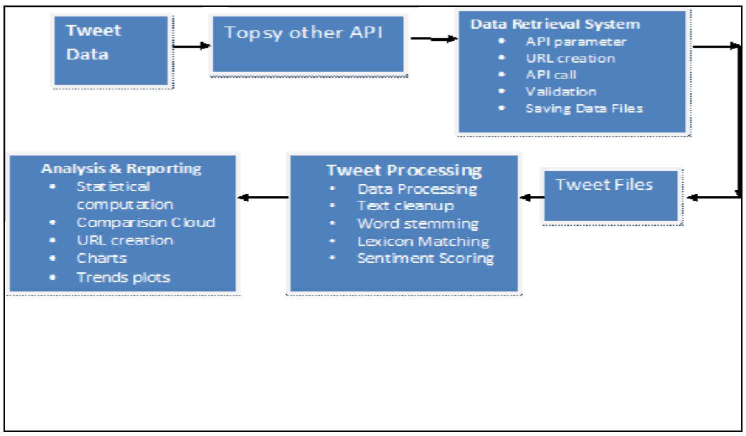



Figure 4: Methodology

## 9. APPROACH

In machine learning, our solution is to use unique classifiers and feature extractors. The classificators of machine learning are Naive Bayes, Maximum Entropy and Support Vector Machines (SVM). Functionality extractors are unigrams, bigrams, unigrams, and bigrams, as well as unigrams with a section of voice labels. We're developing a system that handled classifiers as two different components and features extractors. This system helps us to make it fast various classifiers and feature extractor combinations. Sentiment Analysis methods can be divided into four groups [10].

***Keyword Spotting* -** This technique emphasizes unequivocal words with clear meanings such as happy, sad, bored, etc.

***Lexical Affinity*** It senses words as well as assigning a specific generic emotion to certain arbitrary words.

***Statistical methods* –** This approach is based on mastering the machine and uses a foundation of knowledge often known as a bag of words.

***Concept-level techniques*** - This technique attempts to be more precise about the feeling of a sentence by detecting both the sentiment holder and the target.

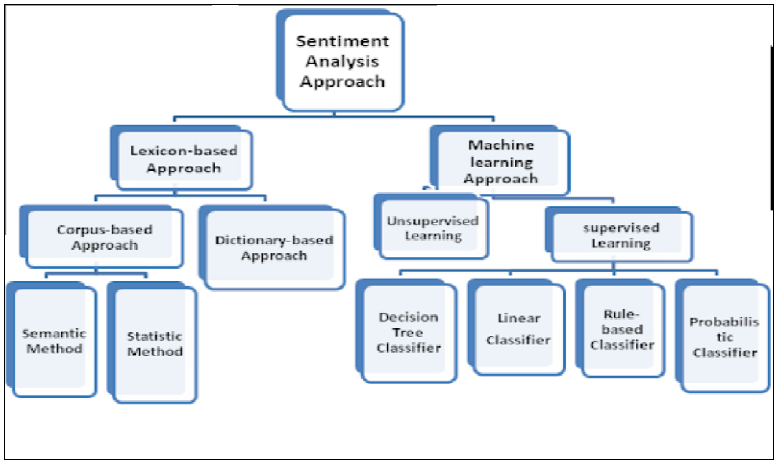



Figure 5: The approach of SENTIMENT ANALYSIS of Social Media

## MACHINE LEARNING METHODS

We check various classifiers: keyword-based, Naive Bayes, maximum entropy, and support vector machines [8, 15].

### Supervised Algorithm

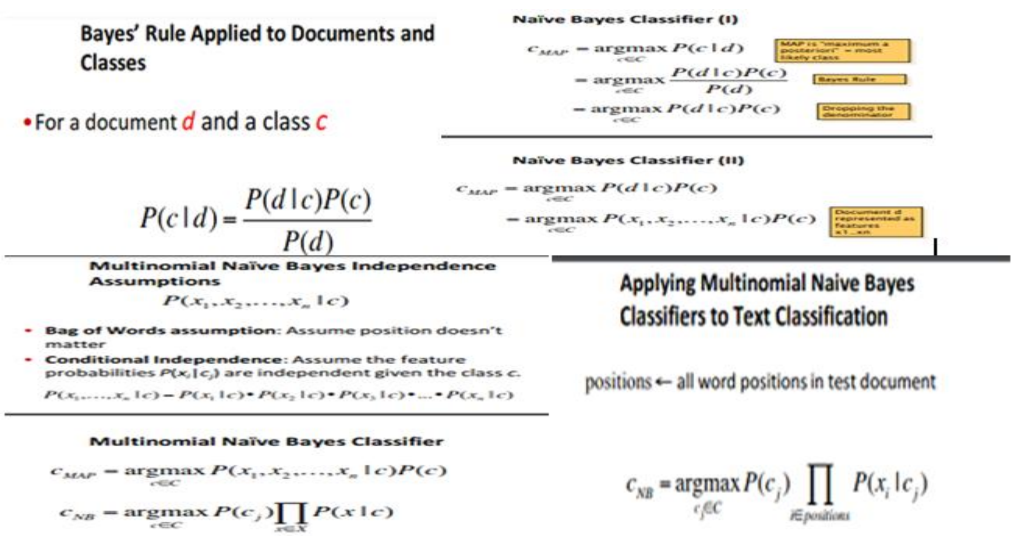
***A collection of inputs (text phrase) and the correct output class (Positive or Negative) must be given for those inputs. Using those inputs the learning algorithm will prepare. Before that a new instance can be listed***

### Baseline

Twitter is a website which analyzes tweet thought. A list of positive and negative keywords is to be used in approach. We use the keyword list on twitter as a reference which is free to the public. This list comprises 174 positive and 185 negative terms. We measure the number of negative keywords and positive keywords which appearing for each tweet. This classification returns the greater amount of polarity. If a tie occurs, the polarity returns positive.

### Naive Bayes

Naive Bayes is a fundamental model that fits perfectly in categorizing text. We make use of a version of the Naïve Bayes multinomial.



A. Maximum Entropy

Maximum Entropy models are designed to favor feature-based models to the most stable models that meet a given MaxEnt constraint. Using a distribution over the groups is the same for a two-class case as using the logistic regression. Like Naive Bayes, MaxEnt makes no statements about its features regarding freedom. What does it mean that we can add features like bigrams and sentences to MaxEnt without having to worry about surpassing functions? The platform is composed of:

$$\frac{1}{|\mathcal{D}|}\sum_{d\in\mathcal{D}} f_i(d, c(d)) = \sum_{d} \mathrm{P}(d)\sum_{c} \mathrm{P}(c|d) f_i(d, c). \quad (1)$$

We use the training data in maximum entropy to configure constraints on conditional distribution. Every restriction expresses a train data feature that should be present in the learned distribution as well. We let any document's real-valued function and the class is a component, fi (d, c). Maximum entropy allows one to limit the distribution of the model in order to have the exact value to be desired for this function as seen from preparation results, D. We therefore stipulate that the acquired conditional distribution P (c|d) shall have the property

## 10. FUTURE OF SENTIMENT ANALYSIS

Techniques of machine learning work well in classifying sentiment in tweets. We assume there is still scope to boost the accuracy. The future of sentiment analysis consists in overcoming the problems faced and in creating an appropriate method for sentiment analysis with the following additions [11, 14].

***Harnessing the wisdom of crowds*** **-** Let the tool scale and learn, until appropriate data are accessible across different platforms.

***Flexibility*** **-**Ensure sure the method adapts through different time domains and for the years to come due to the changing needs of the brands. Don't stop at positive or negative feelings :- Generate a resource that adds various other feelings besides existing ones. To the current feelings add expectations, reservations, similarities and preferences.

***Detection of biased messages-*** The method must be capable of detecting misleading advertising used to present the brand in a positive manner. Those messages are created by the machine.

***Used for the various platforms-***The terminology the finance industry uses is distinct from that used by the food industry. Analysis of sentiments should adapt to the target industry and produce the production accordingly.

***Handling neutral tweets*** **-** Negative tweets can't be dismissed in real world applications. Neutral perception needs to be given due consideration

***Internationalization*** **-** We just concentrate on English sentences but there are many foreign users on Twitter. Our method should be useful in classifying sentiment into another languages.

***Utilizing emoticon data in the test set-*** Our data on training strips out the emoticons. This implies that if our test results involve an emoticon function, the classifier for a class is not affected. It should be debated, because the emoticon's features are of great significance.

## 11. LIMITATION OF SENTIMENT ANALYSIS

Organizations gather employee input to gain insight into employee attitudes, actions, and motivation and use that insight to develop cultures of committed, motivated, creative, and successful workers. Sentiment analysis, the technology historically used for collecting positive, negative, and neutral categories of such learning, sport specific, human-generated data. This method is too simplistic to reflect human thought by misrepresenting a large portion of data and limits predictive capacity. These computational findings are neither accurate nor actionable. Because of web API's fair use policy, not all tweets were available [12, 13].

- There are no re-tweet counts, regional locations and so on.
- Cannot be used to search and stream APIs on Facebook.

## 12. CONCLUSION

Thus, feeling determination is a crucial step towards converting unstructured content into structured content so that people can see trends and patterns within the content. Therefore, with the growing impact of social media, sentiment analysis lays a powerful foundation in the study of public opinions. This paves the way for brand tracking, mark marketing, consumer preferences research, etc. Therefore, with the help of numerous examples, we explained and explored the value of sentiment analysis for a brand's development on the social media platform. We also spoke about how a company can make effective use of sentiment analysis to revive or enervate its market. The review paper also aims to inform all its readers with the different difficulties they face during their implementation, and also addressed their varied applications. Finally, we discussed the issues faced and the appropriate improvements required to be made in order to create an efficient tool for analyzing sentiments. We demonstrate that the use of emoticons to train data as noisy labels is an effective way to conduct remote supervised learning. Usage of this approach of machine learning (Naive Bayes, maximum entropy classification, and vector machine support) algorithms will achieve high precision. Through twitter tweets have features in comparison to other companies, machine learning have shown strong success in understanding tweeting sensations.

## REFERENCES

[1] Anmol Jain, BhavyaSanghavi, RishabhRathod, KhushaliDeulkar, International Journal of Engineering and Technical Research (IJETR) ISSN: 2321-0869, Volume-3, Issue-2, February 2015.

[2] John N. Fotis, The use of Social Media and its impacts on consumer behavior: the context of holiday travel, May 2015 Bournemouth University.

[3] T. Karnal, G.Seenuvasan2, Sentiment Analysis in E-Commerce using a recommendation system, IJCSMC, August 2017.

[4] Federico Alberto Pozzi, The role of emojis in sentiment analysis, July 2018.

[5] Joel Brynielsson, Fredrik Johansson, Carl Jonson & Anders Westling , Emotion classification of social media posts for estimating people's reactions to communicate alert messages during crises, August 2014.

[6] Subhabrata Mukherjee, Pushpak Bhattacharyya, Sentiment Analysis: A Literature Survey, April 2013.

[7] AbinashTripathy, AnkitAgrawal, Santanu Kumar Rath, Classification of Sentiment Reviews using N-gram Machine Learning Approach, March 2016.

[8] AyushiDalmia, AyushiDalmia, Arpit Kumar Jaiswa, ChinthalaTharun Reddy, Sentiment Analysis in Twitter.

[9] Federico Neri,CarloAliprandi,FedericoCapeci,MontseCuadros, Sentiment Analysis on Social Media, August 2012.

[10] Alec Go, RichaBhayani, Lei Huang, Twitter Sentiment Classification using Distant Supervision, 2016.

[11] Harsh Thakkar and Dhiren Patel Department of Computer Engineering, National Institute of Technology,

Surat - 395007, India, Approaches for Sentiment Analysis on Twitter: A State-of-Art study.

[12] Frank Namugera, Ronald Wesonga &Peter Jehopio , Text mining and determinants of sentiments: Twitter social media usage by traditional media houses in Uganda, April 2019.

[13] van Leeuwen, J. (ed.): Computer Science Today, Recent Trends and Developments. Lecture Notes in Computer Science, Vol. 1000. Springer-Verlag, Berlin Heidelberg New York (1995).

[14] Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs in 3rd edition. Springer-Verlag, Berlin Heidelberg New York (1996).

[15] T. Joachims, Making large-scale support vector machine learning practical in B. Scholkopf, C. J.C.Burges, and A. J. Smola, editors, Advances in kernel methods: support vector learning, pages 169-184.MIT Press, Cambridge, MA, USA