Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 **ISSN: 2321-8363**
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)

Sentiment Analysis on Twitter Data: A Comparative Approach

**Subhadip Chandra[1]; Randrita Sarkar[2]; Sayon Islam[3];
Soham Nandi[4]; Avishto Banerjee[5]; Krishnendu Chatterjee[6]**

[1]Assistant Professor, Information Technology, B. P. Poddar Institute of Management and Technology, Kolkata, West Bengal, India, subhadipchandra@gmail.com
[2,3,4,5,6]Student, Information Technology, B. P. Poddar Institute of Management and Technology, Kolkata, West Bengal, India
DOI: 10.47760/ijcsma.2021.v09i10.001

Abstract
Sentiment analysis is the methodical recognition, extraction, quantification, and learning of affective states and subjective information using natural language processing, text analysis, computational linguistics, and biometrics. People frequently use Twitter, one of numerous popular social media platforms, to convey their thoughts and opinions about a business, a product, or a service. Analysis of tweet sentiments is particularly useful in detecting if people have a good, negative, or neutral opinion. This study assesses public opinion about an individual, activity, commodity, or organization. The Twitter API is utilised in this article to directly get tweets from Twitter and develop a sentiment categorization for the tweets. This paper has used Twitter data for two separate approaches, viz., Lexicon & Machine Learning. Lexicon based approach further categorized in Corpus-based and Dictionary-based. And various Machine learning-based approaches like Support Vector Machine (SVM), Naïve Bayes, Maximum entropy are used to analyse Twitter data. Neural Network (NN), Decision tree-based sentiment analysis is also covered in this research work, to find out better accuracy of the approaches in the various data range. Graphs and confusion matrices are used to visualise the results of the analysis for positive, negative, and neutral remarks regarding their opinions.
Keywords: Sentiment Analysis, Lexicon, Machine Learning, Support Vector Machine, Neural Network.

1. Introduction
Social Media has been the go-to tool to approach the world with your views about political or non-political aspects in the era of the World Wide Web. They choose multiple social media platforms to connect with people like Facebook, Twitter, LinkedIn, and others. They also express their opinions on various local and global issues. It is related to their daily life and their needs to use such online social platforms.
Giving rise to a pile load of data on these web portals suitable for many conditions. Corporate houses and marketing sectors are using these data for further analysis and upgradation of their product quality and services, to reach their consumers with various recommendations and approachable methods. Consumers are also be benefited from the generated reviews and personalized recommendation catering to their requirements.
But the quantity of the data is too huge to search by any individual for further assessment, as more than 80% of the data available over the internet is unstructured. The situation where the process of automation has its role to play to make matters easier.

Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 ISSN: 2321-8363
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)

In today's generation, Twitter is one of the most frequently used social media portal. With about an average of 6000 tweets every second, 350,000 tweets every minute, and 500 million tweets each day, and around 200 billion tweets in a year. There are 330 million monthly active users who use Twitter and tweet their views on any topic they want. Twitter is an American microblogging and social networking site where people interact with other people by using a text of 140 characters which is known as Tweet. Chaturvedi, I., et al. [1] mentioned that the information delivered can be in the form of arguments, news, views, opinions and other several types of sentences. This makes the Twitter data, rich in the text.
Information processing technique from text consists of a few steps, like searching proper keywords, and further analysis of retrieved text to produce useful information or facts. The facts comprising of some purposeful material, but on the other hand of the textual content indicates thematic features. These subjective characteristics are mainly emotions, opinions, views, appraisal, and sentiments. All of these forming Sentiment Analysis (SA). It offers huge challenging scopes to develop a new application by multiple online data sources. For example, the recommendation suggesting the upliftment of things that might be of trending interest, which are often estimated by positive or negative opinions about those items taken under consideration, by making use of sentiment analysis.
Twitter sentiment analysis is a process by which we try to get the mentality, the opinion and the frame of mind of any person, who is behind the tweet by analyzing the Twitter text. Twitter API is a great way to get Twitter data set from the source. After tweet retrieval, we need to pre-process it. It is a data cleaning process, in which unnecessary information is reduced from the tweet text.
Sentiment analysis having few steps. The below figure is depicted to describe the flow of the architecture of Twitter data sentiment analysis.

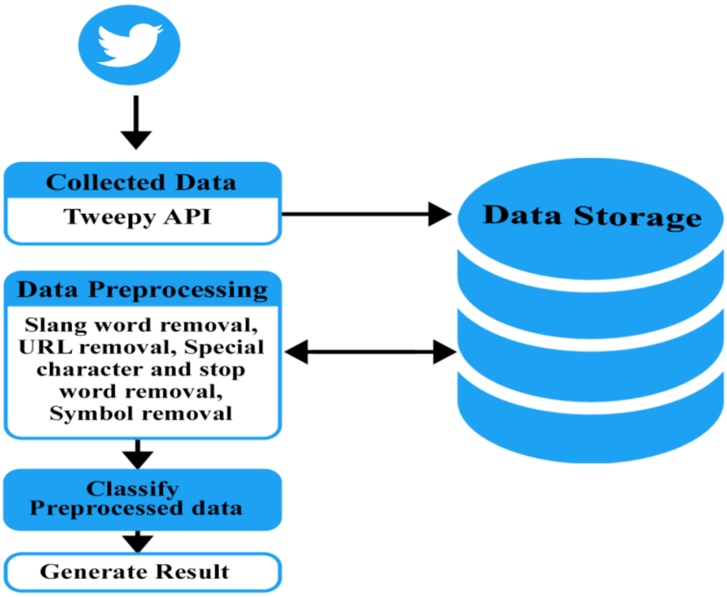


Figure 1: Sentiment analysis workflow of Twitter data

In the next section of this article, the description of sentiment analysis and its categories are described briefly. For sentiment analysis, we fetch raw data from Kaggle (Twitter_Data.csv) based on political data set on Narendra Modi. Next, various pre-processing methods used by the earlier researchers, to clean up the unnecessary part from the raw data text, are described. In the next part of the article, utilization of available classifiers applied by others is described along with their drawbacks and challenges faced. And at the end of this paper, various approaches were tested on a dataset of size 45,000 data. Accuracy results are represented in tabular format, graphically and using a confusion matrix.

2. Sentiment Analysis

People have a wide range of emotions – positive or negative, interested or uninterested and sad or happy. To capture this wide range of emotions, various sentiment analysis models are available.

Let's take a look at some of the most common types of sentiment analysis.

2.1 Fine Grains

This sentiment analysis model aids in the calculation of polarity precision. Sentiment analysis can be performed using the polarity categories of extremely positive, positive, neutral, negative, or very negative. For the study of reviews and ratings, fine-grained sentiment analysis is useful.

On a scale of one to five, one represents a very unfavourable situation and five represents a highly favourable situation. On a scale of one to five.

2.2 Aspect based

The aspect-based analysis goes deeper than fine-grained analysis in determining the overall polarity of your customer evaluations. It assists you in determining which components of the conversation are being discussed.

For example, “In artificial lighting conditions camera suffers.”

Using aspect-based analysis, you can understand that the reviewer has commented on something “negative” about the “camera.”

2.3 Emotion Detection

Emotion detection, as the name implies, aids in the identification of emotions. Anger, sadness, happiness, frustration, fear, concern, panic, and other emotions are examples. Emotion detection systems frequently employ lexicons, which are collections of words that express specific emotions. Robust machine learning (ML) algorithms are also used by some advanced classifiers.

Humans express emotions in a variety of ways, ML is preferred over lexicons. Take, for example, this line: "This product is about to kill me." This statement could be used to show anxiety and panic.

A related phrase, "This product is killing it for me," has a completely different and positive connotation. In the lexicon, however, the term "kill" may be connected with terror or panic. This may lead to false emotion detection.

2.4 Intent Analysis

Companies may save time, money, and effort by accurately determining consumer intent. Businesses frequently find themselves chasing customers who have no intention of purchasing in the near future. This problem can be solved with accurate intent analysis.

The intent analysis can assist you to figure out whether a customer is looking to buy anything or is just looking around.

If a customer is willing to make a purchase, you can monitor them and market to them. You can save time and money by not advertising to customers who aren't ready to buy.

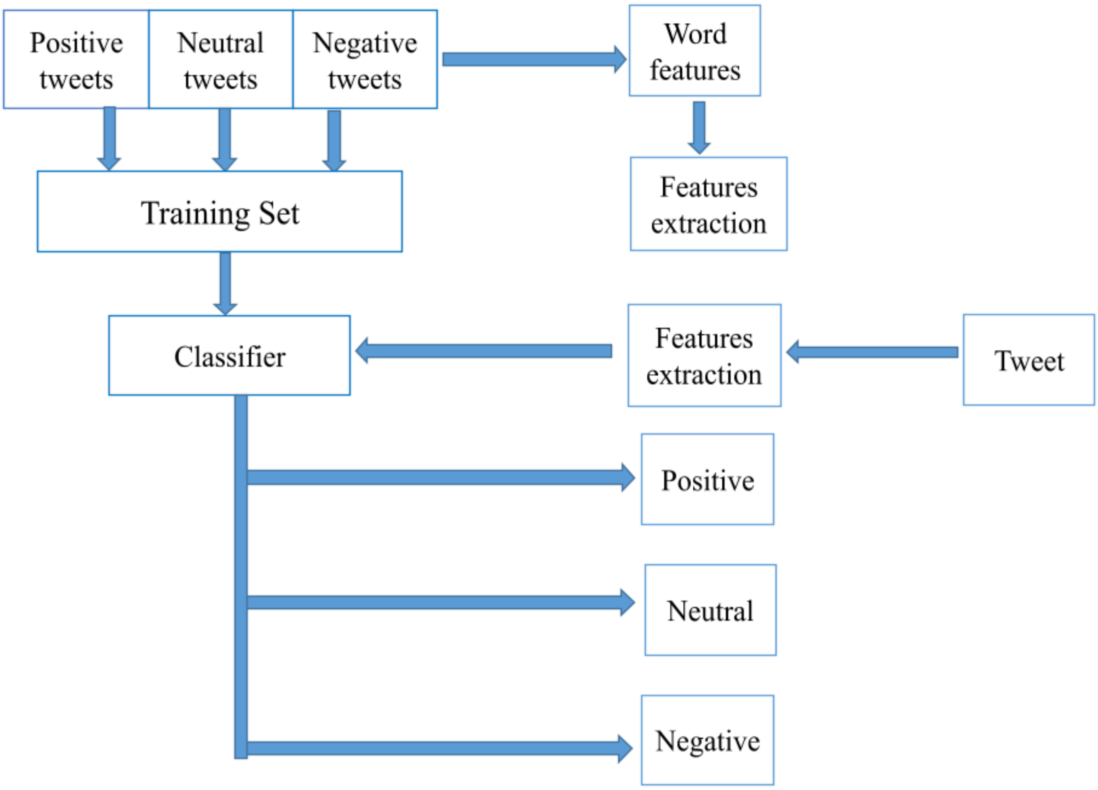


Figure 2. A schematic model of Twitter data sentiment analysis.

3. Data Pre-processing

Data pre-processing can be referred to as the process of initial refinement in an approach similar to the terms of 'Data-Mining'. Thereby cleaning and entered raw data into short, simple, and manageable versions converting to a meaningful and efficient format as a result of removing unnecessary scraps of data. A perfectly pre-processed data aids in the generation of a better result in terms of accuracy and thereby eases the process of classification in the further steps.

In the pre-processing part, Ikoro V, et al. [2] removed the incomplete noisy and inconsistent data from the raw data. (removal of re-tweets → removing URLs, special characters, punctuations, numbers etc. → removing stop words → stemming → tokenization). Agarwal A, et al. [3], introduced two new dictionaries (i) an emoticon dictionary (ii) an acronym dictionary. In the emoticon dictionary, they listed 170 emoticons with their emotional state from Wikipedia. They classify the emoticons as positive, extremely positive, negative, extremely negative and neutral in these categories. They get the acronym dictionary from http://www.noslang.com. They remove the emoticons from their polarity from the dictionary and replace URLs with the tag || U ||, replace usernames or mentions by using || T ||. Update negations with the tag "NOT". They use a stop word dictionary (www.webconfs.com/stop-words.php) to identify and remove the stop words from the tweet.

Kusrini and Masuri, M. [4] apply tokenization to the raw data in the first phase, which means break down the entire tweet into specific separate words. They created a database of slang words with their synonyms, if it matches with the raw data, then it will be replaced with its proper synonym. They also remove stop words, which have no particular meaning and at last, after stemming (converting words to their root words) they complete their pre-processing work and make it ready for further classification.

Sindhu, P. V. et al. [5] also used the tokenization process on the tweet text. The entire text is broken down into pieces like phrase words, keywords, symbols and other elements. Unwanted words like hashtags, mentions, numbers, URLs, other stops words also removed. So the size of the tweet will be shortened and easy to analyze.
Wongkar, M. and Angdresey, A. [6] use a crawler to fetch Twitter data from the web. After that, to break the entire statement into several words, tokenization has been applied & these words are further analyzed and categorized by the basis of their polarity, i.e, they are positive, negative or neutral.
Raw data is highly susceptible to inconsistency and full of redundancies. The polarity will not be fully appropriate. So Gautam, G. and Yadav, D. [7] used a Labelled dataset. It includes removal of punctuations, repeating words, extra words, special characters etc. They extract the adjectives from every tweet, by which they show the polarity of the sentence, whether it's positive, negative or neutral.
All the URLs, hashtags, usernames, mentions etc. firstly removed by Kharde, V. A. and Sonawane, S. [8]. They correct all the spellings and delete repeated characters. They also remove all the punctuations, symbols, digits. Modify all the emoticons with their polarity and expand all the acronyms, like bff = best friends forever as well as remove all non-English tweets.

4. Classification

The classification methods used in the different researches can be further broken into:-

1. Lexicon Based Techniques.
2. Machine Learning Techniques.
3. Hybrid Techniques

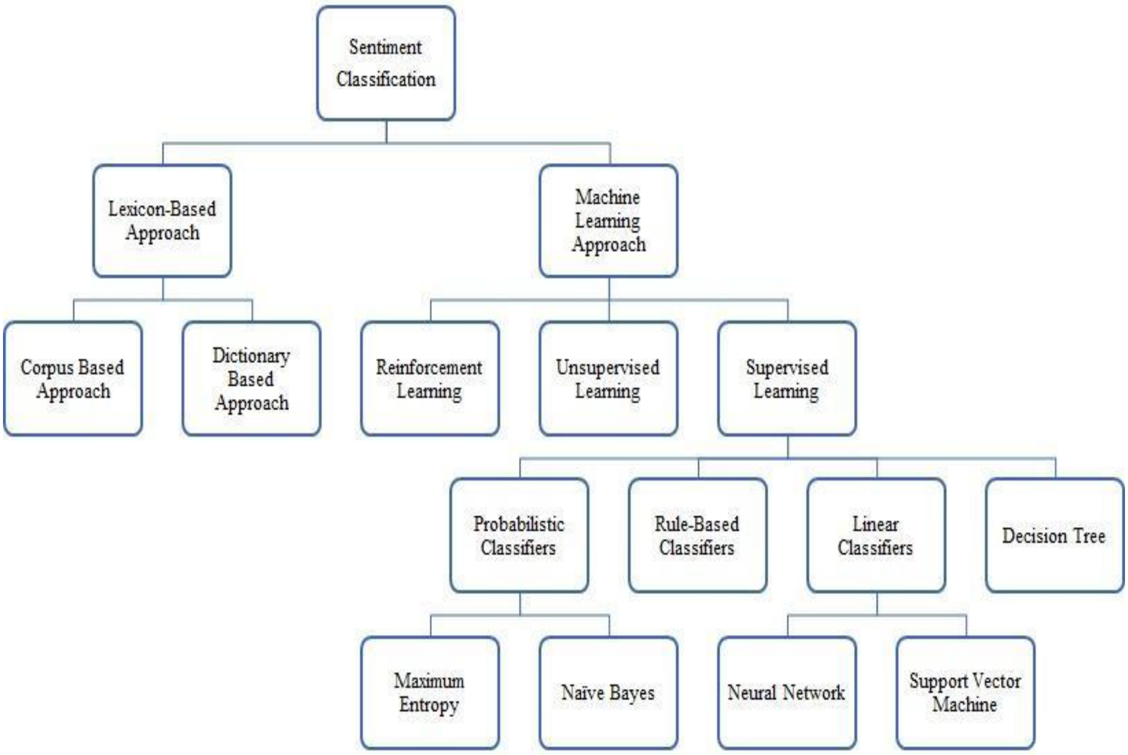


Figure. 3: Classification of Sentiment Analysis Techniques

IJCSMA
Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 **ISSN: 2321-8363**
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)

Lexicon Based Approach

In [4], authors worked on Sentiment Analysis using Lexicon Based Polarity and Polarity Multiplication. They pre-processed the data using Natural Language Processing. Then classified the tweets into three categories: - positive, negative and neutral tweets. In this paper, researchers showed that the accuracy of the Lexicon Based Polarity is less than that of Machine Learning Techniques as used Lexicon Dictionaries suffered by lack of adjectives. They had calculated an accuracy of 68.33% for lexicon Based Approach.

Park, C. W. and Seo, D. R. [9] had worked on Sentiment Analysis of Twitter Corpus Related to Artificial Intelligent Assistants. The classification was done into positive, negative and neutral using the lexicon VADER (Valence Aware Dictionary and sentiment Reasoner). They conducted their analysis through samples using T-test, Kruskal Wallis and Mann-whitney Tests.

Tweets using the lexicon-based dictionary Bag of Words (BOW) was proposed in [5]. If the tweets contained words with the positive sentiment then it is classified as positive, similarly, if the tweets contained negative sentiments it is classified as negative. And if the total score of the positive and negative sentiments in a tweet is 0, then it would be termed as neutral. R language is used for the whole process.

Machine Learning Approaches

Machine Learning Techniques are further sub-divided into Supervised Machine Learning and Unsupervised Machine Learning.

Pak, A. and Paroubek, P. [10] experimented with their data with Multinomial Naive Bayes and SVM classifiers with unigram, bigram and trigram features. They suggested that Multinomial Naive Bayes with Bigram features performed better than the other combinations.

Kharde, V. A. and Sonawane, S. [8] proposed a detailed study on Twitter Sentiment Analysis. A detailed study on SVM, Maximum Entropy and Naive Bayes with their results of accuracy for unigram, bigram and trigram features has been depicted. They recorded that SVM results better than Maximum entropy and Naive Bayes.

Go, A. et, al. [11] used Naive Bayes, Maximum Entropy and SVM with unigrams, bigrams and POS features. It has been found that Maximum Entropy with both unigram and bigram features outperformed the rest.

In [12] authors categorized the Twitter Sentiment Analysis into 7 categories as follows:-

1. Strongly Positive
2. Positive
3. Weakly Positive
4. Neutral
5. Weakly Negative
6. Negative
7. Strongly Negative

They conducted the classification using the Multinomial Naive Bayes Classifier. After calculating the coefficients of the Multinomial Bayes Classifier they fitted it into a linear classifier of the y = mx+b and then calculated the sentiment of the twitters which gives an accuracy of 78.38%.

Gautam, G. and Yadav, D. [7] performed a Sentiment Analysis experiment utilizing a machine learning methodology and semantic analysis. Their research mainly focused on SVM, Naive Bayes and Maximum Entropy. They found that Naive Bayes gave better results than the other two and upon using unigram the result became better. In order to improve the accuracy furthermore, they had subjected the Naive Bayes classifier to WordNet Semantic Analysis which resulted in 89.9% from 88.2%.

Wongkar, M. and Angdresey, A. [6] provided a detailed comparison study between KMN, SVM and Naive Bayes and classify the tweets just between Positive and Negative. Their research has shown that

IJCSMA

Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 **ISSN: 2321-8363**
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)

Naive Bayes performs with 80.90% accuracy, whereas the accuracy of SVM & KNN are 75.58% and 63.99% respectively.

Gamallo, P. and Marcos, G. [13] employed a class of Naive Bayes Classifier to do Twitter Sentiment Analysis, and they found that using a binary classifier that only classifies tweets into positive and negative categories yielded the best results. They also discovered that when a polarity lexicon and multiword are employed instead of the Naïve Bayes classifier, the result is better, with an F-score of 63%.

Dhawan, S. et, al. [14] has used Python to execute Sentiment Analysis of Tweets and has proposed an algorithm for doing so. Using that method, they also able to categorise the tweets as positive, negative, or neutral.

Alsaeedi A, Khan M Z [15] performed thorough research on various methods used in Twitter sentiment analysis. They explained the Naive Bayes Classifier and SVM classifier and then shown a full study of Supervised, unsupervised, ensemble, lexicon-based and hybrid methods used in sentiment analysis of Twitter data in a tabular form along with the accuracy percentage and the added advantage of all the methods.

In paper [16], authors researched on e-commerce based Twitter data for Sentiment Analysis. They compared decision tree, KNN and Naïve Bayes method on this dataset, where decision based approach outperforms over the other approaches.

In paper [17], authors proposed a fine-grained bi-sense emoji embedding technique which produces complex semantic and sentiment information along with attention-based LSTM networks. It depicts from the result that this method's accuracy outmatch others.

5. Challenges and Drawbacks

Subjective and objective parts of the text

According to [8], there are some of the tweets has a particular word that can be treated as an objective as well as subjective. So it may create an issue while analyzing its sentiment. As an example (i) it was a bad day for our team. (ii) Yesterday I saw a fantastic movie named “Bad Boy”. In the first example word ‘bad’ acts as an opinion, and the polarity should be negative. But in the second example, it acts as the subject which holds positive polarity.

Sarcasm Detection

Sarcasm Detection is a unique way to express negative thoughts in a very positive way. Patil, H. P. and Atique, M. [18], detected the sarcasm from the text & identified its actual sentiment as a tough thing to do. For example – “WOW! Great! It is really a genius choice to quit your job in this pandemic situation”. In this sentence ‘Wow’ and ‘great’ these words are extremely positive words, but the actual meaning of the sentence is negative. This actual inner meaning needs to be identified by the analyzer.

Domain Dependence

As per Peddinti, V.M.K. and Chintalapoodi, P. [19], the same phrase or the same word can be treated as positive and sometimes negative. As an example (i) “Your test result is really unpredictable. Nobody knows whether you pass or not”. (ii) “What a match! Ending was really unpredictable”. In the first example, the word unpredictable acts as a negative sentiment whereas in the second example the same word holds a positive sentiment.

Internationalization

Pan, S. J. et al. [20] and Pak, A. and Paroubek, P. [10], shown that all sentiment analysis work based on English tweet. But Twitter has millions of users worldwide. So it should be solved anyhow for a better result for other languages. However, in paper [3] they show that raw tweet can be translated into English by using Google translator.

Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 **ISSN: 2321-8363**
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)

Grammatically Incorrect

The setback of detecting perfect sentiments of a text is curbed in terms of results by the infiltration of grammatical errors and incorrect spellings in the data obtained in the form of a textual tweet. Making the process bear less fruit.

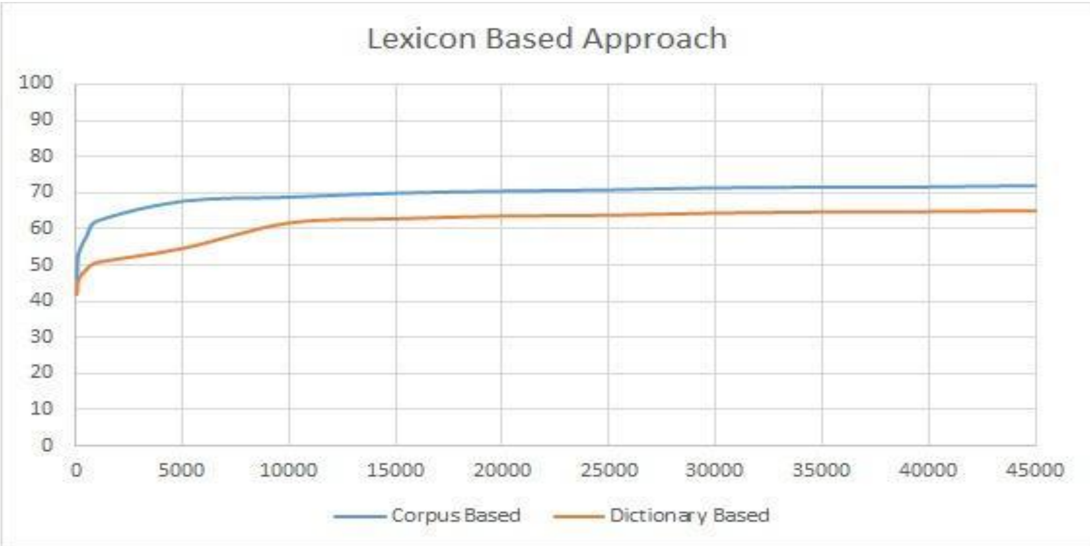
6. Result Analysis

Sentiment analysis was done for the Twitter dataset in two separate part. In the first part, we have applied Lexicon based approach on the data for both Corpus-based and Dictionary-based. From the table, it is understandable that the Corpus-based approach is more powerful than the Dictionary-based model. Similarly, the same dataset is used by multiple common Machine Learning-based approaches. In all the cases, 70% of total data are used for training purpose and the rest 30% for testing purpose. Table 1 reflects that Neural Network based approach is best over all other techniques, but it provides better result when the size of the data amount is high. For a small range of data, SVM & Naïve Bayes gives better result, though it cannot reach the accuracy given by the decision tree-based approach for a larger dataset.

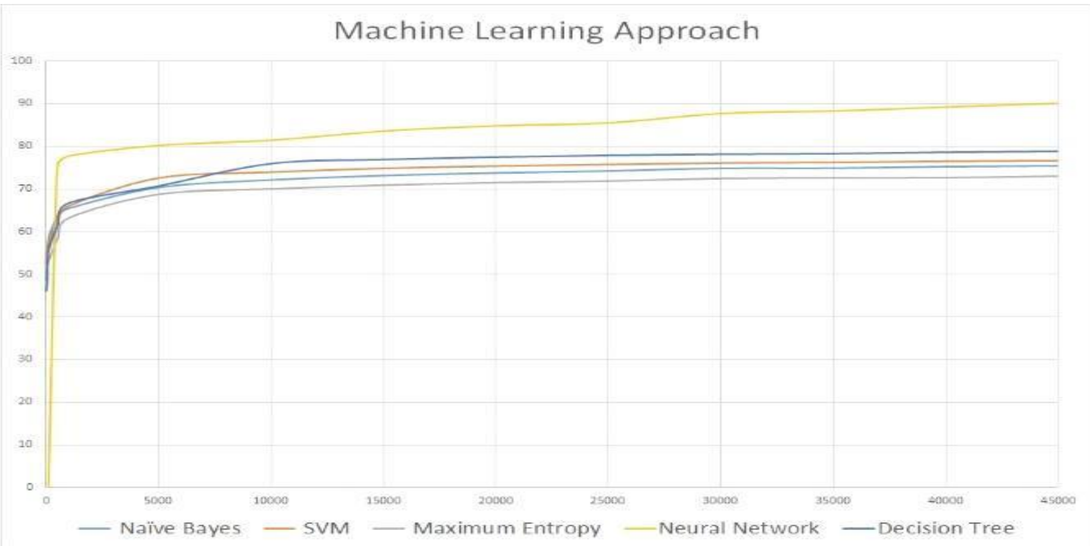
Table 1: Accuracy metric on Twitter Sentiment Analysis

Approach ⇨	Lexicon-Based Approach		Machine Learning Approach				
Data range ⇩	Corpus Based	Dictionary Based	Naïve Bayes	SVM	Maximum Entropy	Decision Tree	Neural Network
10	0.4596	0.4179	0.4869	0.5254	0.5125	0.4615	-
50	0.4829	0.4287	0.5284	0.5505	0.5404	0.4896	-
100	0.5294	0.4593	0.5815	0.5697	0.5491	0.5538	-
500	0.5811	0.4894	0.6343	0.6261	0.6019	0.6129	0.7536
1000	0.6218	0.5075	0.6543	0.6604	0.6412	0.6665	0.7767
5000	0.6763	0.5465	0.7034	0.7262	0.7066	0.7067	0.8015
10000	0.6875	0.6177	0.7211	0.7398	0.7192	0.7591	0.8136
15000	0.6985	0.6291	0.7313	0.7489	0.7281	0.7686	0.8352
20000	0.7042	0.6367	0.7374	0.7542	0.7342	0.7743	0.8475
25000	0.7075	0.6389	0.7421	0.7581	0.7377	0.7784	0.8543
30000	0.7132	0.6451	0.7483	0.7613	0.7438	0.7811	0.8765
35000	0.7152	0.6483	0.7487	0.7628	0.7452	0.7823	0.8821
40000	0.7159	0.6492	0.7523	0.7655	0.7456	0.7858	0.8915
45000	0.7191	0.6514	0.7541	0.7668	0.7493	0.7879	0.9001

IJCSMA
Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 ISSN: 2321-8363
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)



A

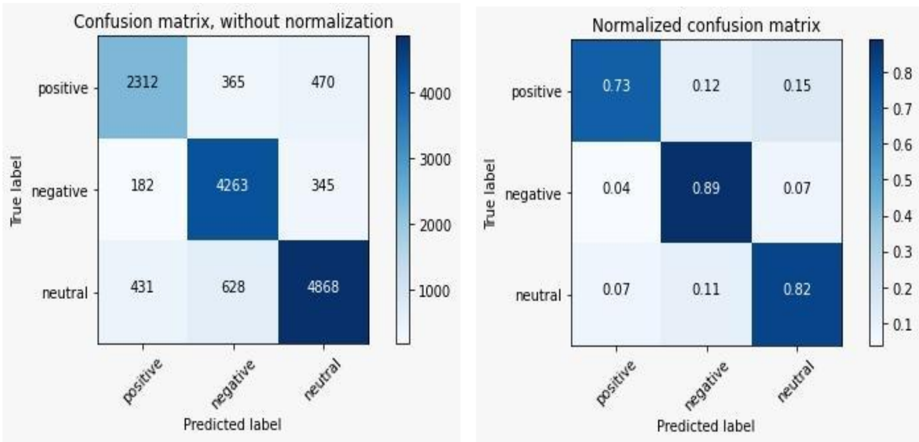


B

Figure 4: (A) Graphical representation of accuracy for Corpus-based and Dictionary-based approach. (B) Graphical representation of accuracy for Machine Learning based approaches (Naïve Bayes, SVM, Maximum Entropy, Neural Network and Decision Tree).

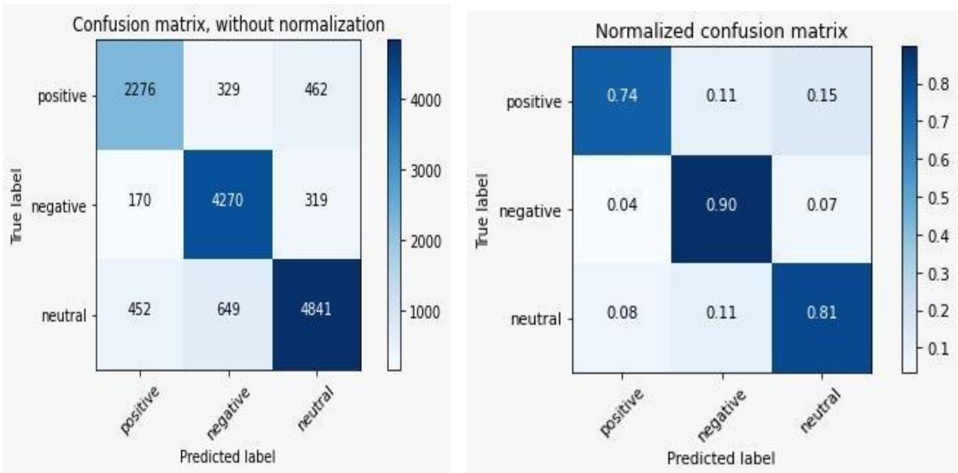
Figure 4(A) shows the graphical representation of accuracy for Corpus-based & Dictionary based approach and figure 4(B) indicates the graphical representation of accuracy for different approaches namely Naive Bayes, SVM, Maximum Entropy, Neural Network Decision Tree etc.

Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 **ISSN: 2321-8363**
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)



A B

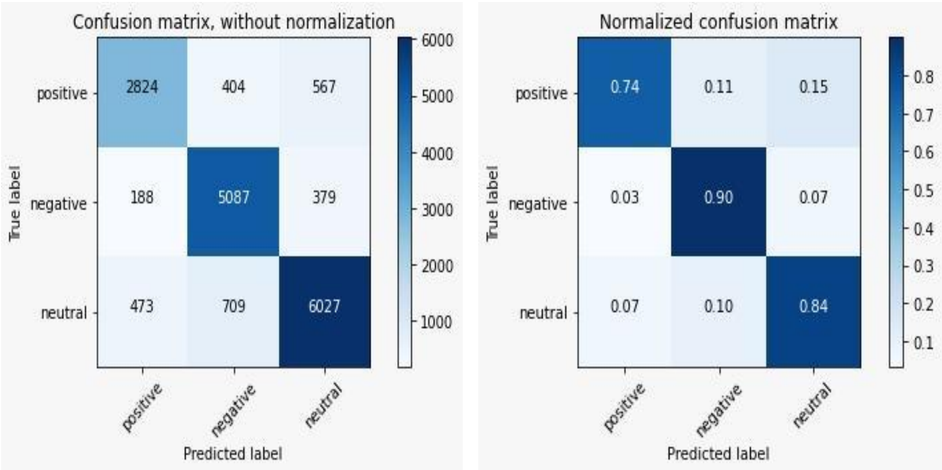
Figure. 5: Confusion matrix for Corpus based approach.



A B

Figure. 6: Confusion matrix for SVM based approach.

Subhadip Chandra *et al*, International Journal of Computer Science and Mobile Applications,
Vol.9 Issue. 10, October- 2021, pg. 01-12 **ISSN: 2321-8363**
Impact Factor: 6.037
(An Open Accessible, Fully Refereed and Peer Reviewed Journal)



A **B**

Figure. 7: Confusion matrix for Neural Network-based approach.

In the above figures (5 – 7), part A represents the confusion matrix for Corpus-based, SVM & Neural Network-based approaches. In these above figures, rows are indicating the true label of the data whereas column headings represents the predicted positive, negative and neutral labels. Part B of the figure indicates the normalized values of part A.

7. Conclusion

The frustration, happiness, happening and updates amount up to the maximum on social media platforms. Making it available as information in the form of an open-source. The method and topic portrayed through the representation of this review paper brings forward the ideation and a detailed overview of a comparative approach to the process of Sentiment Analysis.

Giving vivid descriptions of Sentiment analysis, the techniques and processes involved in the procedure of procuring results as well as the classification models with their respective accuracies. The numerable challenges faced in the evolution of a successful mechanism as well as the unsolved drawbacks are discussed thoroughly. With models based on both Lexicon and Machine learning approaches as parameters for comparison it is rightly concluded that Machine Learning models tend to have a greater accuracy score as compared to other models and methods.

References

[1] Chaturvedi, Iti, Erik Cambria, Roy E. Welsch, and Francisco Herrera. "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges." Information Fusion 44 (2018): 65-77.

[2] Ikoro, Victoria, Maria Sharmina, Khaleel Malik, and Riza Batista-Navarro. "Analyzing sentiments expressed on twitter by UK energy company consumers." In 2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 95-98. IEEE, 2018.

[3] Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca J. Passonneau. "Sentiment analysis of twitter data." In Proceedings of the workshop on language in social media (LSM 2011), pp. 30-38. 2011.

[4] Mashuri, Mochamad. "Sentiment Analysis In Twitter Using Lexicon Based and Polarity Multiplication." In 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT), pp. 365-368. IEEE, 2019.

[5] Prakruthi, V., D. Sindhu, and S. Anupama Kumar. "Real time sentiment analysis of Twitter posts." In 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), pp. 29-34. IEEE, 2018.

[6] Wongkar, Meylan, and Apriandy Angdresey. "Sentiment analysis using Naive Bayes Algorithm of the data crawler: Twitter." In 2019 Fourth International Conference on Informatics and Computing (ICIC), pp. 1-5. IEEE, 2019.

[7] Gautam, Geetika, and Divakar Yadav. "Sentiment analysis of twitter data using machine learning approaches and semantic analysis." In 2014 Seventh International Conference on Contemporary Computing (IC3), pp. 437-442. IEEE, 2014.

[8] Kharde, Vishal, and Prof Sonawane. "Sentiment analysis of twitter data: a survey of techniques." arXiv preprint arXiv:1601.06971 (2016).

[9] Park, Chae Won, and Dae Ryong Seo. "Sentiment analysis of Twitter corpus related to artificial intelligence assistants." In 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), pp. 495-498. IEEE, 2018.

[10] Pak, Alexander, and Patrick Paroubek. "Twitter as a corpus for sentiment analysis and opinion mining." In LREc, vol. 10, no. 2010, pp. 1320-1326. 2010.

[11] Go, Alec, Richa Bhayani, and Lei Huang. "Twitter sentiment classification using distant supervision." CS224N project report, Stanford 1, no. 12 (2009): 2009.

[12]Reddy, A. B., Vasundhara, D. N. and Subhash, P. Sentiment Research on Twitter Data. International Journal of Recent Technology and Engineering. 8(2019): 1068-1070.

[13] Gamallo, Pablo, and Marcos Garcia. "Citius: A naivebayes strategy for sentiment analysis on english tweets." In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014. 2014.

[14] Dhawan, Sanjeev, Kulvinder Singh, and Priyanka Chauhan. "Sentiment Analysis of Twitter Data in Online Social Network." In 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC), pp. 255-259. IEEE, 2019.

[15] Alsaeedi, Abdullah, and Mohammad Zubair Khan. "A study on sentiment analysis techniques of Twitter data." International Journal of Advanced Computer Science and Applications 10, no. 2 (2019): 361-374.

[16] Bayhaqy, Achmad, Sfenrianto Sfenrianto, Kaman Nainggolan, and Emil R. Kaburuan. "Sentiment analysis about E-commerce from tweets using decision tree, K-nearest neighbor, and naïve bayes." In 2018 international conference on orange technologies (ICOT), pp. 1-6. IEEE, 2018.

[17] Chen, Yuxiao, Jianbo Yuan, Quanzeng You, and Jiebo Luo. "Twitter sentiment analysis via bi-sense emoji embedding and attention-based LSTM." In Proceedings of the 26th ACM international conference on Multimedia, pp. 117-125. 2018.

[18] Patil, Harshali P., and Mohammad Atique. "Sentiment analysis for social media: a survey." In 2015 2nd International Conference on Information Science and Security (ICISS), pp. 1-4. IEEE, 2015.

[19] Peddinti, Viswa Mani Kiran, and Prakriti Chintalapoodi. "Domain adaptation in sentiment analysis of twitter." In Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence. 2011.

[20] Pan, Sinno Jialin, Xiaochuan Ni, Jian-Tao Sun, Qiang Yang, and Zheng Chen. "Cross-domain sentiment classification via spectral feature alignment." In Proceedings of the 19th international conference on World wide web, pp. 751-760. 2010.

A Brief Author Biography

Subhadip Chandra – is currently working as an Assistant Professor in the Department of Information Technology at B. P. Poddar Institute of Management & Technology. He received his Master of Technology in Information Technology degree in 2009. His research interests include Machine Learning, Deep Learning, Computer vision. He has published his research work in International Conferences and book chapters.

Randrita Sarkar – is currently working a student in the Department of Information Technology at B. P. Poddar Institute of Management & Technology.

Sayon Islam – is currently working a student in the Department of Information Technology at B. P. Poddar Institute of Management & Technology.

Soham Nandi – is currently working a student in the Department of Information Technology at B. P. Poddar Institute of Management & Technology.

Avishto Banerjee – is currently working a student in the Department of Information Technology at B. P. Poddar Institute of Management & Technology.

Krishnendu Chatterjee – is currently working a student in the Department of Information Technology at B. P. Poddar Institute of Management & Technology.