

## Sentiment Analysis with Tweets Behaviour in Twitter Streaming API

Kuldeep Chouhan<sup>1</sup>, Mukesh Yadav<sup>2</sup>, Ranjeet Kumar Rout<sup>3</sup>, Kshira Sagar Sahoo<sup>4</sup>, NZ Jhanjhi<sup>5,\*</sup>, Mehedi Masud<sup>6</sup> and Sultan Aljahdali<sup>6</sup>

<sup>1</sup>Computer Science and Engineering, I. T. S Engineering College, Greater Noida, 201310, India

<sup>2</sup>Computer Science and Engineering, DPG Institute of Technology and Management, Gurgaon, 122004, India

<sup>3</sup>Computer Science and Engineering, National Institute of Technology Srinagar, Jammu and Kashmir, India

<sup>4</sup>Department of Computer Science and Engineering, SRM University, Amaravati, Andhra Pradesh, 522240, India

<sup>5</sup>School of Computer Science SCS, Taylor's University, Subang Jaya, 47500, Malaysia

<sup>6</sup>Department of Computer Science, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif, 21944, Saudi Arabia

\*Corresponding Author: NZ Jhanjhi. Email: noorzaman.jhanjhi@taylors.edu.my

Received: 03 April 2022; Accepted: 08 June 2022

**Abstract:** Twitter is a radiant platform with a quick and effective technique to analyze users' perceptions of activities on social media. Many researchers and industry experts show their attention to Twitter sentiment analysis to recognize the stakeholder group. The sentiment analysis needs an advanced level of approaches including adoption to encompass data sentiment analysis and various machine learning tools. An assessment of sentiment analysis in multiple fields that affect their elevations among the people in real-time by using Naive Bayes and Support Vector Machine (SVM). This paper focused on analysing the distinguished sentiment techniques in tweets behaviour datasets for various spheres such as healthcare, behaviour estimation, etc. In addition, the results in this work explore and validate the statistical machine learning classifiers that provide the accuracy percentages attained in terms of positive, negative and neutral tweets. In this work, we obligated Twitter Application Programming Interface (API) account and programmed in python for sentiment analysis approach for the computational measure of user's perceptions that extract a massive number of tweets and provide market value to the Twitter account proprietor. To distinguish the results in terms of the performance evaluation, an error analysis investigates the features of various stakeholders comprising social media analytics researchers, Natural Language Processing (NLP) developers, engineering managers and experts involved to have a decision-making approach.

**Keywords:** Machine learning; Naive Bayes; natural language processing; sentiment analysis; social media analytics; support vector machine; Twitter application programming interface



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1 Introduction

Sentiment analysis, also known as opinion mining, indicates the utilization of Natural Language Processing (NLP), text mining and computational linguistics to recognize and evaluate emotional states [1]. Sentiment analysis is an approach to assess a user in correspondence with a customer support. The current analytical tools are competent to squeeze with incredible measurements of customer criticism constantly and accurately. The sentiment analysis efforts to control the temper of a speaker or opponent might be an assessment of enthusiastic responses. Sentiment mining is a task that analyzes a widespread number of archives to accumulate the sentiments of remarks postured [2] and integrates various approaches, including computational and information retrieval [3].

Twitter is a social network where rapid user sentiments have leading social media analytical frequencies. It generates billions of Application Programming Interface (API) screams where Twitter sentiments have a massive extent to comprehend user's perceptions and provide early warnings occurrence [4,5]. Twitter key sentiment inputs have various sentiment divergence (positive, negative or neutral) that allows users to communicate with other users and share their ideas, feelings, etc., with steadiness and approaches [6]. A posted tweet can be visible by default to the users unless they revolute their confidentialities to make it understandable by the followers on Twitter incline [7]. The eminence of tweet sentiment polarity classifications has information and emerging analytics applications [8]. In this research work, a novel approach is proposed to represent expressions in the semantic direction, based on data abstraction, pre-processing of extracted data and classification on Twitter.

Twitter sentiment analysis tools developed in social media include API-based tools subscription that comprises feasible stand-alone tools and integrated trained tools [9]. API-based tools can be downloaded as system applications appropriate to unlabelled forms to evaluate Chatterbox, Sentiment, TextAnalytics, Anonymous, SentimentAnalyzer, TextProcessing, Machine Learning Analyzer, etc. through API [10]. Sentiment analysis uses an accomplished machine learning classifier created on a vast Twitter *corpus* of positive and negative tweets [11], which are based on the presence of emoticons.

### *Objectives and Motivation*

Research is structured speculation of a delinquent to gain a solution to a real problem. In this work, the objective is vibrant and transitory, providing direction to examine the variables. The research objective of this work is specified as outlined to achieve the work that includes,

- Study sentiment analysis.
- Analyze the sentiment techniques to understand the user's behaviour that can decide the mental situation going on.
- Find the accuracy of the user's tweets status and measure the perception using python programming.
- Awareness of how to deal with tweet data.
- Develop a computational tool for sentiment analysis.
- Relate the sentiment analysis knowledge with real-time life experience.
- Prospect to study advanced properties that may respond something triumph to do definite fertile to the researchers.

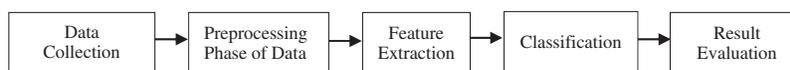
Nowadays, the furthest distinguished area among the people is to analyze user's sentiments can have conspicuous study to have attracted the future planning and creation of innovative concepts. The political ideas or campaigns, people's opinions or perceptions on the social media are the informal approaches to accessing the sentiments through the tweets, which can have predictions for various events on social media.

## 2 Related Work

Yadav et al. [12] presented various techniques to analyze the behaviour of the tweets as positive, negative or neutral for the evaluation and satisfy the sentiments. Asghar et al. [13] presented a unified framework for classifying tweets to enhance the performance of Twitter-based sentiment analysis that classifies precise accuracy. Abbasi et al. [14] proposed using sentiment analysis to classify web environment opinions in various languages to predict the sentiments. Anber et al. [15] presented information exploration to analyze different hashtags, Twitter's network topology and identification of influence sentiment. Kumar et al. [16] proposed approaches for binding Twitter data to discover solutions to complicate searches. Da Silva et al. [17] presented the uses of sentiment analysis to search the products to monitor the public sentiment toward their brands. Devika [18] presented a sentiment analysis tool to extract the information to competence the collective sentiments of the reviews. Chenaghlu et al. [19] presented information epoch consequences of data in various forms and modalities to discover instructive information. Ho et al. [20] proposed a data analytic approach consisting of machine learning algorithms and combinatorial fusion. Tan et al. [21] observed the emergent issues within the sentiment variation related to the causes behind the deviations. Dubey et al. [22] presented text mining-based tweets generated for various personalities on different opinions and issues. Fu et al. [23] proposed a novel crawler that incorporates sentiment information to enhance the collection of rich web content for a particular issue. Marcel et al. [24] presented the performance of a social network that allows users to distribute real-time hypermedia and data to localize prospective attention. Mostafa et al. [25] presented machine learning classifier algorithms are applied on crawled Twitter data to encode techniques to accomplish the accuracy. Hassan et al. [26] proposed a sentiment to elaborate bootstrapping ensemble and depictive abundance issues. Bermingham et al. [27] proposed classifying the sentiment in microblogs as more relaxed than in blogs and relating to supervised learning research. Goodchild et al. [28] presented that data quality generates trusted data.

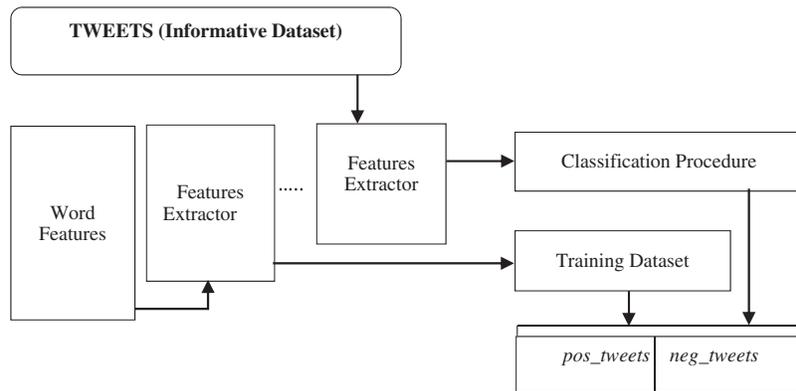
## 3 Proposed System

In this study, the Twitter dataset collected tweets to anticipate the valuation to find the characteristics of tweets. The tweet collection phase as an input dataset uses the beginning of tweets that have to pre-process to extract the phases to receive estimation and analysis of the data classification. The framework of the tweet data analysis is shown in Fig. 1 represents the process from data collection to evaluate the results.



**Figure 1:** Framework of Tweet analysis mechanism

Tweets through web-based networking are imperative opinion can break down individuals' sensitivity and sentiments by such type of social media. In this work, tweet datasets can compose the data and be detached into the testing and training to the datasets that measure the performance calculation of the developed model. This model can categorize newly generated tweets that can be elevated from the Twitter API development account into the tweets to categorize the modules [18], as shown in the proposed flowchart in Fig. 2.



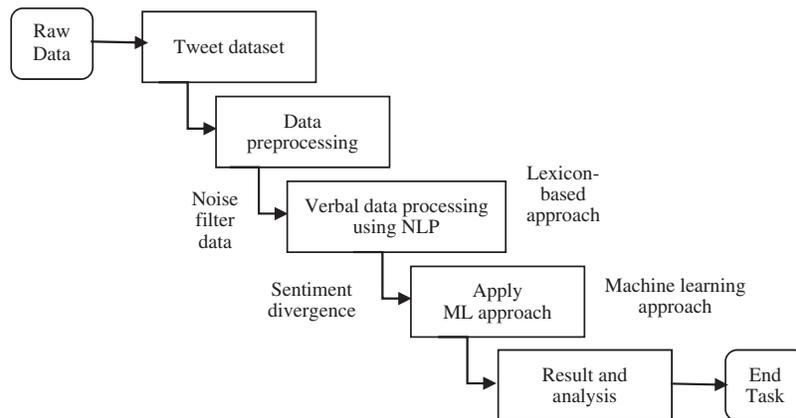
**Figure 2:** Proposed flowchart based on a generated tweets dataset

**4 Methodology**

The tweet preprocessing filters the dataset and reduces the clutter from it. The vital data processing algorithm of verbal uses NLP that includes;

- (i) Assigned sentiment variance to the tweets [19] with a lexicon-based approach [20,21] and
- (ii) The dataset is trained such as Naive Bayes and Support Vector Machine (SVM) to measure the accuracy [22].

The subsequent collectives are having lexicon for sentiment analysis to represent a framework that explains the procedure from the level of raw data which is capable to analyze the tweets at various levels and apply a machine learning approach to provide appropriate results at the end as shown in Fig. 3.

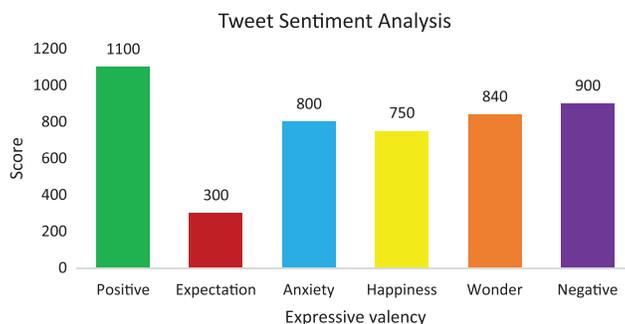


**Figure 3:** Framework of sentiment analysis using various approaches

**5 Sentiment Analysis on Twitter**

This study analyses users’ estimation, sentiment, assessment, impudence, and emotions from transcribed language [23] with the help of NLP and artificial intelligence analysis, categorized as positive, negative, expectation, anxiety, happiness, and wonder, as shown in Fig. 4.

Here, the proposed text dataset notch for positive tweets are reasonably developed other sentiments whereas anticipated as negative and cannot be categorized known as neutral tweets. People’s views on the level of arrangement on a particular topic, product, service or even political elections are imminent under sentiment analysis [23]. The increasing distinctiveness of social media has become informal and more accessible to crawl the user’s feedback [24,25] and analyze their sentiments as positive, negative or neutral [26–28] tweet approaches. There are two specific approaches are employed to acquire the sentiment analysis comprises,



**Figure 4:** Various emocation using Twitter sentiment analysis

- (i) *NLP Approach:* This approach is the interface between computers and human languages to assess users’ sentiment. In this work, various features are used effectively to interpret the sentiments of users where a sentence has two different meanings in which one entity is interpreted with positive and another entity is interpreted with a negative sentiment tag showing a tweet has a varied sentiment.
- (ii) *Machine Learning Approach:* Machine learning regulates the formation of the sentiment concept to study algorithms [29,30] and learn from various datasets to analyze [31] the sentiment of Twitter users [32].

### 5.1 Semantic Features for Sentiment Analysis

The semantic feures integrate into the sentiment analysis method that extracts from tweets to measure the overall correlation of entities along with sentiment polarity [33,34]. It helps to recognize tweet sentiment containing any of the entities which are incapable of looking in the training dataset and composed with their possibilities in positive and negative tweets [35]. Various methods are possible with sentiment analysis to Naive Bayes probabilistic classifier [36,37] where a scheme of a sentiment class  $c$  can be classified with a specified tweet  $tw$  which is calculated and given as,

$$\hat{C} = \arg \max_{c \in C} P(c|tw) = \arg \max_{c \in C} P(c) \prod_{1 \leq i \leq N_{tw}} P(tw_i|c) \tag{1}$$

where,

$N_{tw}$  signifies number of tweet words,

$P(c)$  denotes a tweet presence in class  $c$ ,

$P(tw_i|c)$  is the preventive prospect of the tweet word  $tw_i$  (tweet class as  $c$ ).

In polynomial Naive Bayes,  $P(c)$  estimated by,

$$P(c) = N_c = N$$

where,

$N_c$  represents the number of tweets and

$N$  denotes a comprehensive number of tweets.

$P(tw_i|c)$  is appraised with Laplace flattening which is given as,

$$P(tw|c) = \frac{N(tw, c) + 1}{\sum_{tw' \in L} N(tw'|c) + |L|} \quad (2)$$

where,  $N(tw, c)$  denotes the frequency of tweet word  $tw$  with dataset training and  $|L|$  represents the number of word counting in the lexis.

### 5.2 Tweet Classification Framework

Twitter sentiments framework that process with contemplates tweet collection and tweet analysis that is forwarded by users (as hashtag) to swift their sentiments [38] and accumulates the retrieved tweets to pre-process the dataset. To understand the sentiments of the tweets at an optimum level, need to find the maximum accuracy for polarity and bias calculator to use sentiment analyzers that include,

- (i) **Textblob**: The essential features of NLP are used to analyze the polarity and bias calculation of tweets [39].
- (ii) **SentiWordNet**: Accessible analyzer that is extracted from a wordnet database of the English language that comprehends the sentiments.
- (iii) **Word Sense Disambiguation**: Able to find the accurate word sense within a definite framework for vibrant senates for the polarity calculation. It is effective other than NLP framework recognition functions due to a problem identification discrepancy.

Each classifier is enabled to obtain the training dataset which is tested on a classification model using Naive Bayes and SVM classifier [40] to provide an optimal accuracy [41]. The apparatuses and procedures maintained in sentiment analysis are represented as a framework shown in Figs. 5a and 5b.

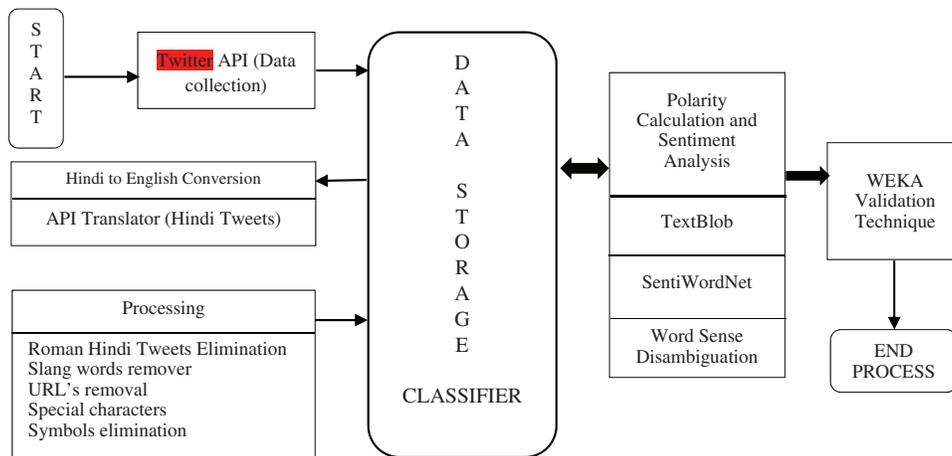
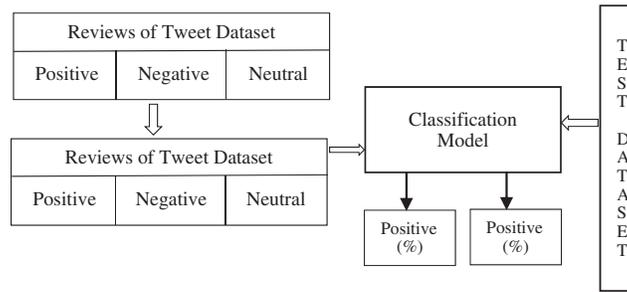


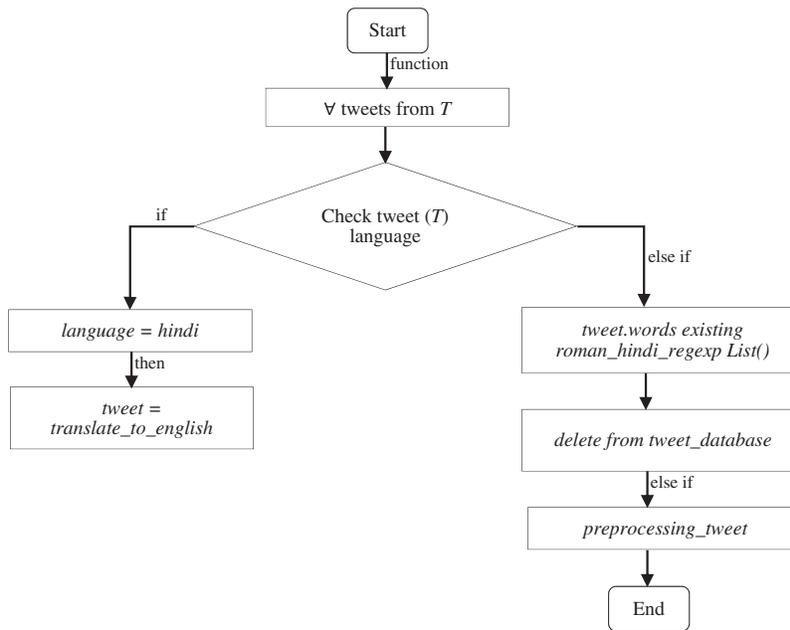
Figure 5: (Continued)



**Figure 5:** (a) Twitter API classification using data classifier with sentiment analysis (b) Framework of tweet classification and sentiment analysis

**5.3 Polarity Calculation and Sentiment Analysis**

In social media platforms, the sentiment analysis has precious sensitivities with a level of the dataset that characterizes in unstructured performance of sense with sentiments. Though using polarity, each tweet allocates a notch (-1 to 1) with words in sentences representing a positive sentiment while the ‘0’ value is measured as a neutral sentiment. While adding the Hindi tweets converts into English accomplishes additional tasks for pre-processing as flowchart is shown in Fig. 6.



**Figure 6:** Tweet conversion and tasks for pre-processing approach

Each sentiment analyzer uses bias and polarity of processed tweets (sentiment of polarity such as  $S_p$ ) with the Python code to determine the sentiment as given in Tab. 1.

where,

**PT:** Pre-processed tweets from the database.

**B:** Biased control of the statements.

**P1 and S1:**  $S_p$  computed by TextBlob.

**P2 and S2:**  $S_p$  computed by SentiWordNet.

**P3 and S3:**  $S_p$  computed by WSD reference library.

**Table 1:** Sentiment analysis using bias and polarity with various attributes

Input attributes	Output attributes
B	(SO $\geq$ 0.5 && (Positive: P1 > 0, Neutral: P1 = 0, Negative: P1 < 0)
P1	(SO $\geq$ 0.5 && (Positive: P2 > 0, Neutral: P2 = 0, Negative: P2 < 0)
P2	(SO $\geq$ 0.5 && (Positive: P3 > 0, Neutral: P3 = 0, Negative: P3 < 0)
P3	

## 6 Various Approached Used in Sentiment Analysis

This work regulates the solution of tweet segmentation in positive or negative classification which is essential to applying a lexicon-based approach along with machine learning for sentiment analysis on Twitter.

### 6.1 Lexicon-Based Approach

This approach is the sentiment of a text which is resolute by its leading sentiment words that contain phrases in a verbal situation which has different approaches includes;

- (i) *Simple Word Count (SWC)*: A sentiment lexicon  $l$ , a document  $d = \{w_1, w_2, \dots, w_n\}$  where  $w_i$  ( $1 \leq i \leq n$ ) represents the  $i^{\text{th}}$  word in the document has the existence of positive words in  $d$  and negative ( $l, d$ ) represent the existence of words in  $d$  that is calculated as,

$$\text{sum}(l, d) = \text{pos}(l, d) - \text{neg}(l, d) \quad (3)$$

The sentiment alignment of  $d$  (where, 1 denotes  $\rightarrow$  “positive” and  $-1$  denotes  $\rightarrow$  “negative”) can be defined as,

$$s_{SWC}(l, d) = \begin{cases} 1, & \text{sum}(l, d) > 0 \\ -1, & \text{sum}(l, d) < 0 \\ RC(d), & \text{otherwise} \end{cases} \quad (4)$$

To find appropriate the problem, necessary to allocate an arbitrary label to  $d$  when the sentiment word sum is ‘0’.

- (ii) *Feature Counting (FC)*: A sentiment lexicon  $l$  having enumeration function of a feature  $f$  plots a feature to a real-valued number that is described as,

$$\text{score}(l, f) = \begin{cases} > 0, & \text{if positive,} \\ = 0, & \text{if neutral,} \\ <, & \text{if negative} \end{cases} \quad (5)$$

The enumeration function not only outlines the polarity of the sentiment features, but also has a degree of freedom. The data with  $d = \{f_1, f_2, \dots, f_n\}$  where,  $f_1 = \{(1 \leq i \leq n)\}$  represents the  $i^{\text{th}}$  feature in  $d$  and comprehensive sentiment sum of  $d$  can be calculated as given in equation as,

$$sum(l, d) = \sum_{i=1}^n score(l, f_i) \tag{6}$$

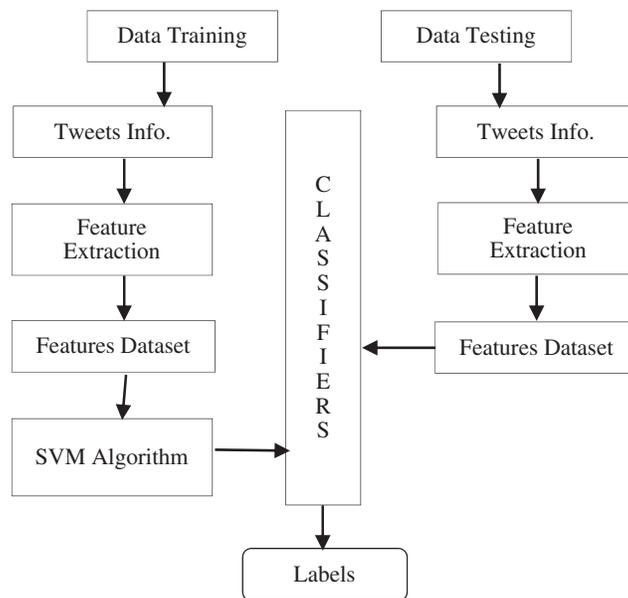
By choosing an edge  $\delta \approx 0$ , the sentiment positioning of  $d$  can be defined as,

$$sf_c(d) = \begin{cases} 1, & sum(l, d) > \delta, \\ RC(d), & -|\delta| \leq sum(l, d) \leq |\delta|, \\ -1, & sum(l, d) < \delta \end{cases} \tag{7}$$

To apt the problem,  $d$  is the sum value between threshold intervals where the enumeration process extracts the features to describe an effective enumeration function and find a precise edge of accuracy.

### 6.2 Twitter Sentiment Analysis Using Supervised Approaches

In this approach, the labeled datasets are dependent and use the machine learning process during the training to train the models and achieved considerable results. Twitter sentiment analyzes the performance to classify [42] and train data to extract the feature sets [43]. The supervised machine learning approach for Twitter sentiment analysis is shown in Fig. 7 stipulates the data training and data testing procedure along with labels as classifiers.



**Figure 7:** Supervised machine learning algorithms with tweet sentiment analysis

Tweet data uses prediction algorithms (such as SVM and Bayesian classifier) to resolve the sentiment polarity of tweets’ performance. Binary classification of the dataset is having sentiment analysis problem to organize the dataset into positive and negative tweets based on the regulation approach with emoticons which is applied in Twitter API to have tweet citations.

## 7 Results and Discussion

### 7.1 Polarity Class

Polarity state defines the positivity or negativity of text is focused on emotions expressed in the statement. It depends on the assessment which can be easily counted and applied Naive Bayes algorithm with the positive and negative words. A sentiment analysis task depends on the words and divides the text into a single word so that evaluation or the computation makes informal then only, Naive Bayes algorithm search for words that can express sentiment in a tweet. The Naive Bayes algorithm works with a bag of words and prerequisite to find out all the positive or negative words which are illustrated and listed below,

Positive words	Prosper, flourishes, applause, praise, cooperative, accommodating, accomplished, truthful, precisely, accurately, reachable, attainment, accomplishment, successes, possible
Negative words	anomalous, eradicate, offensive, atrociously, abhor, outrage, abort, abandon, scrape, harsh, unexpected, brusquely, nonattendance, absconder, irrational, mishandling

Those bags of words are used to find out the text, whether the tweet is positive or negative, each tweet will get 1 indentation that is,

Tweet data indentation = tweets positive – tweets negative

For e.g., a tweet “here is a pretty good climate.”

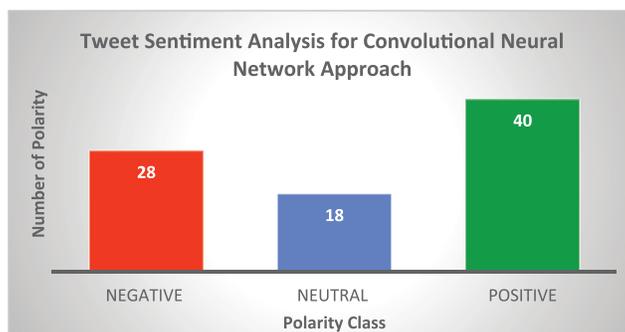
For this we achieve an indentation: positive tweets = pretty (1) + good (1) = 2;

Negative tweets = 0

So, the tweet indentation for this tweet is;

Tweet data indentation = 2 – 0 = 2

All tweets having positive and negative indentation that contains tweet value ‘0’ includes under a neutral tweet. The positive and negative words analyze the difference and get tweet scores to create a histogram plot of various tweets. Now, to track the polarity class database, a polarity class bar diagram is shown in Fig. 8 to analyze the sentiment of tweets.



**Figure 8:** Polarity class bar diagram for tweet sentiment analysis

### 7.2 Influences of Tweet Data in Unambiguous Mode

#### 7.2.1 Tweet Performance on Real-Time Dataset

In this research work, this execution contributes to representing an authenticate performance of the targeted Twitter account where the primary task provides authentication by the Twitter API account after

receiving a confirmation mail from Twitter. The positive and negative tweets (in %) are shown recent few tweets (data taken while running the python code and auto-generated output used recent last 50000 tweets). To generate tweet percentage, need to use various keys which are auto-generated by the API developer account and used in python code. The generated output is shown as,

Authenticated

Positive tweets percentage: 26.08695652173913 %

Positive tweets:

RT @anisha\_Singh07: CELEBRATING SUCCESS OF BBB3

Trending at no 2... WOW ♥️👍...

Negative tweets:

RT @justina\_cha: 이제 돌려주세요🙏

Hello @Twitter @TwitterSupport We have been following @jiminimy for a long time and found this account suspend...

RT @TeamSaath: Our reach has been curtailed since long.

What's happening, @Twitter?

Requesting you all to please Like, Retweet & #FollowT...

@bonngbudi Istimewa 🙏

Lebih baekk @Twitter down seminggu daripada Indodax ataupun IPOT.

RT @fingerofjimin: 🙏👩🏻🙏매일같이 도와주셔서 감사해요🙏👩🏻🙏

Hello @Twitter @TwitterSupport

We have been following @fingerofjimin for a long time and...

RT @FirstSnow\_JM: 이전 돌려줄때 됐다🙏 돌려주세요!!

Hello @Twitter @TwitterSupport

We have been following @kk2010633 for a long time and found this acc...

While using Twitter graphical view is required to have authentication as shown below;

1. import tweepy
2. consumer\_key = 'EoJKizEYz8aud3TGlhKvjP4jG'
3. consumer\_secret = 'JcxWYvqdfX199UwPvSZM18Ru2tHQmKqcCsGPzN5TTApXVXOo3g'

---

4. access\_token = '2701416409-U8I93Za4BubFSFnJ7mGJJmKM9Q9PgNzJFwjVIB'
5. access\_secret = 'FiOOSKJr6py6jZ21WIUFERc0HZrwzAAXNPF13wIXRRj10'
6. auth = tweepy.OAuthHandler(consumer\_key, consumer\_secret)
7. auth.set\_access\_token(access\_token, access\_secret)
8. our\_api = tweepy.API(auth)
9. print('Authenticated',our\_api)

---

Authenticated <tweepy.api.API object at 0x00000191A05FFD00>

natgeo\_tweets [0]

('@dadanhhrn not until @CraigWeekend says so', 59, 1269)

It is to be informed that the above complete output is taken, just for a demonstration purpose in this research work that delivers the strength of work.

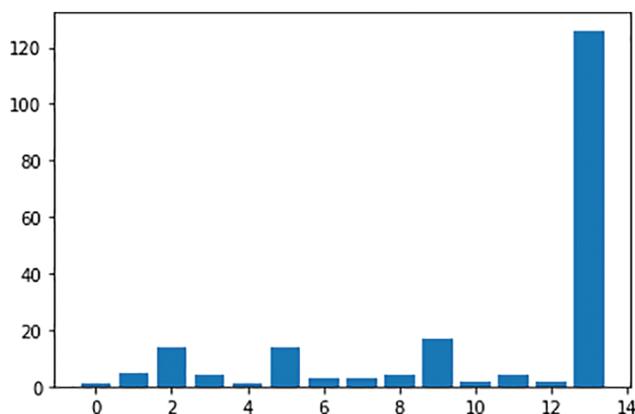
### 7.2.2 Positive and Negative Influences of Tweet Data

The influence of tweet data is presented as a positive and negative response that determine the status of tweets in the bar chart method and acquires more accuracy to comprehend the prediction of the intellectual position of the users as given in [Tab. 2](#).

**Table 2:** Apprehend the real-time status of tweet data code design

1	<code>import matplotlib.pyplot as plt</code>
2	<code>home_cursor = tweepy.Cursor(our_api.home_timeline)</code>
3	<code>tweets = [i.author.screen_name for i in home_cursor.items(200)]</code>
4	<code>unq = set(tweets)</code>
5	<code>freq = {uname: tweets.count(uname) for uname in unq}</code>
6	<code>plt.bar(range(len(unq)), freq.values())</code>
7	<code>plt.show()</code>

The result shows the performance status of the initiated Twitter account that demonstrates the positive and negative status of tweets on their account as shown in [Fig. 9](#).



**Figure 9:** Performance status of tweets in Twitter account

### 7.3 Analysis of Resolute Accuracy with Polarity Calculation

In this section, the detailed experiment is conducted with sentiment assignment to justify the results that determine the accuracy which is evaluated from various sentiment analyzers. The sentiment analyzers have various distinctions for comparison of their accuracies. [Tab. 3](#) represents tweets evaluation to transmit the data with a similar type of sentiments with the total number of tweets such as 2076, as experimented below,

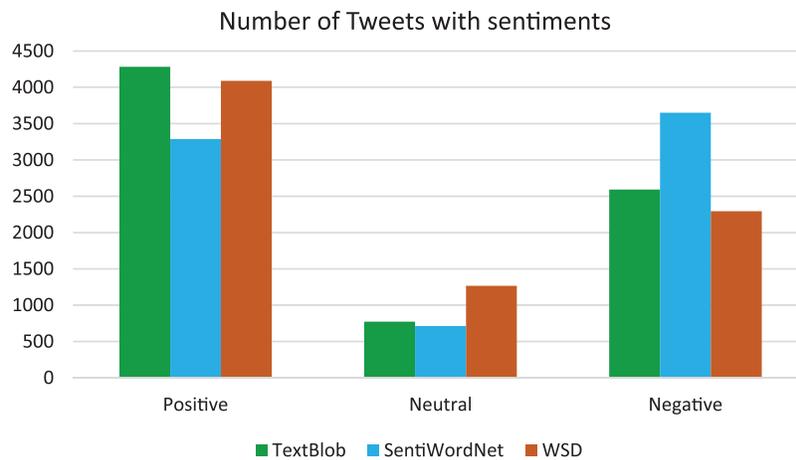
The tweet data is gathered from users' accounts only where preprocessing with 7650 tweets among the various sentiment analyzers and compared with various tweet rates which have been found the TextBlob had the uppermost positive tweets rate 4283 (in number), SentiWordNet produced 3650 the maximum negative sentiment tweet rate are given in [Tab. 4](#) and represented in [Fig. 10](#) respectively.

**Table 3:** Individual number of tweets to analyze sentiment data

Sentiment	Number of Tweets
Positive	1256
Neutral	37
Negative	783

**Table 4:** Polarity calculation using various sentiment analyzers (accuracy in numbers)

Sentiment analyzer	Positive	Neutral	Negative
TextBlob	4283	774	2593
SentiWordNet	3287	713	3650
WSD	4089	1267	2294



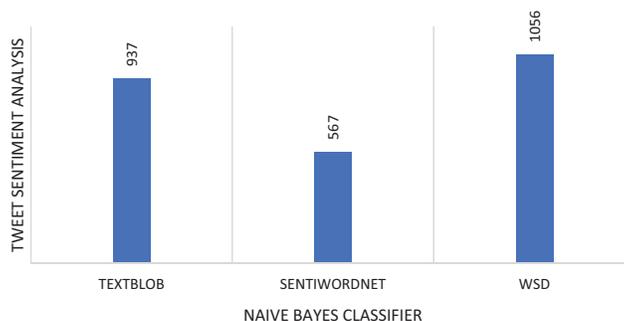
**Figure 10:** Sentiment analysis of tweet data using polarity computation

The validation of tweet data is extracted from various sentiment analyzers are used Naive Bayes classifier on a 2560 training set of 1280 positive and 1280 negative tweets analyzed by respective sentiment analyzers as given in [Tab. 5](#).

**Table 5:** Tweet validation report using sentiment analyzer with Naive Bayes classifier

Sentiment analyzer	Naive Bayes classifier		
	TextBlob	SentiWordNet	WSD
2560	937 (36.60%)	567 (22.14%)	1056 (41.25%)

As a result, Word Sense Disambiguation (WSD) had the highest accuracy rate are found to have a predicted determination sentiment analysis using the Naive Bayes classifier to provide accurate instances which are shown in Fig. 11.



**Figure 11:** Sentiment analyzer with the Naive Bayes classifier (inaccuracy)

## 8 Conclusion

In this research work, we investigate the sentiment analysis that occurs while communicating with tweet data that represent the happening moment around the world. Also, this work analyzed to predict sentiment in Twitter that contributes outperform features and kernel models to accomplish the sentiments using machine learning algorithms. This research is focused on the various sentiment analyzers with the utmost accuracy rate for learning about tweet sentiments on social media. The results are examined using the polarity method which is calculated based on the semantic score of a certain word using a lexicon-based approach. Though, the approach of machine learning is essentially intended to categorize the text by using Naive Bayes and SVM algorithms that provided evaluation between sentiment lexicons to adopt sentiment analysis. A hybrid approach for sentiment analysis has learned to use word sense which can accumulate tweet data that afford statistical sentimental structure of the search keyword and obtain high-efficiency results on social media. The expected consequences of this research are,

- (i) Search tweets using keywords of a particular account or random statistics,
- (ii) Users can perceive trending sizzling issues,
- (iii) Enable to analyze data themselves from twitter and
- (iv) Outputs are possibly be observed as histogram plots or bar diagrams or chart basis.

In the future scope, need to compare Twitter analysis with other spheres that prerequisite further research to know how to apply and solve the real-time problems of the society in terms of elections, product reputation evaluation, etc. would dive into the Twitter sentiment analysis aspects in near future. In sentiment analysis, researchers are working on the accuracy of the algorithm and used in commercial aspects or to do further research to find out more customer satisfaction based on tweets.

**Funding Statement:** This work was supported by Taif University Researchers Supporting Project (TURSP) under number (TURSP-2020/73), Taif University, Taif, Saudi Arabia.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] A. Alsaedi and M. Z. Khan, "A study on sentiment analysis techniques of Twitter data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, pp. 361–374, 2019.

- [2] M. Anjaria and R. M. R. Guddeti, "Influence factor-based opinion mining of Twitter data using supervised learning," in *Proc. 6th Int. Conf. on Communication Systems and Networks*, Bangalore, India, pp. 1–8, 2014.
- [3] S. Umer, P. P. Mohanta, R. K. Rout and H. M. Pandey, "Machine learning method for cosmetic product recognition: A visual searching approach," *Multimedia Tools and Applications*, vol. 80, no. 28, pp. 34997–35023, 2021.
- [4] A. DuVander, "Which APIs are handling billions of requests per day?," *Programmable Web*, 2012. <https://www.programmableweb.com/news/which-apis-are-handling-billions-requests-day/2012/05/23>.
- [5] K. N. Alam, M. S. Khan, A. R. Dhruva, M. M. Khan and J. F. Al-Amri, "Deep learning-based sentiment analysis of COVID-19 vaccination responses from Twitter data," *Computational and Mathematical Methods in Medicine*, vol. 2021, no. 1, pp. 1–15, 2021.
- [6] H. U. Khan, S. Nasir, K. Nasim, D. Shabbir and A. Mahmood, "Twitter trends: A ranking algorithm analysis on real-time data," *Expert Systems with Applications*, vol. 164, no. 3, pp. 113990, 2021.
- [7] A. Hasan, S. Moin, A. Karim and S. Shamshirband, "Machine learning-based sentiment analysis for Twitter accounts," *Mathematical and Computational Applications*, vol. 23, no. 1, pp. 1–15, 2018.
- [8] K. Chouhan, A. Singh, A. Srivastava, S. Agrawal, B. D. Shukla *et al.*, "Structural support vector machine for speech recognition classification with CNN approach," in *Proc. 9th Int. Conf. on Information Technology for Cyber and IT Service Management*, Bengkulu, Indonesia, Inspec no. 21297945, 2021.
- [9] Developers, "Twitter rest API," 2018. [Online]. Available: <https://dev.twitter.com/rest/public>.
- [10] I. Twitter, "Second quarter report," 2016.
- [11] K. P. Murphy, *Naive Bayes Classifiers*, vol. 18, Vancouver, BC V6T 1Z4, Canada: University of British Columbia, 2006.
- [12] N. Yadav, O. Kudale, S. Gupta, A. Rao and A. K. Shitole, "Twitter sentiment analysis using machine learning for product evaluation," in *Proc. 5th Int. Conf. on Inventive Computation Technologies*, Coimbatore, India, pp. 181–185, 2020.
- [13] M. Z. Asghar, F. M. Kundi, S. Ahmad, A. Khan and F. Khan, "T-SAF: Twitter sentiment analysis framework using a hybrid classification scheme," *Expert Systems*, vol. 35, no. 1, pp. 42–55, 2018.
- [14] A. Abbasi, H. Chen and A. Salem, "Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums," *ACM Transactions on Information Systems*, vol. 26, no. 3, pp. 1–34, 2008.
- [15] H. Anber, A. Salah and A. A. A. El-Aziz, "A literature review on Twitter data analysis," *International Journal of Computer and Electrical Engineering*, vol. 8, no. 3, pp. 241–249, 2016.
- [16] S. Kumar, F. Morstatter and H. Liu, *Twitter data analytics*, New York: Springer, 2014.
- [17] N. F. F. Da Silva and E. R. Hruschka, "Tweet sentiment analysis with classifier ensembles," *Decision Support Systems*, vol. 66, no. C, pp. 170–179, 2014.
- [18] M. D. S. C. Devika, "Sentiment analysis: A comparative study on different approaches," *Procedia Computer Science*, vol. 87, no. 4, pp. 44–49, 2016.
- [19] M. A. Chenaghlu, M. R. F. Derakhshi, L. Farzinvas, M. A. Balafar and C. Motamed, "Topic detection and tracking techniques on Twitter: A systematic review," *Hindawi Complexity*, vol. 21, pp. 1–15, 2021.
- [20] J. Ho, D. Ondusko, B. Roy and D. F. Hsu, "Sentiment analysis on tweets using machine learning and combinatorial fusion," in *Int. Conf. on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing*, Fukuoka, Japan, Cyber Science and Technology Congress, 2019.
- [21] S. Tan, Y. Li, H. Sun, Z. Guan, X. Yan *et al.*, "Interpreting the public sentiment variations on Twitter," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 5, pp. 1158–1170, 2013.
- [22] G. Dubey, S. Chawla and K. Kaur, "Social media opinion analysis for Indian political diplomats," in *Proc. 7th Int. Conf. on Cloud Computing, Data Science and Engineering-Confluence*, Noida, India, pp. 681–686, 2017.
- [23] T. Fu, A. Abbasi, D. Zeng and H. Chen, "Sentimental spidering: Leveraging opinion information in focused crawlers," *ACM Transactions on Information Systems*, vol. 30, no. 4, pp. 1–30, 2012.

- [24] K. Marcel, H. Safa and M. Masud, "WhatsAppNow: Urban social application with real-time peer-to-peer ambient and sensory data exchanges," *Multimedia Tools and Applications*, vol. 75, no. 21, pp. 13349–13374, 2016.
- [25] G. Mostafa, I. Ahmed and M. S. Junayed, "Investigation of different machine learning algorithms to determine human sentiment using Twitter data," *International Journal of Information Technology and Computer Science*, vol. 2, pp. 38–48, 2021.
- [26] A. Hassan, A. Abbasi and D. Zeng, "Twitter sentiment analysis: A bootstrap ensemble framework," in *Proc. ASE/IEEE Int. Conf. on Social Computing*, Alexandria, VA, USA, pp. 357–364, 2013.
- [27] A. Birmingham and A. Smeaton, "Classifying sentiment in microblogs: Is brevity an advantage?," in *Proc. 19th Int. Conf. on Information and Knowledge Management and Co-located Workshops*, Toronto, Ontario, Canada, pp. 1833–1836, 2010.
- [28] M. F. Goodchild and J. A. Glennon, "Crowd sourcing geographic information for disaster response: A research frontier," *International Journal of Digital Earth*, vol. 3, no. 3, pp. 231–241, 2010.
- [29] S. Hossain, S. Umer, V. Asari and R. K. Rout, "A unified framework of deep learning-based facial expression recognition system for diversified applications," *Applied Sciences*, vol. 11, no. 19, pp. 1–26, 2021.
- [30] S. H. Biradar, J. V. Gorabal and G. Gupta, "Machine learning tool for exploring sentimental analysis, on Twitter data," *Materials Today Proceedings*, vol. 56, pp. 1927–1934, 2022.
- [31] S. Umer, R. K. Rout, C. Pero and M. Nappi, "Facial expression recognition with trade-offs between data augmentation and deep learning features," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, pp. 721–735, 2021.
- [32] S. Bahri, P. Bahri and S. Lal, "A novel approach of sentiment classification using emoticons," in *Int. Conf. on Computational Intelligence and Data Science*, Gurugram, India, pp. 1–10, 2018.
- [33] A. Abirami and V. Gayathri, "A survey on sentiment analysis methods and approach," in *Proc. 8th Int. Conf. on Advanced Computing*, Chennai, India, pp. 72–76, 2017.
- [34] Sentiment classifier. 2022. [Online]. Available: [https://github.com/kevincobain2000/sentiment\\_classifier](https://github.com/kevincobain2000/sentiment_classifier).
- [35] N. C. Dang, M. N. M. Garcia and F. D. L. Prieta, "Sentiment analysis based on deep learning: A comparative study," *Electronics*, vol. 9, no. 3, pp. 1–29, 2020.
- [36] V. Sahayak, V. Shete, A. Pathan and A. Pathan, "Sentiment analysis on Twitter data," *International Journal of Innovative Research in Advanced Engineering*, vol. 2, no. 1, pp. 178–183, 2015.
- [37] S. Jawale and S. D. Sawarkar, "Interpretable sentiment analysis based on deep learning: An overview," in *IEEE Pune Section Int. Conf.*, Pune, India, 2020.
- [38] W. Medhat, A. Hassan and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [39] P. Ray and A. Chakrabarti, "A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis," *Applied Computing and Informatics*, vol. 18, no. 1/2, pp. 163–178, 2022.
- [40] S. Umer, R. Mondal, H. M. Pandey and R. K. Rout, "Deep features based convolutional neural network model for text and non-text region segmentation from document images," *Applied Soft Computing*, vol. 113, no. 3, pp. 107917, 2021.
- [41] K. Chouhan, A. Kumar, A. K. Chakraverti and R. R. Cholla, "Human fall detection analysis with image recognition using convolutional neural network approach," in *Proc. Int. Conf. on Trends in Computational and Cognitive Engineering*, Singapore, Springer Nature, vol. 376, pp. 95–106, 2022.
- [42] R. K. Rout, S. S. Hassan, S. S. Umer, K. S. Sahoo and A. Gandomi, "Feature-extraction and analysis based on spatial distribution of amino acids for SARS-CoV-2 protein sequences," *Computers in Biology and Medicine*, vol. 141, pp. 105024, 2022.
- [43] S. S. Hassan, R. K. Rout, K. S. Sahoo, N. Jhanjhi and S. Umer, "A vicenary analysis of SARS-CoV-2 genomes," *Computers, Materials and Continua*, vol. 69, no. 3, pp. 3477–3493, 2021.