IJCRM

International Journal of Contemporary Research In Multidisciplinary

Research Article

A Survey on Analysis of Sentiment Using Twitter Dataset

Ashish Kharb [1*], Karambir Bidhan [2], Atul Sharma [3]

[1] Department of Computer Science and Engineering, UIET, Kurukshetra, Haryana, India
[2] Department of Computer Science and Engineering, UIET, Kurukshetra, Haryana, India
[3] Department of Computer Science and Engineering, UIET, Kurukshetra, Haryana, India

Corresponding Author: *Ashish Kharb

DOI: https://doi.org/10.5281/zenodo.15476092

Abstract

In this modernizing era, the elevation of web technology and its proliferation have built a web network on which lies a vast volume of data for the users of internet, and a lot of data is being formed. The Internet has progressed as a modernized framework for exchanging ideas, skill development, and communicating opinions through online mediums. Social networking sites like Instagram, Twitter, and Facebook have gained an immense amount of popularity due to their features that allow people to interact, express their points of view, share beliefs about certain topics, have discussions, or post messages, images, and videos across the internet, connecting the globe. There has been enormous work done in the area of Twitter data analysis for sentiment. This research work focuses mainly on emotion classification of Twitter data, which helps in analyzing the information shared through tweets in which opinions can be extremely heterogeneous, unstructured, and also can be positive or negative, or neutral in some scenarios. In this study, we will supply the survey and a performance comparative analysis of present processes for analyzing sentiment, like Machine Learning, Deep Learning, and lexicon-based methods, by evaluating their performance metric, i.e., accuracy. Evaluation of techniques will include Machine Learning algorithms like Support Vector Machine, Naive Bayes, Decision Tree, Random Forest, and, along with the deep learning algorithms which include Recurrent Neural Networks, Long Short-Term Memory Networks, and Hybrid Approaches, and many more. After evaluation of past techniques, we will provide the best technique based on the accuracy of existing opinion mining models, comprising traditional methods and recent enhancements in AI.

Case Report Information

- **ISSN No:** 2583-7397
- **Received:** 26-04-2025
- **Accepted:** 17-05-2025
- **Published:** 20-05-2025
- **IJCRM:**4(3); 2025: 114-123
- **©2025, All Rights Reserved**
- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

How to Cite this Case Report

Kharb A, Bidhan K, Sharma A. A survey on the analysis of sentiment using the Twitter dataset. Int J Contemp Res Multidiscip. 2025;4(3):114–123.

Access this Article Online

www.multiarticlesjournal.com

KEYWORDS: Sentiment Analysis, Machine Learning, Deep Learning, Social Media, Twitter Data.

INTRODUCTION

Emotion is considered one of the basic instincts of human beings. Emotion detection plays a crucial role in the field of textual analysis. Currently, humans' expressions and emotional states have become a leading topic for research work. Detection of Emotion and Recognition of the meaning behind particular emotions from texts are trending fields of research that lie under the category of Emotion Analysis. Emotion Analysis targets towards detecting and understanding feelings through the expressions from sentences which might include anger, surprise, joy, fear, sad, love, etc. Sentiment analysis refers to the method of analyzing text to identify its intrinsic emotional tone. As social media platforms which including Twitter, have procured popularity, sentiment analysis has emerged as a crucial resource for companies, organizations, and governments aiming to understand public opinion and develop informed viewpoints. Natural Language Processing (NLP) techniques are widely utilized for analyzing sentiment because they allow machines to understand and analyze the language of human. NLP methods assess tweets instantly, recognize the expressed sentiment in tweets, and offer information about current patterns and trends in public opinion. ML algorithms, categorized within NLP, can learn from comprehensive datasets and precisely determine the sentiment of fresh tweets. The primary aim of this paper is to inspect current sentiment analysis techniques used on Twitter datasets and offer theoretical comparisons of the leading methods.

Sentiment Analysis

Sentiment analysis is often referred to as opinion mining, which is a sub-part of ML execution where the user requirement is to find out the simplified sentiment from the provided document. NLP and ML techniques are applied for extracting subjective information attained by documents and classifying concerning polarity, which might be positive, neutral, or negative. It's an effective method for analysis as we can conclude the overall opinion about movie reviews, or predict stock markets for a given company, such as when most people think positively about it, then there is a high possibility that its stock market will advance, and many more. Sentiment analysis in real life is a very complicated task to answer because of the complicated nature of languages (subjectivity or objectivity, vocabulary, negation, grammar); however, it is also the intriguing reason for generating interest in working on. In this article, by developing the model based on probabilities, we will be performing the machine learning techniques for classifying tweets from Twitter, which might be "positive" or "negative" sentiment. Twitter is an open world, microblogging website, and is free to use, where people share their emotions and feelings instantly by posting a tweet using 140 characters. We directly address a tweet to other users by using the sign "@" or by utilizing the hashtag "#" on a tweet to participate in an online debate on a topic. Twitter has become an excellent source of data that helps understand the present opinion among all users about anything.

A) Types of Sentiment Analysis

The primary types of sentiment analysis include:

1. Binary Sentiment Analysis

It is the type of analysis where the text is classified: positive or negative. It is straight to the point and typically achieves accuracy at a higher point due to its simple structure, but it lacks granularity.

2. Multi-class Sentiment Analysis

In this method text is categorized into multiple sentiment classes, unlike binary analysis, such as positive, neutral, negative and very negative. It requires a lot more training data, and it is complex to implement, whereas it gives increased, nuanced insights.

3. Granular (Fine-Grained) Sentiment Analysis

In granular sentiment analysis the scoring is done on sentiments with continuous scale (e.g., 1 to 5 stars or 0 to 100), facilitating a more comprehensive understanding of the feeling expressed in a textual manner. This is commonly used in customer satisfaction surveys.

4. Aspect-Based Sentiment Analysis

For this type in sentiment analysis the method has focus on particular quality of product or service, finding whether the opinion for that quality is positive or negative. For example, it is used for mining customer feedback in some specific characteristic of a mobile phone instead of feedback about the whole product.

5. Emotion Detection Analysis

In emotion detection analysis, as suggested by its name, the feeling written or expressed in the document (text) might contain emotions like surprise, joy, sad, fear, or anger. In this method, the analysis goes beyond labelling text as negative or positive. This method gives a deep insight into the human sentiment communicated through the text document.

6. Intent-Based Sentiment Analysis

In this approach of analyzing sentiment, target is to fully comprehend the reason behind writing text, as well as, analyze the the emotion. For example, this method is used when we require to find out the main reason behind the complaint from the customer about the service or product is about little frustration or is it a resolution demanding.

RELATED WORK

The former prominent related research work on sentiment analysis will be discussed, as well as the datasets on which they performed the tasks. The comparison criteria here are the different techniques' accuracy working on different datasets, but here all datasets must include tweets only. We will consider the metrics, such as accuracy, for performance evaluation. Communication has been, in recent decades, shifted towards online platforms such as Instagram, Twitter, Facebook, etc., hence providing a lot of data for processing and analyzing to help understand the exact emotion and meaning behind that particular written data. For instance, Sanjiban Sekhar Roy *et al.* [1] concluded that the precision is 0.98, recall is 0.99 and F1 score is 0.98 is achieved by long short term memory(LSTM) where accuracy was 97% and its was found that this accuracy surpassed other models accuracy such as XGBoost (XGB), support vector machine (SVM), K-nearest neighbor (k-NN), Naïve Bayes (NB), Artificial Neural Network (ANN), logistic regression (LR), Bidirectional encoder representations from transformers (BERT) and random forest (RF) models used in detecting hate speech. When acting upon a dataset from Twitter API for analyzing sentiment for hateful tweets, it was found that techniques got accuracy LSTM 97% (highest among others), NB 92%, LR 93%, SVM 93%, XGBoost 95%, RF 92%, kNN 94%, ANN 94%, BERT 93%. In another work, NIKLAS BRAIG *et al.* [2] focused on ML based analysis of sentiment procedures where comparison of the most effective classification algorithms using five databases, where data is from Twitter which is COVID-19 related, which are: ACM DL, ScienceDirect, IEEE Xplore DL, AIS Electronic Library and SpringerLink is done. This work concluded with the ensemble models, which consisted of several ML classifiers being best best-performing models in terms of accuracy. Specifically, RoBERTa models and BERT (Bidirectional Encoder Representations from Transformers).

Thotakura Venkata Sai Krishna *et al.* [3] applied several ML algorithms which includes NB, logistic regression (LR), decision tree (DT), RF and DL algorithms which are Recurrent Neural Network (RNN), LSTM, and GRU separately and put forth an innovative ensemble comprising with ML algorithms and DL algorithms for classification of sentiment which attained higher accuracy by comparing to the established work. They used an ensemble novel method for sentiment classification on the Twitter dataset, achieving an accuracy with TextBlob (average): NB 72.3%, LR 90.3%, DT 87.5%, RF 87.5%, Recurrent Neural Networks (RNN) 88.5%, LSTM 90.86%, Short LSTM (SLSTM) 90.75%, as well as with VADER (average): NB 75.7%, LR 90.3%, DT 87%, RF 87%, RNN 88.13%, LSTM 89.7%, SLSTM 90.06%. Another work, Vijay K *et al.* [4] used classification algorithms including SVM, RF, LSTM, Ensemble Machine Learning, ANN, Bidirectional LSTM (Bi-LSTM) where approaches such as filtering, removal of stopwords, tokenization, lemmatization and stemming were implemented for pre-processing the tweet API, later the preprocessed input is provided as input for Bag of Words and Term Frequency-Inverse Document Frequency (TF-IDF) to vectorize. Followed by classification being implemented with the aforementioned models. Bidirectional LSTM proved itself as the most accurate with an accuracy rate of 98.39.% %

and 98.14% in vectorizing techniques, including Bag of Words and TF-IDF, respectively, making the tool crucial for conducting voice analyses on the platform.

In Saadat M. Alhashmi *et al.* [5], a recommended hybrid approach is proposed, which uses various classifiers which are used for dealing with different problems in order to enhance the accuracy. The suggested research strategy is the fusion-based tactic of SVM & BFTAN (SVM and Bayes Factor Tree Augmented Naive Bayes), which achieved the highest accuracy among other techniques such as BFTAN, TAN, Naïve Bayes (NB), SVM, and Random Forest (RF).

However, in another work, Rasika Wagh *et al.* [6] analyzed the sentiment approaches used to accomplish the extraction of sentiment from the tweets and stated that WordNet is succeeded by ML techniques, which includes Maximum entropy, SVM, and Naïve-Bayes with more accuracy.

In Vishal A Kharde *et al.* [7], it was found that when techniques like Baseline, SVM, Maximum Entropy, and NB are implemented on datasets HASH, EMOT, ISIEVE, Columbia univ, and Stanford, then these techniques achieve accuracy of 73.65%, 74.56%, 76.68%, and 74.93% respectively.

In Mantasha Khan *et al.* [8], it was noted that when sentiment analysis was performed on a dataset from Kaggle, then the techniques gained accuracy for XGB with Gensim 59%, XGB with CountVectorizer 70%, Vader 57%, RF with Gensim 68%, RF with CountVectorizer 69%, Bi-LTSM 73%, and Single LSTM 71%. In Bhumika Gupta *et al.* [9] techniques like Random Forest Classifier, SVM, DAN2, Bayesian Logistic Regression, Neural Network, Ensemble classifier, Maximum Entropy, Naïve Bayes are implemented on dataset from Twitter API and the accuracy was noted 86.06%, 85%, 74.84%, 66.24%, 87.5%, 89.93%, 90%, 90% respectively. In other work, Vaashini Palaniappan *et al.* [10] it was investigated that the accuracy for techniques where (before embedding sentiment polarity) K-Nearest Neighbours (KNN) 64%, RF 44%, Decision Tree 84%, NB 80%, Logistic Regression 83%, Support Vector Machines (SVM) 84% and (After embedding sentiment polarity) Random Forest (RF) 45%, KNN 76%, NB 77%, Decision Tree 91%, LR 85%, Support Vector Machines (SVM) 93% when action on dataset from github (https://github.com/ananyasarkertonu/Twitter-Dataset.)

In the work of J. C. Pereira-Kohatsu *et al.* [11], it was observed that a hybrid method for identifying and monitoring hate speech on Twitter attained an accuracy of 82.80% with LSTM with Multilayer Perceptron (LSTM+MLP)

Other work, S. Agarwal *et al.* [12] implemented the techniques LSTM and RNN for analyzing sentiment on the Twitter dataset by Twitter API with the accuracy of 88.99%.

In S. Sadiq *et al.* [13], it was discovered that an accuracy of 91.30% is achieved when finding aggression from data of Twitter by using techniques Convolutional Neural Networks (CNN), LSTM, MLP, Bidirectional LSTM (BiLSTM).

In O. Araque *et al.* [14] accuracy of 90.63% is gained through an ensemble method of LSTM and SVM for detecting online hate speech working on a dataset from twitter API.

A. S. Alammary *et al.* [15] used TF-IDF with Arabic questions and achieved accuracy for models NB 78.72%, Logistic Regression (LR) 77.95%, SVM 77.78%, KNN 76.65%, Decision Tree 59.71%

In B. Raufi *et al.* [16], mobile applications were considered for hate speech and for dataset twitter API is accessed to help achieve accuracy of 94-95% by using technique Artificial Neural Networks (ANN).

In P. K. Roy *et al.* [17] deep convolutional neural network is used for hate speech identification on twitter API dataset where the technique of SVM and LSTM are used and both techniques attained accuracy of 80% and 97% (LSTM with highest accuracy).

In A. S. Imran *et al.* [18], deep learning techniques are used, such as GloVe(Twitter) + Long Term Short Memory (LSTM), which has 81.9% accuracy, and other techniques, FastText + LST, M implemented, which got 82.4% accuracy on a self-collected Covid-19 tweet dataset. P. Gupta *et al.* [19] found the COVID-19 vaccine opinion of the people from self-collected Twitter data by applying the technique of Unigram + LinearSVC, which performed the task with an accuracy of 84.40%

In study of L.-A. Cottas *et al.* [20] it was found out that an accuracy of 78.94% is achieved when the social media platform of twitter is tracked during the period of COVID-19 and the data used here is the self-collected dataset from the twitter. Techniques applied here transformer model that is BERT.

In L. Miao *et al.* [21] studied about the period when there was lockdown in New York during Covid-19, where the technique Glove + LSTM is performed on the dataset from Chen *et al.*2020 and attained the accuracy of 66%.

Other work, S. Malla *et al.* [22] studied the dataset WNUT 2020 (Nguyen *et al.*2020) from the COVID-19 period by applying the MVEDL Ensemble model consisting of Roberta, BER Tweet, and CT-BERT and achieved an accuracy of 91.75%.

Another work, K. Chakraborty *et al.* [23], proved that popularity affected the accuracy in social media by acting on the dataset, which is self-collected from the COVID-19 period. Used technique Trigrams and TF-IDF score + LR and got accuracy 81.40%. In other work, M. Mahdikhani *et al.* [24] analyzed different stages of Covid-19 pandemic tweets were analyzed by using the technique Crystal Feel Assemble voting classifier, Random Forest (RF), (SGD), and Logistic Regression (LG) on a self-collected COVID-19 dataset and gained an accuracy of 95.04%. In another work, M. S. Satu *et al.* [25] used a novel ML classification method labelled as TClust VID acting on the COVID-19 data. Trust VID is an Ensemble method with DL Models with accuracy >90%. M. Y. Kabir *et al.* [26] found out that after applying BiLSTM on the data of self-collect Covid-19 dataset accuracy of 89.51% can be achieved for opinion mining.

In S. Behl *et al.* [27] used the technique Word2Vec + MLP (Multilayer Perceptron) on a dataset Self self-collected Covid-19 data and Nepal and the Italian earthquake 2015 for disaster relief and hazard crises via sentiment analysis, gaining 83% accuracy.

Other work M. M. Rahman *et al.* [28] used various exploration techniques Bag of Word (BOW), Digit Terrain Model (DTM), Part of Speech (POS), Differentially Privacy (DP) and n-grams, logit model for analysis reopening of US during Covid-19 period and data used here is self-collect Covid-19 dataset with 56.18% accuracy. H. Lyu *et al.* [29] mentioned the XLNet technique for studying the opinions of the public about COVID-19 vaccines worked on self-collected COVID-19 data from Twitter and achieved 63% accuracy. In other work, J. Choudrie *et al.* [30] it was found that a novel deep learning model, RoBERTa can achieve 80.33% accuracy when acting on the dataset which is a self-collected Covid-19 dataset and emotion in text by Crowdflower. In B. N. Ramya *et al.* [31] it was observed that an accuracy of 92.49% (short tweet characters<70) and 60.56% (long tweets characters<150) is achieved when developing a Smart Simon Bot for sentiment analysis using technique Naïve Bayes on the dataset of self-collected data from tweets. In A. Alsayat *et al.* [32], it was discovered that when using an ensemble deep learning model (Ensemble Classifier (LSTM+ FastText, BERT….)), accuracy is 92.65% when working on Self collected COVID-19 dataset, Crowdflower and Yelp dataset.

Another work, M. Singh *et al.* [33] this study was held on opinion mining about impact of Covid-19 on social life through deep learning model (BERT) which provide accuracy of 93.89% on self-collected Covid-19 data from twitter.

In other work C. Caliskan *et al.* [34] focused on the analysis of Covid-19 pandemic in Ohio where techniques like already trained GloVe+ DL models (RNN, CNN) being implemented on dataset containing tweets from self-collected COVID-19 datasets, providing accuracy of 71%.

R. Goel *et al.* [35] study was held for analyzing leaders during the periods of COVID-19 by implementing the technique TF-IDF Random Forest, giving accuracy of 96% when implemented on self-collected Covid-19 dataset of tweets.

Misbah Ul Hoque *et al.* [36] found that the technique Vader worked better than some DL based prediction models for tweets analysis on data of COVID-19 when working on types of Twitter APIs: Academic Research Track and Historical PowerTrack.

In another work, M.-Y. Cheng *et al.* [37] used the hybrid approach that is Hybrid SGRU (Gated Recurrent Unit) for text mining on construction site mis happenings which gained the highest precision among others by working on construction accident narrative historical data.

A. Subasi *et al.* [38] used the hybrid approach where they combined the technologies PSO (particle swarm optimization) and GA (genetic algorithm) on a dataset of EEG signal data and achieved 99.38% accuracy.

In M. Loey *et al.* [39] gave a hybrid DL model with ML method to detect the face mask during the Covid-19 period by techniques SVM, Decision Tree and ensemble algorithm gaining accuracy of 99.76% for the decision tree and 100% for the support vector machine.

In other work, M. Anjaria *et al.* [40] checked what effect will the supervised learning will have on opinion mining when acting on data from US Presidential Elections 2021 and Karnataka state election 2013 where the accuracy is 88% and the technique used is SVM with PCA (principal component analysis). In A. Go *et al.* [41] attained the accuracy of 80% when applying the techniques Naïve Bayes, Maximum Entropy & SVM for emotion classification on Twitter data from a dataset of microblogging services. In another work, S.Zhu [42] found that the accuracy is 62.90% when applied the technique of support vector machines (SVM) on the dataset from Sina microblog data.

In M. Al-Ayyoub *et al.* [43] used Arabic tweets for lexicon-based opinion mining by applying the technique support vector machine (SVM) and attained an accuracy of 86.89% when working on manually named data including 300 +ve, -ve, and neutral tweets. In X. Wang *et al.* [44] used technique support vector machine (SVM) for graph-based hashtag opinion classification with accuracy of 84.13% when performing on one-week tweets and hashtags. In other work, Ankit *et al.* [45] informed that when using the technique SVM, RF, and NB accuracy of 75.81% is achieved when working on the dataset Stanford Sentiment140 corpus and HCR. In P. Melville *et al.* [46] Naïve Bayes technique is used for analysis by combining lexical information and text classification, on the dataset Internet Movie Database, and then the 81.42% of accuracy is achieved. In another work, H. Ghorbel *et al.* [47] worked on enhancements in intelligent web mastering by experimenting on French movie reviews. They used the Support Vector Machine (SVM) technique on a dataset of French movie reviews and attained 93.25% accuracy. V. K. Singh *et al.* [48] mentioned techniques NB, SVM, on the dataset movie review and attained an accuracy of 81.14% by using unsupervised and sentiwordnet approaches.

S.Tan *et al.* [49] conducted research on Chinese text and used techniques centroid, K-nearest neighbor, window, NB, and SVM, with an efficiency the information gain is 0.9. They worked on a dataset of 1021 documents with a Chinese sentiment corpus. In B. Gokulakrishnan *et al.* [50] done sentiment analysis was done using the techniques NB, Random Forest (RF), and SVM on microblogging data with 72.70% accuracy. In another work, E. Boiy *et al.* [51] worked on multilingual web texts from a 2000 movie reviews dataset. They used techniques SVM and NB, which achieved 86.35% accuracy.

Table 1: Summary Table for Research Articles

Paper	Technique used	Efficiency	Dataset
Vishal A Kharde *et al.* [7]	Baseline, NB, Max Entropy, and SVM	Baseline 73.65%, Naïve Bayes 74.56%, Support VM 76.68%, Maximum Entropy 74.93%	HASH, EMOT, ISIEVE, Columbia Univ, Stanford
NIKLAS BRAIG *et al.* [2]	Traditional ML techniques like VADER, NB, RF, SVM, and DL Techniques like RoBERTa, BERT, Bi-LSTM	RoBERTa and BERT show high efficiency among all	Twitter COVID-19 dataset
Rasika Wagh *et al.* [6]	SVM, Maximum entropy and Naïve-Bayes	Accuracy increased by 4% to 5% when using the Hybrid approach.	Twitter
Mantasha Khan *et al.* [8]	Vader, XGBoost with Gensim, XGBoost with CountVectorizer, Random Forest with Gensim, Random Forest with CountVectorizer, Bidirectional LTSM, Single LSTM,	Vader 57%, XGBoost with Gensim 59%, XGBoost with CountVectorizer 70%, Random Forest with Gensim 68%, Random Forest with CountVectorizer 69%, Bidirectional LTSM 73%, Single LSTM 71%	Kaggle
Bhumika Gupta *et al.* [9]	DAN2, Bayesian Logistic Regression, SVM, Neural Network, NB, Ensemble classifier, RF Classifier, Maximum Entropy,	DAN2 86.06%, Bayesian Logistic Regression 74.84%, SVM 85.0, Neural Network 89.93%, NB 66.24%, Ensemble classifier 90.0%, RF Classifier 87.5%, Maximum Entropy 90.0%	Twitter
Vaashini Palaniappan *et al.* [10]	RF, SVM, KNN, NB, DT, LR,	before embedding sentiment polarity RF 44%, SVM84%, KNN 64%, NB 80%, DT 84%, LR 83% After embedding sentiment polarity RF 45%, SVM 93%, KNN 76%, NB 77%, DT 91%, LR 85%	https://github.com/ananyasarkertonu/Twitter-Dataset.
Sanjiban Sekhar Roy *et al.* [1]	LSTM, NB, LR, SVM, XGBoost, RF, kNN, ANN, BERT	LSTM 97%, NB 92%, LR 93%, SVM 93%, XGBoost 95%, RF 92%, kNN 94%,ANN 94%,BERT 93%	Twitter
J. C. Pereira-Kohatsu *et al.* [11]	LSTM+MLP	82.8%	Twitter
S. Agarwal *et al.* [12]	RNN AND LSTM	88.99%	Twitter
S. Sadiq *et al.* [13]	CNN, LSTM, MLP, BiLSTM	91.3%	Twitter
O. Araque *et al.* [14]	LSTM, SVM	90.63%	Twitter
A. S. Alammary [15]	NB, LR, SVM, KNN, Decision Tree	NB 78.72%, LR 77.95%, SVM 77.78% , KNN 76.65%, Decision Tree 59.71%	Twitter
B. Raufi *et al.* [16]	ANN	94-95%	Twitter
P. K. Roy *et al.* [17]	SVM, LSTM	SVM 80% LSTM 97%	Twitter
Thotakura Venkata Sai Krishna *et al.* [3]	NB, LR, DT, RF, RNN, LSTM, SLSTM	With TextBlob (average): NB 72.3%, LR 90.3%, DT 87.5%, RF 87.5%, RNN 88.5%, LSTM 90.86%, SLSTM 90.75%. With VADER (average): NB 75.7%, LR 90.3%, DT 87%% %, RF 87%, RNN 88.13%, LSTM 89.7%, SLSTM 90.06%	Twitter
Vijay K *et al.* [4]	SVM, Ensemble machine learning, Random Forest, ANN, Bi-LSTM, LSTM	SVM Accuracy 85.3% Precision 84.3% F1 score 85.01%, Random Forest Accuracy 88.45% Precision 89.32% F1 score 89.45%, Ensemble machine learning Accuracy 89.45% Precision 9.32% F1 score 90.75%, ANN Accuracy 93.54% Precision 92.45% F1 score 93.09%, LSTM Accuracy 95.33% Precision 94.35% F1 score 95.78%, Bi-LSTM Accuracy 98.14% Precision 97.45% F1 score 98.54%	Twitter
A. S. Imran *et al.* [18]	1. Glove(Twitter)+LSTM 2. Fast Text +LSTM	1. 81.9% 2.82.4%	Self-collected Covid-19 dataset
P. Gupta *et al.* [19]	Unigram + LinearSVC	84.4%	Self-collected COVID-19 dataset
L.-A. Cotfas *et al.* [20]	BERT	78.94%	Self-collected COVID-19 dataset
Saadat M. Alhashmi *et al.* [5]	Word2Vec+ BFTAN	82.8%	Covid-19 dataset and Expo2020 dataset
L. Miao *et al.* [21]	GloVe + LSTM	66%	Chen *et al.*2020
S. Malla *et al.* [22]	MVEDL Ensemble model consisting of RoBERTa, BERTweet, and CT-BERT	91.75%	WNUT 2020 (Nguyen *et al.*2020)

K. Chakraborty *et al.* [23]	Trigrams and TF-IDF score + LR	81.4%	Self-collected COVID-19 dataset
M. Mahdikhani [24]	CrystalFeel Ensemble voting classifier (RF, SGD, LG)	95.04%	Self-collected COVID-19 dataset
M. S. Satu *et al.* [25]	TcrustVID (Ensemble method with DL Models)	>90%	Covid-19 dataset
M. Y. Kabir *et al.* [26]	BiLSTM	89.51%	Self-collected COVID-19 dataset
S. Behl *et al.* [27]	Word2Vec +MLP	83%	Self-collected COVID-19 dataset and Nepal and Italian earthquake 2015
M. M. Rahman *et al.* [28]	Various exploration techniques used: BOW (bag of words), DTM (digit terrain model), POS (part of speech), DP (Differentially private), and n-grams, Logit model	56.18%	Self-collected COVID-19 dataset
H. Lyu *et al.* [29]	XLNet	63%	Self-collected COVID-19 dataset
J. Choudrie *et al.* [30]	RoBERTa	80.33%	self-collected COVID-19 dataset and emotion in text by Crowdflower
B. N. Ramya *et al.* [31]	Naïve Bayes	92.49% (short tweet characters<70) 60.56% (long tweets characters<150)	Self-collected COVID-19
A. Alsayat [32]	Ensemble Classifier (LSTM+ FastText,BERT….)	92.65%	Self-collected COVID-19 dataset, Crowdflower, and Yelp dataset
M. Singh *et al.* [33]	BERT	93.89%	Self-collected COVID-19 dataset
C. Caliskan [34]	Pre-trained GloVe+ Deep learning models (RNN, CNN)	71%	Self-collected COVID-19 dataset
R. Goel *et al.* [35]	TF-IDF Random Forest	96%	Self-collected COVID-19 dataset
Misbah Ul Hoque *et al.* [36]	Vader	works positively as compared to some state-of-the-art DL-based prediction models for analysing data of tweets from COVID-19	two types of Twitter APIs: Academic Research Track and Historical PowerTrack
M.-Y. Cheng *et al.* [37]	Hybrid SGRU	Highest precision among others	Construction accident narrative historical data
A. Subasi *et al.* [38]	GA and PSO	99.38%	EEG signal data
M. Loey *et al.* [39]	Decision Tree, SVM, and ensemble algorithm	DT 99.76% and 100% for SVM	face masked datasets
M. Anjaria *et al.* [40]	SVM with PCA	88%	US Presidential Elections 2021 and Karnataka State Election 2013
A. Go *et al.* [41]	Naïve Bayes, Maximum Entropy & SVM	80%	microblogging services
S.Zhu *et al.* [42]	SVM	62.9%	Sina microblog data
M. Al-Ayyoub *et al.* [43]	SVM	86.89%	manually tagged data of 300 +ve , -ve and neutral tweets
X. Wang *et al.* [44]	SVM	84.13%	One week of tweets and hashtags
Ankit *et al.* [45]	NB, RF and SVM	75.81%	Stanford Sentiment140 corpus and HCR
P. Melville *et al.* [46]	NB	81.42%	Internet Movie Database
H. Ghorbel *et al.* [47]	SVM	93.25%	French movie reviews
V. K. Singh *et al.* [48]	NB, SVM	81.14%	3 movie review dataset
S.Tan *et al.* [49]	centroid, K-nearest neighbor, window, NB, and SVM	For SVM information gain is 0.9	size of 1021 documents with the Chinese sentiment corpus
B. Gokulakrishnan *et al.* [50]	NB, RF, and SVM	72.7%	microblogging data
E. Boiy *et al.* [51]	NB and SVM	86.35%	2000 movie reviews

I. Gathering data:

For gathering data, there are many options. In some of the previous research papers, a program is built to collect automatically a vast range of tweets using twitter api techniques. While some build their self-made dataset of tweets (by collecting and annotating them manually, which is hectic and tedious). Whereas some used the datasets created by other researchers or organizations to reduce the hassle of collecting data. Additionally, when gathering a corpus of tweets, we are required to have a balanced dataset; we must have an equal number of negative and positive tweets. The larger the data, the more training data will be there for training purpose of classifier and ultimately enhancing accuracy.

II. Analyzing Sentiment of Twitter Data

The sequence of operation for analyzing sentiment is shown in Figure 1. The setup comprises four main stages: data collection module, data processing module, classifier module, and output analysis.

a. Collect data from Twitter: Tweets are of unstructured, semi-structured, and structured types. In Emotion Analysis research work, tweets are collected by implementing numerous programming languages such as Python or R by using the twitter API as well as the datasets from different websites such as GitHub, Kaggle and self collected Databases of tweets.

b. Pre-Processing Data: Data pre-processing is the process by which the data is collected it goes through filtering to eliminate the noisy, incomplete, and inconsistent data. The steps involved in this stage are:

- Stemming
- Removing Punctuations, Special characters, URLs, Numbers, etc.
- Retweets Removal for avoiding repetition
- Tokenization
- Removing Stop words

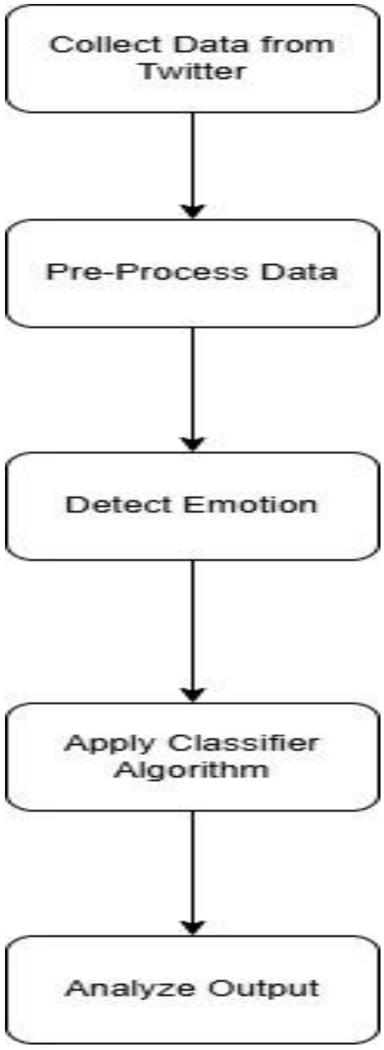

Fig.1: Work Flow for Analyzing Sentiment

c. Detect Emotion:

Identifying feeling words is a crucial job in multiple products of opinion mining and analyzing sentiment, which includes tweet classification and mining tweets. Emotion words are defined in three parts: Negative, Neutral, and Positive words. The cardinal work of Analyzing Emotion is to classify a given tweet's polarity. The polarity can be of three parts, i.e., Negative, Neutral, and Positive. To identify polarity, various lexicons (some examples: Bing Lui sentiment lexicon, Sent WordNet, etc.) are used, which helps in calculating sentiment score, sentiment strength, etc.

Apply Classifier Algorithm:

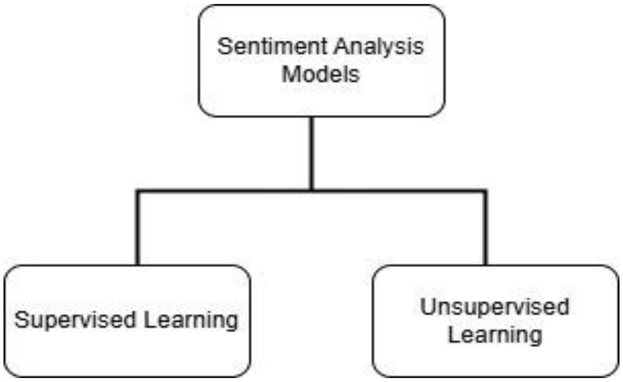

Fig 2: Sentiment Analysis Classifier Algorithms

There are two essential ways for analyzing sentiment, i.e., Unsupervised Learning Approach and Supervised Learning Approach, as shown in Figure 2. By implementing supervised machine learning approaches, which include NB, Maximum-Entropy, SVM, etc., the classification of sentiment on Twitter data is done. Classifier's efficiency depends on which dataset has been used by which classifier. For training a classification model in the Supervised ML approaches, the dataset (training) is utilized to help in the classification of test data.

c. Analyze Output:

The basic thinking behind analyzing sentiment is to transform data into meaningful or understandable data. After completing the analysis, the outcomes are shown and represented in Table 1.

CONCLUSION

For mining opinion or sentiment, Twitter data analysis is performed from diverse points of view. In the paper, opinion mining and the concept of analyzing sentiment are defined concerning various phases in analyzing sentiment. Various approaches and techniques for analyzing sentiment are reviewed within the survey paper. When we are dealing with sentiment analysis of Twitter, it is crucial to understand Twitter's working, concepts, and also about extricating tweets from Twitter, its architecture, and significance. The survey paper provides a brief overview of tweets. So, the requisite knowledge needed for performing the analysis of the sentiment of Twitter is extensively analyzed in the paper. A comparison study of various techniques for analyzing sentiment by working on the Twitter datasets can be found in the survey paper. The examination of literature exhibits the various types for analyzing sentiment and different approaches used to accomplish the extracting of sentiment from the tweets. Here, the study proves that when analyzing, semantic methods are succeeded by ML techniques, the accuracy will be improved. Furthermore, when implementing the Hybrid method, the accuracy will be raised by 5% to 7%.

REFERENCES

1. Roy SS, Roy A, Samui P, Gandomi M, Gandomi AH. Hateful sentiment detection in real-time tweets: An LSTM-based comparative approach. IEEE Trans Comput Soc Syst. 2024;11(4).
2. Braig N, Buettner R, Benz A, Voth S, Breitenbach J. Machine learning techniques for sentiment analysis of COVID-19-related Twitter data. IEEE Access. 2023;10:DOI:10.1109/ACCESS.2023.3242234.
3. Krishna TVS, Krishna TSR, Kalime S, Krishna CVM, Neelima S, Rao PBV. A novel ensemble approach for Twitter sentiment classification with ML and LSTM algorithms for real-time tweets analysis. Indones J Electr Eng Comput Sci. 2024;34(3):1904–14.
4. Vijay K, Samuel P, Krishna BV, Manikandan J. Exploration of sentiment analysis in Twitter propaganda: a deep dive. Multimed Tools Appl. 2024;83:44729–51.
5. Alhashmi SM, Khedr A, Arif I, El Bannay M. Using a hybrid-classification method to analyze Twitter data during critical events. IEEE Access. 2021;9:1023–141035.
6. Wagh R, Punde P. Survey on sentiment analysis using Twitter dataset. In: Proc 2nd Int Conf Electron Commun Aerosp Technol (ICECA); 2018.
7. Kharde VA, Sonawane SS. Sentiment analysis of Twitter data: a survey of techniques. Int J Comput Appl. 2016;139(11).
8. Khan M, Srivastava A. Sentiment analysis of Twitter data using machine learning techniques. Int J Eng Manag Res. DOI:10.5281/zenodo.10791485.
9. Gupta B, Negi M, Vishwakarma K, Rawat G, Badhani P. Study of Twitter sentiment analysis using machine learning algorithms on Python. Int J Comput Appl. 2017 May.
10. Palaniappan V, Mustapha A, Amin R, Pillay KG, Shahir M, Omar S. Twitter sentiment-based topic classification using machine learning algorithms. J Adv Res Appl Sci Eng Technol. Available from: https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index

11. Pereira-Kohatsu JC, Quijano-Sánchez L, Liberatore F, Camacho-Collados M. Detecting and monitoring hate speech in Twitter. Sensors. 2019;19(21):4654.
12. Agarwal S, Chowdary CR. Combating hate speech using an adaptive ensemble learning model with a case study on COVID-19. Expert Syst Appl. 2021;185:115632.
13. Sadiq S, Mehmood A, Ullah S, Ahmad M, Choi GS, On B-W. Aggression detection through deep neural model on Twitter. Future Gener Comput Syst. 2021;114:120–9.
14. Araque O, Iglesias CA. An ensemble method for radicalization and hate speech detection online empowered by sentic computing. Cogn Comput. 2022;14(1):48–61.
15. Alammary AS. Arabic questions classification using modified TFIDF. IEEE Access. 2021;9:95109–22.
16. Raufi B, Xhaferri I. Application of machine learning techniques for hate speech detection in mobile applications. In: Proc Int Conf Inf Technol (InfoTech); 2018. p. 1–4.
17. Roy PK, Tripathy AK, Das TK, Gao X-Z. A framework for hate speech detection using deep convolutional neural network. IEEE Access. 2020;8:204951–62.
18. Imran AS, Daudpota SM, Kastrati Z, Batra R. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. IEEE Access. 2020;8:181074–90.
19. Gupta P, Kumar S, Suman RR, Kumar V. Sentiment analysis of lockdown in India during COVID-19: A case study on Twitter. IEEE Trans Comput Soc Syst. 2021;8(4):992–1002.
20. Cotfas L-A, Delcea C, Roxin I, Ioanas C, Gherai DS, Tajariol F. The longest month: Analyzing COVID-19 vaccination opinions dynamics from tweets. IEEE Access. 2021;9:33203–23.
21. Miao L, Last M, Litvak M. Tracking social media during the COVID-19 pandemic: The case study of lockdown in New York state. Expert Syst Appl. 2022;187:115797.
22. Malla S, Alphonse PJA. COVID-19 outbreak: An ensemble pretrained deep learning model for detecting informative tweets. Appl Soft Comput. 2021;107:107495.
23. Chakraborty K, Bhatia S, Bhattacharyya S, Platos J, Bag R, Hassanien AE. Sentiment analysis of COVID-19 tweets by deep learning classifiers—A study to show how popularity is affecting accuracy. Appl Soft Comput. 2020;97:106754.
24. Mahdikhani M. Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of COVID-19 pandemic. Int J Inf Manage Data Insights. 2022;2(1):100053.
25. Satu MS, Khan MI, Mahmud M, Uddin S, Summers MA, Quinn JMW, et al. TClustVID: A novel machine learning classification model to investigate topics and sentiment in COVID-19 tweets. Knowl Based Syst. 2021;226:107126.
26. Kabir MY, Madria S. EMOCOV: Machine learning for emotion detection, analysis and visualization using COVID-19 tweets. Online Social Netw Media. 2021;23:100135.
27. Behl S, Rao A, Aggarwal S, Chadha S, Pannu HS. Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. Int J Disaster Risk Reduct. 2021;55:102101.
28. Rahman MM, Ali GG, Li XJ, Samuel J, Paul KC, Chong PHJ, et al. Socioeconomic factors analysis for COVID-19 U.S. reopening sentiment with Twitter and census data. Heliyon. 2021;7(2):e06200.
29. Lyu H, Wang J, Wu W, Duong V, Zhang X, Dye TD, et al. Social media study of public opinions on potential COVID-19 vaccines: Informing dissent, disparities, and dissemination. Intell Med. 2022;2(1):1–12.
30. Choudrie J, Patil S, Kotecha K, Matta N, Pappas I. Applying and understanding an advanced, novel deep learning approach: A COVID-19 text-based emotions analysis study. Inf Syst Front. 2021;23(6):1431–65.
31. Ramya BN, Shetty SM, Amaresh AM, Rakshitha R. Smart Simon Bot with public sentiment analysis for novel COVID-19 tweets stratification. Social Netw Comput Sci. 2021;2(3):227.
32. Alsayat A. Improving sentiment analysis for social media applications using an ensemble deep learning language model. Arab J Sci Eng. 2022;47(2):2499–511.
33. Singh M, Jakhar AK, Pandey S. Sentiment analysis on the impact of coronavirus in social life using the BERT model. Soc Netw Anal Mining. 2021;11(1):33.
34. Caliskan C. How does ‘A bit of everything America’ state feel about COVID-19? A quantitative Twitter analysis of the pandemic in Ohio. J Comput Social Sci. 2022;5(1):19–45.
35. Goel R, Sharma R. Studying leaders and their concerns using online social media during the times of crisis: A COVID case study. Soc Netw Anal Mining. 2021;11(1):46.
36. Hoque MU, Lee K, Beyer JL, Curran SR, Gonser KS, Lam NSN, et al. Analyzing tweeting patterns and public engagement on Twitter during COVID-19 recognition period: A study of two U.S. states. IEEE Access. 2022;DOI:10.1109/ACCESS.2022.3189670.
37. Cheng M-Y, Kusoemo D, Gosno RA. Text mining-based construction site accident classification using hybrid supervised machine learning. Autom Constr. 2020;118:103265.
38. Subasi A, Kevric J, Canbaz MA. Epileptic seizure detection using hybrid machine learning methods. Neural Comput Appl. 2017;31(1):317–25.
39. Loey M, Manogaran G, Taha MHN, Khalifa NEM. A hybrid deep transfer learning model for face mask detection during COVID-19. Measurement. 2021;167:108288.
40. Anjaria M, Guddeti RMR. Influence factor-based opinion mining of Twitter data using supervised learning. In: Proc 6th Int Conf Commun Syst Netw (COMSNETS); 2014. p. 1–8.
41. Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision. Stanford Univ Tech Rep CS224N. 2009;1(12).
42. Zhu S, Xu B, Zheng D, Zhao T. Chinese microblog sentiment analysis based on semi-supervised learning. In: Semantic Web and Web Science. Springer; 2013. p. 325–31.

43. Al-Ayyoub M, Essa SB, Alsmadi I. Lexicon-based sentiment analysis of Arabic tweets. Int J Social Netw Mining. 2015;2(2):101–14.
44. Wang X, Wei F, Liu X, Zhou M, Zhang M. Topic sentiment analysis in Twitter: A graph-based hashtag sentiment classification approach. In: Proc 20th ACM Int Conf Inf Knowl Manage (CIKM); 2011. p. 1031–40.
45. Ankit, Saleena N. An ensemble classification system for Twitter sentiment analysis. Procedia Comput Sci. 2018;132:937–46.
46. Melville P, Gryc W, Lawrence RD. Sentiment analysis of blogs by combining lexical knowledge with text classification. In: Proc 15th ACM SIGKDD Int Conf Knowl Discov Data Min (KDD); 2009. p. 1275–84.
47. Ghorbel H, Jacot D. Further experiments in sentiment analysis of French movie reviews. In: Mugellini E, Szczepaniak PS, Pettenati MC, Sokhn M, editors. Adv Intell Web Mastering 3. Springer; 2011. p. 19–28.
48. Singh VK, Piryani R, Uddin A, Waila P, Marisha. Sentiment analysis of textual reviews: Evaluating machine learning, unsupervised and SentiWordNet approaches. In: Proc 5th Int Conf Knowl Smart Technol (KST); 2013. p. 122–7.
49. Tan S, Zhang J. An empirical study of sentiment analysis for Chinese documents. Expert Syst Appl. 2008;34(4):2622–9.
50. Gokulakrishnan B, Priyanthan P, Ragavan T, Prasath N, Perera A. Opinion mining and sentiment analysis on a Twitter data stream. In: Proc Int Conf Adv ICT Emerg Regions (ICTer); 2012. p. 182–8.
51. Boiy E, Moens M-F. A machine learning approach to sentiment analysis in multilingual web texts. Inf Retr. 2009;12(5):526–58.

Ashish Kharb is a final-year M.Tech student in Computer Engineering at UIET, Kurukshetra University. With a B.Tech in the same field (CGPA 8.07), he also gained practical experience at Infosys. His core interests lie in Machine Learning and AI, where he actively engages in projects, research, and continuous learning.