International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 17 • Issue 2 pp. 440-454 March 2026-Sept 2026 www.csjournals.com

Analysis of Negative Public Opinion on Twitter Using Sentiment

Sunaina[1], Dr. Vishal Verma[2], Sandeep Rathi[3]
MCA Student[1], Professor[2], Asst. Prof.[3]
Deptt. of Comp. Sc. and Application, Ch. Ranbir Singh University, Jind, Haryana, India

ABSTRACT

The huge volume of unstructured and expressive text data that comes from the exponential growth of social media platforms, particularly Twitter (now X), can be extremely valuable for businesses, researchers, and politicians. The use of sentiment analysis, a technique that leverages Natural Language Processing (NLP) to automatically classify text based on its sentiment, has become an essential way to extract valuable insights from such a massive amount of data.

The purpose of this research is to introduce a machine learning-based binary classification system to tackle the problem of detecting positive sentiment in tweets. The suggested approach simplifies the problem to a binary one: either positive or not positive, rather than trying to do a difficult multiclass classification procedure. This specific formulation minimizes computational overhead and will have direct applications in areas such as customer feedback analysis, and brand monitoring.

Three supervised machine learning algorithms (Multinomial Naive Bayes, Support Vector Machine (Linear SVM) and Logistic Regression) are developed and rigorously analyzed using the Sentiment140 dataset (1,600,000 tweets). TF-IDF feature extraction is a systematic pre-processing of the data set that involves the elimination of stop words, the cleaning of the text with regular expressions and the application of Porter stemming. The models trained on an 80-20 stratified train-test split are evaluated based on accuracy, precision, recall and F1-score.

Based on the experimental results, it is observed that Multinomial Naive Bayes, Linear SVM and Logistic Regression show the classification accuracy of 77.43%, 79.54%, 80.14% respectively with the F1-scores of 0.7637, 0.7977, 0.8045 respectively. The comparison analysis reveals that the discriminative linear models outperform the simpler and more traditional probabilistic Naive Bayes model on high-dimensional sparse representations of Twitter text using TF-IDF.

International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 17 • Issue 2 pp. 440-454 March 2026-Sept 2026 www.csjournals.com

INTRODUCTION

The digital platforms are not one and the same. Twitter is a popular microblogging service that allows people to post short status updates (also known as "tweets") that typically include thoughts, emotions, and reactions to products, events, institutions, and public figures [1]. There is a never-ending stream of tweets.Tweets, tweets, tweets, tweets, tweets, tweets, tweets, tweets, tweets.

Opinion mining or sentiment analysis (also referred to as sentiment classification) is a part of Natural Language Processing (NLP) that relies on automated techniques to classify the emotionality of a text into positive, negative or neutral sentiment [2]. Previously, sentiment analysis was applied in longer text such as news articles and product reviews, but it is now increasingly applied to shorter text, such as tweets and microblogging. Breviations of breviations of breviations of the use of breviations of breviations of breviations of breviations of the use of the breviations of the.

NLP is used to analyze the structure, meaning, and context of text, aiding a robot's ability to understand and generate human language. In NLP, there are three aspects that are commonly identified: pragmatics is how language is used and understood in a particular context; semantics is what is literally said; and syntax is the structure of words into grammatically correct sentences [3].

In this paper, attention is paid to examine the use of supervised machine learning models to detect positive sentiment in tweets. The study does not try to classify the problem as a multiclass class problem (positive class, negative class and neutral class) but reduces it to a binary classification problem (positive class vs non positive class). This formulation maintains direct actionability in the applications such as brand monitoring or consumer happiness tracking, and at the same time lowers the complexity of the model and the cost of computation.

1.1.1Machine Learning Introduction

Machine Learning (ML) is an area of Artificial Intelligence (AI) that allows computer systems to self-learn from data and get better at some tasks when it has not been programmed. In supervised learning, as used in this study, an algorithm is trained on a labelled data set, where the algorithm learns to translate input features to labels. Then the method is applied to classify the new data that is unknown.

In text classification, like sentiment analysis, text is typically represented using methods like Bag-of-Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), or word embeddings (Word2Vec, GloVe) and then used by machine learning algorithms.

Types of Machine Learning

Machine learning can be categorized into three types:

1. Supervised Learning

2. Unsupervised Learning

3. Reinforcement Learning

1.1 Research Objectives

- To correctly mark and label negative tweets in the Twitter data.
- To analyze the negative perception of events, products and social issues through the application of machine learning tools.
- To process difficult situations such as sarcasm, slang, abbreviations, short text in tweets using ML techniques.
- To make more accurate and efficient prediction.

1.2 Related Work

The study of distant supervision was conducted by Go, Bhayani & Huang (2009).Go, Bhayani & Huang (2009) did the study on distant supervision.

This study examines how brands can be evaluated for reputation.In this study, the authors explore the assessment of brands for reputation.

Agba et al. [6] used support vector machine (SVM) to classify the sentiment of consumers towards a skincare brand on Twitter. The results reveal a positive sentiment trend, with about 69% of the samples reporting positive attitudes towards the brand being studied. In this report, the research direction directly relates to data from this work, which illustrates the feasibility of binary or near-binary sentiment classification for brand monitoring in the specific context of the domain.

Rasool et al. [7] explored the difficulties in analyzing consumer opinions on Twitter for apparel brands, given the vast amount of information available. The study emphasized that automatic extraction of feedback and applied the Naive Bayes classification algorithm and a lexicon dictionary method. The authors emphasized the need for automated solutions due to the volume and velocity of the Twitter data and the benefits of using multiple classification strategies to increase overall robustness.

Lin (2017) — Enhanced Sentiment Analysis using Custom Language Models Lin [8] talked about the vast amount of data available on social media like Twitter, and how to use that information to figure out who are influential users to help businesses make better decisions. The study proposed a Naive Bayes model with an additional custom-built language model, and showed that it outperforms the baseline language model classifiers relying on standard lexicon words, suggesting the importance of integrating probabilistic classifiers with domain-specific language modelling.

1.3 Dataset and Datasources

B. Reading and Explanation

The Twitter Sentiment Analysis Dataset, originally made public on the Kaggle website and based on the Sentiment140 dataset created by Stanford University researchers, was used in this research. This dataset comprises of ~1.6 million tweets, which were automatically labeled by distant supervision methods with emoticons, without manual labeling.

Several important reasons justified the choice of this secondary dataset: large scale, public accessibility, wide use in research on sentiment analysis, and appropriateness for text classification using machine learning. The data set is a realistic sample of social media user-generated data, including non-standard linguistic features like slang, abbreviations, hashtags, emojis and contextually appropriate expressions that can be commonly found in the Twitter social media environments. This means it is very suitable for public opinion analysis and negative sentiment detection.

Because of restrictions with the Twitter API, time constraints and data availability policies, primary data collection via the Twitter API was not taken into account for this study. Thus, reliability, reproducibility and consistency can be achieved in the experimental evaluation using a validated and widely accepted benchmark dataset. Previous sentiment analysis studies have widely used such similar datasets, which allows us to make comparative performance analysis.

B. Dataset Description

The dataset used in this research contains of approximately 1,600,000 Twitter posts (tweets) collected from public Twitter users. The data was preprocessed and cleaned before being utilized for sentiment classification of positive and negative sentiment. Sentiment labels and related textual data for machine learning models training/evaluation is included in each tweet record.

The original data set comprises six attributes, the tweet text being the main input attribute and the sentiment label being the target attribute. Before training the model, the text was preprocessed by removing stop words, stemming, cleaning the text, and using the TF-IDF vectorization method to perform feature engineering.

Table I presents the description of the dataset attributes used in this study.

TABLE I

DESCRIPTION OF DATASET ATTRIBUTES

Attribute Name	Description	Data Type
Target	Sentiment label (0 = Negative, 4 = Positive)	Integer
IDs	Unique identifier assigned to each tweet	Integer
Date	Date and time of tweet posting	DateTime
Flag	Query information or metadata	String
User	Twitter username of the author	String
Text	Original tweet content used for sentiment analysis	

The final experimental data set was divided into training and testing data sets in the conventional manner of machine learning techniques. The tweet dataset is large enough to be used for diverse and reliable performance evaluation of various sentiment classification algorithms like Logistic Regression, Support Vector Machine (SVM), Multinomial Naive Bayes.

1.4 METHODOLGY

A. Research Framework

This study uses a typical supervised machine learning approach for analyzing sentiment in Twitter data. The proposed framework is divided into five major steps: data acquisition, data preprocessing, feature extraction, model development, and the evaluation of the performance. The first step was to gather a public sentiment data set from Twitter, and prepare it for analysis. The tweets were then subjected to text preprocessing techniques to clean up and standardize the content of tweets. The textual data has been cleaned and then the numerical representation has been obtained by the TF-IDF feature extraction method. Various machine learning algorithms like Logistic Regression (LR), Linear Support Vector Machine (SVM) and Multinomial Naive Bayes (MNB) were trained and tested with the processed data set. Last, the models were tested

based on traditional performance measures and statistical analysis techniques for their ability to capture negative public opinions.

The sequence of proposed research framework is shown below:

Data is collected and then preprocessed, followed by feature extraction using TF-IDF and with the help of this data a model is trained and tested, and finally the performance of the model is evaluated and statistical analysis is performed.

B. Preprocessing Pipeline

The informal language, abbreviations, emojis, URLs, mentions, and hashtags found in Twitter data make it inherently noisy and unstructured. Thus, a detailed data preprocessing pipeline was developed to enhance the quality of textual data prior to model training.

The pre-processing steps are:

1. Data Loading
2. Text Cleaning
3. Lowercase Conversion
4. Tokenization
5. Stop-word Removal
6. Stemming/Lemmatization
7. Feature Extraction
8. Train-Test Split

This is a pre-processing pipeline which can be used to reduce dimensionality, eliminate noise and enhance classification accuracy.

C. Machine Learning Models

In this research, three supervised machine learning algorithms are used to classify the sentiment and detect the negative opinions.

1. Logistic Regression (LR)

Logistic Regression is a probabilistic classification algorithm which is frequently used in binary classification problems. It is predicting whether a tweet will fall into one of the two classes (positive or negative sentiment). Logistic Regression is a simple, efficient and easily interpretable model and is thus used as a benchmark for sentiment analysis.

2. Linear Support Vector Machine (Linear SVM)

An SVM with a linear classifier is a unique kind of classifier that produces an optimal separating hyperplane that has the largest margin between the different classes of sentiments. It is useful on

high dimensional text classification applications and has shown good performance on sentiment analysis applications.

3. Multinomial Naive Bayes (MNB)

Multinomial Naive Bayes is a classification algorithm that uses naive Bayes techniques, but is most suited to text classification problems. It derives the probability distribution of words in documents, and classifies sentiments through Bayesian principles. It is very efficient in computation, and hence is widely used in large-scale sentiment analysis studies.

Confusion Matrix

The classification performance was then visualised using a confusion matrix, where True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) were shown.

Furthermore, model comparison was carried out based on the classification accuracy, computational efficiency, and prediction consistency.

E. Statistical Hypothesis Tests

The obtained results were tested using the statistical hypothesis testing techniques to verify the significance of the results. These tests are used to determine if there are statistically significant differences in performance between the machine learning models, or if the differences are due to chance. The following hypotheses were proposed:

Null Hypothesis (H_0)

There is no significant difference between the performance of Logistic Regression, Linear SVM and Multinomial Naive Bayes models in predicting the sentiment of tweets.

H0:μLR=μSVM=μMNBH_0:\mu_{LR}=\mu_{SVM}=\mu_{MNB}H0:μLR=μSVM=μMNB

Alternative Hypothesis (H_1)

At least one of the sentiment classification models performs significantly different.

H1:μi≠μjH_1:\mu_i\neq\mu_jH1:μi□=μj

Depending on the experimental design, statistical significance tests, like the Paired t-test, McNemar's Test or Analysis of Variance (ANOVA) may be used to evaluate results of the experiments. A significance level of $\alpha = 0.05$ was considered for hypothesis testing. The calculated p-value is smaller than 0.05, null hypothesis is rejected, significant statistical differences among the models evaluated.

International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 17 • Issue 2 pp. 440-454 March 2026-Sept 2026 www.csjournals.com

1.5 Results

Class Distribution and Balancing

The results of experiments conducted with the Sentiment 140 data on three machine learning classifiers (Logistic Regression, Linear SVM, Multinomial Naive Bayes) are presented in this chapter. Every result in this chapter has been obtained from the execution of the model on the test set of 320,000 tweets that were not used for training the model.

5.1 Performance Metrics

The performance of the models were assessed using four common classification metrics:

Accuracy: Percentage of correctly classified instances divided by the number of instances.

Accuracy = (TP + TN) / (TP + TN + FP + FN).

Precision: The percentage of correct predictions of a positive outcome given that there was a positive outcome. Precision = TP / (TP + FP). Few false positives, indicating high precision.

Recall (Sensitivity): Percentage of the correct identification of all the true positive samples. Recall = TP / (TP + FN). A high recall means that there are few false negatives.

Precision x Recall / (2 x Precision x Recall) = F1-Score. In this study, the F1 score is used as the key measure to compare since it takes into account both precision and recall with a class balanced dataset.

TP = True Positives, TN = True Negatives, FP = False Positives and FN = False Negatives.

5.2 Results

5.2.1 Logistic Regression Results

To be trained for a maximum of 400 iterations on TF-IDF features, Logistic Regression reached an accuracy of 80.14% and an F1-score of 0.8045 on the held-out test-set of 320,000 tweets.

The logistic regression classification report for the held-out test set is shown below.Logistic regression classification report for the held-out test set is presented below.

Logistic Regression showed good results for both classes, with a slightly higher recall for the positive class (0.82) compared to the not-positive class (0.79). This means that the model is slightly more likely to classify the positive class when the class separation is not very clear, and has almost equal precision (0.81 vs. 0.79) for the two classes.

5.2.2 Linear SVM Results

The Linear Support Vector Machine (SVM) classifier had an accuracy of 79.54% and an F1 score of 0.7977 on the same held-out test set.

Linear SVM was competitive with LR, with an accuracy of around 0.6 percentage points lower. This is consistent with the theoretical relationship between these two algorithms: both are linear, discriminative classifiers directly fitted on the high-dimensional TF-IDF feature space, but differ in the loss function (Logistic Regression: log-loss, Linear SVM: hinge loss) and the optimisation criterion for the decision boundary (Linear SVM: maximum margin, Logistic Regression: maximum likelihood).

5.2.3 Results of Multinomial Naive Bayes

The lowest of the three models tested was the Multinomial Naive Bayes with an accuracy of 77.43% and an F1 score of 0.7637 on the held-out test set.

The fact that Naive Bayes is not as effective as the other classifiers is actually expected; one of its key assumptions is that given the class label, features are conditionally independent, which is not true of natural language text, where the relationship between pairs of words reveals important sentiment information that Naive Bayes cannot directly capitalize on. However, as noted by Go et al. [1] and confirmed here, Naive Bayes is still a fairly good and extremely compute-light baseline model for high dimensional text classification.

5.3 Comparative Analysis

The comparative results show that the performance is in the following order: Logistic Regression > Linear SVM > Multinomial Naive Bayes. Logistic Regression is 0.6% more accurate than Linear SVM, and 2.7% more accurate than Multinomial Naive Bayes. Both discriminative linear models (Logistic Regression and SVM) outperform the generative, probabilistic Naive Bayes classifier, as their relative performance was reported by Go et al. [1] with a smaller manually labelled test set, where SVM and Logistic-Regression based classifiers outperformed Naive Bayes on a similar distantly supervised Twitter corpus.

All three models (a difference of only 2.71 percentage points in accuracy) were not far behind each other, meaning that the selection of a particular linear or probabilistic classification algorithm was less important than the choice of TF-IDF features and the pipeline for preprocessing the data, for the dataset and the feature representation used in this case. This conclusion is consistent with the general literature, where it has been concluded that after a good feature representation is developed, there is not much benefit to be gained from choosing different algorithms [14].

5.4 Discussion of Findings

The experimental results given in this chapter indicate several general observations:

This is where discriminative models outperform generative models: Logistic Regression and Linear SVM models, which explicitly model the class separation surface, also performed better than Multinomial Naive Bayes, which assumes a strong conditional independence between features and models every class feature distribution separately.

As the Sentiment 140 data is perfectly balanced (50% positive, 50% not positive) and stratified across the train-test split, the accuracy is a reliable and not misleading measure in this study as compared to several real world imbalanced sentiment datasets.

The preprocessing pipeline (regex cleaning, stop word removal, Porter stemming) and the features matrix (TF-IDF) were the same for all three classifiers, so there is no difference in the features representation that would explain the differences in performance that were observed.

The models differ in terms of their accuracy; Naïve Bayes is the least accurate, but is also the fastest to train, requiring only a fraction of the time of Linear SVM on a dataset of this size. The compromise between accuracy and CPU usage becomes significant when deciding which model to use in a resource-limited or real-time application.

Headroom for improvement: An accuracy ceiling of around 80% for all three classical models is typical when dealing with the known limitations of bag-of-words / TF-IDF representations, and their inability to model word order, scope of negation, or contextual meaning. This sparks us to consider the transformer-based models in the future scope of Chapter 6.

Overall, the results presented in this chapter reveal that Logistic Regression model is the most accurate model among the three models tested for the sentiment classification of Twitter, with all three models showing the practicability of the proposed binary classification method for the Sentiment140 data set.

1.6 Discussion

There is a selection of machine learning algorithms used in this study because of their successful performance in text classification and sentiment analysis problems. Logistic Regression (LR), Linear Support Vector Machine (Linear SVM), and Multinomial Naive Bayes (MNB) were selected as the supervised machine learning algorithms because they are interpretable, efficient in computation and have been found to be successful in high dimensional text data.

Logistic Regression was selected as a baseline classification model as it is simple, fast to train and make probabilistic outputs. Works well with low dimensional representations like TFIDF-vectors.

Linear Support Vector Machine (Linear SVM) was added since it has outperformed other classifiers in text classification tasks in all situations. It is well suited for sentiment classification tasks with large textual data sets due to its capacity of creating the best separating hyperplane in high dimensional feature spaces.

In view of the probabilistic nature and the computational efficiency in word frequency distributions, multinomial Naive Bayes (MNB) was chosen as the model. Although it is assumed to be feature independent, it is still one of the most popular sentiment analysis algorithms.

These models' performances were compared using various performance measures such as accuracy, precision, recall, and F1-score. The results of the experiments show that the performance of the models is different in identifying negative public opinions, because of the difference in their learning mechanisms.

Overall, Linear SVM has a better classification performance as it can efficiently handle high dimensional sparse data. Logistic Regression is a method that performs well compared to other methods and has a high interpretability, while Multinomial Naive Bayes is a method that has a short calculation time but may have problems with the contextual relationships that exist in social media texts.

Comparative analysis of more than one machine learning algorithms can be used to choose the most appropriate model for large scale sentiment classification and negative opinion detection from Twitter.

B. Feature importance and Clinical significance.

Due to the nature of machine learning models, they can't directly process the raw textual information, and feature extraction is a crucial part of sentiment analysis. In this study, text Tweets were converted into numerical feature vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) method, while retaining the relative importance of words related to sentiment.

TF-IDF “weights” words that are more common in a single tweet than in the entire collection of tweets. This allows the machine learning models to detect key terms that are linked to negative sentiment expressions, such as dissatisfaction, criticism, complaints, anger, fear, and frustration.

Feature importance analysis showed that some sentences containing sentiment specific keywords, emotional expressions and contextual keywords are having significant impact on the classification performance. Negative emotion-related words are often given higher TF-IDF values, making the models more effective to distinguish public opinion with negative emotions.

Practical implications of the findings of this study exist in numerous fields. Sentiment analysis systems can be used to measure customer dissatisfaction, assess the sentiment of users of their products or services, analyze political opinions, track social concerns, and detect emerging

concerns in real time. Auto-detecting negative public sentiment can be beneficial when making business decisions, handling reputation issues, monitoring for crises, and assessing policies.

Moreover, the research resulted in verifying the effectiveness of Machine Learning based Sentiment Analysis algorithm for large-scale social media data, laying the groundwork for future advancements in real-time sentiment analysis, multilingual sentiment analysis, and enhanced deep learning methods.

C. Limitations

There are several limitations that should be noted in the present study:

The Sentiment140 training set was collected from 2009, and the vocabulary and slang of Twitter have changed significantly over the years. Models trained on this corpus may not be applicable to modern tweets.

The neutral sentiment class is omitted: This study does not consider a neutral class, which makes it more difficult to apply directly to tasks that need to classify sentiment across all three classes (positive/negative/neutral).

No dedicated mechanism for sarcasm/irony in the preprocessing and modelling pipeline: Sarcasm and irony are known areas of misclassification in sentiment analysis in general and in Twitter data in particular, but there is no dedicated mechanism to deal with it in the preprocessing and modelling pipeline.

1.7 Conclusion

In this report, we have presented the design, implementation, and evaluation of a machine learning based system for sentiment classification in tweets based on the binary classification approach, that classifies the tweets as positive or negative. The study tackled the particular challenges of informal language on Twitter, while at the same time designing a clean, reproducible experimental pipeline for the Sentiment140 dataset of 1.6 million distantly supervised tweets.

The main results achieved in this work are as follows:

A binary classification framework for Twitter sentiment analysis that makes the traditional multiclass problem to one of positive versus not positive to reduce the model complexity without sacrificing its applicability to real-world scenarios like brand monitoring, customer feedback analysis, etc.

A fully reproducible preprocessing pipeline, optimized for the informal nature of Twitter text, with preprocessing steps of regular-expressions to clean the text, stop-word removal, and Porter stemming.

The three supervised machine learning algorithms (Logistic Regression, Linear SVM, and Multinomial Naive Bayes) are trained and tested on the same TF-IDF feature representation, and on an identical train-test split by 80:20 stratified sampling, so that they can be directly and fairly compared.

The main result of the research is that the approach of supervised machine learning with TF-IDF feature extraction seems to be a reliable and computationally efficient method for binary sentiment classification on Twitter data, which provides a basis for the central hypothesis. All three classifiers achieved over 77% accuracy on a held-out set of 320 thousand tweets, confirming the potential of the proposed method to be used in practical applications such as brand sentiment monitoring, product feedback analysis, and mass opinion monitoring.

Future Work

The findings of this research have several avenues for future research:

Deep Learning and Transformers Models..

One of the most viable options for enhancing the classification accuracy over this ~80% threshold identified in this study is the use of a language model like Bidirectional Encoder Representations from Transformers (BERT) or RoBERTa that has been pretrained on a large corpus of unstructured data. These models learn word semantics in context and have demonstrated state of the art performance on the majority of NLP benchmarks. Additionally, the fine-tuning of a pre-trained transformer model using the Twitter sentiment data may further enhance the accuracy and robustness to sarcasm which are compared to the classical TF-IDF based models used in this study. Real-Time Sentiment Monitoring System

Current system is based on a dataset that is collected and is static. A useful practical extension might be an actual sentiment monitoring pipeline, consuming live tweets from the Twitter API, and visualizing sentiment trends on a live dashboard and applying the trained classifier. It will allow applications like real-time tracking of brand sentiment during product launches, and live public opinion during major events.

Identify sarcasm and irony.

One of the most challenging open problems of sentiment analysis is the detection of sarcasm. One of the major limitations found in Section 6.2 is the lack of a dedicated sarcasm-detection module in the pipeline prior to the sentiment classification. This could be incorporated in the future as a pre-processing stage before the sentiment classification.The key limitation that was identified in Section 6.2 is the absence of a dedicated sarcasm-detection module in the pipeline before the sentiment classification. This could be implemented in the future as a pre-processing step before sentiment classification.

Multiclass classification task and a fine-grained sentiment classification task as well.

The current binary classification could be expanded into a three-class (positive/negative/neutral) or fine-grained emotion classification (e.g., joy, anger, sadness, fear) to give more insightful and detailed information in customer experience analysis, sentiment analysis in politics, and social media monitoring for public health.

Multilingual Sentiment Analysis

This is limited to English language tweets. It is worth noting that the proposed approach would significantly expand the applicability of global content monitoring use cases across social media if it was enhanced to support multilingual content, for instance by using multilingual transformer models like mBERT or XLM-RoBERTa.

Word Embeddings and Big Data Integration.

Use of dense word embeddings such as Word2Vec, GloVe, or contextual embeddings, could help to improve the ability of the model to capture semantic similarity between related sentiment terms. This pipeline could be expanded to much larger tweet datasets, real-time, by integrating with distributed big data frameworks like Apache Spark alongside the examples used here.

References

[1]. Go, R. Bhayani and L. Huang, "Twitter Sentiment Classification using Distant Supervision," CS224N Project Report, 2009, Stanford University.
[2]GeeksforGeeks, Natural Language Processing (NLP), [Online]. Available: https://www.geeksforgeeks.org/nlp/introduction-to-natural-language-processing-nlp/. [Accessed: Jun. 2026].
[3] Naive Bayes Classifiers, GeeksforGeeks, [Online]. Available: https://www.geeksforgeeks.org/machine-learning/naive-bayes-classifiers/. [Accessed: Jun. 2026].
[4]GeeksforGeeks, Understanding Logistic Regression [Online]. Available: https://www.geeksforgeeks.org/machine-learning/understanding-logistic-regression/. [Accessed: Jun. 2026].
[5]Random Forest Algorithm in Machine Learning, [Online]. [5] Available at: https://geeksforgeeks.org/random-forest-algorithm-in-machine-learning/. Available: https://www.geeksforgeeks.org/machine-learning/random-forest-algorithm-in-machine-learning/. [Accessed: Jun. 2026].
[6] M. W. Agba et al., "Memanfaatkan Analisis Sentimen Twitter untuk Mengelola Reputasi Merek," vol. 4, no. 3, Aug. 2025, doi: 10.38035/jim.v4i3.1196.
[7] A. Rasool, R. Tao, K. Marjan, and T. Naveed, “Twitter Sentiment Analysis: A Case Study for Apparel Brands,” J. Phys.: Conf. Ser., 2015, doi: 10.1088/1742-6596/1176/2/022015.
[8] A. Lin, "Improving Twitter Sentiment Analysis using Naive Bayes and Custom Language Model," Computation and Language, Nov. 2017.

[9] M. Z. Zakaria, T. B. Kurniawan, Misinem, "Twitter Sentiment Analysis on Automotive Companies," 2022, doi: 10.61453/jods.v2022no06.
[10] Y. Bae and H. C. Lee, “Sentiment Analysis of Twitter Audiences: Measuring the Positive or Negative Influence of Popular Twitterers,” Korea University, 2012, doi: 10.1002/asi.22768.
[11] G. R. Patil, R. Shrivastava and A. Manhar, "Twitter Sentiment Analysis," Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol. (IJSRCSEIT), vol. 9, no. 3, pp. 520–528, May–Jun. 2023, doi: 10.32628/CSEIT23903117.
[12] The business insights and consumer behaviour trends in the USA are indicated by the sentiment analysis of social media data, as stated by M. A. Al Montaser et al., “Sentiment Analysis of Social Media Data: Business Insights and Consumer Behavior Trends in the USA” (2018) in Edelweiss Appl. Sci. Technol., vol. 9, no. 1, pp. 545–565, 2025, doi: 10.55214/25768484.v9i1.4164.
[13] H. Thakkar, D. Patel, “Approaches for Sentiment Analysis on Twitter: A State-of-the-Art Study”, National Institute of Technology Surat, India, 2016. [Online]. Available: arXiv:1512.01043.
[14] Opinion Mining and Sentiment Analysis" (Found., 2014). Trends Inf. Retr., vol. 2, no. 1–2, pp. 1–135, 2008.
[15] Thumbs Up? Bangaru Pillai B., Lee L., Vaithyanathan S., A discussion of Sentiment Classification Using Machine Learning Techniques can be found in Proc. EMNLP, 2002, pp. 79–86.