



Sentiment Analysis of Twitter Data

**Mr. Peesala Ilanna¹, Pamulaparthi Sai Nithin², Muniganam Akshitha³, Godishala Venkataramana⁴,
Gundoju Abhiram⁵, Bodiga Sai Charan⁶**

Assistant Professor, Department of CSE (Data Science), ACE Engineering College,

Hyderabad, Telangana, India

III B.Tech. Students, Department of CSE (Data Science), ACE Engineering College,

Hyderabad, Telangana, India.

How to Cite this Article:

Nithin, P. S., Akshitha, M., Venkataramana, G., Abhiram, G. & Charan, B. S. (2026). Sentiment Analysis of Twitter Data. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04).

<https://doi.org/10.55041/ijcope.v2i4.200>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i4.200>

Abstract

Nowadays, social media platforms like Twitter are widely used by people to express their opinions on different topics such as products, politics, and social issues. Every day, a large amount of data is generated in the form of tweets, which are usually short and unstructured.

Analyzing this data manually is very difficult and time-consuming. So, in this project, we use sentiment analysis to automatically identify whether a tweet is positive, negative, or neutral.

First, the data is cleaned by removing unwanted elements like special characters, URLs, and stop words. Then, the text is converted into numerical form using TF-IDF technique.

After that, machine learning algorithms are applied to classify the tweets into different sentiment categories. This system helps in understanding public opinion easily and can be useful for businesses and researchers.



INTRODUCTION

Sentiment analysis is a technique used to understand emotions or opinions from text data. In simple terms, it tells whether a piece of text is positive, negative, or neutral.

Twitter is one of the most popular platforms where people share their thoughts in real time. These tweets contain valuable information about public opinion. However, tweets are usually short and include slang, emojis, and informal language, which makes analysis a bit difficult.

In this project, we try to build a system that can analyze tweets and identify their sentiment. The main idea is to process the tweet data and classify it into different categories.

This kind of system can be useful in areas like market analysis, social research, and decision-making.

EXISTING SYSTEM

The existing systems for sentiment analysis mainly depend on basic machine learning techniques.

- They use simple preprocessing methods
- They apply models like Naive Bayes and SVM
- They mainly focus on text-based analysis
- They ignore user behavior and metadata information
- They struggle with short and noisy text like tweets
- They are sensitive to data imbalance problems

Limitations

- Cannot understand sarcasm properly
- Poor handling of emojis and slang
- Limited contextual understanding
- Lower accuracy in real-world data

LITERATURE SURVEY

The following table presents a summary of previous research works related to sentiment analysis of Twitter data, including the techniques used and their limitations.

Title Name	Authors	Technologies Used	Limitations	Year
Twitter (X) as an Electoral Barometer: Systematic Evidence from Sentiment Analysis of Twitter Data	M. Asokere, A. Wusu, O. Olabanjo	Machine Learning, NLP, Sentiment Classification	Dataset bias and limited validation with real election results	2025
Twitter Sentiment Analysis Using Machine Learning and Deep Learning Techniques	Safa Khalid Al Sarayrah et al.	SVM, Logistic Regression, LSTM, Deep Learning	Difficulty handling sarcasm and emojis; high computational cost	2025
Sentiment Analysis: A Hybrid Approach on Twitter Data	C. G. Maurya	Hybrid Machine Learning, Opinion Mining, NLP	Complex processing pipeline and difficult implementation	2024



BERT Model Adoption for Sarcasm Detection on Twitter Data	T. Javed et al.	BERT Transformer Model	Limited accuracy in sarcasm detection	2024
A BERT Framework for Sentiment Analysis of Tweets	A. Bello	BERT, CNN, BiLSTM	Requires high computational resources and training time	2023
Improving Sentiment Classification Using a RoBERTa-Based Hybrid Model	N. A. Smary et al.	RoBERTa, LSTM, CNN Hybrid	Requires large amount of training data	2023
Sentiment Analysis Based on User Opinions on Twitter Using Machine Learning	Jorge Cordero, José Bustillos	TF-IDF, Naïve Bayes, SVM	Limited contextual understanding of text	2022
Sentiment Analysis on COVID Tweets Using COVID-Twitter Data	H. Y. Lin et al.	Deep Learning, NLP, Text Classification	Domain-specific and limited generalization	2021

PROPOSED SYSTEM

In this project, we propose an improved sentiment analysis system using machine learning techniques.

- The system collects Twitter data
- It preprocesses and cleans the data
- It converts text into numerical form using TF-IDF
- It uses machine learning algorithms for classification

Key Features

- Better accuracy compared to existing systems
- Efficient handling of large data
- Automatic sentiment classification
- Easy to implement and use

METHODOLOGY

- In this project, the first step is collecting the dataset containing tweets with different sentiments.
- After collecting the data, we clean it by removing unwanted elements such as special characters, URLs, and stop words. This step is important because raw data usually contains noise.
- Next, we convert the text into numerical form using TF-IDF, since machine learning models cannot understand text directly.
- Then, the dataset is divided into training and testing sets. The model is trained using training data and tested using unseen data.
- Finally, the model predicts the sentiment of tweets and classifies them into positive, negative, or neutral.



MODEL EVALUATION

The performance of the model is evaluated using the following metrics:

- **Accuracy** → Overall correctness
- **Precision** → Correct positive predictions
- **Recall** → Ability to detect actual positives
- **F1-score** → Balance between precision and recall

The model shows good performance in classifying tweet sentiments.

RESULT

The developed system successfully performs sentiment analysis on Twitter data and presents the results in a visual format using a dashboard.

The output includes multiple visualizations that help in understanding the sentiment patterns clearly.

Figure: Twitter Sentiment Analysis Dashboard:





CONCLUSION

In this project, we developed a sentiment analysis system for Twitter data using machine learning.

The system processes tweets and classifies them into different sentiment categories. It helps in understanding public opinion in an easy way.

This project shows how machine learning and NLP can be applied in real-world problems like social media analysis.

FUTURE ENHANCEMENT

- Improve sarcasm detection
- Use deep learning models like LSTM and BERT
- Perform real-time sentiment analysis
- Include emoji-based sentiment analysis

ACKNOWLEDGMENT

This section is optional. Acknowledge funding agencies, institutions, or individuals who contributed to the work but are not listed as authors.

REFERENCES

- Bing Liu, Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, 2012.
- Bo Pang and Lillian Lee, "Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval, 2008.
- Alec Go, Richa Bhayani, and Lei Huang, "Twitter Sentiment Classification using Distant Supervision," Stanford University, 2009.
- Jurafsky, D., & Martin, J. H., Speech and Language Processing, Pearson Education, 2019.
- Basics of Python programming: <https://www.python.org/>
- Dataset: Twitter Dataset was extracted from Kaggle.
- Natural Language Toolkit: <https://www.nltk.org/>
- Random Forest Model: https://en.wikipedia.org/wiki/Random_forest