International Journal of Scientific Research and Engineering Development-– Volume 8 Issue 2, Mar-Apr 2025

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

# MoodScope An AI-Powered Exploration of Twitter Sentiments Using Frameworks

Mrs.S.Surya, Sr Yarlagadda Reshma Chowdary, Challa Pavani Ratna Sri Santhoshi, Chintala Sai Rama Krishna, Balasadi Chandu Praveen, Aithabathula Jeswanth Sai

Emails: suryasri.s@pragati.ac.in, yarlagaddareshmachowdary302@gmail.com, challasanthoshi86@gmail.com, sairamakrishna882@gmail.com, chandhupraveen123@gmail.com, aithabathulatensing@gmail.com

----------**********----------

## Abstract:

Sentiment analysis, the computational study of opinions, sentiments, attitudes, and emotions expressed in written language, has gained prominence in recent years due to its wide applications and challenging research problems. This paper focuses on evaluating sentiment analysis using Logistic Regression and Support Vector Machine (SVM) algorithms. The analysis was performed with metrics including accuracy (0.941), precision (0.615), recall (0.466), and F1 score (0.530). The dataset comprises 319 true positives (TP), 8705 true negatives (TN), 200 false positives (FP), and 365 false negatives (FN). Sentiment analysis has evolved alongside the growth of social media, becoming integral to various fields beyond computer science, including management and social sciences. The paper explores mainstream sentiment analysis research and recent advancements in modeling comments, discussions, and debates, highlighting the algorithms' effectiveness in capturing nuanced sentiments and opinions.

*Keywords* — **Logistic Regression, Sentiment Analysis, Support Vector Machine (SVM), Accuracy, F1 Score, Opinion Mining, Precision, Recall**

----------**********----------

## I. INTRODUCTION

The advancements in machine learning, sentiment analysis has evolved to achieve higher accuracy in understanding opinions and emotions expressed online. Several algorithms, including Logistic Regression, Support Vector Machines (SVM), Decision Trees, and Gradient Boosting, have been employed to classify sentiment effectively. Identifying influential users on social media further enhances the value of sentiment analysis, as it helps businesses and policymakers understand public perception and target key opinion leaders.[1]

This research explores sentiment analysis using machine learning techniques while simultaneously identifying influential users based on engagement metrics such as retweets and follower counts. A case study on airline-related tweets demonstrates the effectiveness of various machine learning models in classifying sentiments and ranking users by influence. The findings contribute to the growing body of knowledge on social media analytics and offer practical implications for organizations seeking to optimize their digital strategies. [2]

The rise of social media has significantly transformed digital communication, allowing users to express their opinions and share information in real time. Among the various platforms, Twitter stands out as a major source of public sentiment, where users engage in discussions on diverse topics such as politics, business, entertainment, and customer experiences. With the increasing volume of tweets generated daily, analyzing user sentiments and identifying influential figures has become

  

crucial for businesses, researchers, and policymakers. [3]

By combining sentiment analysis with influencer identification, this study contributes to the growing field of social media analytics, providing valuable tools for businesses, marketers, and researchers to better understand digital interactions and leverage public opinion for decision-making.[4]

## II.LITURETURE SURVEY

[1] Essaidi, Abdessamad, Dounia Zaidouni, and Mostafa Bellafkih propose a novel method for measuring the influence of Twitter users in their 2020 study. Presented at the Fourth International Conference on Intelligent Computing in Data Sciences (ICDS), the research focuses on evaluating user impact beyond traditional metrics by considering engagement patterns and social interactions. Their findings contribute to the growing field of social media analytics, aiding businesses and researchers in identifying key influencers who shape online discussions and trends.

[2] Qi, Yuxing, and Zahratu Shabrina compare lexicon-based and machine-learning-based approaches for sentiment analysis using Twitter data. Published in the journal Social Network Analysis and Mining in 2023, their study evaluates the strengths and weaknesses of both methodologies in classifying sentiments effectively. The research highlights the importance of selecting the right technique based on data characteristics and application requirements, offering insights into improving sentiment analysis accuracy in real-world scenarios.

[3] Bhatnagar, Sarvesh, and Nitin Choubey explore sentiment analysis on closely related topics using Twitter data. Their 2021 study, published in Social Network Analysis and Mining, investigates how tweets on similar subjects can be differentiated based on sentiment patterns. By applying machine learning techniques, the authors demonstrate the challenges and effectiveness of sentiment classification in social media discussions, contributing to the refinement of sentiment analysis models for better contextual understanding.

[4] Gupta, Bhumika, et al. examine Twitter sentiment analysis using various machine learning algorithms in their 2017 study published in the International Journal of Computer Applications. The research explores how algorithms such as Logistic Regression, Support Vector Machines, and Decision Trees perform in classifying sentiment in tweets. Their findings provide a comparative analysis of different machine learning techniques, offering valuable insights for researchers and practitioners looking to enhance sentiment classification accuracy in Python-based implementations.

[5] Hasan, Ali, et al. present a machine learning-based approach for sentiment analysis of Twitter accounts in their 2018 study published in Mathematical and Computational Applications. Their research focuses on leveraging classification models to analyze sentiment trends across different user accounts, helping identify patterns in public opinion. By incorporating various machine learning techniques, the study contributes to the development of more effective sentiment analysis frameworks for social media platforms.

## III. Proposed System

The system aims to address key challenges such as sarcasm detection, context awareness, and real-time sentiment classification while ensuring accurate identification of influential users based on engagement metrics. By leveraging multiple classification algorithms and influence ranking methods, this system provides a comprehensive framework for analyzing public opinion and digital influence.

### 3.1. System Architecture:

  

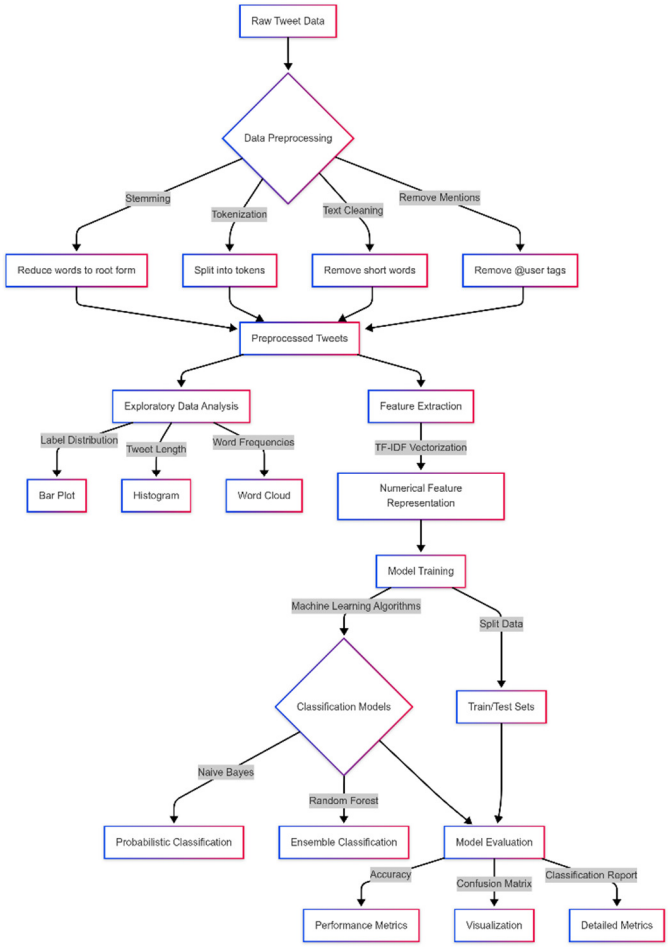


***Fig 1. System architecture***

The architecture depicted in the diagram represents a tweet classification pipeline, starting from raw tweet data and progressing through various preprocessing, analysis, feature extraction, training, and evaluation stages.The process begins with data preprocessing, where raw tweets undergo multiple transformations, including stemming (reducing words to their root form), tokenization (splitting text into tokens), text cleaning (removing short words), and mention removal (eliminating "@user" tags). The resulting preprocessed tweets serve as the foundation for further analysis.

Next, exploratory data analysis (EDA) is performed to gain insights into the dataset. This includes analyzing label distribution (visualized using a bar plot), tweet length (represented with a histogram), and word frequencies (illustrated using a word cloud). Parallelly, feature extraction is conducted using TF-IDF vectorization, converting text into a numerical feature representation suitable for machine learning.

Following feature extraction, the dataset is prepared for model training. The data is split into train and test sets, and various machine learning algorithms are applied. The classification models considered include Naïve Bayes (for probabilistic classification) and Random Forest (for ensemble classification).

The final phase involves model evaluation, where the trained models are assessed based on performance metrics like accuracy, and tools like the confusion matrix and classification report help visualize results and provide detailed metrics.

This structured pipeline ensures efficient preprocessing, robust model training, and comprehensive evaluation, making it suitable for classifying tweets based on their textual content.

## 3.2. Evaluation Matrix:

Sentiment classification models are evaluated using standard performance metrics, including Accuracy, Precision, Recall, and F1-Score.

### 3.2.1 Accuracy:

Accuracy measures the proportion of correctly classified tweets among all classified tweets.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Where:

- TP (True Positives): Correctly classified positive sentiments
- TN (True Negatives): Correctly classified negative sentiments
- FP (False Positives): Incorrectly classified negative tweets as positive
- FN (False Negatives): Incorrectly classified positive tweets as negative

  

### 3.2.2.Precision:

Precision calculates how many tweets predicted as positive are actually positive.

$$Precision = \frac{TP}{TP + FN} \quad (2)$$

A higher precision value indicates fewer false positives in sentiment classification.

## III. CONCLUSIONS

In conclusion, sentiment analysis has emerged as a vital tool for understanding opinions, sentiments, and emotions expressed in written language, particularly in the era of social media. This study evaluates the effectiveness of Logistic Regression and Support Vector Machine (SVM) algorithms in sentiment classification, demonstrating their capability to capture nuanced sentiments with a notable accuracy of 0.941. Additionally, identifying influential users based on engagement metrics such as retweets and follower counts enhances the applicability of sentiment analysis in business, research, and policymaking. The case study on airline-related tweets highlights the practical implications of combining sentiment classification with

## ACKNOWLEDGMENT

I would like to express my sincere gratitude to everyone who contributed to the successful completion of this research on sentiment analysis using machine learning techniques. I extend my appreciation to my mentors and colleagues for their valuable guidance, insightful discussions, and continuous support throughout this study. Special thanks to the authors and researchers whose works have provided a strong foundation for this research, enabling a deeper understanding of sentiment classification and influencer identification on social media. I am also grateful for the availability of datasets and computational resources that facilitated the implementation and evaluation of machine learning models. Lastly, I acknowledge the role of technological advancements in the field of artificial intelligence, which have significantly contributed to the evolution of sentiment analysis, making it an essential tool for businesses, researchers, and policymakers.

## REFERENCES

[1] Essaidi, Abdessamad, Dounia Zaidouni, and Mostafa Bellafkih. "New method to measure the influence of Twitter users." 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS). IEEE, 2020.

[2] Qi, Yuxing, and Zahratu Shabrina. "Sentiment analysis using Twitter data: a comparative application of lexicon-and machine- learning-based approach." Social Network Analysis and Mining 13.1 (2023): 31.

[3] Bhatnagar, Sarvesh, and Nitin Choubey. "Making sense of tweets using sentiment analysis on closely related topics." Social Network Analysis and Mining 11.1 (2021): 44..

[4] Gupta, Bhumika, et al. "Study of Twitter sentiment analysis using machine learning algorithms on Python." International Journal of Computer Applications 165.9 (2017): 29-34.

[5] Hasan, Ali, et al. "Machine learning-based sentiment analysis for twitter accounts." Mathematical and computational applications 23.1 (2018): 11.

  