ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-3s/779-785
C Jabilli *et. al.*, / International Journal of Engineering & Science Research

Social Media Sentiment Analysis

Syeda Fatima[1], C Jabilli[2], S Sravani[3], T Sri Nikitha[4]
[1]Assistant Professor, Department of CSE(AI&ML), Bhoj Reddy Engineering College for Women, India
[2,3,4,5]B.Tech Students, Department of CSE(AI&ML), Bhoj Reddy Engineering College for Women, India

ABSTRACT

In the digital era, social media has become a powerful platform where users express their opinions and emotions on various topics such as politics, brands, products, and current events. This project, Social Media Sentiment Analysis, aims to analyze and interpret these user sentiments to classify them as positive, negative, or neutral using techniques from Natural Language Processing (NLP) and Machine Learning (ML).

The system collects data from platforms like Twitter, Facebook, or Instagram, preprocesses the textual content to remove noise, and applies sentiment classification models such as Naïve Bayes, Support Vector Machines, or deep learning models like BERT. The analyzed results are visualized through charts and graphs to help organizations, researchers, and businesses gain actionable insights from public opinion. This project highlights the importance of sentiment analysis in brand monitoring, crisis detection, market research, and decision-making.

Introduction

This Mini-project delves into the world of social media sentiment analysis, a vital tool for understanding online opinions. We will begin by gathering data from platforms like Twitter, using APIs to access relevant posts. Preprocessing techniques will then be applied, cleaning and structuring the text for analysis. Natural language processing (NLP) will play a key role in extracting meaningful features from the text. Machine learning models will be trained to classify sentiments as positive, negative, or neutral. Our goal is to create a system that accurately reflects public sentiment on chosen topics. This analysis offers valuable insights for businesses, researchers, and anyone seeking to understand public perception. By exploring these methods, we gain practical experience in data science and its applications to real-world social media data.

Problems in Existing System

- **Lower Accuracy**

Existing models often struggle with contextual understanding, especially in the presence of sarcasm, slang, abbreviations, or emojis commonly used on social media.

- **Limited Scalability & Increased Development Time**

Traditional sentiment analysis systems may not efficiently handle large-scale real-time data from multiple platforms.

- **Feature Limitation & Maintenance Burden**

Many existing systems lack advanced features such

as emotion detection, and also sarcasm handling, topic extraction, or real-time.

Proposed System

Our project proposes a focused sentiment analysis system. It will gather targeted social media data, process it through a custom NLP pipeline, and utilize machine learning for sentiment classification. The system will then visualize and report sentiment trends, providing insights within a defined area of interest. This approach allows for customization and a deeper understanding of the sentiment analysis process.

. DESIGN

Software Architecture

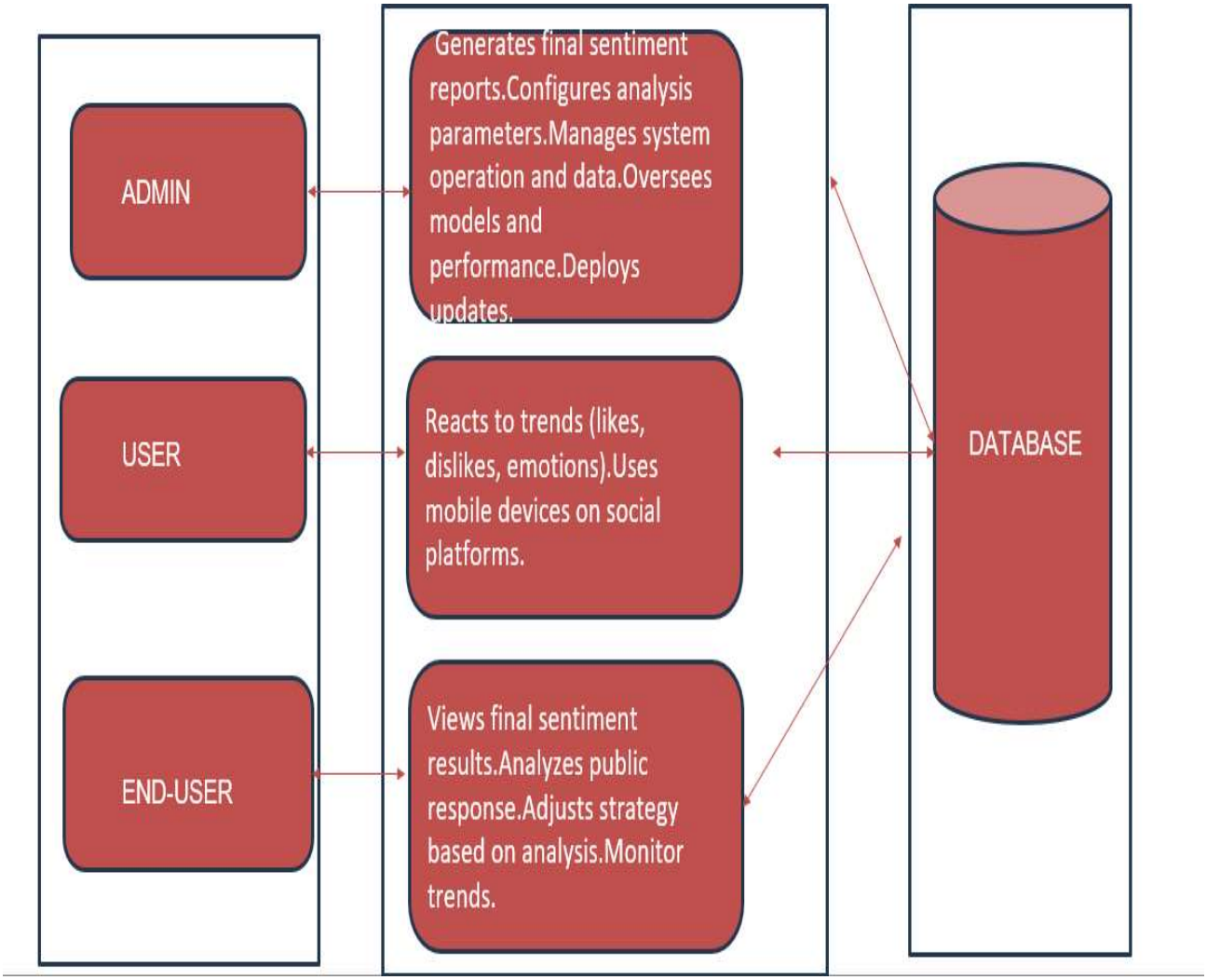

Fig 3.1.1: Software Architecture

IMPLEMENTATION

Technologies

1. Data Collection
    a. Use Twitter API (Tweepy) or scraping tools to gather posts.
    b. Collect fields like text, timestamp, user info, hashtags.
    c. Choose a relevant topic, hashtag, or keyword for filtering data.
2. Data Preprocessing
    a. Clean text: remove URLs, mentions, hashtags, emojis, and special characters.
    b. Normalize: convert to lowercase, remove stopwords.
    c. Lemmatize or stem: reduce words to their base/root form.
3. Text Vectorization
    a. Convert text to numbers using TF-IDF, Bag of Words, or word embeddings.
    b. Ensure consistent vector length for model input.
    c. Choose Word2Vec or BERT for deep learning-based models.
4. Sentiment Labeling
    a. Use pre-labeled datasets or lexicon tools (e.g., VADER, TextBlob).
    b. Label text as positive, negative, or neutral.
    c. Manually label if dataset is small or specific.
5. Model Training
    a. Split data into training and testing sets.
    b. Train with models like Logistic Regression, Naive Bayes, or BERT.
    c. Use cross-validation to improve model reliability.
6. Model Evaluation
    a. Evaluate using accuracy, precision, recall, and F1-score.
    b. Use a confusion matrix to understand classification.
    c. Fine-tune hyperparameters to improve performance.
7. Visualization
    d. Create bar/pie charts for sentiment distribution.
    e. Use word clouds to highlight frequent terms.
    f. Plot time-series trends to track sentiment changes over time.

Test cases:-

Test Case ID	Input Email Content	Expected Output	Actual Output	Result
01	"I love this product so much!"	Positive	Positive	Pass
02	"This app is terrible and slow."	Negative	Negative	Pass
03	"It's okay, nothing special."	Neutral	Neutral	Pass
04	"Je t'aime ce produit"	Unsupported Lang	Unsupported Lang	Pass
05	"Great job ruining everything!"	Negative	Negative	Pass
06	"❤️❤️❤️"	Positive	Positive	Pass
07	"This is AMAZING!!! #love #awesome"	Positive	Positive	Pass
08	"I don't know what to feel about this."	Neutral	Neutral	Pass

Fig No. 5.1.1.Test cases

RESULTS

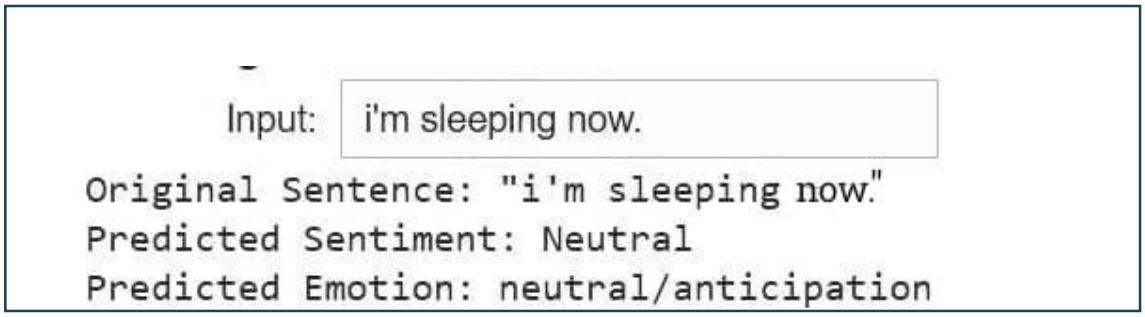

Fig 6.1 Output 1

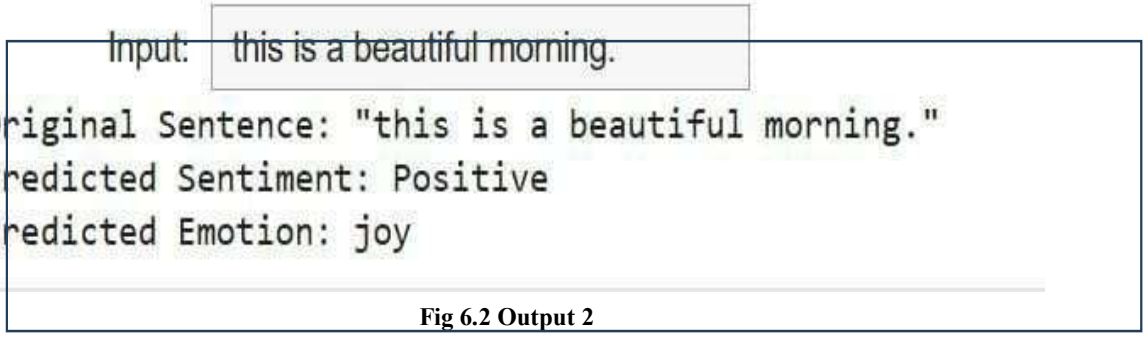

Fig 6.2 Output 2

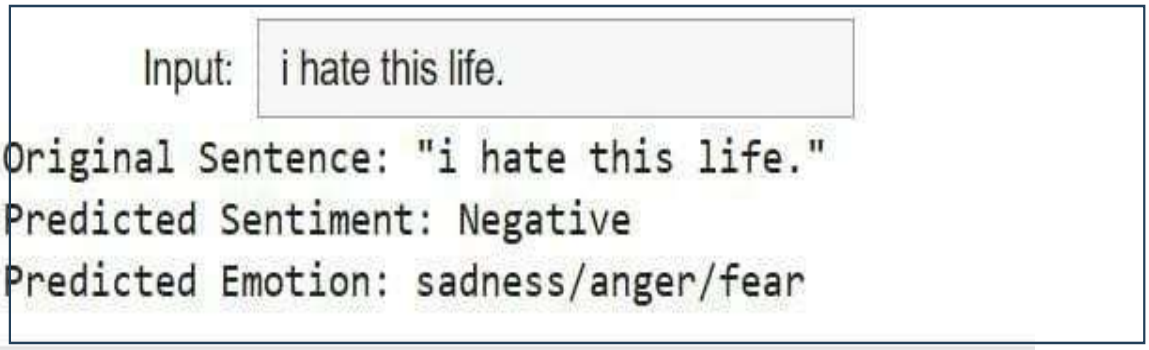

Fig 6.3 Output

Fig 6.4 Positive tweets

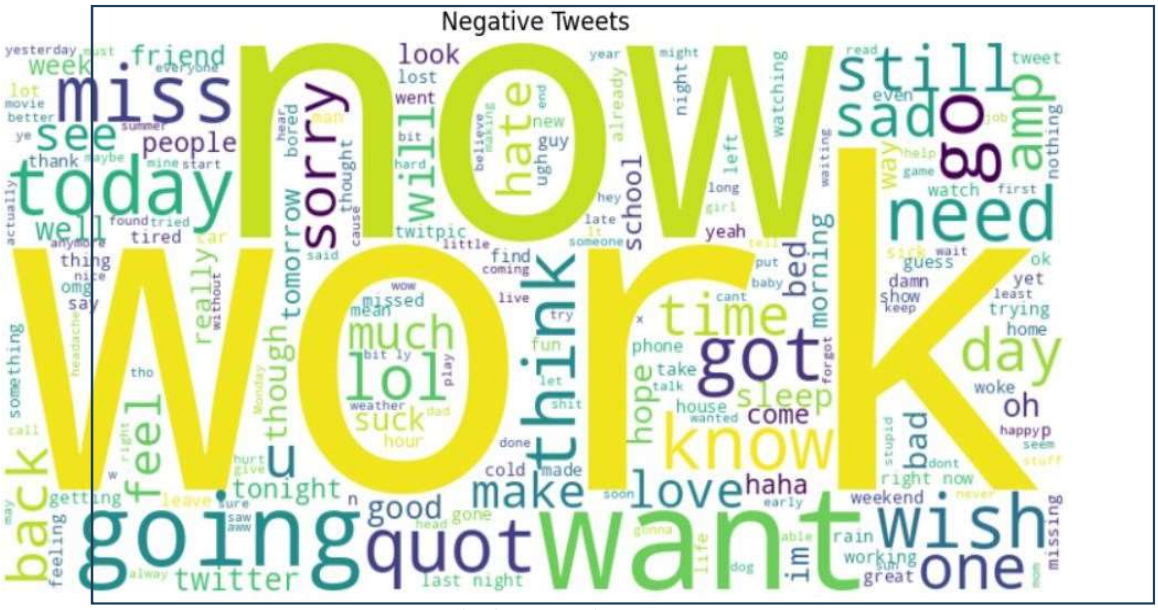

Fig 6.5 Negative tweets

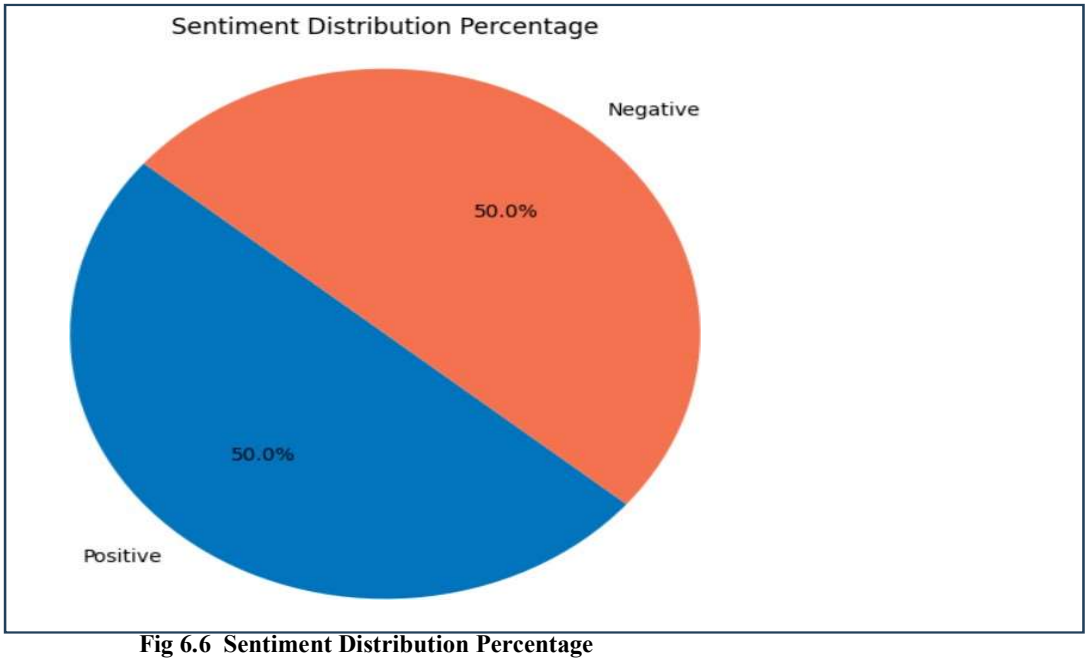

Fig 6.6 Sentiment Distribution Percentage

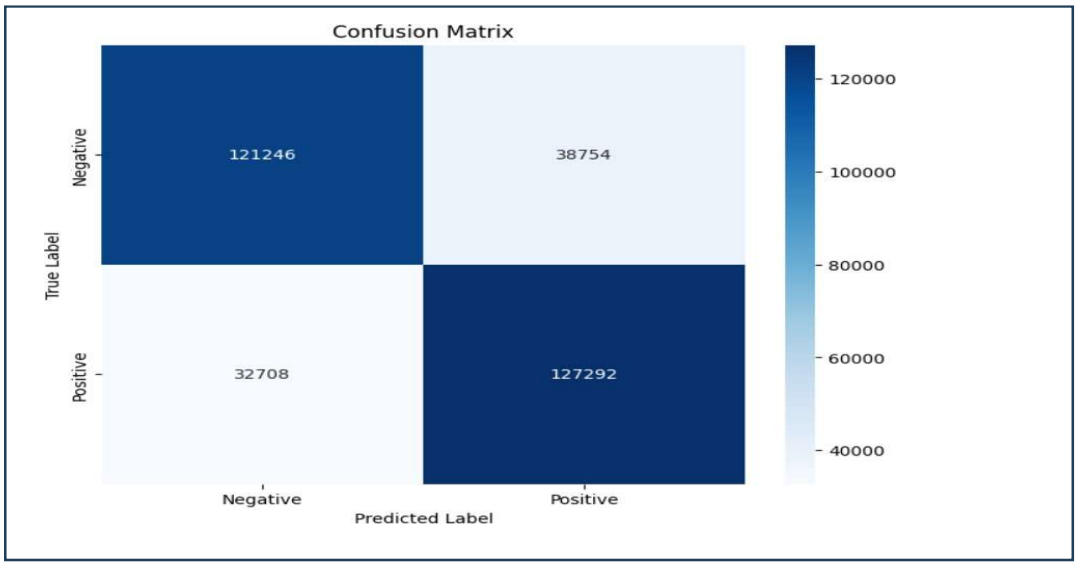

Fig 6.7 Confusion Matrix

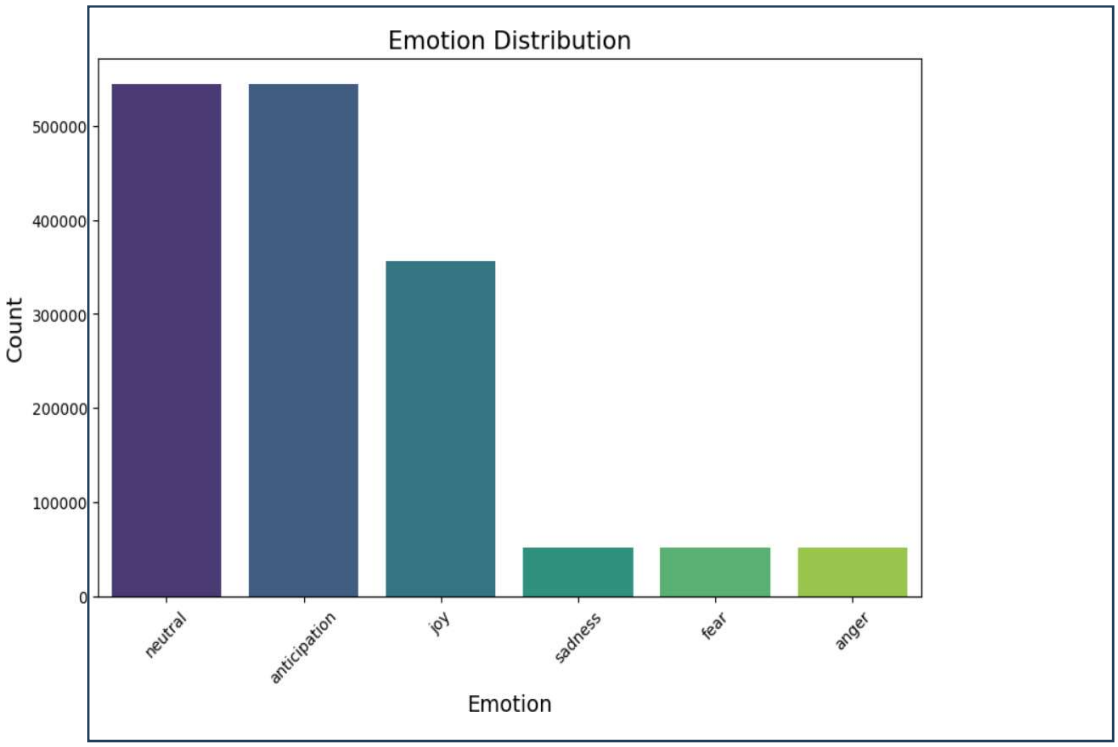

Fig 6.8 Emotion Distribution

CONCLUSION & FUTURE SCOPE

Conclusion:

This social media sentiment analysis project using NLP effectively classifies user comments into positive, negative, and neutral sentiments. By analyzing text data, it helps in understanding public opinion, tracking sentiment trends, and supporting decision-making processes for marketing, customer feedback, and brand reputation management.

Future Scope:

- This project uses NLP and ML to classify social media sentiments in real time, aiding businesses, researchers.
- It processes noisy text, including slang, hashtags, and emojis, for accurate sentiment analysis.
- A trained logistic regression model delivers reliable predictions, including neutral sentiment handling.
- It supports both real-time and batch analysis via CSV/API for broader use.
- The system can integrate with external tools (e.g., CRM, feedback platforms) through API support.

➢ Results can be exported for use in analytics, reports, or dashboards.

REFERENCES

References

1) Advances in Sentiment Analysis - Techniques, Applications, and Challenges, K.Victor Raja, January 10, 2024.

2) Deep Learning Research Applications for Natural Language Processing, Multiple contributors (e.g., Kundu R., Singh S.) 2023.

3) Research Anthology on Implementing Sentiment Analysis Across Multiple Disciplines, Multiple contributors (e.g., Veerappa B. Pagi) 2024, Veerappa B. Pagi.