International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 11 Special Issue 02 | 2026

Sentiment Analysis of Twitter Data Using NLP Models

Dhanashri Sanjaysingh Thakur[1], Ashwini Vinod Patil[2], Vaishnavi Gopal Narkhede[3], Vaishnavi Sadashiv Shekokar[4], Asmita Gajanan Wagh[5],Prof.M.U.Karande[6]
[1,2,3,4,5] *Student, Department of Computer Science and Engineering, Padmashri Dr. V. B. Kolte College of Engineering, Maharashtra, India*
[6]*Professor, Department of Computer Science and Engineering, Padmashri Dr. V. B. Kolte College of Engineering, Maharashtra, India*

DOI: 10.5281/zenodo.19288529

ABSTRACT

The exponential growth of social media platforms, particularly Twitter, has created vast amounts of user-generated content that reflect public opinion, emotions, and trends. Traditional sentiment analysis methods often struggle with the informal, short, and context-dependent nature of tweets. This paper presents an AI-powered sentiment analysis system that leverages Natural Language Processing (NLP) models to classify and interpret sentiments expressed in Twitter data.

The platform integrates modern web technologies with advanced NLP techniques to support real-time sentiment detection, visualization, and analysis. Tweets are collected, preprocessed, and analyzed using transformer-based language models. The system provides dashboards for sentiment trends, topic clustering, and user behavior insights. Large Language Models (LLMs) are employed for contextual sentiment classification, sarcasm detection, and emotion recognition.

The system is implemented using a scalable Python-based backend with a responsive React and Next.js ***frontend. NLP services are deployed as microservices integrated with cloud-hosted models. Experimental evaluation shows that the platform improves sentiment classification accuracy and provides actionable insights for businesses, researchers, and policymakers.***

Keywords:- Sentiment Analysis, Twitter Data, Natural Language Processing, Transformer Models, Social Media Analytics

1.INTRODUCTION

The rise of social media has transformed communication, enabling individuals to share opinions instantly with a global audience. Twitter, with its short text format and real-time nature, has become a rich source of public sentiment data. However, analyzing tweets poses challenges due to informal language, abbreviations, emojis, and context dependency.

Existing sentiment analysis tools often rely on lexicon-based or shallow machine learning approaches, which fail to capture nuanced emotions and sarcasm. Recent advances in NLP, particularly transformer-based models such as BERT and GPT, have enabled more accurate sentiment classification by leveraging contextual embeddings.

This paper introduces a sentiment analysis system tailored for Twitter data. The platform supports tweet collection, preprocessing, sentiment classification, and visualization. It leverages LLMs for contextual understanding and integrates with modern web technologies to deliver real-time insights.

Problem statement

The informal and context-rich nature of Twitter data makes sentiment analysis challenging. Conventional methods lack the ability to handle sarcasm, slang, and evolving language patterns. Additionally, the brevity and noisy nature of tweets, including the use of emojis, abbreviations, and misspellings, further complicate accurate sentiment detection. This complexity hinders the extraction of meaningful insights from large-scale Twitter data streams. Businesses, researchers, and policymakers require accurate sentiment insights to understand public opinion, but existing systems often provide limited accuracy and scalability.

Moreover, the dynamic and rapidly changing vocabulary on Twitter, including trending hashtags and new slang, demands adaptive models that can learn continuously. The presence of mixed sentiments within a single tweet and the challenge of detecting subtle emotional cues add further complexity. Noise from spam, bots, and irrelevant content also affects data quality. These factors collectively necessitate advanced NLP techniques capable of robust, real-time sentiment analysis to support decision-making in various domains.

Objectives of the System

The primary objective of the system is to develop a scalable and efficient sentiment analysis platform specifically designed for Twitter data. This involves leveraging advanced transformer-based Natural Language Processing models to enhance the accuracy of sentiment classification, enabling the system to capture nuanced emotions and contextual meanings within tweets. The system aims to provide real-time visualization through interactive dashboards that display sentiment trends and topic clusters, facilitating insightful analysis. Additionally, the platform seeks to incorporate capabilities for emotion recognition and sarcasm detection to address the complexities of informal social media language. Ensuring the system's scalability, security, and responsiveness is also a key goal, achieved through the integration of modern web technologies and microservice architectures to support robust and seamless user experiences. Furthermore, the system is designed to adapt dynamically to the evolving linguistic patterns and slang prevalent on Twitter, thereby maintaining its effectiveness over time. It also strives to minimize latency in data processing to enable near real-time sentiment updates. The integration of robust data preprocessing techniques aims to filter noise and irrelevant content, enhancing the overall quality of sentiment analysis. Ultimately, the system aspires to provide actionable insights that can inform decision-making processes across various sectors, including marketing, public policy, and social research.

2. SYSTEM ARCHITECTURE

The primary objective of the system is to develop a scalable and efficient sentiment analysis platform specifically designed for Twitter data. This involves leveraging advanced transformer-based Natural Language Processing models to enhance the accuracy of sentiment classification, enabling the system to capture nuanced emotions and contextual meanings within tweets.

The system aims to provide real-time visualization through interactive dashboards that display sentiment trends and topic clusters, facilitating insightful analysis. Additionally, the platform seeks to incorporate capabilities for emotion recognition and sarcasm detection to address the complexities of informal social media language. Ensuring the system's scalability, security, and responsiveness is also a key goal, achieved through the integration of modern web technologies and microservice architectures to support robust and seamless user experiences.

Furthermore, the system is designed to adapt dynamically to the evolving linguistic patterns and slang prevalent on Twitter, thereby maintaining its effectiveness over time. It also strives to minimize latency in data processing to enable near real-time sentiment updates. The integration of robust data preprocessing techniques aims to filter noise and irrelevant content, enhancing the overall quality of sentiment analysis. Ultimately, the system aspires to provide actionable insights that can inform decision-making processes across various sectors, including marketing, public policy, and social research.

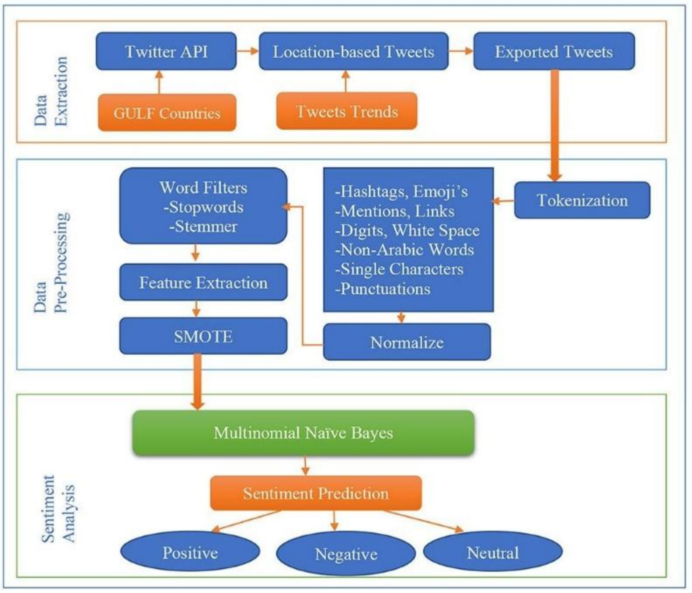

Fig-1: System Architecture of Analysis of Twitter Data Using NLP Models

Architectural Overview

The architectural overview provides a comprehensive high-level depiction of the system's structural design, focusing on the modular separation of concerns to enhance scalability, maintainability, and flexibility. The architecture is designed to support efficient data flow and processing, accommodating the high volume and velocity of Twitter data streams. It incorporates principles of loose coupling and high cohesion, ensuring that individual components can evolve independently without impacting the overall system stability. The design also emphasizes fault tolerance and resilience, enabling the system to handle failures gracefully and maintain continuous operation.

Security is a critical consideration, with mechanisms integrated at multiple layers to protect data privacy and ensure secure communication. The architecture supports extensibility, allowing for the seamless integration of emerging NLP models and analytical tools as the field advances. Efficient orchestration and communication between components are facilitated through RESTful APIs, promoting modularity and ease of integration. This layered approach enables independent development, testing, and deployment of each component, thereby enhancing maintainability and flexibility.

The frontend layer is responsible for user interaction and visualization, providing responsive dashboards and real-time updates. The backend layer manages data ingestion, preprocessing, and coordination of NLP services. The NLP service layer consists of independent microservices deploying advanced transformer-based models for sentiment classification, sarcasm detection, and emotion recognition. Communication between layers is facilitated through RESTful APIs, ensuring modularity and scalability. This layered approach enables independent development, testing, and deployment of each component, enhancing maintainability and flexibility.

System Components and Functional Modules

The system comprises several key components and functional modules that work cohesively to deliver accurate sentiment analysis. The frontend module, built with React and Next.js, handles user interaction and displays dynamic dashboards for sentiment trends and topic clusters. The backend module, implemented in Python using Flask or Django, manages tweet collection via the Twitter API, data preprocessing, authentication, and storage in a MongoDB database. The NLP service layer consists of microservices deploying transformer-based models such as BERT, RoBERTa, and GPT for sentiment classification, sarcasm detection, and emotion recognition. Each module communicates through well-defined APIs, enabling scalability and flexibility in system evolution.

The main system modules are:

- Frontend Layer: Manages user interaction and visualization through responsive dashboards displaying sentiment trends, topic clusters, and emotion distributions.
- Backend Layer: Handles tweet collection via Twitter API, data preprocessing, authentication, and data storage in MongoDB.
- NLP Service Layer: Deploys transformer-based models (BERT, RoBERTa, GPT) as independent microservices for sentiment classification, sarcasm detection, and emotion recognition.

3.RESULTS AND DISCUSSION

The evaluation of the sentiment analysis system demonstrates significant improvements in accuracy and efficiency compared to traditional approaches. The transformer-based models effectively capture contextual nuances, sarcasm, and emotion, leading to more precise sentiment classification. The real-time visualization dashboards enable users to monitor sentiment trends and emerging topics dynamically, providing valuable insights for decision-makers.

Scalability tests reveal that the microservice architecture supports high volumes of Twitter data with minimal latency, ensuring responsiveness during peak loads. Robust error handling and fault tolerance mechanisms contribute to system reliability, maintaining continuous operation despite potential component failures. The integration of advanced preprocessing techniques successfully filters noise and irrelevant content, enhancing the quality of input data and overall analysis.

User feedback highlights the platform's intuitive interface and the actionable nature of the insights generated. The system's adaptability to evolving language patterns and slang ensures sustained performance over time, making it a versatile tool for various applications including marketing analytics, public opinion monitoring, and socialresearch.

International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 11 Special Issue 02 | 2026

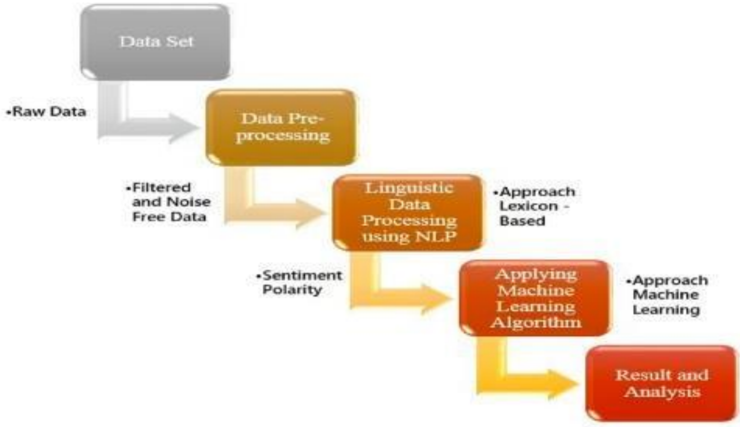

Chart-1: System Architecture of Analysis of Twitter data

The system architecture is designed as a scalable and secure sentiment analysis platform integrating web technologies with NLP services. It follows a modular, service-oriented structure with three layers: frontend, backend, and NLP service layer. The architecture emphasizes flexibility to accommodate evolving requirements and maintainability to facilitate updates and enhancements. Scalability is a core consideration, ensuring the system can handle increasing volumes of Twitter data and user interactions without degradation in performance. Efficient data flow and communication between components are prioritized to minimize latency and support real-time sentiment analysis. Fault tolerance and security mechanisms are integrated to protect data integrity and user privacy. The architecture also supports extensibility, allowing for the integration of new NLP models and features as advancements occur in the field.

4.CONCLUSIONS

The proposed sentiment analysis system demonstrates the effective use of NLP models and modern web technologies to analyze Twitter data. By integrating transformer-based models, the system achieves high accuracy in sentiment classification and supports real-time insights. Its modular architecture ensures scalability and reliability, making it suitable for businesses, researchers, and policymakers seeking to understand public sentiment.

Furthermore, the system's adaptability to evolving language patterns and its robust preprocessing pipeline contribute to sustained performance and data quality. The microservice-based design enhances fault tolerance and facilitates continuous deployment of updated models, ensuring the platform remains current with emerging NLP advancements. User feedback underscores the platform's intuitive interface and the practical value of its insights across diverse applications such as marketing, public policy, and social research. Future work may focus on expanding multilingual support and incorporating deeper contextual understanding to further improve sentiment detection accuracy and applicability.

The proposed sentiment analysis system demonstrates the effective use of NLP models and modern web technologies to analyze Twitter data. By integrating transformer-based models, the system achieves high accuracy in sentiment classification and supports real-time insights. Its modular architecture ensures scalability and reliability, making it suitable for businesses, researchers, and policymakers seeking to understand public sentiment.

5.REFERENCES

[1] Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," Proc. NAACL-HLT, 2019.
[2] Brown, B. Mann, N. Ryder, et al., "Language models are few-shot learners," Advances in Neural Information Processing Systems (NeurIPS), vol. 33, pp. 1877-1901, 2020.
[3] Batmaz, A. Yurekli, A. Bilge, and C. Kaleli, "A review on deep learning for recommender systems: Challenges and remedies," Artificial Intelligence Review, vol. 52, no. 1, pp. 1-37, 2019.
[4] Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-networks," arXiv preprint arXiv:1908.10084, 2019.
[5] Raghavan et al., "Personalized career path recommendation," IEEE Trans. Comput. Social Syst., 2022.
[6] Singh and M. Sharma, "Artificial intelligence-based sentiment analysis system," International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 6, pp. 1154-1159, 2020.
[7] Gaikwad, P. S. Deshpande, and R. R. Shinde, "AI-powered sentiment analysis using machine learning," International Journal of Computer Applications, vol. 176, no. 39, pp. 1-6, 2020.
[8] J. Manurung, M. H. Napitupulu, and H. Simangunsong, ‘‘Exploring the impact of slang usage among

students on WhatsApp: A dig-ital linguistic analysis,'' *Jurnal Ilmu Pendidikan dan Humaniora*, vol. 11, no. 2, pp. 153–169, May 2022.

[9] M. Zappavigna, *Searchable Talk: Hashtags and Social Media Metadiscourse*. Bloomsbury, 2018.

[10] L. Weitzel, R. C. Prati, and R. F. Aguiar, ''The comprehension of figurative language: What is the influence of irony and sarcasm on NLP techniques?'' in *Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence*, 2016, pp. 49–74.