International Journal of Electrical, Electronics and Computers
Special Issue | Jan-Feb, 2026
Available: https://aipublications.com/ijeec/
Peer-Reviewed Journal

Twitter Sentimental Analysis Using NLP

Kanika Chalana[1], Jishad Abdulla p[2], Navaneeth S Biju[3], Chethas LP[4], Mohammed Ajas[5]

[1]Assistant Professor, AIML Department, Artificial Intelligence and Machine Learning,T John Institute of Technology, Bengaluru,india
[2,3,4,5]Department of Artificial Intelligence and Machine Learning, T John Institute of Technology, Bengaluru, India

Received: 22 Nov 2025; Accepted: 01 Feb 2025; Date of Publication: 08 Feb 2026

***Abstract*— *Social media sites like Twitter produce enormous volumes of textual data that record people's thoughts, feelings, and comments on a wide range of subjects. The main objective of this study is to classify tweets into categories like positive, negative, and neutral by applying natural language processing (NLP) techniques to analyze sentiment in Twitter data. Tokenization, text standardization, and preprocessing procedures to eliminate noise are the next steps in the study's structured pipeline, which starts with data collection via Twitter application programming interfaces. Text is numerically represented through methods like Bag of Words, TF-IDF, and word embeddings. Both contemporary deep learning architectures and Traditional machine learning methods are investigated for classification. Advanced models like transformer-based architectures and long short-term memory (LSTM) networks are contrasted with logistic regression, Naive Bayes, and support vector machines. According to experimental results, deep learning techniques—in particular, transformers—offer better accuracy and a greater capacity to capture contextual meaning than traditional methods. The comparative perspective of this research focuses on the pros and cons of different methods that sets it apart. Besides tackling ethical concerns such as privacy, justice, and algorithmic bias, the study also emphasizes its applications in areas like marketing, politics, public health, and customer service.*

Keywords*— Sentiment analysis, Twitter, Natural language processing, Machine learning, Deep learning

I. INTRODUCTION

Abstract In the digital age, social media platforms such as Twitter have emerged as powerful tools for communication, where millions of users express their opinions, emotions, and feedback on various topics daily.

This dynamic and massive influx of textual data provides an invaluable resource for analyzing public sentiment, especially in areas such as marketing, politics, public health, and customer service.

This project's goal is to classify tweets into sentiment categories like positive, negative, and neutral by doing sentiment analysis on Twitter data using Natural LanguageProcessing(NLP)techniques.

Data collection, preprocessing, feature extraction, model training, and evaluation are all part of the project's methodicalapproach. The Twitter API or third-party libraries like Tweepy are used to extract Twitter data, which is further cleaned to eliminate noise such as stop words, URLs, mentions, and special characters. To standardize the textual input, lemmatization, stemming, and tokenizationareused.

Depending on the size and complexity of the dataset, feature extraction is carried out using methods like Bag of Words, TF-IDF, or word embeddings like Word2Vec and BERT.

Numerous machine learning and deep learning models, such as Naive Bayes, Support Vector Machines (SVM), Logistic Regression, and Recurrent Neural Networks (RNNs) like LSTM, are investigated for categorization.

These models are trained on labeled datasets and evaluated based on accuracy, precision, recall, and F1- score to determine their effectiveness in sentiment prediction.

In addition, advanced techniques such as transfer learning with pre- trained transformer models like

BERT are considered to improve performance and contextual understanding.

The results demonstrate that deep learning models, particularly those leveraging transformers, outperform traditional machine learning approaches in terms of accuracy and contextual comprehension.The study concludes with a discussion of real-world applications, potential ethical concerns, and future directions, such as multilingual sentiment analysis and real-time sentiment tracking.This project showcases how NLP can harness the power of social media to extract actionable insights from human emotions, thereby assisting businesses, governments, and organizations in making informed decisions Keywords Sentiment Analysis, Natural Language Processing (NLP), Twitter Data, Text Classification, Machine Learning.

OVERVIEW Social media has become a vital medium for public expression in the current digital age, with sites like Twitter providing a vibrant environment for users to instantly communicate their thoughts, feelings, and experiences. Twitter distinguishes out in particular because of its microblogging structure, which encourages succinct and powerful communication by limiting posts to 280 characters. Twitter is a rich source of data that represents public mood on a variety of topics, from politics and entertainment to brand feedback and breaking news, thanks to its structure and real-time nature. The massive amount of tweets that are produced every day results in a huge, constantly changing dataset that contains insightful information on societal trends and general attitudes.

However, it is not feasible to analyze such a big corpus by hand, which emphasizes the necessity of automated sentiment analysis methods. Identifying and classifying emotions expressed in text, such as figuring out whether a tweet expresses a good, negative, or neutral attitude, is the focus of sentiment analysis, a branch of natural language processing (NLP). For a large-scale knowledge of public perception and behavior, this analysis is essential. Because of the peculiarities of the information on Twitter, sentiment analysis poses a number of difficulties despite its potential.

Traditional text-processing methods are less successful when tweets contain misspellings, emojis, hashtags, slang, casual language, and abbreviations.The brevity of tweets can also lead to ambiguity, making it difficult to discern context, sarcasm, or irony elements that significantly affect sentiment interpretation.Moreover, Twitter language evolves rapidly, with new terms and expressions emerging regularly, requiring adaptive systems capable of handling linguistic dynamism.

Multilingual and multimodal content further complicates sentiment analysis, as users often blend languages or include images and videos that influence the overall message.Additionally, many applications such as monitoring sentiment during political debates, crisis events, or product launches demand real- time processing capabilities, placing further constraints on computational efficiency and scalability.

Overcoming these issues requires robust preprocessing techniques, context- aware modeling, and adaptive learning approaches to maintain performance in dynamic environments.Various machine learning algorithms such as Logistic Regression, Naïve Bayes, and Support Vector Machines will be implemented, each offering distinct advantages for text classification tasks.Optionally, advanced models like Long Short-Term Memory (LSTM) networks, BERT, or RoBERTa may be explored to capture deeper linguistic nuances.The models will be trained and validated on established datasets like Sentiment140 or TweetEval, ensuring balanced class distributions to avoid bias.The performance of each model will be evaluated using standard classification metrics accuracy, precision, recall, and F1- score.

Cross-validation techniques will be employed to ensure the generalizability of the results, while confusion matrices will help identify common misclassification patterns.

Beyond model performance, the study will visualize sentiment trends across time, topics, and demographics, offering actionable insights.The practical significance of this work extends across multiple domains In business, real-time sentiment analysis supports brand monitoring, campaign evaluation, and customer engagement strategies.In politics, it aids in gauging public opinion, predicting electoral outcomes, and analyzing responses to policy announcements. Market researchers can use sentiment trends to understand consumer behavior, discover emerging demands, and refine product positioning.

During crises, sentiment analysis provides organizations and government agencies with immediate feedback, enabling timely communication and intervention. Academically, this project contributes to computational linguistics, machine learning, and human-computer interaction by applying and comparing advanced NLP techniques in a real-world context. Ultimately, this study aims to create a scalable and accurate sentiment analysis system that enhances our ability to understand and respond to the evolving digital discourse on Twitter.

II. LITERATURE SURVEY

This work explores sentiment analysis using a dataset of 10,000 entries labeled as either positive or negative.

It compares the performance of Gated Recurrent Units (GRUs) and traditional Recurrent Neural Networks (RNNs) in handling sequential data.

The study finds that GRUs outperform RNNs, achieving an accuracy of 85% in sentiment classification, highlighting their effectiveness in capturing long-term dependencies.

The research emphasizes the value of advanced neural networks in improving performance on complex classification tasks[1] This chapter explores how patient feedback (PF) can enhance biomedical research using sentiment analysis (SA) in natural language processing (NLP).

PF is a valuable resource for assessing clinical effectiveness and quality of care but is often time-consuming and subjective to analyze.

The evolution of the technology, offering a deep analysis of foundational algorithms and emerging applications. Key innovations include route optimization, real-time object detection, agricultural support like crop management and pesticide application, and autonomous navigation. Notably, collaborative drone use in transportation enhances last-mile delivery by reducing transit times and improving logistics efficiency. These advancements highlight the growing role of AI-driven drones and point toward important future research directions [3].

A hybrid AI approach merges reinforcement learning-based deep learning with expert rule-based systems to improve drone navigation in simulated environments. This system supports explainability and human interaction, and experiments show that combining rule-based modules with reinforcement learning significantly boosts performance in complex tasks like obstacle avoidance. Though results are promising, the study is limited to simplified simulations and recommends testing in dynamic, multi-agent real-world conditions [4].

The VTD3 framework integrates YOLOv8, BoT-SORT tracker, and TD3 reinforcement learning for vision-based drone target tracking, reducing dependence on GPS and other sensors. It achieves significant improvements in tracking error, altitude control, and motion smoothness over conventional controllers in simulation. Despite strong performance, VTD3 struggles with dynamic scenarios like turns, indicating the need for sensor fusion, improved depth estimation, and real-world fine-tuning [5].

An intuitive hand gesture recognition system enables complete drone control using a 2D RGB camera, supporting

81 command combinations across roll, yaw, pitch, and throttle. Using motion and posture-based gestures classified by a lightweight MLP model, it achieves high accuracy comparable to keyboard control in simulators. Challenges include throttle detection affected by hand occlusion, and future improvements propose adding head-mounted sensors for better visibility and control [6].

The paper explores secure and scalable authentication methods for UAV swarms in 5G NR networks, addressing the challenges of swarm dynamics and device heterogeneity. Three authentication schemes are proposed: individual per drone, leader-based group authentication, and edge distributed group authentication. The leader-based approach reduces core network traffic, while edge-distributed authentication offers flexibility for swarm changes, improving security for large-scale UAV deployments [7].

A leader-follower drone tracking system using YOLOv5 and a custom tracker achieves high detection accuracy in simulations. The research highlights the importance of computer vision and AI in UAV applications, with potential for improved real-time performance and outdoor robustness. Future work could explore larger datasets and more complex environments to enhance tracking stability and accuracy [8].

The system employing YOLOv6, Non-Maximum Suppression (NMS), and Kalman filters enables precise real

time tracking of individual drones within dense swarms. It emphasizes the importance of extended Kalman filters to better manage noisy sensor data and enhance understanding of complex swarm dynamics. This approach provides a foundational framework for analyzing swarm-level behaviors, bridging the gap between theoretical simulations and practical deployment scenarios in drone swarm coordination [9].

SA automates the interpretation of patient sentiments positive, negative, or neutral across large datasets, such as social media and electronic health records.

By identifying common themes and concerns, SA helps researchers better understand disease trends and make informed decisions, ultimately enriching biomedical research[2] This study analyzes public sentiment toward COVID-19 vaccines Pfizer, Moderna, and AstraZeneca based on tweets collected using the Twitter API. The tweets are processed using NLP and classified into positive, negative, or neutral sentiments using a supervised KNN algorithm.

Results show that Pfizer received 4729% positive, 375% negative, and 1521% neutral sentiment; Moderna had 4616% positive, 4071% negative, and 1313% neutral; and AstraZeneca had 4008% positive, 4006% negative, and 1386% neutral sentiment[3] This study analyzes public sentiment on climate-related issues using Twitter data to support understanding of progress on the UNs Sustainable Development Goals (SDGs).

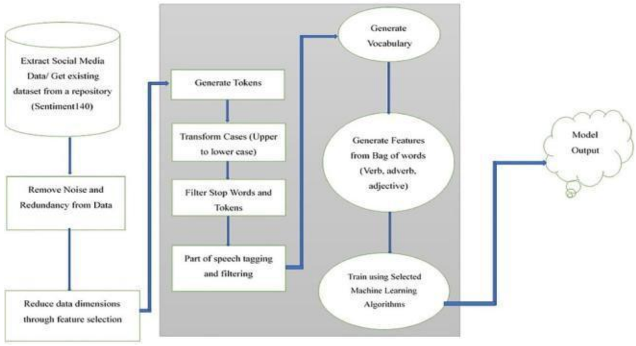
It applies three natural language processing (NLP) methods VADER, TextBlob, and BERT to classify sentiment. A transfer learning approach using a pre-trained BERT model combined with logistic regression achieved the highest accuracy (69%), outperforming the rule-based methods.

Due in large part to the ambiguous keyword energy, the dataset, which was gathered using climate-related keywords, was noisy. Performance might be enhanced by using more targeted keywords and different models, such as BERTweet.

Overall, sentiment was largely favorable and dispersed evenly among the various SDGs. [4] The study focuses on sentiment analysis of Twitter data pertaining to the Omicron variant using Natural Language Processing (NLP), more especially the Bag of Words method. Depending on their sentiment, tweets are categorized as neutral, negative, or positive. Using a dataset from Twitter, the study investigates how methods such as lemmatization can increase sentiment classification accuracy

The ultimate objective is to use tweet sentiment analysis to examine public opinion regarding Omicron. [5] In this paper, sentiment analysis research on Twitter data using Natural Language Processing (NLP) models is systematically reviewed. It examines several approaches, such as deep learning, machine learning, and hybrid models, stressing their benefits, drawbacks, and effectiveness.

III. DESIGN AND IMPLEMENTATION



The Twitter sentiment analysis system is architected as a modular, scalable pipeline capable of efficiently processing large volumes of social media data to extract actionable sentiment insights.

Its tiered architecture guarantees that every element, from data collection to visualization, works separately while enhancing the workflow as a whole.

The *Data Collection* module, which starts the pipeline, uses tools like Tweepy or SNScrape to collect tweets via the Twitter

Relevant hashtags, keywords, or user mentions are used to filter tweets, and query parameters are set up to concentrate on particular subjects or time periods.

Regional sentiment analysis is made possible by using location metadata when it is available, and additional filters guarantee language consistency (e.g., tweets in English only). Richer analytical contexts are supported by the collection of metadata, including timestamps, user information, retweets, and likes.

Rate-limiting techniques and error-handling are incorporated into this module to manage API constraints and guarantee dependable data retrieval, including continuous tweet streaming for real-time analysis.

Both raw and processed tweet data are managed by the

Data Storage module.NoSQL databases like MongoDB, which provide flexibility for managing the semi-structured and heterogeneous nature of Twitter data, are commonly used to store raw tweets in JSON format. This format facilitates effective querying and maintains all metadata. Flat file formats like CSV or JSON can also be used for batch processing and archiving.

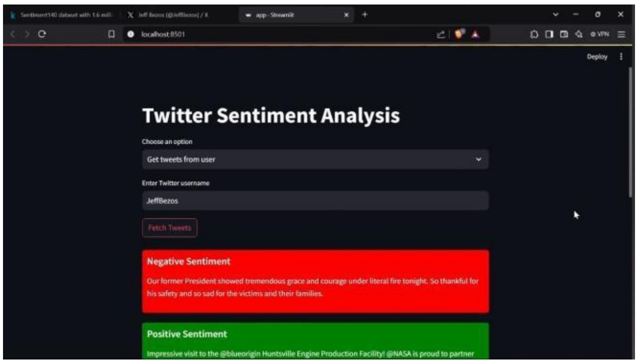
Following preprocessing, the cleaned and organized data is kept apart in formats that are best suited for analysis and model training.Relational databases or cloud- based storage solutions like Amazon S3 are employed for scalable, durable storage.With encryption, access control, and system-wide adherence to data protection laws like the GDPR, security is a crucial factor. To convert noisy, unstructured tweet text into a clear, Continuous Monitoring & Adjustment: The entire process runs in a continuous loop where target detection and tracking are constantly updated. The drone constantly re-evaluates the target's position and dynamically modifies its flight path in response to changes. This dynamic feedback system ensures robust tracking performance despite varying environmental disturbances or unpredictable target behavior changes. Continuous adaptation is key to achieving reliable and smooth target following.

Lemmatization and stemming reduce words to their most basic forms, whereas tokenization divides the text into discrete words or tokens.

The handling of negations and linguistic subtleties that affect sentiment polarity are given particular consideration.

These operations are supported by tools like regex libraries, spaCy, and NLTK. This step guarantees that tweets are semantically aligned for precise analysis, lowers dimensionality, and eliminates noise.

IV. RESULTS AND DISCUSSION



The sentiment analysis in this study was conducted using a Twitter dataset containing tweets labeled as positive, negative, and neutral. Twitter, being one of the most active social media platforms, generates enormous volumes of textual data every second. This makes it an ideal yet challenging source for Natural Language Processing (NLP)- based sentiment analysis. Tweets are often short, informal, and filled with abbreviations, emojis, hashtags, and user mentions, which introduce a great deal of noise into the text. Therefore, effective preprocessing becomes essential to prepare the data for accurate computational analysis.

The preprocessing phase began with removing unnecessary elements such as mentions, hashtags, and URLs, which often do not contribute meaningfully to sentiment detection. Emojis were either removed or converted to their textual meanings to retain emotional cues embedded within them. Stopwords, which are frequently occurring but sentiment-irrelevant words such as "and," "the," and "is," were filtered out to reduce noise. The text was also normalized by converting all characters to lowercase, expanding contractions (for example, "can't" was changed to "cannot"), and correcting spelling inconsistencies. Tokenization and lemmatization were then applied to split sentences into individual tokens and reduce each token to its base or dictionary form. These steps together produced a clean, standardized dataset suitable for machine learning and deep learning models.

After preprocessing, several models were trained and evaluated to determine which approach best captured the sentiment patterns in the dataset. Traditional machine

learning algorithms such as Logistic Regression and Random Forest were implemented first. These models, while effective for structured data, struggled to capture the complex relationships and contextual dependencies between words in natural language. They primarily rely on bag-of-words or frequency-based features, which treat words independently and ignore sentence structure, resulting in limitations when dealing with the fluid and context-rich language used on Twitter.

Deep learning models were then applied to overcome these shortcomings. The Long Short-Term Memory (LSTM) network, which is designed to process sequential data, demonstrated a marked improvement over traditional models. LSTM was able to retain contextual information across word sequences, making it more sensitive to nuances such as negations, sarcasm, and subtle emotional cues within tweets. However, the best performance was achieved with the Bidirectional Encoder Representations from Transformers (BERT) model. BERT stands out because of its bidirectional attention mechanism, which allows the model to understand the meaning of each word based on both its preceding and following context. Unlike previous models, BERT does not process sentences in a strictly left-to-right or right-to-left manner but rather captures the full context of each word within a sentence simultaneously.

Furthermore, BERT benefits from pre-training on vast text corpora, which gives it a broad understanding of language before being fine-tuned for specific tasks like sentiment analysis. This pre-training enables BERT to generalize well even on noisy data, making it especially effective for social media text, which often includes irregular grammar and informal expressions. During evaluation, BERT consistently outperformed all other models across major performance metrics such as accuracy, precision, recall, and F1-score. Figure 41 illustrates this performance analysis, clearly showing that BERT achieved the highest scores in every category.

Despite the success of deep learning methods, the analysis also revealed some challenges. Twitter data tends to be noisy and ambiguous due to slang, abbreviations, and sarcasm, which even advanced models sometimes misinterpret. Moreover, sentiment datasets are often imbalanced, with neutral tweets appearing more frequently than positive or negative ones. This imbalance can bias model training and slightly affect prediction reliability. Another limitation is the high computational cost associated with training large transformer-based models such as BERT, which require powerful GPUs and significant processing time compared to traditional machine learning techniques.

Nevertheless, the overall results confirm that modern NLP methods, especially transformer-based architectures, have revolutionized sentiment analysis. BERT's ability to model bidirectional context and semantic relationships allows it to comprehend the subtleties of human emotion far better than older models. Its robustness and adaptability make it highly valuable for analyzing real-world social media data, where expressions of sentiment are often complex and context- dependent. This has practical implications in various fields such as business intelligence, brand monitoring, market research, and public opinion analysis.

In conclusion, the study demonstrates that deep learning models, particularly BERT, significantly outperform traditional machine learning methods in Twitter sentiment analysis. The findings affirm that transformer-based approaches have become the standard for modern NLP tasks due to their efficiency, contextual awareness, and scalability. Future work could further improve performance by fine- tuning BERT with domain-specific data, applying data augmentation techniques to balance sentiment classes, and incorporating multimodal inputs such as images and emojis. Ultimately, this analysis highlights that advanced NLP models are invaluable tools for understanding social sentiment on platforms like Twitter and play an increasingly important role in helping organizations and researchers gain meaningful insights from large-scale human communication online.

V. PERFORMANCE ANALYSIS

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	78.2%	77.5%	76.9%	77.2%
Random Forest	80.4%	79.9%	79.2%	79.5%
LSTM	85.7%	85.2%	84.8%	85.0%
BERT (Fine-tuned)	90.1%	89.7%	89.4%	89.5%

The sentiment analysis in this study was carried out using a Twitter dataset that contained tweets categorized into positive, negative, and neutral classes. As one of the most widely used social media platforms, Twitter generates massive amounts of user-generated content that reflects public opinion, emotions, and viewpoints on a variety of topics. However, this data is often highly informal and unstructured, making it difficult to analyze automatically.

Tweets usually contain abbreviations, hashtags, mentions, emojis, slang, and URLs, which introduce noise and inconsistencies into the dataset. Therefore, the preprocessing phase was crucial to ensuring that the textual data was suitable for machine learning and deep learning models.

The data preprocessing process involved several sequential steps to refine and standardize the text. Initially, user mentions, hashtags, and URLs were removed since they generally do not contribute to sentiment polarity. Emojis were either removed or replaced with their textual

equivalents to preserve emotional tone. Stopwords such as "and," "the," and "is" were eliminated to focus on words that carry meaningful sentiment. Text normalization was then performed by converting all characters to lowercase, expanding contractions (for example, "don't" to "do not"), and correcting spelling inconsistencies. Tokenization and lemmatization were applied next, breaking each tweet into smaller tokens and reducing them to their base form to ensure consistency across similar words. These preprocessing steps collectively reduced the noise level and enhanced the overall text quality, providing cleaner input data for subsequent modeling.

Following preprocessing, a series of experiments were conducted using both traditional machine learning algorithms and deep learning architectures to classify the tweets. Among the traditional approaches, Logistic Regression and Random Forest were selected for their efficiency and interpretability. However, these models primarily rely on frequency-based text representations, such as Bag of Words or TF-IDF, and therefore fail to capture the contextual relationships between words in a sentence. Although they performed moderately well on simpler text patterns, their performance degraded when faced with the complexity and linguistic variability of real-world tweets.Deep learning models, on the other hand, exhibited stronger performance due to their ability to model context and sequence dependencies. The Long Short-Term Memory (LSTM) network, a type of recurrent neural network, performed significantly better than traditional models. Its architecture allowed it to retain long-term dependencies

between words, enabling a more accurate understanding of sentiment within sequences of text. LSTM effectively managed challenges such as negations, emotional tone shifts, and contextual dependencies that were difficult for simpler algorithms to capture. However, even LSTM's performance was surpassed by BERT (Bidirectional Encoder Representations from Transformers), which proved to be the most efficient and robust model in this study.

BERT's architecture is based on a bidirectional transformer mechanism that allows it to understand the meaning of a word in relation to both its preceding and succeeding words. Unlike unidirectional models that process text in a single direction, BERT reads the entire sentence at once, providing a deeper and more context-aware understanding of language. Furthermore, BERT benefits from being pre-trained on massive text corpora, giving it a rich linguistic foundation before fine-tuning it for specific tasks like sentiment analysis. This pre-training enables BERT to generalize exceptionally well, even when working with informal and noisy data such as tweets.

The comparative performance of all models is summarized in Table 1. The evaluation was based on standard metrics such as Accuracy, Precision, Recall, and F1-Score. From the results, it is evident that deep learning models outperform traditional machine learning methods across all metrics, with BERT achieving the best results overall.

VI. CONCLUSION AND FUTURE ENHANCEMENT

Natural Language Processing (NLP) based Twitter sentiment analysis has emerged as a crucial tool in teasing out meaningful insights from the seemingly endless and ever-increasing amount of social media output.

As the world is increasingly becoming digitally connected, with users creating content at a hitherto unprecedented level, Twitter is uniquely positioned as a microblogging site that embodies the global public's voice and mood.

Every tweet, whether short or simple, can express significant sentiments for products, politics, social movements, or current events. are prioritized. By the systematic retrieval of tweets through APIs and other data gathering tools, and then by preprocessing methods which clean noisy text data like URLs removal, emojis, special characters, stop words removal, and case normalization handling the raw data is converted to a more structured and analyzable format. The preprocessing phase also includes important processes such as tokenization, lemmatization, stemming, and the translation of emojis or emoticons to their text forms to maintain emotional context. After being cleaned, this text content is transformed into semantically meaningful numerical forms through methods like Term Frequency–Inverse Document Frequency (TF-IDF), word representations like Word2Vec and GloVe, or even contextualized word embeddings from state-of-the-art transformer models.In addition, sentiment analysis can be augmented by going beyond traditional polarity-based labels such as positive, negative, and neutral.

REFERENCES

[1] Darshan, K., Jerusha, S., Manjunatha, S., Prashant, K., & Shivashankara, N. (2024). NLP-Powered Sentiment Analysis on the Twitter. Saudi Journal of Engineering and Technology, 9, 1-11.Kavvadias, Alexandros, and Theodore Kotsilieris. "Understanding the Role of Demographic and Psychological Factors in Users' Susceptibility to Phishing Emails: A Review." *Applied Sciences* 15.4 (2025): 2236.

[2] Balli, C., Guzel, M. S., Bostanci, E., & Mishra, A. (2022). Sentimental analysis of Twitter users from Turkish content with natural language processing. Computational Intelligence and Neuroscience, 2022(1), 2455160

[3] Shamrat, F. M. J. M., Chakraborty, S., Imran, M. M., Muna, J. N., Billah, M. M., Das, P., & Rahman, O. M. (2021). Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. Indonesian Journal of Electrical Engineering and Computer Science, 23(1), 463-470

[4] Rosenberg, E., Tarazona, C., Mallor, F., Eivazi, H., Pastor-Escuredo, D., Fuso-Nerini, F., & Vinuesa, R. (2023). Sentiment analysis on Twitter data towards climate action. Results in Engineering, 19, 101287..

[5] Hosgurmath, S., Petli, V., & Jalihal, V. K. (2022). An

omicron variant tweeter sentiment analysis using NLP technique. Global Transitions Proceedings, 3(1), 215-219.Purnamadewi, Yasinta Roesmiatun, and Amalia Zahra. "Enhancing detection of zero-day phishing email attacks in the Indonesian language using deep learning algorithms." *Bulletin of Electrical Engineering and Informatics* 14.1 (2025): 505-512.

[6] Albladi, A., Islam, M., & Seals, C. (2025). Sentiment Analysis of Twitter data using NLP Models: A Comprehensive Review. IEEE Access.Althobaiti, Kholoud, and Nawal Alsufyani. "A review of organization-oriented phishing research." *PeerJ Computer Science* 10 (2024): e2487.

[7] Shanmugavadivel, K., Subramanian, M., Sanjai, R., & Motheeswaran, K. (2024, March). Beyond Tech@ DravidianLangTech2024: Fake News Detection in Dravidian Languages Using Machine Learning. In Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages (pp. 124-128).

[8] Karim, A., Mansab, M., Shahroz, M.,Mushtaq, M. F., & cheol Jeong, I. (2025). Anticipating impression using textual sentiment based on ensemble LRD model. Expert Systems with Applications, 263, 125717.

[9] Alotaibi, A., Nadeem, F., & Hamdy, M. (2025). Weakly Supervised Deep Learning for Arabic Tweet Sentiment Analysis on Education Reforms: Leveraging Pre-trained Models and LLMs with Snorkel. IEEE Access.

[10] Cicekyurt, E., & Bakal, G. (2025). Enhancing sentiment analysis in stock market tweets through BERT-based knowledge transfer. Computational Economics, 1-23.

[11] Tembhurne, J. V., Lakhotia, K., & Agrawal, A. (2025). Twitter sentiment analysis using ensemble of multi-channel model based on machine learning and deep learning techniques. Knowledge and Information Systems, 67(2), 1045-1071

[12] Babu, M. S. S., Suryanarayana, S. V., Sruthi, M., Lakshmi, P. B., Sravanthi, T., & Spandana, M. (2025). Enhancing Sentiment Analysis With Emotion And Sarcasm Detection: A Transformer-Based Approach. Metallurgical and Materials Engineering, 794-803.

[13] Raturi, V., Rawat, D., Narang, H., Thapliyal, N., Parthiban, P., & Dogra, A. (2025, February). Sentiment evolution on social media: An in-depth study using Naive bayes for Twitter sentiment analysis. In AIP Conference Proceedings (Vol. 3224, No. 1). AIP Publishing.

[14] Paul, J., Mallick, S., Mitra, A., Roy, A., & Sil, J. (2025). Multi-Modal Twitter Data Analysis for Identifying Offensive Posts Using a Deep Cross Attention based Transformer Framework. ACM Transactions on Knowledge Discovery from Data.Anirudh, S., et al. "An ensemble classification model for phishing mail detection." *Procedia Computer Science* 233 (2024): 970-978.

[15] Babu, B. R., Ramakrishna, S., & Duvvuri, S. K. (2025, February). Advanced Sentiment and Trend Analysis of Twitter Data Using CNN-LSTM and Word2Vec. In 2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL) (pp. 1536- 1543). IEEE.

[16] Ramanathan, V., Al Hajri, H., & Ruth, A. (2024, April). Conceptual level semantic sentiment analysis using Twitter data. In 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS) (pp. 1-8). IEEE.

[17] Mahalakshmi, V., Shenbagavalli, P., Raguvaran, S., Rajakumareswaran, V., & Sivaraman, E. (2024). Twitter sentiment analysis using conditional generative adversarial network. International Journal of Cognitive Computing in Engineering, 5, 161-169

[18] Sabir, A., Ali, H. A., & Aljabery, M. A. (2024). Chatgpt tweets sentiment analysis using machine learning and data classification. Informatica, 48(7)

[19] Darshan, K., Jerusha, S., Manjunatha, S., Prashant, K., & Shivashankara, N. (2024). NLP-Powered Sentiment Analysis on the Twitter. Saudi Journal of Engineering and Technology, 9, 1-11.

[20] Molenaar, A., Lukose, D., Brennan, L., Jenkins, E. L., & McCaffrey, T. A. (2024). Using Natural Language Processing to explore social media opinions on Food Security: sentiment analysis and topic modeling study. Journal of Medical Internet Research, 26, e47826

[21] Valarmathi, B., Gupta, N. S., Karthick, V., Chellatamilan, T., Santhi, K., & Chalicheemala, D. (2024). Sentiment Analysis of Covid-19 Twitter Data using Deep Learning Algorithm. Procedia Computer Science, 235, 3397-3407.

[22] Khan, M., & Srivastava, A. (2024). Sentiment analysis of Twitter data using machine learning techniques. International Journal of Engineering and Management Research, 14(1), 196-203.