Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
Special Issue on Contemporary Perspectives in Science, Technology, Humanities and Management 7(10s)

A COMPARATIVE SAAS FRAMEWORK FOR REAL-TIME SOCIAL MEDIA SENTIMENT ANALYSIS USING MULTI-MODEL APPROACH

Papiya Mukherjee [1], Parul Saini [2], Priyanka Tyagi [3], Shweta Singh [4], Varsha Srivastava [5], Mamunur Islam [6]

[1] School of Computer Science and Engineering, IILM University, Greater Noida, India
[2] School of Computer Science and Engineering, IILM University, Greater Noida, India
[3] School of Computer Science and Engineering, IILM University, Greater Noida, India
[4] School of Computer Science and Engineering, IILM University, Greater Noida, India
[5] School of Computer Science and Engineering, IILM University, Greater Noida, India
[6] School of Computer Science and Engineering, IILM University, Greater Noida, India

Received 26 February 2026
Accepted 28 April 2026
Published 18 May 2026

Corresponding Author
Papiya Mukherjee,
papiya.mukherjee@iilm.edu

DOI
10.29121/shodhkosh.v7.i10s.2026.8187

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2026 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License.

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

OPEN ACCESS

ABSTRACT

The process of examining a piece of text to identify whether the underlying sentiment expressed is positive, negative, or neutral is known as sentiment analysis. While it may seem straightforward, sentiment analysis is currently playing a vital role in comprehending how people perceive and interact on social media platforms, which is critical information for businesses and content creators. By gauging the sentiments conveyed through textual data, companies and individuals creating online content can gain valuable insights into the emotional responses and engagement levels of their target audiences. This paper presents our platform "Analytix" which is designed for analyzing sentiments across YouTube, Facebook, Reddit, Twitter and WhatsApp etc. It supports 197 languages and can handle multiple languages at once. The core of our platform is its user-friendly dashboard that pulls out the sentiments (positive, negative, or neutral) from posts across different social media platforms. It offers users the flexibility to choose from three different model options for sentiment analysis: GPT, transformer models, and a custom proprietary sentiment analysis model developed inhouse. It deeply analyzes each post to identify the overall sentiment people have towards it.

Keywords: Sentiment Analysis, LLM, Supervised Machine Learning, Openai, Web Scraping, Data Dashboarding

1. INTRODUCTION

In the rapidly evolving social media landscape, content creators play a pivotal role in shaping narratives, driving conversations, and engaging with audiences across various platforms. However, navigating the vast expanse of user-generated content can be a daunting task, making it challenging for creators to effectively gauge the sentiment and

engagement levels of their audiences. This research proposes a holistic software-as-a-service (SaaS) solution that empowers content creators with social media sentiment analysis capabilities. The proposed SaaS solution serves as a comprehensive toolkit, streamlining the entire process of real-time data collection, processing, and visualization, enabling content creators to gain valuable insights into their audience's sentiments and engagement levels in real time. By harnessing the power of natural language processing (NLP) and sentiment analysis techniques, this solution aims to provide a deep understanding of user reactions, opinions, and emotions expressed on social media platforms. One of the key objectives is to establish a robust data collection and preprocessing pipeline capable of efficiently gathering and structuring social media data from multiple sources in real time, including handling complexities such as informal language, slang, and multilingual content [13], and ensuring accurate and reliable data processing.

This research focuses on developing state-of-the-art sentiment analysis models and algorithms tailored specifically for the social media domain. These models leverage advanced NLP techniques, including deep learning [11] and lexicon-based approaches [14], to identify and extract sentiment information from user comments, posts, and interactions accurately. A unique feature is the incorporation of a prompt-based interface, which allows content creators to communicate with data using natural language prompts, enabling them to explore specific aspects of their audience engagement, ask targeted questions, and receive insightful visualizations and sentiment summaries in response.

By presenting the sentiment analysis results through inter- active and visually appealing dashboards, the solution aims to provide content creators with a comprehensive understanding of their audience's sentiments, enabling them to make data-driven decisions, refine their content strategies, and foster deeper connections with their audiences. However, there are certain limitations, such as potential latency issues arising from large volumes of social media data being processed in real-time, and the model's inability to accurately detect humorous comments, emojis, and neutral sentiments, which can impact the overall sentiment analysis accuracy.

The scope of this project is focused on real-time data collection and processing, eliminating the need for storage space as the data is processed and discarded after analysis. This approach ensures efficient resource utilization and minimizes the overhead associated with maintaining large databases. The research paper will delve into the technical aspects of the proposed SaaS solution, including the architecture, data processing pipelines, sentiment analysis models, and visualization techniques. Additionally, it will explore strategies to mitigate latency issues and improve the model's performance in detecting complex sentiments. Realworld applications, case studies, and potential impact on the content creation industry will also be discussed, highlighting the value of this solution in enhancing audience engagement and driving successful content strategies.

2. LITERATURE REVIEW

This section provides an overview and analysis of past research undertaken by other researchers on sentiment analysis. Several studies have used sentiment analysis techniques to investigate public debates and sentiments across social media platforms. Federico Neri et al. [1] assessed the sentiment expressed in over 1,000 Facebook posts regarding news broadcasts and compared it to Rai (an Italian public broadcasting service) and the growing commercial enterprise La7 (Neri, 2018). This study compared its findings with those of Osservatorio di Pavia, an Italian research institute that specializes in media analysis at both the theoretical and empirical levels, with a specific emphasis on studying political communication in mass media. Notably, this study combined sentiment analysis of social media data with quantitative audience data provided by the Auditel, a company that records viewership statistics. The researchers wanted to acquire full knowledge of public perception and interaction with these news providers by comparing Facebook sentiment analysis with quantitative audience numbers.

To analyze large amounts of multimedia data, this study used the I Syn Semantic Center, a content-enabling system that provides deep semantic information access and dynamic categorization features (Neri et al., 2018). This approach allowed the researchers to quickly examine and analyze the attitude indicated in the gathered Facebook postings. The integration of sentiment analysis approaches, third-party data sources, and specialized research tools demonstrated in this study exemplifies the comprehensive approach that is frequently required to glean significant insights from social media. By merging sentiment analysis with other data streams and utilizing advanced analysis techniques, researchers can gain a more comprehensive picture of public opinion and sentiments within certain fields.

Anukarsh et al. [2] created a technique to discriminate be-tween sarcastic and non-sarcastic tweets using vocabulary and emoticons. This study considered slang and emoji values from dictionaries. Sarcasm in tweets from the Twitter

Streaming API is identified using a variety of classification approaches, including Random Forest, Gradient Boosting, Adaptive Boost, Gaussian Naive Bayes, Logistic Regression, and Decision Tree. To achieve maximum efficiency, the best classification algorithms were selected and used with various preprocessing and filtering techniques, such as emoji and slang dictionary mapping.

Mondher et al. [3] proposed a novel pattern-based approach for recognizing sarcastic language on Twitter. This study used four feature sets. The feature sets identify different types of sarcasms and grade their relevance. Feature sets were used to classify tweets based on their sarcastic polarity. Using the ensemble method, the resulting model achieved 83.1% accuracy and 91.1% precision.

More recently, Naseem et al. [10] proposed a transformer-based deep intelligent contextual embedding model for Twitter sentiment analysis, demonstrating the effectiveness of modern transformer architectures in social media text classification. Liu et al. [9] introduced RoBERTa, a robustly optimized BERT pretraining approach that has since become a widely adopted baseline for various NLP tasks including sentiment analysis. Furthermore, Chen et al. [12] and Biswas et al. [15] have explored the capabilities of large language models such as GPT and ChatGPT for sentiment analysis, highlighting both their strengths in handling multilingual and nuanced text and their limitations in terms of cost and latency.

3. PROPOSED WORK

The proposed study aims to develop a comprehensive software-as-a-service (SaaS) solution that empowers content creators with real-time social media sentiment analysis capabilities.

The following key objectives were addressed.

1) **Real-time data collection and pre-processing pipeline:** Establish a robust and scalable data collection and preprocessing pipeline capable of efficiently gathering and structuring social media data from multiple sources in real-time. This pipeline will handle complexities, such as informal language, slang, and multilingual content, ensuring accurate and reliable data processing.
2) **State-of-the-Art Sentiment Analysis Models:** Develop advanced sentiment analysis models tailored specifically to the social media domain. These models leverage cutting-edge natural language processing (NLP) techniques, including deep learning and lexicon-based approaches, to accurately identify and extract sentiment information from user comments, posts, and interactions.
3) **Prompt-based Interface:** Incorporate a prompt-based interface that allows content creators to communicate with data using natural language prompts. This interface enables them to explore specific aspects of their audience engagement, ask targeted questions, and receive insight-ful visualizations and sentiment summaries in response.
4) **Interactive and visually appealing dashboard:** Present sentiment analysis results through interactive and visually appealing dashboards provide content creators with a comprehensive understanding of their audience's sentiments. These dashboards will enable data-driven decision making, content strategy refinement, and deeper connections with audiences.
5) **Real-time Data Processing and Visualization:** Implement real-time data processing and visualization capabilities, ensuring that content creators can access up-to-date sentiment analysis results and promptly make informed decisions.
6) **Latency Mitigation and Performance Optimization:** Explore strategies to mitigate potential latency issues arising from large volumes of social media data processed in real-time. Optimize the sentiment analysis models and data-processing pipelines to improve the performance and scalability.
7) **Sentiment-complexity handling:** Investigate techniques to enhance the model's ability to accurately detect com-plex sentiments, such as humorous comments, emojis, and neutral sentiments, thereby improving overall sentiment analysis accuracy.
8) **Efficient Resource Utilization:** Implement a stream-lined approach to data processing and visualization, eliminating the need for permanent storage space by processing and discarding data after analysis. This approach ensures efficient resource utilization and minimizes the overhead associated with maintaining large databases.

9) **Real-world Applications and Case Studies:** Conduct real-world case studies and evaluate the impact of the proposed SaaS solution on the content-creation industry. Demonstrate the value of the solution in enhancing audience engagement and driving successful content strategies through data-driven insights.
10) **Technical Documentation and Implementation:** Com-prehensive technical documentation including de-tailed descriptions of the architecture, data processing pipelines, sentiment analysis models, and visualization techniques employed in the SaaS solution.

4. METHODOLOGY

Our social media sentiment analysis dashboard was developed using an organized process based on the Software Development Life Cycle (SDLC), which consists of six stages: planning, analysis, design, implementation, testing, and documentation. Every stage was essential for ensuring that the system was successfully delivered.

Figure 1

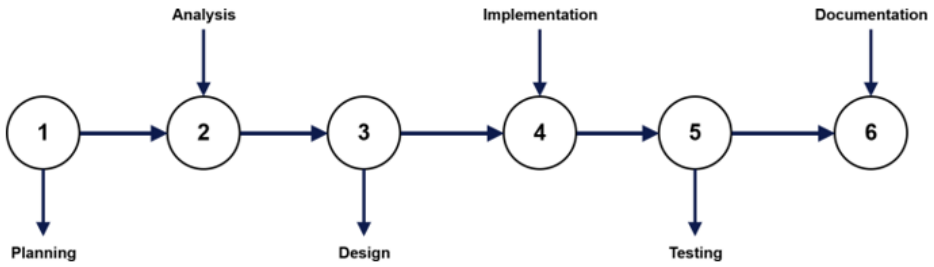

Figure 1 Various Phases of Development Used

4.1. PLANNING

The research team established the project's goals, scope, and objectives during this phase, which dealt with setting the project's foundation. Setting deadlines and assigning resources were important tasks. This stage served as a blueprint for the development's later phases.

4.2. ANALYSIS

The research team established the project's goals, scope, and objectives during this phase, which dealt with setting the project foundation. Setting deadlines and assigning resources are important tasks. This stage served as a blueprint for the later phases of development.

4.3. DESIGN

The research team thoroughly evaluated the functions and user requirements of the system throughout this phase. This requires an analysis of the current procedures. Gaining a thorough grasp of user expectations and needs was the aim.

4.4. IMPLEMENTATION / DEVELOPMENT

In this phase, the research team translated the design specifications into working codes for a Social Media Sentiment Analysis Web Application. Utilizing suitable programming languages and development tools, the features and functionalities of the application are built according to the defined requirements. Continuous communication and collaboration among team members facilitated the smooth execution of the implementation process, ensuring the successful deployment of a robust sentiment analysis platform for social media data.

4.5. TESTING

Once the implementation phase was completed, rigorous testing was conducted on the social media Sentiment Analysis web application to validate its functionality and performance. The goal is to identify and resolve any defects or inconsistencies in the application before deployment. Tadesse (2026)

4.6. DOCUMENTATION

Throughout the development process, comprehensive documentation was created to provide guidance for users and developers. Clear and concise documentation ensured that authorities could understand and utilize the system effectively.

The complete workflow of our project is described below using Figure 2.

1) **Collect Post URL:** The initial step in which the URLs of the posts or sources from which comments need to be analyzed are collected. These URLs can be from various platforms such as social media, forums, blogs, or any other online source where user comments are available.

Figure 2

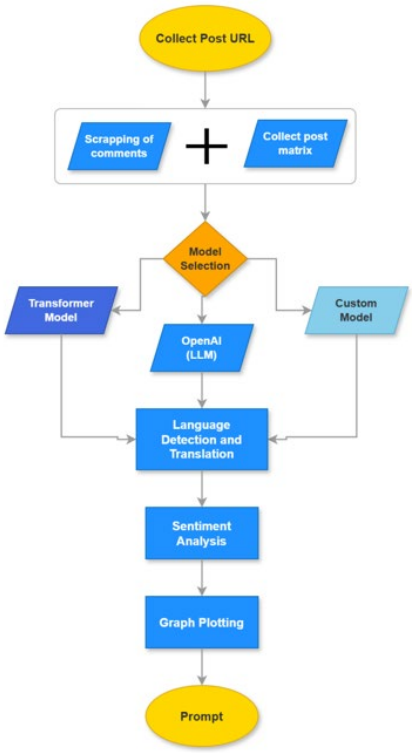

Figure 2 Workflow of the Project

2) **Scrapping of Comments:** Once post URLs are collected, a scraper API [4] is used to extract or scrape comments from these sources. The scraper API is a tool or library that can navigate through web pages and retrieve the desired data; in this case, user comments.
3) **Collect Post Matrix:** In this step, scraped comments are organized and collected into a structured format, often referred to as the post matrix. This matrix may include additional metadata or information related to the posts and comments, such as timestamps, user information, and other relevant details.
4) **Selection of Model:** Here user or system selects the appropriate model for performing sentiment analysis on the collected comments. Three model options are available.

- **Transformer Model [7]:** This is a type of language model based on the transformer architecture, which is effective for natural language processing tasks such as sentiment analysis. Transformer models are generally fast, and can handle a single language.
- **OpenAI Model (LLM):** This refers to OpenAI language models, which are large language mod-els capable of handling multiple languages (190+ languages, as mentioned). These models are versa-tile but may be more resource-intensive than other options.
- **Custom Model:** This option allows the user or system to utilize a custom-built model specifically designed or trained for sentiment analysis

5) **Language Detection and Translation:** Because the collected comments can be in various languages, this step involves detecting the language of each comment and translating it to a common language, in this case, English. This step ensures that sentiment analysis can be performed consistently across comments in different languages.
6) **Sentiment Analysis:** After comments are translated into English, sentiment analysis is performed on them. This step involves analyzing the text of the comments and predicting whether the sentiment expressed is positive or negative. The sentiment analysis model selected in the previous step is used for this task.
7) **Graph Plotting:** Based on the sentiment analysis results, a graph or visual representation of the sentiment scores was generated. This graph provides an overview of the sentiment distribution across the analyzed comments, helping identify trends or patterns.
8) **Prompt:** In addition to the primary workflow, the model features an additional component called "Prompt," which enables users to enter their own specific requirements or prompts. The system then generates sentiment scores based on these inputs. This functionality is especially helpful for tailoring sentiment analysis and investigating various scenarios.

The user interface of our social media sentiment analysis web application seems as follows:

The landing page used for selecting the sentiment analysis model is shown in fig. 3(a). The options displayed were "GPT-4", "transformers-sentiment-analysis", and "Analitix". The dropdown menu at the top is currently set to "GPT-4". [8] The background is dark, giving it a sleek, modern appearance typical of coding or data analysis tools and interfaces. The dashboard is created using streamlit [4].

After the selection of the model, the next page of our web application, shown in Figure 3(b), displays video details or statistics for a YouTube [6] or similar video titled "Pro Chef Reacts. To Chef Ranveer Brar's EGGLESS Omelette?". The information provided includes the total number of views (106967), likes (5112), and comments (781). Additionally, it shows the ratio of likes to views (0.05), and the ratio of comments to views (0.01). There is also a button labeled "description," which likely reveals more details about the video content

Figure 3

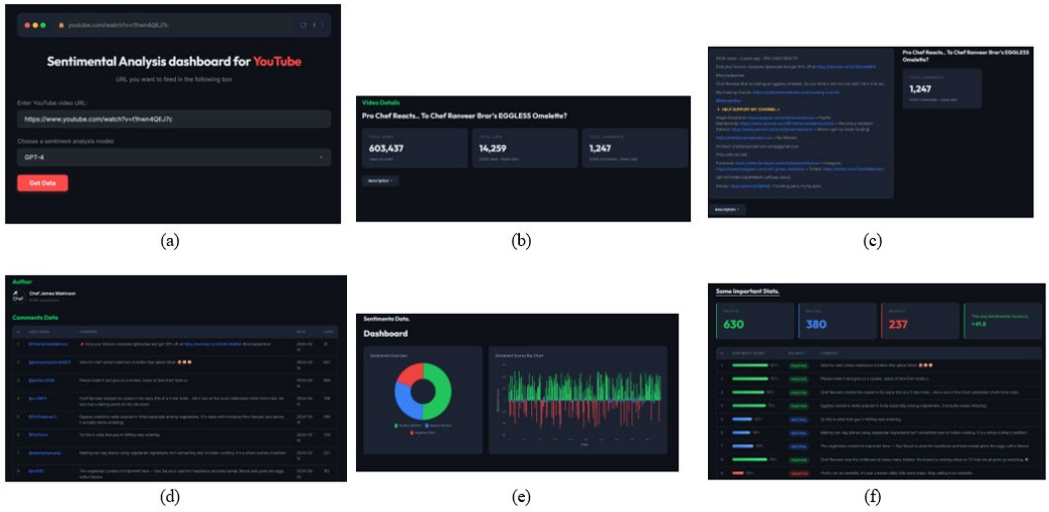

Figure 3 (A) The Main Dashboard (B) Video Details(C) Description of the Video (D) Comments Data (E) Graphical Representation (F) Sentiment Score.

Figure 3(c) shows the description section of a video titled "Pro Chef Reacts. To Chef Ranveer Brar's EGGLESS Omelette?". The description includes information about the cooking course, links to the video on YouTube[6], ways to support the channel through donations and memberships, links to the creator's social media accounts and merchandise, and mentions of kitchen equipment used by the creator.

In addition, the image displays the total number of comments (781) and the ratio of comments to views (0.01) for this video. In addition, the image displays the total number of comments (781) and the ratio of comments to views (0.01) for this video. Figure 3(d) shows information about the author "Chef James Makinson" and a table displaying comments data for a video, likely the one titled "Pro Chef Reacts. To Chef Ranveer Brar's EGGLESS Omelette?". The table includes columns for username, comment text, date, and likes, including Chef James Makinson's pinned comment promoting a cookware discount, users discussing the eggless omelette concept itself, one user suggesting Chef James should try to make and review the eggless omelette, and another mentioning an unrelated incident about an order lacking proper ingredients.

While the comments touch on various topics, they appear to relate mostly to the video's subject matter of reacting to or discussing an eggless omelette prepared by another chef.

Figure 3(e) displays a sentiment analysis dashboard with two visualizations: a pie chart and bar chart. The pie chart titled "Sentiments Overview" shows the distribution of the sentiment scores across the data. This indicates that 50.5% of the data have positive sentiment, 30.5% are neutral, and 19% have negative sentiment. The bar chart titled "Sentiment Scores Bar Chart" plots the individual sentiment scores for each data point along the x axis. The bars are color-coded, with green representing positive scores, red representing negative scores, and blue representing neutral or near-zero scores. The y-axis of the bar chart shows the sentiment score values, which range from approximately −50 to 50. The chart reveals that while there is a mix of positive, negative, and neutral sentiment scores, the majority of the scores cluster around the neutral/near-zero range.

Figure 3(f) displays some important statistics related to sentiment analysis. At the top, it shows the counts for positive (199), neutral (120), and negative (75) sentiments. The average sentiment score was as +41.8. Below, there is a table with rows displaying the sentiment score on a scale, the associated polarity (positive, neutral, or negative), and a corresponding comment or text snippet. The table provides examples of comments that have been analyzed and categorized based on their perceived sentiment or emotional tone, ranging from highly positive to highly negative scores.

5. RESULT

Table I highlights the key capabilities and performance characteristics of the three models: Analytix, Transformer, and GPT. The following metrics are compared: multilingual support, bilingual support, latency, accuracy, irony identification, positive/negative polarity detection, and neutral polarity detection.

The Analytix and Transformer models lack multilingual and bilingual capability, but the GPT model contains a checkmark (✓).In terms of latency, the GPT model has the lowest (0.8 ms), followed by Analytix (3 ms) and Transformer (9 ms).

Table 1

Table 1 Comparison of the Three Models

Parameter	Analytix	Transformer	GPT
Multilingual Support	×	×	✓
Bilingual Support	×	×	✓
Latency	3 ms	9 ms	0.8 ms
Accuracy	82%	72%	96%
Irony Detection	×	×	×
Positive Polarity Detection	✓	✓	✓
Negative Polarity Detection	✓	✓	✓
Neutral Polarity Detection	✓	✓	✓

Accuracy is stated differently in all three methods. Analytix has an accuracy rate of 82%, Transformer has 72%, and GPT has 96%. All three models lack Irony Detection capabilities, as shown by a cross (×). Positive, negative, and neutral polarity detections are all recorded.

Our model integrated two social media networks, YouTube[7] and Reddit, but excluded additional platforms such as Reddit [5], Facebook, Instagram, WhatsApp conver-sations, and Linkdin, as shown in Table 2.

6. CONCLUSION

In the ever-evolving landscape of social media, content creators face the constant challenge of understanding and effectively engaging with their audiences. This research pro-poses an innovative Software-as-a-Service (SaaS) solution that empowers content creators by providing real- time sentiment analysis capabilities, enabling them to gain valuable insights into their audiences' emotions, opinions, and engagement levels across various platforms. Through the development of robust data collection pipelines, advanced sentiment analysis models tailored for social media, and an intuitive prompt-based interface, the proposed SaaS solution offers a comprehensive toolkit for content creators. By presenting sentiment analysis results through interactive and visually compelling dashboards, this solution facilitates data driven decision-making, content strategy refinement, and deeper connections with audiences.

While addressing challenges such as latency issues and accurately detecting complex sentiments like humour and emojis, this research demonstrates the potential impact of the proposed SaaS solution on the content creation industry. By providing content creators with real-time, data-driven insights, this solution paves the way for more informed, effective, and audience centric content strategies, driving successful engagement and fostering meaningful connections in the dynamic social media landscape.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

The authors wish to thank IILM University, Greater Noida, Uttar Pradesh Campus for their support in completing this research.

Table 2

Table 2 Accessibility for Different Social Media Platforms

Social Media Platform	Integration	Accessibility
Youtube	✓	✓
Reddit	✓	✓
Facebook	×	×
Instagram	×	✓
Whatsapp Chats	×	×
Linkdin	×	×

REFERENCES

"API Reference." Streamlit. https://docs.streamlit.io/develop/ api-reference (accessed April 10, 2024).
"Documentation" gpt-4-turbo-and-gpt-4. https://platform.openai.com/ docs/models/gpt-4-turbo-and-gpt-4 (accessed March 30).
"Reddit API Documentation" Reddit. https://docs.streamlit.io/develop/ api-reference (accessed date April 15, 2024).
"Transformers" Hugging face. https://huggingface.co/transformers/v3.5. 1/quicktour.html (accessed April 7, 2024).
"Youtube Data API" References. https://developers.google.com/youtube/ v3/docs/search/list (accessed April 5, 2024).
Alqahtani, A., Ilyas, H., Naeem, M. W., and Bashir, S. (2023). "A Comprehensive Survey on Multilingual Sentiment Analysis", IEEE Access, vol. 11, pp. 30048–30071. IEEE.

Anukarsh, G. P., Sanjana, S., Skanda M. B., and Harish, B. S. (2017). "Sentiment Analysis for Sarcasm Detection on Streaming Short Text Data", 2nd International Conference on Knowledge Engineering and Applications, IEEE, 2017.

Biswas, E., Karmaker Santu, S. K., and Chakraborty, S. (2023). "Ex-amining the Impact of ChatGPT and LLMs on Sentiment Analysis and Opinion Mining", Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 3214–3229. ACL.

Bonta, V., Kumaresh, N., and Janardhan, N. (2022). "A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis", Asian Journal of Computer Science and Technology, vol. 8, no. S2, pp. 1–6.

Chen, Z., Zhang, Y., and Liu, B. (2022). "Sentiment Analysis in the Era of Large Language Models: A Reality Check", arX iv preprint arXiv:2305.15005, 2023.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2020). "RoBERTa: A Robustly Optimized BERT Pretraining Approach", Proceedings of the International Conference on Learning Representations (ICLR 2020).

Mondher, B., and Tomoaki, O. (2017). Apatternbased approach for sarcasm detection on twitter. IEEE. Transl. 2017

Naseem, U., Razzak, I., Musial, K., and Imran, M. (2020). "Transformer Based Deep Intelligent Contextual Embedding for Twitter Sentiment Analysis", Future Generation Computer Systems, vol. 113, pp. 58–69. Elsevier.

Neri, F., Crisci, A., Agresta, S., Ambrosini, L., Cucchiarelli, A., & Stramaglia, S. (2018). Analyzing Sentiments in Social Media: A Case Study on News Broadcasts. In Proc. of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2017) (pp. 419-428). SCITEPRESS.

Tadesse, Y. (2026). Ai-Powered Security Strategies for The Osi Model. ShodhAI: Journal of Artificial Intelligence, 3(1), 9–19. https://doi.org/10.29121/shodhai.v3.i1.2026.62

Yadav, A., and Vishwakarma, D. K. (2020). "Sentiment Analysis Using Deep Learning Architectures: A Review", Artificial Intelligence Review, vol. 53, no. 6, pp. 4335–4385. Springer.