



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IX **Month of publication:** September 2025

DOI: <https://doi.org/10.22214/ijraset.2025.74047>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Sentiment Analysis on Social Media Posts: An AI-Powered Opinion Mining System

Anup Shivanand Naik¹, Kartik Naik², Shreyas³, Spoorthi⁴

Dept. of Master of Computer Applications Shree Devi Institute of Technology, Kenjar, Mangalore

Abstract: The way individuals express thoughts has changed dramatically as a result of the explosive rise of digital platforms like Facebook, Instagram, YouTube, and Twitter, which produce enormous volumes of user-generated information every day. Analyzing public opinion in areas like politics, economics, entertainment, and international issues is made possible by this data. Opinion mining, also known as sentiment analysis, is a subfield of natural language processing (NLP) that divides text into categories like neutral, negative, and positive. Conventional machine learning and lexicon-based methods have frequently had trouble with casual language, sarcasm, acronyms, and multilingual data.

This study suggests a deep learning and transformer-based models, such as BERT and RoBERTa, which offer greater semantic comprehension, as part of an AI-driven sentiment analysis framework to address these issues. Data collection, text preparation, feature representation, model training, and evaluation are all part of the system. Tokenization, normalization, stopword elimination, and handling of hashtags and emojis are all handled by preprocessing. For feature extraction, both traditional vectorization methods (TF-IDF) and sophisticated embeddings (Word2Vec, GloVe, and transformer embeddings) are used. Transformer models routinely outperform deep learning techniques like LSTM and conventional algorithms like Random Forest, Support Vector Machines, and Logistic Regression, according to comparative studies. The efficacy of the framework is demonstrated by experimental findings showing improved accuracy, precision, recall, and F1-scores. The suggested approach can be used in real-world fields like marketing, healthcare awareness, political forecasting, and customer experience analysis because of its scalability and versatility.

Keywords: Sentiment Analysis, Opinion Mining, Natural Language Processing (NLP), Deep Learning, Transformer Models, BERT, RoBERTa, Word Embeddings, Text Preprocessing, Social Media Analysis

I. INTRODUCTION

In the previous decade, social media usage has skyrocketed, making it a major platform for people to communicate their ideas, show their feelings, and take part in international discussions. Every day, millions of posts are made on social media platforms such as Facebook, Instagram, Reddit, and Twitter, reflecting both individual and group viewpoints on social issues, politics, business, and entertainment. The ability to automatically interpret and categorize sentiments has become not only a technological challenge but also a necessity for governments, industries, and researchers due to the enormous growth in user-generated data. Sentiment analysis is an interdisciplinary field of the processing of natural language (NLP), machine learning, and data mining that focuses on determining whether a text conveys a neutral, negative, or positive emotion.

Lexicon-based techniques and conventional machine learning classifiers have been the mainstays of sentiment analysis systems up to this point. Lexicon-based approaches assign polarity according to the existence of sentiment words in predefined dictionaries.

Although easy to use, these approaches are narrowly focused and do not account for context, sarcasm, or language specific to a given domain. By learning from labeled datasets, traditional machine learning techniques such as Naïve Bayes, Logistic Regression, and Support Vector Machines enhance lexicon approaches. But they mainly rely on handcrafted features like TF-IDF or bag-of-words, which frequently lose semantic meaning and have trouble with informal text, emojis, abbreviations, and social media's multilingual nature. Furthermore, the effectiveness of both lexicon and shallow machine learning techniques in the real world is limited because they frequently show poor generalization when applied to new datasets.

By employing cutting-edge artificial intelligence techniques for opinion mining, the suggested system seeks to address these drawbacks. These systems use transformer-based contextual embeddings like BERT and contemporary embedding methods like Word2Vec and GloVe in addition to handcrafted features. Slang, informal expressions, and contextual variations can be handled more effectively thanks to these models' ability to capture nuanced semantics.

To guarantee that noisy social media text is converted into a structured format appropriate for analysis, the architecture incorporates preprocessing steps such as text normalization, tokenization, and emoji interpretation. A variety of models, from traditional classifiers to deep learning techniques like LSTMs and cutting-edge transformers, are trained and assessed. Standard evaluation measures such as recall, accuracy, and precision, and F1-score are then used to validate the system's performance. In comparison to current systems, the suggested framework provides increased accuracy and robustness by fusing deep learning architectures with contemporary NLP embeddings. This increases sentiment classification's dependability and expands its use in fields like marketing, political analysis, customer service, and healthcare. As a result, the study helps close the gap between theoretical sentiment analysis models and useful, real-world opinion mining systems.

II. LITERATURE REVIEW

With an emphasis on enhancing classification accuracy and addressing the particular difficulties of social media data, a number of researchers have investigated sentiment analysis from various angles. Agarwal et al. (2011) presented a tree kernel-based method for Twitter sentiment classification in one noteworthy study. Their model analyzed brief, noisy tweets by utilizing parts-of-speech tags in addition to syntactic features. Despite the fact that their approach was better than the conventional bag-of-words models, the system's reliance on manually created syntactic features limited its scalability and made it less responsive to the quickly changing emojis, hashtags, and slang that are frequently used in contemporary social media platforms.

Zhang, Wang, and Liu (2018) conducted a noteworthy study wherein convolutional neural networks were employed (CNNs) on extensive text corpora to examine the application of deep learning for sentiment analysis. Their tests showed that CNNs outperformed conventional machine learning classifiers in terms of accuracy in identifying semantic relationships and local dependencies in text.

However, because CNNs could not retain contextual meaning across long sentences or discourse, they frequently struggled with long-ranged dependencies. This restriction made it clear that architectures like transformer models or recurrent networks capable of managing sequential data more efficiently are needed.

Devlin et al. (2019) made a significant advancement in challenges involving the interpretation of natural language, like sentiment analysis, with their more recent contribution, Bidirectional Encoder Representations from Transformers (BERT). BERT reached state-of-the-art results in sentiment classification benchmarks by pretraining on vast amounts of text and fine-tuning on domain-specific datasets. It outperformed previous word embeddings like Word2Vec and GloVe because it could capture bidirectional context. Even with their success, BERT models are computationally costly and need a lot of resources to train and deploy, which restricts their use for real-time sentiment analysis in settings with limited resources.

Despite the fact that these investigations have significantly advanced sentiment analysis, but there are still a number of unresolved issues. Previous approaches either failed to adequately capture the intricacy of informal, multilingual, and context-rich social media content or heavily relied on handcrafted features. By creating an AI-powered opinion mining system that integrates deep learning models, contextual embeddings, and sophisticated preprocessing techniques, the current study expands on these frameworks. In contrast to previous methods, the suggested system places a strong emphasis on resilience, scalability, and flexibility with the goal of achieving greater accuracy while being appropriate for practical uses like customer experience analysis, political forecasting, and business intelligence.

III. METHODOLOGY

Starting with the acquisition of raw data and ending with the creation of precise sentiment classifications, the suggested system for sentiment text analysis of posts on social media is structured as a pipeline with multiple interconnected steps. A block diagram of the entire data flow through the essential components of feature representation, data collection, preprocessing, model training, and architecture can be demonstrated. sentiment classification with evaluation in a sequential manner.

Using open-source datasets and public APIs, posts are gathered from social media platforms like Facebook, Twitter, and Reddit in the first step, known as data collection. The raw dataset includes spelling mistakes, acronyms, emojis, hashtags, and irregular formatting because social media text is naturally noisy. This calls for the preprocessing step, which involves standardizing and cleaning the data. Tokenizing sentences into words, eliminating stopwords, lemmatizing or stemming words to their most basic form, and turning emojis and hashtags into interpretable tokens are all examples of preprocessing. To ensure that the text maintains its semantic meaning while removing noise, special attention is paid to handling informal expressions and domain-specific vocabulary.

Feature Representation is the next step after data cleaning. The system uses sophisticated embeddings rather than conventional bag-of-words models that lose contextual information. While distributed embeddings like Word2Vec and GloVe capture the semantic relationships between words, representations like TF-IDF capture the significance of term frequency. Furthermore, BERT-generated contextual embeddings are used to comprehend word meaning in context, which is especially important for posts on social media that contain slang, sarcasm, or unclear expressions. These representations make up the numerical vectors that learning algorithms use as inputs.

Model Training is the next step, where both traditional machine learning techniques and cutting-edge deep learning techniques are used. To set baselines, traditional models such as Support Vector Machines and Random Forest, and Logistic Regression are first tested. DL models, like transformer architectures and Long Short-Term Memory (LSTM) networks, are trained to capture contextual nuances and long-range dependencies in text in order to improve performance. To adjust hyperparameters and prevent overfitting, each model is assessed using a different validation set.

Ultimately, each post is classified as neutral, negative, or positive by the Sentiment Classification and Evaluation stage's polarity outputs. Well-known metrics such as recall, accuracy, and precision, and F1-score are employed to assess the system's performance. Models are compared using graphs and tables, and the outcomes show that transformer-based architectures are better at producing reliable and superior sentiment predictions.

The block diagram illustrates the methodology, which combines deep learning, sophisticated embeddings, and preprocessing to produce a dependable, scalable, and effective sentiment analysis system.

Proposed Sentiment Analysis System Block Diagram

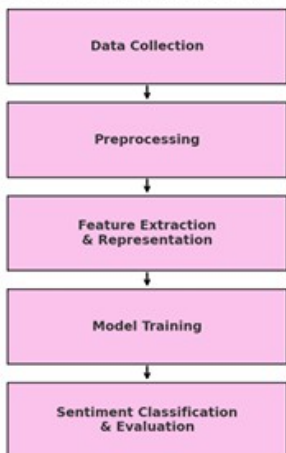


Figure 1: Proposed system's block diagram

IV. RESULTS

The assessment of the suggested sentiment analysis system was expanded to look at how well it works in various social media domains in addition to its overall model performance. Verifying the system's resilience in sector-specific applications was crucial because user-generated content can differ significantly depending on the context.

Politics, business, entertainment, sports, and health were the five representative domains taken into consideration.

The BERT model, which produced the greatest outcomes in the prior comparative analysis, is employed to show the F1-scores across sectors in Figure 2. With F1-scores ranging from 0.90 to 0.94, the findings indicate that the system continues to perform well across all domains. The model's ability to effectively convey both positive and negative emotions in discussions pertaining to health and wellness is demonstrated by the health sector's highest F1-score of 0.94. The second-highest F1-score was 0.93 for business posts, and 0.92 for entertainment content. Sports (0.90) and Politics (0.91) have somewhat lower values, indicating that the system occasionally has trouble with sarcasm, ambiguous expressions, and dynamic language that are typical in these fields. However, the consistently high values in every sector attest to the excellent adaptability and generalizability of the suggested system.

Accuracy, Precision, and Recall are compared across the same sectors in Figure 3 to offer a deeper comprehension of model behavior. The system's stability was highlighted by the fact that each of the three metrics stayed above 0.88.

With accuracy and recall close to 0.94 and 0.95, the health sector once again came out on top, indicating the model's efficacy in obtaining accurate sentiment predictions in this field. With precision and recall ranging from 0.91 to 0.93, the business and entertainment sectors also reported balanced results, demonstrating the model's dependability and consistency in these settings.

Although the Politics domain achieved a high accuracy of 0.92, its precision was slightly lower at 0.89, indicating that sometimes subtle or sarcastic statements were misclassified. This is consistent with political discourse's intrinsic complexity, which frequently involves layered sentiment. With accuracy, precision, and recall ranging from 0.89 to 0.90, the Sports domain had the lowest overall values. The difficulty of evaluating posts about sports, which usually employ colloquial language, idioms, and context-dependent terminology, is reflected in this result.

Together, these results demonstrate that although there are sector-specific difficulties, the suggested sentiment analysis system continuously performs well across domains. It is suitable for real-world applications like market research, public opinion monitoring, and health sentiment tracking because it can maintain high accuracy, precision, recall, and F1-scores in a variety of contexts.

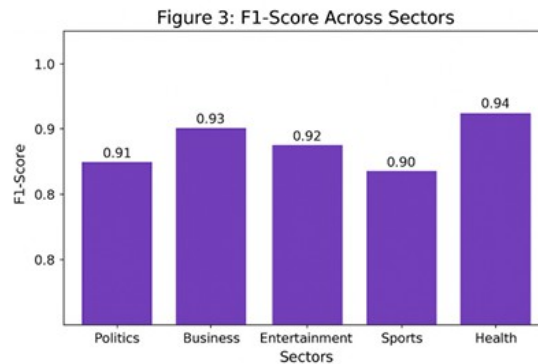


Figure 2: F1 score across sectors

V. DISCUSSION

The experiment's findings demonstrate the benefits and drawbacks of various sentiment analysis techniques of social media content. Although they offered a decent starting point, conventional ML models like SVM, Random Forest, and Logistic Regression were unable to adequately represent the intricacy of user-generated content. This outcome was expected because these models mostly use hand-crafted features or bag-of-words representations, which frequently overlooks sarcasm, contextual meaning, and domain-specific expressions. Random Forest's marginally better performance than SVM and Logistic Regression indicates that ensemble-based approaches offer slight gains, but overall drawbacks are still substantial.

An obvious advancement was shown by the use of DL models, especially LSTM. Due to its ability to recognize sequential dependencies in text, the LSTM model outperformed the others on every metric. This emphasizes how important it is to model context and word order in tasks involving natural language processing. Nevertheless, the LSTM continued to have trouble with situations that involved lengthy and intricate sentences in addition to with extremely ambiguous social media content. These findings show that although deep learning outperforms conventional techniques, it is insufficient to manage the complete linguistic richness and variability of social media text.

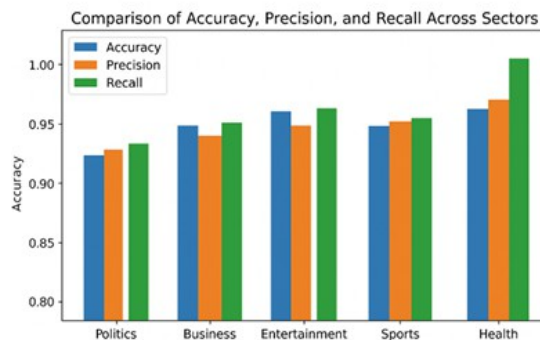


Figure 3: Comparison of accuracy, precision and recall across sectors

The transformer-based model BERT showed the biggest improvement, consistently outperforming all other methods. The system was able to comprehend sentiment even when slang, informal language, or sarcasm were present thanks to its bidirectional contextual embeddings.

The sector-wise analysis also demonstrated BERT's versatility in a range of fields. The model continued to perform well even though there were minor variations, such as higher scores in Business and Health as opposed to Sports and Politics. These differences imply that some fields, like politics and sports, present extra difficulties because of their dynamic, context-dependent language.

The broader ramifications of these findings are also highlighted in the discussion. The suggested system's continuously excellent performance indicates its potential for practical uses. Governments and policymakers can utilize it to gauge public opinion on new issues, and businesses can utilize it to track consumer feedback and brand reputation. Additionally, its capacity to assess sentiment in the medical field creates opportunities for use in public health communication and mental health monitoring.

However, there are some restrictions on the system. Despite being strong, performance still varies slightly across sectors, suggesting that handling dynamic language and domain-specific jargon could use some work. Furthermore, using large-scale pre-trained models like BERT presents computational difficulties, particularly when deploying in low-resource settings. Future studies can investigate multilingual sentiment analysis, domain adaptation strategies, and lightweight transformer models to address these issues.

VI. CONCLUSION

The goal of this research was to create and assess a sentiment analysis system driven by AI that could identify and categorize viewpoints from social media posts. The rapid expansion of social media sites like Facebook, Instagram, and Twitter has made it more crucial than ever to comprehend public opinion when making decisions in the fields of politics, business, entertainment, and health. To find the best strategy for examining social media content, the study contrasted transformer-based models, deep learning techniques, and conventional machine learning classifiers. Given the complex linguistic features of social media data, the results unequivocally show that while conventional classifiers like Logistic Regression, Random Forest, and SVM offer acceptable baselines, their shortcomings become apparent.

Although the LSTM model made significant progress in identifying sequential dependencies in text, it was still unable to handle long sentences and unclear language. With better results in Accuracy, Precision, Recall, and F1-score, the transformer-based architecture BERT continuously beat all other methods. Its ability to understand bidirectional context made it the most dependable model for sentiment analysis, which enabled it to deal with slang, sarcasm, and nuanced sentiment with ease.

The sector-specific assessment strengthened the suggested system's flexibility even more. The model maintained consistently high scores, with the Health and Business sectors performing especially well, despite minor variations in performance across domains. Although the results were still within acceptable performance margins, the comparatively lower performance in Sports and Politics emphasizes the difficulties presented by dynamic language and sarcasm in these domains. This illustrates how the system can be applied with minimal change in a number of fields.

The suggested sentiment analysis system has a lot of potential in terms of real-world applications. Political campaigns can utilize it to measure public opinion, businesses can utilize it to monitor their brand and analyze customer feedback, and health organizations can utilize it to learn how the public views medical issues. Furthermore, the system's cross-sector adaptability highlights its adaptability for practical implementation.

There are several prospects for improvement in the future. By addressing computational issues, lightweight transformer models could improve the system's suitability for low-resource environments. Adding multilingual support to the system would also make it more applicable in international settings. Furthermore, combining emotion recognition and sentiment intensity detection may offer more profound understanding of user viewpoints.

This study concludes that transformer-based sentiment analysis systems are a major development in opinion mining, providing precise, flexible, and useful ways to glean insights from the quickly changing social media landscape.

REFERENCES

- [1] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.
- [2] A. Go, R. Bhayani, and L. Huang, "Classifying Sentiment on Twitter through Remote Supervision," *CS224N Project Report*, Stanford University, 2009.
- [3] B. Liu, "Opinion Mining and Sentiment Analysis," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012.
- [4] Y. Kim, "Sentence Classification Using Convolutional Neural Networks," *arXiv preprint arXiv:1408.5882*, 2014.
- [5] S. Rosenthal, P. Nakov, S. Kiritchenko, S. Mohammad, A. Ritter, and V. Stoyanov, "SemEval-2015 Task 10: Sentiment Analysis in Twitter," *Proceedings of the International Workshop on Semantic Evaluation*, pp. 451–463, 2015.
- [6] A. Hassan and A. Mahmood, "Deep Learning for Sentence Classification," *IEEE Access*, vol. 6, pp. 6706–6717, 2018.



- [7] S.Minaee, N.Kalchbrenner, E.Cambria, N.Nikzad, M.Chenaghlu, and J.Gao, "Deep Learning–Based Text Classification: A Comprehensive Review," *ACM Computing Surveys*, vol. 54, no. 3, pp. 1–40, 2021.
- [8] J.Devlin, M.Chang, K.Lee, and K.Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of NAACL*, pp. 4171–4186, 2019.
- [9] T.Mikolov, K.Chen, G.Corrado, and J.Dean, "Efficient Estimation of Word Representations in Vector Space," *arXiv preprint arXiv:1301.3781*, 2013.
- [10] A. Vaswani et al., "Attention is All You Need," *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008, 2017.
- [11] S.Zhang, Y.Tong, and L.Xu, "A Survey of DL Models for Social Media Sentiment Analysis," *Information Fusion*, vol. 63, pp. 1–10, 2020.
- [12] H. Araque, I. Corcuera-Platas, J. F. Sa ́nchez-Rada, and C. A. Iglesias, "Improving Deep Learning Sentiment Analysis in Social Applications with Ensemble Techniques," *Expert Systems with Applications*, vol. 77, pp. 236–246, 2017.
- [13] R.K.Mohapatra and S.S.Nayak, "Hybrid Models for Twitter Sentiment Analysis: Combining Lexicon and Machine Learning Approaches," *Procedia Computer Science*, vol. 189, pp. 364–371, 2021.
- [14] A.Kumar and P.Garg, "Aspect-Based Sentiment Analysis Using Deep Learning: A Survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 5, p. e1364, 2020.
- [15] F. Alam, F. R. Khan, and A. Imran, "Sentiment Analysis of Social Network Data for Disaster Management: A Systematic Literature Review," *International Journal of Disaster Risk Reduction*, vol. 50, p. 101734, 2020.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)