



# Poll Prediction Using Sentiment Analysis: Perspectives from Social Media Posts

Manuraj Tomar\*<sup>1</sup>, Rambir Singh<sup>2</sup>

<sup>1</sup> Department of CSE, MIET Meerut, India

\*manuraj.tomar.mtcs.2024@miet.ac.in

<sup>2</sup> Department of CSE (AIML), MIET Meerut, India

rambir.singh@miet.ac.in

**Abstract.** The exponential development in social media platforms has altered the way people voice their opinion, feelings, and political positions, with the resulting enormous amount of user-generated data that can be used for data-driven decision-making. Sentiment analysis — one of the important techniques used in natural language processing — allows for the systematic extraction and interpretation of public sentiment expressed in texts shared on social media. This research examines the use of sentiment analysis on platforms such as Twitter (X), Facebook and Instagram for the purpose of extracting meaningful information to assist in the prediction of election results and political trends. The goal of the research, which analyzes large-scale social media data, is to infer voter preferences, ideological leanings, and collective political attitudes in near real-time. The study critically analyzes current methods of sentiment analysis, such as lexicon-based, machine learning, and deep learning methods, that are most suitable for the political domain, along with their advantages and disadvantages. Additionally, it deals with important problems such as data noise, sarcasm, misinformation, demographic bias and data privacy related ethical issues. The paper points out that there is a lack of research that could help increase the accuracy of election forecasting by incorporating multimodal data sources, advanced contextual modeling, and sentiment fusion across platforms. Overall, the work lays a scientific basis for improving the reliability of sentiment-driven election prediction systems via social media analytics.

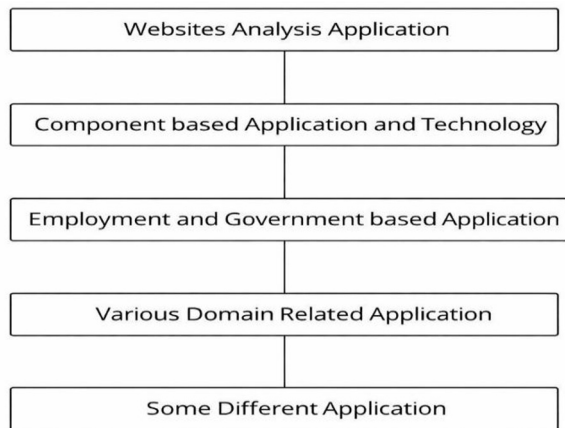
**Keywords:** Sentiment Analysis, NLP, Social Media, Optimization.

## 1 Introduction

In the digital age, social media platforms have taken center stage by giving people a forum to express their thoughts on a variety of subjects, including current affairs, goods, and world news. Because these platforms allow opinions to be shared in real time with a worldwide audience, they are extremely useful instruments for gauging public mood. The goal of this research area, called sentiment analysis, is to determine if textual material expresses a neutral, negative, or positive mood. The more general area of Natural Language Processing (NLP) incorporates it. Word, phrase, document,

aspect, and concept levels are among the levels at which sentiment analysis may be carried out. Sentiment analysis is an evaluation of the word and sentence level polarity of discrete words or subjective phrases. Document-level sentiment analysis, in its turn, examines the overall tone of the entire text. On the other hand, aspect-level sentiment analysis captures sentiments that are related to specific characteristics or aspects of an object such as a product [1].

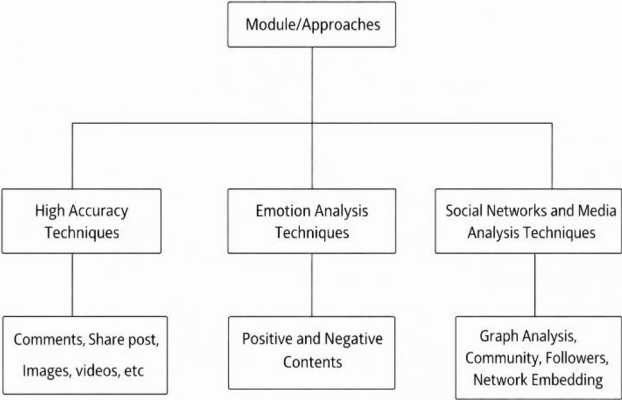
Concept-level sentiment analysis goes beyond surface-level text analysis by considering the semantic association of words. This approach allows a more refined understanding of the text that communicates perspectives by combining words with similar meanings into everyday concepts. One can treat a comment such as “Flat rate was the best thing that happened this year, UBER bring it back” as a positive one because the semantically related word is price. Concept-level analysis offers more information about the sentiment of a document because semantic networks are used to find important phrases and the concepts that they are associated with. This area gives advanced sentiment analysis and opinion mining algorithms that are no longer limited to word-level sentiment analysis [2]. These methods enhance sentiment analysis across various data sets by facilitating a greater and more complex division into classifications by utilizing semantic networks and contextual connections. This is a significant advance in understanding and analysis of complex sentiment data as it transcends to the concept-level analysis.



**Fig. 1.** Key Applications of Sentiment Analysis.

This hierarchical framework in Figure 1 shows various areas of application of a website analysis framework. At the top it has a Website Analysis Application, which is the central system. This system is realized with the help of component-based applications and technologies which allow for development that is modular and scalable. The framework is subsequently implemented in employment and government-based decision support and analytics applications. Moreover, it spreads to other applications that relate to domains indicating diversity across sectors. Lastly, it also incorporates other varied applications, which underline the flexibility of methods

of website analysis across different real-world applications. The volume-based strategy also gives valuable information for the calculation of election results as well as practices such as online followers, likes and shares on postings. To form a perception of the society on whether opinion is positive or negative concerning their preferred parties, this method gathers user interactions. The users of social media will usually engage in discussions of election matters and publicly express their support for political parties [3-5]. These interactions can be identified in a structural graph with centrality measures such as inbound, outbound, or bidirectional links. This paper discusses the use of social media analytics to predict the outcome of elections through a critical evaluation of the past literature and objectives. Through these measurements, the study aims at understanding the relationship between linked behavior and the real voting results as a foundation for further surveys and application in practice.



**Fig. 2.** Predictive Modeling Techniques for Social Media.

Figure 2 demonstrates significant modules and strategies of social media and sentiment analysis systems. It is categorized into three main branches: High Accuracy Techniques, Emotion Analysis Techniques, and Social Networks and Media Analysis Techniques. High accuracy methods handle various user-created material like comments, shared posts, images and videos. Emotion analysis methods are based on the detection of positive and negative feelings in content. Social network and media analysis approaches provide analysis of relationships based on graph analysis, community detection, follower interactions, and network embedding to obtain more profound behavioral and structural information. The paper provides an in-depth discussion of sentiment analysis including its methods, applications, and future trends. It discusses the relationship between social media and election prediction results with a specific focus on the application of sentiment analysis in predicting the outcome of elections. The report identifies gaps in research by examining previous studies and offers understanding of the current state of the topic [6]. The increasing significance of social media in forecasting election outcomes is the driving force behind this study.

The authors were motivated to investigate how social media data may be used to predict election results with accuracy.

The remaining sections of this paper are organized as follows: Section 2 is the literature review, Section 3 is research methodology and case analysis, Section 4 is results and Section 5 is the conclusion of the study including limitations and future research.

## 2 Literature Review

Predicting election results from social media data has been the subject of much research by social analysts and scholars throughout the world. They have given a summary of their research, highlighting the part social media plays in these forecasts and summarizing relevant publications. As seen in Tables 1 and 2, this section highlights important studies in this field.

**Forecasting Social Media Datasets:** This part examines studies that use data from social media to forecast election results. Sentiment profiles of politicians were studied in studies like [7], which provided detailed insights from election campaigns where tweets were a useful tool for predicting outcomes. Similar to this, [8] looked into the possibility of using Twitter data to analyze election results, such as the 2011 Irish election, and found that sentiment-based metrics outperformed volume-based ones. Other studies examined the 2010 U.K. general election using retweet graphs to show party hierarchies and identify considerable campaign segregation, and examined the 2010 U.S. Senate race, highlighting sample anomalies on social media platforms. Real Clear Politics (RCP) data and election forecasts were shown to correlate in research by [9], and social networking and sentiment analysis were studied in Pakistan's 2013 elections, successfully energizing voters. Additionally, other studies provided insightful information on public opinion by using Twitter data to pinpoint dynamic social trends. In order to increase prediction accuracy, some studies combined social media insights with Twitter's structural network data to study the 2012 U.S. presidential election, while others used user tweets and inductive data from social media to investigate election predictions in three EU nations. Last but not least, one study emphasized the significance of political communication by demonstrating emotional evaluations and associations between the number of retweets and candidate tweets, thereby highlighting the function of social media in forming electoral narratives [10].

**Table 1.** Study of Related Work

| Authors | Data Source         | Algorithms      | Outcomes |
|---------|---------------------|-----------------|----------|
| [1] [2] | Twitter             | NRC Lexicon     | Positive |
| [4] [5] | Government Websites | Optimization    | Positive |
| [6]     | Twitter             | RNN             | Positive |
| [8] [9] | Twitter             | Common Approach | Negative |
| [10]    | Twitter             | Optimization    | Positive |

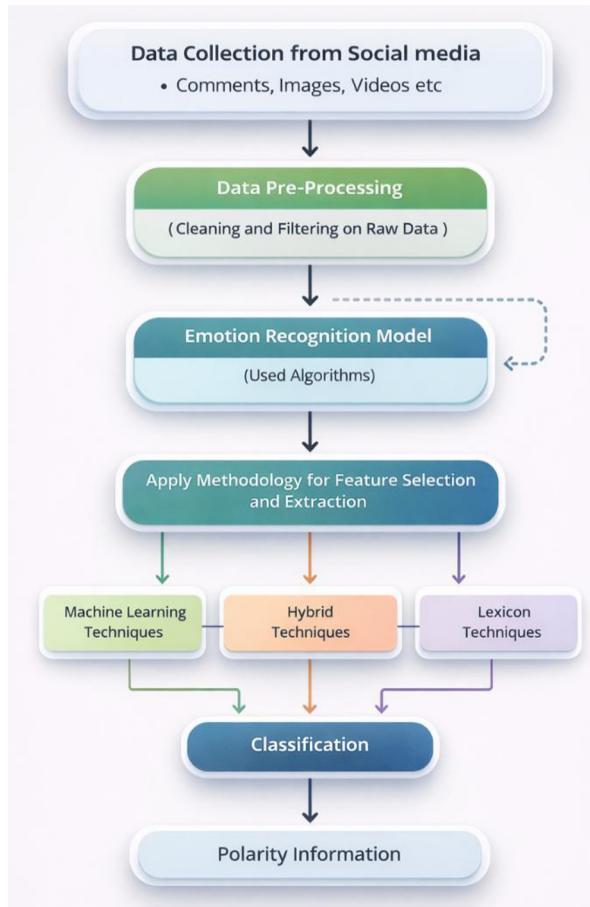
|      |         |     |          |
|------|---------|-----|----------|
| [11] | Twitter | CNN | Positive |
|------|---------|-----|----------|

**Table 2.** Social Media Comments Analysis and Accuracies

| Authors   | Platform     | Techniques             | Accuracy |
|-----------|--------------|------------------------|----------|
| [1] [12]  | Social Media | Multiple Convolutional | 88.31%   |
| [7] [13]  | Social Media | Graph Regularization   | 27.6%    |
| [3] [14]  | Social Media | Maximum Entropy        | 90%      |
| [15] [16] | Social Media | Supervised Learning    | 82.5%    |
| [17]      | Social Media | Deep Learning          | 97.89%   |

Analysis of Social Comments: The field studies the political stances of users of social media by identifying their posts, voting, and participation in online debates. Such moves often help in electioneering and contribute to political rhetoric. The research was capable of forecasting latent political preferences with a high accuracy level. Similarly, some studies analyzed text in search of clues that revealed voting preferences; such clues were often derived from what the author had posted. In another work, [11] divided the population based on gender and age by using Twitter content with particular care to the popular opinion about the Scottish Independence Referendum and defining whether the user was in support of it or against it. Moreover, other studies explained the ideological standpoint of politically neutral groups of users.

### 3 Proposed Methodology



**Fig. 3.** Workflow of the Emotion Analysis Procedure.

Figure 3 shows a sentiment and emotion analysis workflow based on social media. It starts with the gathering of data from social media, in the form of comments, pictures and videos. This raw and unstructured data is then subjected to a data pre-processing phase where cleaning, noise reduction and filtering are done to enhance data quality. The data then undergoes processing to refine it and feed it into an emotion recognition model, which uses the chosen algorithms to detect underlying emotional trends. This is followed by feature selection and feature extraction to obtain the most informative features in the data. These features are further processed through three parallel methods, namely machine learning methods, lexicon-based methods, as well as hybrid methods which integrate the two. The results of these methods are fed into a classification phase whereby the emotions or feelings are classified into pre-determined groups. Finally, the system produces polarity information, indicating the

overall sentiment (such as positive, negative, or neutral), enabling effective opinion mining and sentiment-based decision support.

### **3.1 Data Collection**

Data gathering is the first and most important step in sentiment analysis. Online forums, data archives, chat logs, and social media sites like Facebook and Twitter are just a few of the places where data can be found [13]. Numerous research efforts have cited a condensed list of 20 commonly accepted criteria. Specific keywords like @ and # are employed to retrieve pertinent content while sourcing data. Two methods are commonly used for data collection: the first is scripting to create data from the ground up, and the second is acquiring data from third-party software.

### **3.2 Data Initialization**

This stage ensures the accuracy of the obtained data by preparing it for cleaning and filtering. The main data sources are social media platforms, thus problems like misspelled words, slang, abbreviations, and extra spaces must be fixed. Error correction, data normalization to lowercase, and textual data stemming are all steps in the preprocessing process. To get the data ready for classifier-based sentiment analysis, unwanted material is removed. The preprocessed data is kept for additional examination [14].

### **3.3 Identification of Sentiment**

In order to ascertain sentiments in the data, this phase entails recognizing opinionated terms. Texts are divided into two categories: subjective and objective. Subjective texts are more likely to contain sentiment-related terms and viewpoints. Texts that are objective and devoid of emotion are rejected. “Last month we bought a car, and the windshield wiper fell off,” for example, does not convey any opinion about the car's quality. As shown in Figure 3, machine learning-based methods are used to identify sentiments [15].

### **3.4 Selection and Extraction of Features**

In sentiment analysis, feature extraction is essential. Model performance may suffer if crucial aspects are overlooked or omitted. Semantic and syntactic characteristics are the two categories into which features fall [33]. Part of Speech (POS) tagging and Term Frequency-Inverse Document Frequency (TF-IDF) are common syntactic features. These features use methods such as word vectors and corpus-based global vector representations to indicate feelings, such as positivity or negativity, concerning opinions, concepts, or negations [16].

### **3.5 Classification of Emotions**

Emotion categorization offers insights into user attitudes and establishes the sentiment polarity (positive or negative) of massive datasets. Several research efforts employ deep learning approaches, clustering algorithms and supervised machine learning techniques. Aspects of both explicit and implicit learning phases are highlighted by sentiment analysis utilizing linguistic inquiry techniques. CNN, RNN, and ReCNN are examples of neural network models that have shown efficacy. Modern techniques are also compared in studies. However, the capacity of hybrid techniques to identify natural sentiments may be limited when dealing with complicated sentences or noisy data [17].

### **3.6 Report on Polarity**

Public sentiment and results regarding particular entities are provided by polarity reporting. Studies on sentiment analysis focus on classifying deep feelings and mining emotions, providing insights in a variety of fields.

## **4 Results and Discussion**

In order to forecast election results, researchers have examined a variety of trends, offering important insights into the area. The important findings from the literature are highlighted in a special section that provides a thorough analysis of the major discoveries. The Application Prediction System (APS), which makes use of social media data obtained by crawling techniques, is depicted in Figure 4. A variety of techniques were used in the investigation, such as social network-based studies, volumetric methodologies, and sentiment analysis. Either individually or together, these methods have demonstrated potential to enhance the accuracy of election forecasts, and it is plausible to expect that the application of multiple analytical frameworks together can be useful in understanding the behavior and preferences of voters.

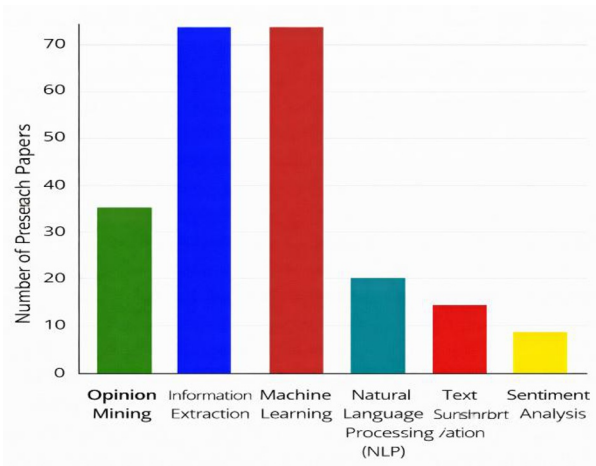


Fig. 4. Visualization of Applicable Techniques.

Figure 4 shows how the research papers are distributed across various domains of analysis pertaining to data mining and sentiment analysis. It indicates that research interest and application of Machine Learning and Information Extraction is high, as it takes over the research landscape. The contribution of Opinion Mining and Natural Language Processing (NLP) is medium, whereas the research publications on Text Summarization and Sentiment Analysis are lower, which points out possible future developments.

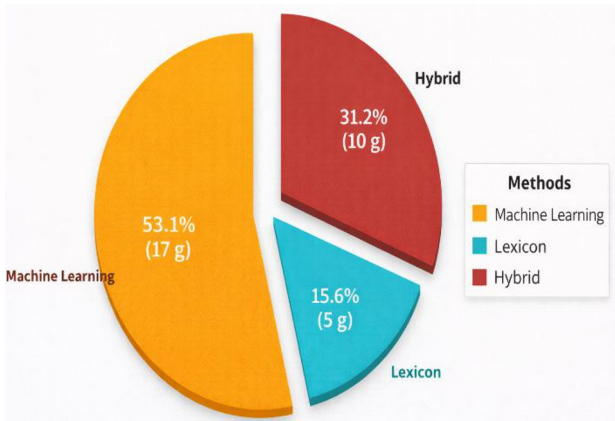
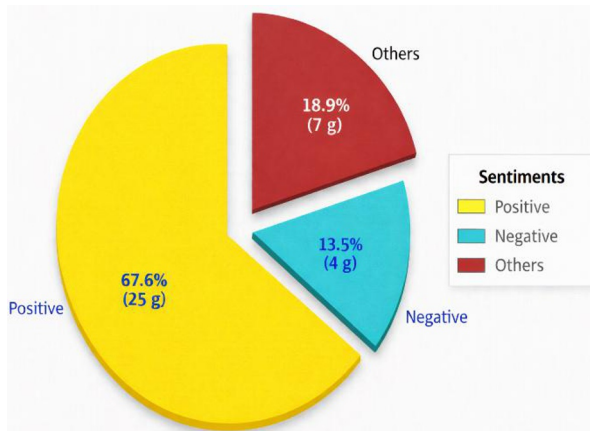


Fig. 5. Prediction identification model.

Figure 5 indicates the distribution of approaches used in sentiment analysis studies. It demonstrates that machine learning methods prevail with the largest percentage, meaning that they are widely used because of their better accuracy and flexibility. Hybrid approaches, with a fusion of machine learning and lexicon-based approaches, comprise a high percentage, with an emphasis on enhancing robustness. The methods

based on lexicon have a smaller share, as they demonstrate lower effectiveness in isolation.



**Fig. 6.** Emotion identification Analysis.

Figure 6 shows the sentiment analysis of the processed data in a pie chart. It divides the total sentiments into Positive, Negative, and Others, which clearly indicates the proportion between them. These outcomes show that there is a predominance of positive sentiment in the dataset, covering 67.6% (25 instances) of the dataset, which implies an overall positive or supportive attitude of the population. The proportion of negative sentiment is smaller at 13.5% (4 instances), which is associated with a relatively low level of dissatisfaction or criticism. The category of Others, which consists of 18.9% (7 instances), consists of neutral, mixed, or unclassified sentiments that do not belong to either the positive or negative categories. On the whole, the diagram is effective in visualizing the sentiment polarity distribution, as it provides a quick understanding of the trends in public opinion and helps evaluate the validity of sentiment classification results for researchers and decision-makers.

## 5 Conclusion and Future Scope

Prediction of elections is a topic of discussion in numerous different areas, and approaches have been established in large numbers to accomplish the same. The new social media that has emerged in recent years has totally altered the mode of forecasting and appraisal-making. Research indicates the importance of sentiment analysis in interpreting the opinion of the group and how it can be employed to estimate the results of polls. This paper examined the methods and approaches that utilize large volumes of information from different social media platforms. The study also indicates that further research is required in order to make predictions more precise. The separation of positive and negative attitudes becomes visible in the new method of communities, providing a view of how the population feels. These

tendencies are also covered in the discussion section, which also suggests the significance of volumetric and sentiment analysis techniques. The existing solutions, including deep learning methods and standard learning models that have been demonstrated to have enormous potential, are addressed in the conclusion of the paper. Nonetheless, in order to increase the level of prediction and reliability further, an unexplored area in this field has to be covered. These gaps need to be filled in order to make future election predictions more reliable and accurate.

**Acknowledgments.** *Not Available*

**Disclosure of Interests.** *Not Available*

## References

1. Paulraj, D., P. Ezhumalai, and Mohan Prakash. "A Deep Learning Modified Neural Network (DLMNN) based proficient sentiment analysis technique on Twitter data." *Journal of Experimental & Theoretical Artificial Intelligence* 36, no. 3 (2024): 415-434.
2. Vatambeti, Ramesh, Srihari Varma Mantena, K. V. D. Kiran, M. Manohar, and Chinthakunta Manjunath. "Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique." *Cluster Computing* 27, no. 1 (2024): 655-671.
3. Modi, Astha, Khelan Shah, Shrey Shah, Samir Patel, and Manan Shah. "Sentiment analysis of Twitter feeds using flask environment: A superior application of data analysis." *Annals of Data Science* 11, no. 1 (2024): 159-180.
4. Ismail, Azlan, Faris Haziq Sazali, Siti Nuraishah Agos Jawaddi, and Sofianita Mutalib. "Stream ETL framework for twitter-based sentiment analysis: Leveraging big data technologies." *Expert Systems with Applications* 261 (2025): 125523.
5. Cam, Handan, Alper Veli Cam, Ugur Demirel, and Sana Ahmed. "Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers." *Heliyon* 10, no. 1 (2024).
6. Mahalakshmi, V., P. Shenbagavalli, S. Raguvaran, V. Rajakumareswaran, and E. Sivaraman. "Twitter sentiment analysis using conditional generative adversarial network." *International Journal of Cognitive Computing in Engineering* 5 (2024): 161-169.
7. Memiş, Erkut, Hilal Akarkamçı, Mustafa Yeniad, Javad Rahebi, and Jose Manuel Lopez-Guede. "Comparative study for sentiment analysis of financial tweets with deep learning methods." *Applied Sciences* 14, no. 2 (2024): 588.
8. Savitri, Putu Rheyana Ananda, I. Made Agus Dwi Suarjaya, and Wayan Oger Vihikan. "Sentiment Analysis of X (Twitter) Comments on The Influence of South Korean Culture in Indonesia." *Journal of Information Systems and Informatics* 6, no. 2 (2024): 979-991.
9. Alam, Md Shah, Md Sabbir Hossain Mrida, and Md Atikur Rahman. "Sentiment analysis in social media: How data science impacts public opinion knowledge integrates natural language processing (NLP) with artificial intelligence (AI)." *American Journal of Scholarly Research and Innovation* 4, no. 01 (2025): 63-100.
10. Mittal, Rohini, Mukesh Rawat, and Vijay Kumar Sharma. "Detection of communal posts with context-aware multilingual sentiment analysis using deep learning." In *Recent Trends in Intelligent Computing and Communication*, pp. 299-303. CRC Press, 2025.

11. Xu, Han, Minghua Xu, Xianjun Deng, and Bang Wang. "Sentiment Diffusion in Online Social Networks: A Survey from the Computational Perspective." *ACM Computing Surveys* (2025).
12. Huang, Yinghui, Mei Li, Fugee Tsung, and Xiangyu Chang. "Mining social media data via supervised topic model: Can social media posts inform customer satisfaction?" *Decision Sciences* (2025).
13. Mohammed Alsekait, Deema, Hana Fathi, Shimaa Abdallah Ibrahim, Ahmed Younes Shdefat, Ahmed Saleh Alattas, and Daa Salama AbdElminaam. "Sentiment analysis: A machine learning utilisation for analyzing the sentiments of facebook and twitter posts." *Intelligent Data Analysis* (2025): 1088467X241301389.
14. Gandy, Lisa M., Lana V. Ivanitskaya, Leeza L. Bacon, and Rodina Bizri-Baryak. "Public health discussions on social media: evaluating automated sentiment analysis methods." *JMIR Formative Research* 9, no. 1 (2025): e57395.
15. Ali, Hafiz Muhammad Usman, Qaisar Farooq, Azhar Imran, and Khalil El Hindi. "A systematic literature review on sentiment analysis techniques, challenges, and future trends." *Knowledge and Information Systems* (2025): 1-68.
16. Sharma, Swati, Vijay Kumar Sharma, Vimal Kumar, and Umang Arora. "Machine Learning Application: Sarcasm Detection Model." In *Artificial Intelligence for a Sustainable Industry 4.0*, pp. 125-138. Cham: Springer International Publishing, 2021.
17. Leandro, Jose Octavian, and Melissa Indah Fianty. "Evaluation of Sentiment Analysis Methods for Social Media Applications: A Comparison of Support Vector Machines and Naïve Bayes." *JOIV: International Journal on Informatics Visualization* 9, no. 2 (2025): 796-807.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

