



# Performance analysis of naive bayes in text classification of islamophobia issues

Faiz Mohammad Ridho <sup>a,1,\*</sup>, Aji Prasetya Wibawa <sup>a,2</sup>, Fachrul Kurniawan <sup>b,3</sup>, Badrudin <sup>b,4</sup>, Anusua Ghosh <sup>c,5</sup>

<sup>a</sup> Universitas Negeri Malang, Malang, Indonesia

<sup>b</sup> State University of Maulana Malik Ibrahim Malang, Malang, Indonesia

<sup>c</sup> University of South Australia, Australia

<sup>1</sup> [faiz.mohammad.1705356@students.um.ac.id](mailto:faiz.mohammad.1705356@students.um.ac.id); <sup>2</sup> [aji.prasetya.ft@um.ac.id](mailto:aji.prasetya.ft@um.ac.id); <sup>3</sup> [fachrulk@ti.uin-malang.ac.id](mailto:fachrulk@ti.uin-malang.ac.id);

<sup>4</sup> [badrudin@hk.uin-malang.ac.id](mailto:badrudin@hk.uin-malang.ac.id); <sup>5</sup> [anusua.ghosh@mymail.unisa.edu.au](mailto:anusua.ghosh@mymail.unisa.edu.au)

\* Corresponding Author

## ARTICLE INFO

## ABSTRACT

### Article history

Received 02 April, 2022

Revised 14 April, 2022

Accepted 03 May, 2022

### Keywords

Naive bayes

Text classification

Islamophobia

In the aftermath of the 2013 Woolwich attack, a disturbing surge in hate crimes against the Muslim community emerged both offline and on social media platforms, prompting concerns about the widespread issue of Islamophobia. To systematically evaluate and quantify the presence of Islamophobic sentiment in online spaces, this study employed sentiment analysis, a robust method for deriving insights from textual data. Two classification models, Bernoulli Naive Bayes and Multinomial Naive Bayes, were selected to conduct a thorough analysis. Bernoulli Naive Bayes, specialized in handling binary data, was used for binary sentiment analysis, while Multinomial Naive Bayes, well-suited for data with multiple occurrences, was applied for more comprehensive analysis. The research encompassed nine meticulously designed test-train data scenarios, ranging from a 10:90 test-train data ratio to a 20:80 ratio. Surprisingly, both models exhibited a maximum accuracy rate of 68% in their respective optimal scenarios, raising intriguing questions about the potential and limitations of sentiment analysis and Naive Bayes models in the complex task of identifying and quantifying Islamophobic content on social media.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## 1. Introduction

In May 2013, the horrific murder of British soldier Lee Rigby in Woolwich, London [1] perpetrated by Michael Adebawale and Michael Adebolajo, sent shockwaves through the United Kingdom. In the wake of this event, a disturbing trend emerged as the public outrage gave rise to a series of hate crimes targeting the Muslim community [2]. These incidents bore witness to the ominous growth of Islamophobia, a deeply entrenched issue that transcends national borders [3]. Even in Indonesia, a nation where the majority of the population adheres to Islam, concerns about Islamophobia persist [4].

The negative perception of Islam has been exacerbated by acts of terrorism carried out by groups claiming to represent the religion, often overshadowing the peaceful majority [5]. While physical attacks against Muslims have been widely reported, the digital realm has also provided a breeding ground for

anti-Muslim sentiment [6]. Individuals have turned to social media as a platform to vent their anger, further propagating Islamophobia [7]. This digital manifestation of prejudice poses a significant challenge, as it can undermine social harmony and the well-being of the Muslim community [8].

To confront the complex issue of online Islamophobia, researchers have turned to sentiment analysis, a powerful technique for automatically extracting sentiment and emotional information from textual data [9]–[14]. Sentiment analysis, or opinion mining, plays a crucial role in natural language processing, computational linguistics, and text mining, enabling the analysis of opinions, attitudes, judgments, and emotions expressed by individuals [15], [16]. By processing and understanding textual data, sentiment analysis helps in identifying and quantifying sentiment on a given topic or issue.

Amidst the multitude of text classification methods available, this study hones in on two Naive Bayes methods: Bernoulli Naive Bayes and Multinomial Naive Bayes [17], [18]. The former excels in binary classification, a vital aspect of sentiment analysis, where texts are categorized as positive or negative [19]. In contrast, Multinomial Naive Bayes considers the frequency of word occurrences, making it an ideal choice for capturing more nuanced sentiment nuances [20].

These methods have been selected for their computational efficiency and historical success in classifying text data with a high level of accuracy. However, their application to the challenging task of classifying sentiment regarding Islamophobia in the realm of Twitter [21] is a subject of keen investigation in this research.

This study seeks to shed light on the efficacy of using the Bernoulli Naive Bayes and Multinomial Naive Bayes methods in classifying sentiment within Twitter discussions surrounding Islamophobia. By delving into the depths of online sentiment analysis, this research contributes to our understanding of the practical tools and methodologies that can be harnessed to tackle the pressing issue of Islamophobia within the digital sphere, a pertinent challenge in today's interconnected world.

## 2. Method

### 2.1. Data Collection

The process of data collection in this research is pivotal as it lays the foundation for subsequent analysis. The dataset under scrutiny comprises Twitter users' opinions concerning Islamophobia during the year 2020, encompassing a total of 1,574 distinct opinions.

Twitter, as a microblogging platform, serves as an abundant source of user-generated content that is a rich ground for understanding contemporary public sentiment. In the context of studying a sensitive and socially relevant issue like Islamophobia, Twitter provides an open forum where users freely express their opinions and emotions. The choice of collecting data specifically for the year 2020 is noteworthy, as it allows for an examination of opinions and sentiments within a defined temporal context, potentially revealing insights into the impact of specific events or developments on public opinion.

The process of data collection in a social media context has its challenges and nuances. The dataset represents a snapshot of discourse on Twitter, but it's essential to acknowledge the inherent biases associated with this platform. Twitter users are a self-selected group, often with distinct demographics and attitudes. Thus, the dataset inherently captures the perspectives of those active on the platform, potentially skewing the sample.

Additionally, the volume of data, in this case, 1,574 opinions, presents opportunities and challenges. It provides a substantial corpus for analysis, offering the potential for meaningful insights into public sentiment. However, the management and analysis of such a dataset necessitates robust methods and tools for efficient and effective processing.

## 2.2. Data Collection and Preparation

The process of data labeling is a crucial step in this research, and it involves assigning each tweet a sentiment label, classifying them as expressing either a positive or a negative opinion regarding Islamophobia. This process was carried out manually and was further validated by an expert, Dr. H. Badruddin, MHI.

**Manual Labeling:** Manual labeling is essential when working with sentiment analysis and classification tasks. It requires human annotators to review and categorize individual pieces of text based on their sentiment. In this case, the sentiment being assessed is related to Islamophobia. The manual aspect of labeling ensures that the process is attentive to context, nuanced expressions, and the subtleties of language that automated sentiment analysis tools may overlook. However, manual labeling is resource-intensive and time-consuming, often requiring multiple human annotators to achieve inter-rater reliability.

**Validation by an Expert:** The additional step of validating the labeling by an expert adds a layer of credibility to the sentiment labels. Dr. H. Badruddin's validation likely involved a thorough review of a sample of the labeled data, checking for consistency and accuracy in sentiment assignments. This validation process addresses the potential for subjectivity or disagreement in sentiment classification that can arise during manual labeling.

**Balancing Sentiment Labels:** The result of this labeling process indicates that the dataset includes 932 tweets expressing a positive sentiment regarding Islamophobia and 642 tweets conveying a negative sentiment. The distribution of sentiment labels is an essential consideration in sentiment analysis research. A balanced dataset, where both positive and negative sentiments are well-represented, is valuable for the development of robust classification models.

**Limitations of Manual Labeling:** While manual labeling, when done diligently, ensures high-quality sentiment labels, it has limitations. Human annotators may bring their biases and interpretations into the process, which can affect the accuracy of labeling. It's essential to provide guidelines and training to annotators to minimize potential bias. Additionally, the size of the dataset and the potential for inter-rater reliability among annotators are factors that must be carefully managed.

## 2.3. Text Pre-Processing

Text pre-processing serves as the foundational stage in the research methodology, encompassing a series of critical techniques aimed at refining the collected and labeled dataset for subsequent sentiment analysis [22]. Each pre-processing step carries significant importance and contributes to the dataset's readiness for classification [23].

**Case Folding:** The process of case folding, an initial pre-processing step, is essential for standardizing letter casing within the dataset [24]. By converting all letters to lowercase, the pre-processing step ensures that variations in letter casing, such as capital and lowercase letters, do not lead to the duplication of words with identical meaning. This standardization not only reduces redundancy but also ensures a consistent representation of words, paving the way for more accurate analysis.

**Punctuation Removal:** Following case folding, punctuation removal emerges as a pivotal pre-processing technique [24]. It involves the systematic elimination of punctuation marks, numbers, and URLs from the text. This step is of paramount importance in maintaining focus on textual content. Removing punctuation marks ensures they do not contribute to the classification process as separate tokens. Furthermore, the elimination of URLs aligns with the goal of maintaining the textual integrity of the dataset.

**Tokenizing:** The tokenization process, a fundamental pre-processing step, involves breaking down sentences into individual words or tokens, typically separated by spaces [25]. Tokenization serves as the cornerstone for converting continuous text into a structured format suitable for analysis. It segments the text into meaningful units, which, in turn, become the basis for subsequent analysis, allowing for the precise measurement of sentiment.

**Stopwords Removal:** Stopwords, commonly occurring words like prepositions (e.g., "at," "to," "from") and conjunctions (e.g., "and," "as," "well"), are prevalent in sentences but often carry limited sentiment or meaning [26]. Therefore, the removal of stopwords is instrumental in enhancing the performance of the classification process. By reducing noise in the data and eliminating frequently occurring, yet semantically uninformative words, this step sharpens the dataset for sentiment analysis.

**Stemming:** The inclusion of stemming in the pre-processing methodology reflects a robust approach to Indonesian text analysis [27]. Stemming, executed using the Sastrawi Python library, plays a pivotal role in data normalization. It reduces words to their root form or stem, aiding in the grouping of words with similar meanings. This process not only facilitates semantic consistency but also ensures that the dataset is finely tuned for the subsequent sentiment analysis, ultimately enhancing the quality and accuracy of the analysis.

In essence, text pre-processing is the cornerstone of effective sentiment analysis. These techniques collectively cleanse and standardize the text data, rendering it fit for classification. It is important to note that text pre-processing is not a one-size-fits-all endeavor and should be tailored to the unique characteristics of the dataset and the language of analysis. In this case, the pre-processing steps are well-adapted for Indonesian text, encompassing case folding, punctuation removal, tokenization, stopwords removal, and stemming. These steps collectively aim to reduce noise, enhance data quality, and elevate the performance of classification models applied in subsequent research stages.

#### 2.4. Text Classification

The text classification phase is a pivotal component of the research methodology, responsible for categorizing the labeled and pre-processed data into sentiments, specifically positive or negative, regarding Islamophobia. This phase encompasses a range of techniques and considerations.

- **Naive Bayes Classification:** The classification process begins with the application of the Naive Bayes algorithm, specifically the Naive Bayes-Bernoulli algorithm and the Naive Bayes Multinomial algorithm [28]. These algorithms are well-suited for text classification tasks, offering distinct approaches. The Naive Bayes-Bernoulli algorithm is designed to handle binary data (0 and 1), which is apt for sentiment classification [29]. In contrast, the Naive Bayes Multinomial algorithm considers the frequency of word occurrences within a class, making it well-suited for tasks where word frequency matters [30]. The selection of these algorithms reflects a thoughtful approach to text classification, leveraging the strengths of Naive Bayes in handling textual data.

- **Data Test and Data Train Comparison:** The classification process is further enriched by exploring various scenarios of data test and data train comparisons. These scenarios encompass different ratios of test data and training data, ranging from 10% test and 90% training to 90% test and 10% training. This systematic variation serves to comprehensively assess the performance of the classification process under different conditions. Each scenario provides a unique perspective on how the model behaves with varying amounts of training and test data. This approach not only ensures robustness but also allows for fine-tuning the model based on the specific needs of the analysis.

**Performance Metrics:** The effectiveness of the classification process is assessed through the utilization of performance metrics, which are instrumental in quantifying the system's accuracy, precision, recall, and overall efficiency [31], [32]. These metrics provide a rigorous evaluation of the sentiment classification models.

- **Accuracy** is a fundamental metric, determining the percentage of data accurately classified by the system (1). It represents the sum of true positives (TP) and true negatives (TN) divided by the sum of TP, TN, false positives (FP), and false negatives (FN). A high accuracy percentage signifies that the model effectively identifies sentiments.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

- **Recall** measures the proportion of correctly predicted positive data compared to the overall positive data value on the actual label. This metric is valuable for understanding the model's ability to capture true positives (2).

$$Recall = \frac{(TP)}{(TP+FN)} \quad (2)$$

- **Precision** evaluates the average precision and recall, offering a holistic assessment of the classification system's effectiveness. It quantifies the proportion of true positives among the total predicted positives (3).

$$Precision = \frac{(TP)}{(TP+FP)} \quad (3)$$

- **F1 Score** presents a balanced measurement of precision and recall. It is calculated as the harmonic mean of precision and recall, offering a comprehensive evaluation of the model's performance (4).

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

### 3. Results and Discussion

This section offers a comprehensive analysis of the classification process, focusing on two Naive Bayes algorithms: Bernoulli Naive Bayes (BNB) and Multinomial Naive Bayes (MNB). The examination encompasses various scenarios of data test and data train comparisons, each shedding light on the performance and behavior of these algorithms. The critical performance metrics evaluated include accuracy (A), precision (P), recall (R), and the F1-score (f1).

- **Variability in Bernoulli Naive Bayes (BNB):** Bernoulli Naive Bayes, tailored for binary data (0 and 1), displays a degree of variability in its accuracy across different data test and data train scenarios.

This variability can be attributed to the binary nature of the algorithm, which may lead to fluctuations in performance. BNB may encounter challenges in accommodating the nuanced sentiment expressions, especially in scenarios with more complex sentiments. The highest accuracy for BNB, observed at 68%, occurs in the 20:80 data test and data train comparison. This scenario also boasts a precision of 71%, recall of 78%, and an F1-score of 74%. While this represents a noteworthy performance, the variability in accuracy highlights the limitations of BNB in capturing nuanced sentiment.

- **Stability and Superiority of Multinomial Naive Bayes (MNB):** Conversely, Multinomial Naive Bayes exhibits a more stable accuracy as the data train size increases. This stability underscores the value of ample training data in achieving consistent results. The highest accuracy for MNB, at 68%, is observed in the 10:90 data test and data train comparison, coupled with an F1-score of 75%. These results suggest that MNB is particularly effective when a substantial portion of the dataset is allocated to training. The performance metrics for MNB consistently reveal accuracy values ranging from 60% to 68%, underscoring its robustness in analyzing sentiments in Twitter users' opinions regarding Islamophobia. For a detailed view of the classification results, please refer to [Table 1](#). [Table 1](#) provides a comprehensive breakdown of the scenarios, algorithm names, data splits, and key performance metrics, offering valuable insights into the performance of BNB and MNB in different test and train scenarios.

**Table 1.** Classification Results for Bernoulli Naive Bayes (BNB) and Multinomial Naive Bayes (MNB) across Various Data Test and Data Train Scenarios. The table provides insights into the performance metrics, including accuracy (A), precision (P), recall (R), and the F1-score (f1), highlighting the capabilities of BNB and MNB in sentiment analysis on the topic of Islamophobia

No	NB	S	A	P	R	f1
1	BNB	0.1	66	69	80	74
2	BNB	0.2	68	71	78	74
3	BNB	0.3	67	67	80	73
4	BNB	0.4	67	68	80	74
5	BNB	0.5	67	67	84	74
6	BNB	0.6	66	65	88	75
7	BNB	0.7	63	62	91	74
8	BNB	0.8	62	61	95	74
9	BNB	0.9	61	61	97	75
10	MNB	0.1	68	72	78	75
11	MNB	0, 2	68	74	71	72
12	MNB	0.3	67	71	70	71
13	MNB	0.4	67	72	68	70
14	MNB	0.5	67	73	68	70
15	MNB	0.6	65	70	68	69
16	MNB	0.7	64	69	68	69
17	MNB	0.8	61	69	63	66
18	MNB	0.9	60	69	59	63

- **Suitability of Naive Bayes Algorithms for Sentiment Analysis:** The findings underscore the suitability of Naive Bayes algorithms for sentiment analysis, even in the context of complex sentiments expressed on Twitter. Both BNB and MNB consistently deliver accuracy values in the range of 60% to 68%, demonstrating their efficacy in analyzing sentiments. However, the

preference between the two algorithms depends on the volume and distribution of training data. MNB's superior stability and accuracy in scenarios with larger training datasets make it a preferable choice in such contexts. The importance of data volume and distribution is evident in the consistency and accuracy of sentiment classification. This observation emphasizes the significance of extensive training data in ensuring the robustness and reliability of sentiment analysis models.

- **Insights for Understanding Public Sentiment on Islamophobia:** The insights gleaned from these results hold substantial significance for understanding public sentiment on the issue of Islamophobia, particularly on Twitter. The choice of this social media platform as the data source aligns with its role as a forum for free expression. The variability observed in BNB highlights the challenge of capturing nuanced sentiments, while MNB's stability underscores the importance of data volume in sentiment analysis. These findings open avenues for further research into the dynamics of sentiment expression on social media and the potential impact of events or discussions on public opinion.

#### 4. Conclusion

In conclusion, this study delves into sentiment analysis concerning Islamophobia, specifically within the realm of Twitter users' opinions. Two Naive Bayes algorithms, Bernoulli Naive Bayes (BNB) and Multinomial Naive Bayes (MNB), were employed across various data test and data train scenarios. BNB demonstrated commendable performance, achieving its highest accuracy at 68% in the 20:80 data test and data train comparison, accompanied by a precision of 71%, a recall of 78%, and an F1-score of 74%. BNB's prowess in handling binary sentiment expressions is evident, although it may encounter challenges when faced with more nuanced sentiments. In contrast, MNB excelled in scenarios featuring larger training datasets, attaining its highest accuracy of 68% in the 10:90 data test and data train comparison, coupled with a precision of 72%, a recall of 78%, and an impressive F1-score of 75%. MNB's strength lies in its ability to consistently maintain accuracy, making it a suitable choice for sentiment analysis, particularly in cases involving extensive training data.

Both BNB and MNB consistently delivered accuracy values within the 60% to 68% range, underscoring their effectiveness in analyzing Twitter users' sentiments concerning Islamophobia. The equivalent highest accuracy values of 68% highlight their proficiency in this context. This study lays the foundation for future research, suggesting that improved accuracy levels can be attained by exploring pre-processing variations and expanding the dataset range. These prospective enhancements will contribute to a deeper understanding of public sentiment on complex and sensitive subjects, such as Islamophobia, and facilitate the development of more robust sentiment analysis methodologies.

#### References

- [1] I. Awan, "Islamophobia and Twitter: A Typology of Online Hate Against Muslims on Social Media," *Policy & Internet*, vol. 6, no. 2, pp. 133–150, Jun. 2014, doi: <https://doi.org/10.1002/1944-2866.POI364>.
- [2] D. M. Jones and M. L. R. Smith, "The Age of Ambiguity: Art and the War on Terror Twenty Years after 9/11," *Stud. Confl. Terror.*, pp. 1–20, Jul. 2021, doi: [10.1080/1057610X.2021.1943813](https://doi.org/10.1080/1057610X.2021.1943813).
- [3] S. L. Perry, "American Religion in the Era of Increasing Polarization," *Annu. Rev. Sociol.*, vol. 48, no. 1, pp. 87–107, Jul. 2022, doi: [10.1146/annurev-soc-031021-114239](https://doi.org/10.1146/annurev-soc-031021-114239).
- [4] P. N. Jain and A. S. Vaidya, "Analysis of Social Media Based on Terrorism — A Review," *Vietnam J. Comput. Sci.*, vol. 08, no. 01, pp. 1–21, Feb. 2021, doi: [10.1142/S2196888821300015](https://doi.org/10.1142/S2196888821300015).

- [5] M. Ozalp and M. Ćufurović, "Religion, Belonging, and Active Citizenship: A Systematic Review of Literature on Muslim Youth in Australia," *Religions*, vol. 12, no. 4, p. 237, Mar. 2021, doi: [10.3390/rel12040237](https://doi.org/10.3390/rel12040237).
- [6] M. Pucelj, "Manifestations of Islamophobia During COVID-19.," *Challenges Futur.*, vol. 7, no. 3, pp. 139-163, 2022. [Online]. Available at: <https://web.p.ebscohost.com/abstract?direct>.
- [7] V. Vidotto, "Italy and Turkey in the Mediterranean: challenges and opportunities for cooperation," p. 123, 2022. [Online]. Available at: <http://dspace.unive.it/handle/10579/21165>.
- [8] A. Emon and N. Hasan, "Under layered suspicion: A review of CRA audits of Muslim-led charities," pp. 1-194, 2021. [Online]. Available at: <https://tspace.library.utoronto.ca/handle/1807/126225>.
- [9] H. A. Al-Shaibani and S. Al-Augby, "Terrorist Tweets Detection using Sentiment Analysis: Techniques and Approaches," in *2022 5th International Conference on Engineering Technology and its Applications (IICETA)*, May 2022, pp. 585-590, doi: [10.1109/IICETA54559.2022.9888461](https://doi.org/10.1109/IICETA54559.2022.9888461).
- [10] N. S. Mullah and W. M. N. W. Zainon, "Advances in Machine Learning Algorithms for Hate Speech Detection in Social Media: A Review," *IEEE Access*, vol. 9, pp. 88364-88376, 2021, doi: [10.1109/ACCESS.2021.3089515](https://doi.org/10.1109/ACCESS.2021.3089515).
- [11] M. Fernandez and H. Alani, "Artificial intelligence and online extremism," in *Predictive Policing and Artificial Intelligence*, Routledge, 2021, pp. 132-162, doi: [10.4324/9780429265365-7](https://doi.org/10.4324/9780429265365-7).
- [12] Y. W. Tama and S. D. Sulistyningrum, "A Systematic Literature Review of Islamophobia on Media: Trends, Factors, and Stereotypes," *Indones. J. Relig. Soc.*, vol. 5, no. 1, pp. 14-23, Jun. 2023, doi: [10.36256/ijrs.v5i1.288](https://doi.org/10.36256/ijrs.v5i1.288).
- [13] C. Bird, E. Ungless, and A. Kasirzadeh, "Typology of Risks of Generative Text-to-Image Models," in *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, Aug. 2023, pp. 396-410, doi: [10.1145/3600211.3604722](https://doi.org/10.1145/3600211.3604722).
- [14] T. Garg, S. Masud, T. Suresh, and T. Chakraborty, "Handling Bias in Toxic Speech Detection: A Survey," *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 1-32, Dec. 2023, doi: [10.1145/3580494](https://doi.org/10.1145/3580494).
- [15] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Syst.*, vol. 226, p. 107134, Aug. 2021, doi: [10.1016/j.knosys.2021.107134](https://doi.org/10.1016/j.knosys.2021.107134).
- [16] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731-5780, Oct. 2022, doi: [10.1007/s10462-022-10144-1](https://doi.org/10.1007/s10462-022-10144-1).
- [17] A. Miller, J. Panneerselvam, and L. Liu, "A review of regression and classification techniques for analysis of common and rare variants and gene-environmental factors," *Neurocomputing*, vol. 489, pp. 466-485, Jun. 2022, doi: [10.1016/j.neucom.2021.08.150](https://doi.org/10.1016/j.neucom.2021.08.150).
- [18] L. Saha, H. K. Tripathy, S. R. Nayak, A. K. Bhoi, and P. Barsocchi, "Amalgamation of Customer Relationship Management and Data Analytics in Different Business Sectors—A Systematic Literature Review," *Sustainability*, vol. 13, no. 9, p. 5279, May 2021, doi: [10.3390/su13095279](https://doi.org/10.3390/su13095279).
- [19] C. Zucco, B. Calabrese, G. Agapito, P. H. Guzzi, and M. Cannataro, "Sentiment analysis for mining texts and social networks data: Methods and tools," *WIREs Data Min. Knowl. Discov.*, vol. 10, no. 1, Jan. 2020, doi: [10.1002/widm.1333](https://doi.org/10.1002/widm.1333).
- [20] F. Arias, M. Zambrano Nunez, A. Guerra-Adames, N. Tejedor-Flores, and M. Vargas-Lombardo, "Sentiment Analysis of Public Social Media as a Tool for Health-Related Topics," *IEEE Access*, vol. 10, pp. 74850-74872, 2022, doi: [10.1109/ACCESS.2022.3187406](https://doi.org/10.1109/ACCESS.2022.3187406).
- [21] S. A. Castaño-Pulgarín, N. Suárez-Betancur, L. M. T. Vega, and H. M. H. López, "Internet, social media and online hate speech. Systematic review," *Aggress. Violent Behav.*, vol. 58, p. 101608, May 2021, doi: [10.1016/j.avb.2021.101608](https://doi.org/10.1016/j.avb.2021.101608).



- 
- [22] K. Machova, M. Mach, and M. Vasilko, "Comparison of Machine Learning and Sentiment Analysis in Detection of Suspicious Online Reviewers on Different Type of Data," *Sensors*, vol. 22, no. 1, p. 155, Dec. 2021, doi: [10.3390/s22010155](https://doi.org/10.3390/s22010155).
- [23] E. Nugraheni, "Indonesian Twitter Data Pre-processing for the Emotion Recognition," in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Dec. 2019, pp. 58–63, doi: [10.1109/ISRITI48646.2019.9034653](https://doi.org/10.1109/ISRITI48646.2019.9034653).
- [24] K. K. Agustiningsih, E. Utami, and H. Al Fatta, "Sentiment Analysis of COVID-19 Vaccine on Twitter Social Media: Systematic Literature Review," in *2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Nov. 2021, pp. 121–126, doi: [10.1109/ICITISEE53823.2021.9655960](https://doi.org/10.1109/ICITISEE53823.2021.9655960).
- [25] G. Yenduri, B. R. Rajakumar, K. Praghash, and D. Binu, "Heuristic-Assisted BERT for Twitter Sentiment Analysis," *Int. J. Comput. Intell. Appl.*, vol. 20, no. 03, Sep. 2021, doi: [10.1142/S1469026821500152](https://doi.org/10.1142/S1469026821500152).
- [26] S. Bagui, C. Wilber, and K. Ren, "Analysis of Political Sentiment From Twitter Data," *Nat. Lang. Process. Res.*, vol. 1, no. 1–2, p. 23, 2020, doi: [10.2991/nlpr.d.201013.001](https://doi.org/10.2991/nlpr.d.201013.001).
- [27] S. K. Trivedi and A. Singh, "Twitter sentiment analysis of app based online food delivery companies," *Glob. Knowledge, Mem. Commun.*, vol. 70, no. 8/9, pp. 891–910, Nov. 2021, doi: [10.1108/GKMC-04-2020-0056](https://doi.org/10.1108/GKMC-04-2020-0056).
- [28] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, "Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN," *Ilk. J. Ilm.*, vol. 14, no. 3, pp. 348–354, 2022, doi: [10.33096/ilkom.v14i3.1505.348-354](https://doi.org/10.33096/ilkom.v14i3.1505.348-354).
- [29] D. C. Neagu, A. B. Rus, M. Grec, M. A. Boroianu, N. Bogdan, and A. Gal, "Towards Sentiment Analysis for Romanian Twitter Content," *Algorithms*, vol. 15, no. 10, p. 357, Sep. 2022, doi: [10.3390/a15100357](https://doi.org/10.3390/a15100357).
- [30] C. Dewi and R.-C. Chen, "Complement Naive Bayes Classifier for Sentiment Analysis of Internet Movie Database," *Nguyen, N.T., Tran, T.K., Tukayev, U., Hong, TP., Trawiński, B., Szczerbicki, E. Intell. Inf. Database Syst. ACIIDS 2022. Lect. Notes Comput. Sci.*, vol. 13757, pp. 81–93, 2022, doi: [10.1007/978-3-031-21743-2\\_7](https://doi.org/10.1007/978-3-031-21743-2_7).
- [31] B. Gaye, D. Zhang, and A. Wulamu, "A Tweet Sentiment Classification Approach Using a Hybrid Stacked Ensemble Technique," *Information*, vol. 12, no. 9, p. 374, Sep. 2021, doi: [10.3390/info12090374](https://doi.org/10.3390/info12090374).
- [32] Y. Fauziah, S. Saifullah, and A. S. Aribowo, "Design Text Mining for Anxiety Detection using Machine Learning based-on Social Media Data during COVID-19 pandemic," in *Proceeding of LPPM UPN "Veteran" Yogyakarta Conference Series 2020–Engineering and Science Series*, 2020, vol. 1, no. 1, pp. 253–261. [Online]. Available at: <http://proceeding.rsfpres.com/index.php/ess/article/view/117>.
-