Public Sentiment Analysis of the Affan Kurniawan Social Issue: A Comparison of Naïve Bayes and SVM Algorithms

[1*]Marsella Iriana Mamusung, [2]Lorna Yertas Baisa, [3]Andreas Leonardo Sumendap
[1,2,3]Department of Informatics Engineering, Faculty of Engineering, University of Papua
Email: [1]marsellairianamamusung@gmail.com, [2]L.baisa@unipa.ac.id, [3]al.sumendap@unipa.ac.id

Article Info

Article history:
Received Jan 20th, 2026
Revised Feb 31th, 2026
Accepted Mar 24th, 2026

Keyword:
Affan Kurniawan Case
Bag-of-words
Count Vectorizer
Naïve Bayes
Sentiment Analysis
Support Vector Machine

ABSTRACT

Social media X is a dynamic public space where opinions on social issues, including the Affan Kurniawan case, spread rapidly. This study aims to analyze sentiment distribution, compare the performance of Multinomial Naïve Bayes and Linear Support Vector Machine (LinearSVC), and evaluate classification consistency under a unified evaluation framework. Indonesian-language posts were collected using keyword-based crawling and cleaned from 10,624 to 7,431 valid records (28 August–2 September 2025). The data were preprocessed through normalization, tokenization, stopword removal, and stemming, and labeled into negative, neutral, and positive sentiments using a lexicon-based approach. The results show a dominance of negative sentiment (50.26%), followed by neutral (30.96%) and positive (18.77%). Using Bag-of-Words features and an 80:20 train–test split, LinearSVC outperformed Naïve Bayes with higher accuracy (0.826 vs 0.745) and macro F1-score (0.759 vs 0.579). This study highlights the effectiveness of SVM as a stronger baseline model for Indonesian sentiment classification on social media data.

Corresponding Author:
Marsella Iriana Mamusung,
Department of Informatics Engineering,
Faculty of Engineering, University of Papua,
Gunung Salju Amban St., Manokwari, West Papua, Indonesia, Postal Code 98314.
Email: marsellairianamamusung@gmail.com

DOI: http://dx.doi.org/10.24014/ijaidm.v9i1.39258

1. INTRODUCTION

Social media X (formerly Twitter) has become an important digital public sphere where public opinion on social issues is rapidly formed and disseminated through large-scale interactions. Its real-time and open characteristics make it a valuable source for analyzing public sentiment dynamics. However, social media data also contain challenges such as informal language, semantic ambiguity, rapid topic shifts, and the presence of bots, all of which may affect the reliability of sentiment signals [1], [2], [6], [7]. Therefore, robust sentiment analysis methods are needed to better understand public responses in online environments.

Sentiment analysis is a branch of text mining and natural language processing that aims to classify opinions or emotional expressions in textual data into categories such as positive, neutral, and negative [1], [2]. In social media studies, sentiment analysis has been widely applied to examine public reactions to political, social, and cultural issues because these platforms provide large volumes of real-time user-generated content [5], [15]. The quality of sentiment classification is influenced by several factors, especially preprocessing, feature representation, and classifier selection [4], [5]. In this context, Bag-of-Words (BoW) with CountVectorizer remains a commonly used representation because of its simplicity, interpretability, and compatibility with classical machine learning algorithms [10], [20].

Among classical classification methods, Multinomial Naïve Bayes and Support Vector Machine (SVM) are widely used for sentiment analysis. Naïve Bayes is often applied as a probabilistic baseline because it is simple and efficient, while SVM is known for its strong discriminative performance on high-dimensional sparse text data [19], [21], [22]. Although previous studies have compared these methods across various sentiment analysis contexts, many were conducted with different datasets, preprocessing pipelines, feature representations, and evaluation settings, making direct comparisons difficult [12], [19], [21], [22]. In addition, limited studies have specifically examined Indonesian-language public sentiment regarding the Affan Kurniawan case, which attracted broad discussion on social media X and reflected diverse public opinions [9].

Based on this gap, this study analyzes public sentiment regarding the Affan Kurniawan case on social media X by comparing Multinomial Naïve Bayes and linear SVM (LinearSVC) on the same dataset, preprocessing pipeline, Bag-of-Words representation, and evaluation framework. The objectives of this study are to identify sentiment distribution, compare the performance of both classifiers, and evaluate their classification consistency across sentiment classes. This study contributes by providing a fair and systematic comparison of probabilistic and margin-based classifiers, and by offering empirical evidence on public sentiment patterns toward a contemporary social issue on social media X [13], [14], [24].

2. RESEARCH METHOD

This study employs a quantitative design with a computational–experimental approach to analyze public sentiment toward the Affan Kurniawan social issue on social media X. This design was selected because it aligns with the objectives of the study, namely to identify sentiment distribution, compare the performance of Multinomial Naïve Bayes and linear SVM (LinearSVC), and assess classification consistency under a uniform evaluation scheme. A machine-learning-based sentiment analysis approach was chosen because it enables large-scale text data to be processed systematically, objectively, and efficiently, making it suitable for examining trends in public opinion on social media platforms [15]. Previous studies also indicate that sentiment analysis on social media is effective for understanding public responses to social issues and public policies [15].

To ensure methodological clarity, transparency, and replicability, the research procedure was organized into several sequential stages, as illustrated in Figure 1. The workflow begins with data collection and cleaning, followed by text preprocessing, sentiment labeling, feature extraction, model development, and performance evaluation. This sequence was designed to ensure that each stage directly supports the research objectives and that the comparison between classification models is conducted under the same experimental conditions.

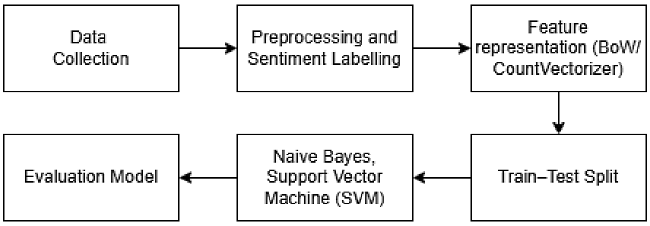

Figure 1.Research Flow

The first stage was data collection and storage. Public posts on social media X related to the Affan Kurniawan case were collected using a keyword-based crawling approach. Only Indonesian-language posts were retained for analysis, and the retrieved data were stored in CSV format. This approach was selected because public posts on X provide real-time, large-scale textual data that can reflect public responses to social issues [10], [17]. In addition, transparency in the data collection process is important because recent changes in Twitter/X data-access policies have implications for sampling, reproducibility, and research limitations [17].

The second stage was initial data cleaning. The raw dataset consisted of 10,624 records. Empty entries and duplicate posts with identical text content were removed using dropna and drop_duplicates on the text column. This step was necessary to improve data quality, reduce redundancy, and increase dataset representativeness before further processing. After cleaning, the dataset was reduced to 7,431 valid records. The cleaned dataset was considered more suitable for sentiment analysis because it excluded non-informative and repeated entries that could bias model learning.

The dataset used in this study consists of textual content and supporting metadata such as timestamps, interaction counts, user identifiers, and tweet URLs. It also includes preprocessing output columns and sentiment labels generated during the analysis pipeline. A summary of the dataset characteristics is presented in Table 1.

Table 1. Dataset overview

Component	Description
Data Source	Public posts on social media X related to keywords for the Affan Kurniawan case.
Dataset file name	affan_kurniawan.csv
Initial (raw) records	10,624 entries.

Component	Description
Cleaning process	Removing duplicate and empty text records (drop_duplicates + dropna on the text column).
Records after cleaning	7,431 entries.
Data collection period	28 August 2025–02 September 2025 (based on the crawling time-window parameters in the notebook).
Language	Indonesian (filtered by the language column).
Main fields	tweet_id, created_time, text, like_count.
Key metadata fields	reply_count, retweet_count, quote_count, tweet_url, user_id, username, image_url, location, reply_to_user.
Preprocessing output columns (in the notebook)	cleaning, case_folding, normalisasi, tokenize, stopword removal, stemming_data.
Sentiment labels	positive/ negative/ neutral (automatic lexicon/rule-based labeling).
Label distribution (after labeling)	Positive: 1,395; Negative: 3,735; Neutral: 2,301 (total 7,431).
Training–testing data split	Train–test split — 80:20 → training: 5,944; testing: 1,487.

Table 1 shows that the dataset consists of 7,431 cleaned records collected from 28 August 2025 to 02 September 2025 and labeled into three sentiment classes: negative, neutral, and positive. The data are imbalanced, with negative sentiment as the dominant class, followed by neutral and positive. Because this imbalance may affect classification performance, the evaluation in this study uses not only accuracy but also precision, recall, and F1-score to provide a more balanced assessment across classes [23].

The next stage was text preprocessing, which was carried out to reduce noise and standardize the data before feature extraction. The preprocessing steps included cleaning, case folding, normalization, tokenization, stopword removal, and stemming using the Sastrawi library. These steps are important in sentiment analysis of social media text because user-generated content often contains informal, abbreviated, and non-standard language [4], [5]. After preprocessing, the texts were automatically labeled into negative, neutral, and positive classes using a rule-based lexicon approach. This method was selected because it enables efficient annotation of large-scale data, though it may also introduce label noise, a limitation recognized in the study [18].

The preprocessed text was then transformed into numerical form using the Bag-of-Words (BoW) model with CountVectorizer. This representation was chosen because it is simple, interpretable, and compatible with classical classification algorithms such as Naïve Bayes and SVM, making it suitable for a consistent comparison of model performance [19], [20].

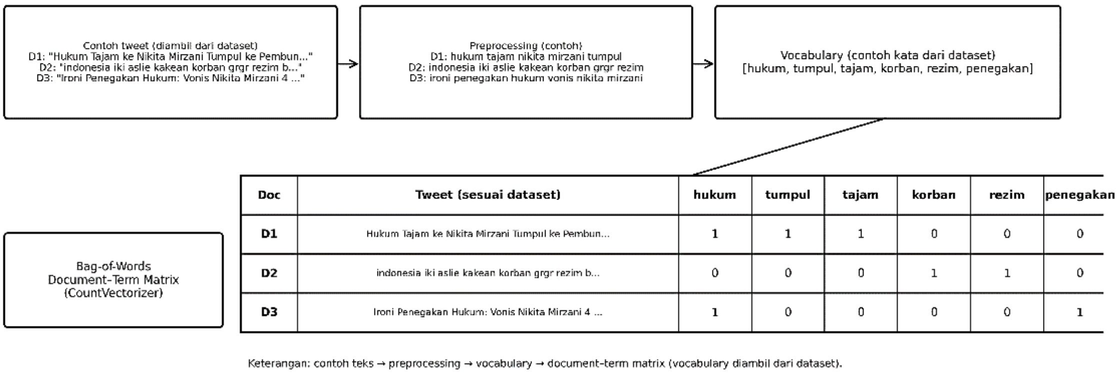

Figure 2. Feature representation using Bag-of-Words

The preprocessed text was converted into numerical vectors using the Bag-of-Words (BoW) model with CountVectorizer. In this representation, each document is transformed into a structured vector based on term frequency, yielding high-dimensional, sparse features suitable for classical text classification methods such as Multinomial Naïve Bayes and LinearSVC [19], [20]. By applying the same BoW representation to both models, this study ensures that the comparison focuses on differences in the classification algorithms rather than differences in feature engineering.

After feature extraction, the dataset was split into training and test sets using an 80:20 train–test stratified split (random_state = 42). Stratification was used to maintain the original class proportions in both subsets, which is important given the dataset's imbalance. This process yielded 5,944 training records and 1,487 test records, enabling the models to be evaluated on unseen data with comparable class distributions.

Two classification algorithms were then implemented and compared: Multinomial Naïve Bayes and linear Support Vector Machine (LinearSVC). Multinomial Naïve Bayes was selected as a probabilistic baseline because of its simplicity, efficiency, and suitability for word-frequency features, while LinearSVC was chosen because of its strong performance on high-dimensional sparse text data [21], [22]. Using the same dataset, preprocessing steps, feature representation, and evaluation framework allowed a fair and controlled comparison between the two methods.

Model performance was evaluated on the test set using accuracy, precision, recall, and F1-score. In addition, macro-average and weighted-average F1-scores were reported to provide a more balanced assessment of performance across imbalanced sentiment classes. Confusion matrices were also used to examine classification errors for each class in greater detail [23]. Overall, the methodology was designed to ensure a fair, valid, and reproducible comparison between the two models.

3. RESULTS AND ANALYSIS

This section presents the results of sentiment classification on public posts related to the Affan Kurniawan case on social media X and provides an analytical interpretation of the findings. The evaluation was conducted on unseen test data obtained from an 80:20 train–test split, while text features were represented using Bag-of-Words (BoW) with CountVectorizer. This feature representation produces high-dimensional vectors that are suitable for classical text classification models such as Naïve Bayes and SVM and is widely used in Twitter/X sentiment analysis studies, particularly when the objective is to compare classifier performance under a consistent representation scheme [10].

Before comparing model performance, it is important to examine the sentiment composition of the dataset. As shown in Table 2, the dataset consists of 7,431 labeled posts, with negative sentiment as the dominant class, followed by neutral and positive sentiment. Specifically, 3,735 posts (50.26%) were labeled as negative, 2,301 posts (30.96%) as neutral, and 1,395 posts (18.77%) as positive. This distribution indicates that public discourse surrounding the Affan Kurniawan case on X was predominantly negative during the observation period. The dominance of negative sentiment suggests that the issue generated stronger critical, emotional, or disapproving responses than supportive or neutral reactions.

Table 2. Sentiment class distribution

Sentiment class	Count (n)	Percentage (%)
Negative	3.735	50.26
Neutral	2.301	30.96
Positive	1.395	18.77
Total	7.431	100.00

The class distribution shown in Table 2 also indicates that the dataset is imbalanced. This condition is important because imbalanced data can bias evaluation results when model performance is assessed solely by accuracy. For this reason, this study emphasizes not only accuracy but also macro-average precision, recall, and F1-score, as well as confusion matrix analysis, to ensure a fairer interpretation of model performance across all sentiment classes [24]. This analytical approach is necessary because a model may appear accurate overall while performing poorly on minority or more ambiguous classes.

The comparative results show that SVM (LinearSVC) outperformed Multinomial Naïve Bayes under the same feature representation and evaluation framework. As presented in Table 3, SVM achieved an accuracy of 0.826, whereas Naïve Bayes reached 0.745. More importantly, SVM also achieved a substantially higher macro F1-score (0.759) than Naïve Bayes (0.579). This difference confirms that SVM not only performed better overall but also achieved more balanced classification across sentiment classes. In contrast, the lower macro F1-score of Naïve Bayes indicates weaker performance in handling minority or less distinct classes, despite showing acceptable results on the dominant class.

Table 3. Modeling Naïve Bayes and Support Vector Machine

Modeling	Class	Accuracy	Precision	Recall	F1-Score
Naïve Bayes (MultinomialNB)	negatif	0.745	0.800	0.889	0.842
	netral		0.581	0.210	0.309
	positif		0.525	0.665	0.586
	macro avg		0.635	0.588	0.579
	weighted avg		0.730	0.745	0.720
SVM (LinearSVC)	negatif	0.826	0.896	0.879	0.887
	netral		0.616	0.681	0.647
	positif		0.757	0.727	0.742
	macro avg		0.756	0.762	0.759
	weighted avg		0.831	0.826	0.828

A closer examination of class-wise results further clarifies the difference between the two models. For the negative class, both models performed relatively well, but SVM still achieved higher precision and competitive recall, resulting in a stronger F1 Score. The most significant gap appears in the neutral class. Naïve Bayes produced very low recall and F1-score for neutral data, indicating that it had difficulty distinguishing neutral posts from negative ones. By contrast, SVM showed a much better ability to classify neutral posts, suggesting that it was more robust in handling sentiment boundaries that are less explicit and more context-dependent. In the positive class, SVM also performed more consistently, with higher precision, recall, and F1-score than Naïve Bayes. These findings demonstrate that the superiority of SVM is not limited to aggregate metrics, but is also reflected at the class level.

The confusion matrices in Figure 3 provide further evidence of these patterns. For Naïve Bayes, the model correctly classified 937 of 1,054 negative instances, but its performance dropped substantially for the neutral class, where only 54 of 257 instances were correctly predicted. Most neutral instances were misclassified as negative, indicating that Naïve Bayes tended to favor the majority class when facing ambiguous or overlapping lexical patterns. The positive class also showed substantial misclassification into the negative class. This pattern suggests that the probabilistic assumptions of Naïve Bayes, particularly the assumption of feature independence, may be less effective for short and noisy social media texts where sentiment expressions are often overlapping and context-sensitive.

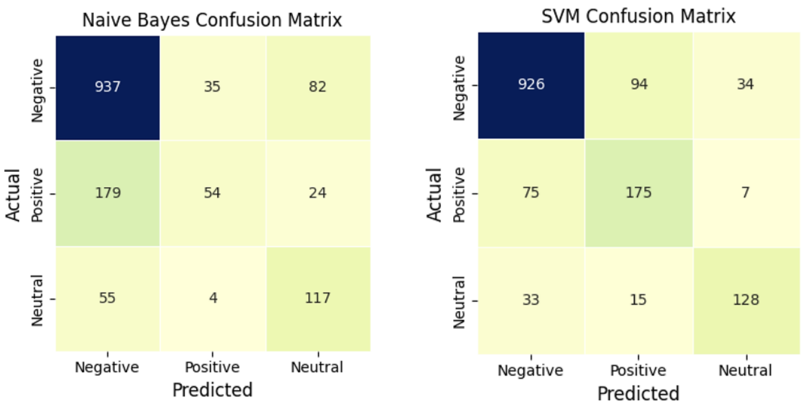

Figure 3 - Confusion matrix Naïve Bayes & SVM

In contrast, SVM produced a more balanced classification pattern, especially for the neutral class. The number of correctly classified neutral instances increased substantially from 54 in Naïve Bayes to 175 in SVM. This improvement is particularly important because neutral sentiment is often the most difficult class to detect in social media data, given that it may contain factual or informational expressions with limited emotional markers. The stronger neutral-class performance of SVM indicates that margin-based classification is more effective than a simple probabilistic baseline for separating subtle class boundaries in high-dimensional text representations. This finding is consistent with previous comparative studies reporting that SVM often outperforms Naïve Bayes in BoW-based sentiment classification tasks [19].

Overall, these findings indicate that under the same BoW–CountVectorizer configuration and evaluation scheme, SVM provides a stronger and more reliable baseline than Multinomial Naïve Bayes for sentiment classification on Indonesian-language social media data. From a methodological perspective, this result highlights that classifier choice has a substantial impact on performance even when the feature representation and dataset are held constant. This is an important contribution because it confirms that performance differences are attributable mainly to the models' learning mechanisms rather than to differences in preprocessing or feature engineering.

From a substantive perspective, the dominance of negative sentiment also provides insight into the public response to the Affan Kurniawan case. The findings suggest that the issue was received more critically than positively, reflecting a digital discourse environment shaped by strong reactions, judgments, and emotional engagement. This implies that sentiment analysis on social media can serve not only as a technical classification task but also as an analytical tool for understanding how social issues are perceived and amplified in online public spaces.

The practical implication of this study is that SVM is a more appropriate baseline model for monitoring public sentiment on social media X, especially in cases involving imbalanced sentiment classes and high lexical variation. For researchers, the results emphasize the importance of using multiple evaluation metrics rather than relying solely on accuracy. For practitioners or institutions monitoring public opinion, the findings suggest that selecting a model with stronger performance on minority classes is essential to avoid misleading interpretations of public sentiment dynamics.

However, the results should be interpreted in light of several limitations. First, the sentiment labels were generated automatically using a rule-based lexicon approach, which may introduce labeling noise [18]. Second, the dataset was limited to a specific time window, so the findings reflect only the observed period and may not capture longer-term sentiment dynamics. Third, the BoW representation cannot capture sarcasm, implicit meaning, or deeper contextual relationships between words. Fourth, public discourse on X may be influenced by bots or inauthentic accounts, which can affect both sentiment distribution and the interpretation of observed opinion patterns [6]. These limitations do not invalidate the findings, but they indicate that the results should be understood as a structured baseline rather than a complete representation of public opinion.

Based on these findings, future research is recommended to improve label quality through manual or hybrid annotation, address class imbalance through resampling or class-weighting techniques, compare BoW with alternative feature representations such as TF-IDF or n-grams, and optimize classifier parameters through hyperparameter tuning and cross-validation [24]. In addition, integrating bot-detection or authenticity-filtering would improve the reliability of sentiment analysis results by reducing discourse distortion from automated accounts [6]. These improvements would strengthen both the methodological robustness and the practical relevance of future sentiment analysis studies on social media.

Figure 4. Wordcloud per sentiment class

Figure 4 complements the classification results by visually presenting dominant terms across sentiment classes. Although word clouds are descriptive rather than inferential, they help illustrate the lexical tendencies associated with negative, neutral, and positive posts. This visualization supports the interpretation that public discourse was dominated by negative expressions, while also containing informational and supportive language across other sentiment categories.

4. CONCLUSION

This study examined public sentiment on social media X regarding the Affan Kurniawan case by applying text preprocessing, lexicon-based sentiment labeling, Bag-of-Words feature representation with CountVectorizer, and classification using Multinomial Naïve Bayes and Support Vector Machine (SVM) [25]. In line with the research objectives, the study first identified the distribution of sentiment in the dataset and found that public discourse was dominated by negative sentiment, followed by neutral and positive sentiment. This finding indicates that the Affan Kurniawan case generated a predominantly critical public response on X, reflecting strong emotional reactions and public disapproval toward the issue and its handling [25].

Second, this study compared the performance of Multinomial Naïve Bayes and SVM under the same dataset, preprocessing pipeline, feature representation, and evaluation framework. The results show that SVM outperformed Naïve Bayes in both overall accuracy and macro F1-score, indicating that SVM provides a stronger, more balanced classification model for Indonesian-language social media sentiment analysis. In particular, SVM performed better at classifying neutral and positive sentiments, whereas Naïve Bayes tended to misclassify minority or ambiguous classes as the dominant negative class [24]. This finding confirms that classifier selection has a substantial effect on sentiment classification performance, even when all other experimental settings are controlled.

Third, this study demonstrated the importance of using a uniform evaluation scheme in order to obtain a fair comparison between classification methods. The use of accuracy, precision, recall, F1-score, and confusion matrices allowed the study to evaluate not only aggregate performance but also class-specific prediction errors [24]. This is especially important for imbalanced social media datasets, where relying solely on accuracy may lead to misleading conclusions about model effectiveness.

The main contribution of this study lies in three aspects. Empirically, it provides evidence of how public sentiment toward a social issue is distributed and expressed on social media X. Methodologically, it offers a transparent comparison between Naïve Bayes and SVM under the same Bag-of-Words representation and evaluation design. In practice, it shows that SVM is more suitable as a baseline model for monitoring public sentiment in fast-evolving, imbalanced social media environments [25].

However, the findings should be interpreted with caution. The sentiment labels were generated automatically using a lexicon-based approach, which may introduce labeling noise. In addition, the dataset was limited to a specific time period and may not fully capture longer-term sentiment dynamics. The possible influence of bots or inauthentic accounts may also affect observed sentiment patterns and interpretations of public discourse [6]. These limitations indicate that the present findings should be understood as a structured baseline rather than a final representation of public opinion.

For future research, several concrete directions are recommended. First, label quality should be improved through manual or hybrid annotation to reduce noise and better capture sarcasm, ambiguity, and implicit meaning [25]. Second, future studies should compare Bag-of-Words with richer feature representations, such as TF-IDF, n-grams, word embeddings, or transformer-based models, to assess whether contextual representations can improve classification performance. Third, methodological robustness can be strengthened through hyperparameter tuning, cross-validation, and class-balancing techniques such as resampling or class weighting [24]. Fourth, future research should incorporate bot detection or authenticity filtering to reduce discourse bias and improve the validity of sentiment interpretation on social media X [6]. Finally, extending the observation period and comparing multiple social issues may provide broader insight into how public sentiment evolves over time and across contexts.

Overall, this study demonstrates that sentiment analysis can serve as both a computational and analytical tool for understanding digital public opinion. By showing that SVM performs more reliably than Naïve Bayes under controlled experimental conditions, this research provides a clear methodological direction for future sentiment analysis studies on Indonesian social media data and contributes to the development of more robust approaches for analyzing public discourse in online environments.

REFERENCES

[1] H. T. Phan, N. T. Nguyen, and D. Hwang, “Sentiment Analysis for Social Media: a Survey,” J. Comput. Sci. Cybern., vol. 37, no. 4, pp. 403–428, 2021, doi: 10.15625/1813-9663/37/4/15892.

[2] Y. Qi and Z. Shabrina, “Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach,” Soc. Netw. Anal. Min., vol. 13, no. 1, pp. 1–14, 2023, doi: 10.1007/s13278-023-01030-x.

[3] S. Lee, Z. Xie, E. Xu, Y. Shao, D. J. Ossip, and D. Li, “Public perceptions of the FDA’s marketing authorization of Vuse on Twitter/X,” Front. Public Heal., vol. 11, no. November 2023, pp. 1–5, 2023, doi: 10.3389/fpubh.2023.1280658.

[4] M. A. Palomino and F. Aider, “Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis,” Appl. Sci. 2022, Vol. 12, Page 8765, vol. 12, no. 17, p. 8765, Aug. 2022, doi: 10.3390/APP12178765.

[5] Q. A. Xu, V. Chang, and C. Jayne, “A systematic review of social media-based sentiment analysis: Emerging trends and challenges,” Decis. Anal. J., vol. 3, p. 100073, Jun. 2022, doi: 10.1016/J.DAJOUR.2022.100073.

[6] L. H. X. Ng and K. M. Carley, “A global comparison of social media bot and human characteristics,” Sci. Rep., vol. 15, no. 1, pp. 1–18, 2025, doi: 10.1038/s41598-025-96372-1.

[7] H. Wan, M. Luo, Z. Ma, G. Dai, and X. Zhao, “How Do Social Bots Participate in Misinformation Spread? A Comprehensive Dataset and Analysis,” pp. 31481–31504, 2025, doi: 10.18653/v1/2025.emnlp-main.1604.

[8] A. Upadhyaya, M. Fisichella, and W. Nejdl, “Towards sentiment and Temporal Aided Stance Detection of climate change tweets,” Inf. Process. Manag., vol. 60, no. 4, p. 103325, Jul. 2023, doi: 10.1016/J.IPM.2023.103325.

[9] A. R. Sembiring and C. K. Dewa, “Sentiment Analysis On Indonesian Tweets about the 2024 Election,” Sinkron, vol. 9, no. 1, pp. 413–422, 2025, doi: 10.33395/sinkron.v9i1.14481.

[10] Y. Wang, J. Guo, C. Yuan, and B. Li, “Sentiment Analysis of Twitter Data,” Appl. Sci., vol. 12, no. 22, pp. 1–14, 2022, doi: 10.3390/app122211775.

[11] H. T. Wijaya and K. Kustiyono, “Sentiment Analysis of Twitter Users Towards the Kartu Prakerja Program Using the Naive Bayes Method,” Int. J. Adv. Data Inf. Syst., vol. 5, no. 2, pp. 242–252, 2024, doi: 10.59395/ijadis.v5i2.1342.

[12] S. Barreto, R. Moura, J. Carvalho, A. Paes, and A. Plastino, “Sentiment analysis in tweets: an assessment study from classical to modern word representation models,” Data Min. Knowl. Discov. 2022 371, vol. 37, no. 1, pp. 318–380, Nov. 2022, doi: 10.1007/S10618-022-00853-0.

[13] A. S. Muliana, D. Lestarini, and S. P. Raflesia, “Analysis of Public Sentiment on Election Results using Naïve Bayes in Social Media X,” Sistemasi, vol. 13, no. 6, p. 2467, 2024, doi: 10.32520/stmsi.v13i6.4592.

[14] Riyantoro, “Sentiment Analysis of Twitter Data on Indonesia’s Cabinet Using Naïve Bayes and Support Vector Machine Algorithms,” vol. 10, no. April, p. 2024, 2024.

[15] J. Y. M. Nip and B. Berthelier, “Social Media Sentiment Analysis,” Encyclopedia, vol. 4, no. 4, pp. 1590–1598, 2024, doi: 10.3390/encyclopedia4040104.

[16] N. Nyoman, A. Sri, I. G. I. Sudipa, N. Nyoman, A. J. Sastaparamitha, and A. Gede, “Public Response on X to the Revocation of Indonesia ’ s 3-Kg LPG Retail Ban : A Support Vector Machine Study,” vol. 6, no. 2, pp. 412–425.

[17] R. Murtfeldt, N. Alterman, I. Kahveci, and J. D. West, “RIP Twitter API: A eulogy to its vast research contributions,” pp. 1–34, 2024, [Online]. Available: http://arxiv.org/abs/2404.07340

[18] P. Paper, H. D. Sharma, and P. Goyal, “An Analysis of Sentiment : Methods , Applications ,” no. Ml, 2023.

[19] J. O. Leandro and M. I. Fianty, “Evaluation of Sentiment Analysis Methods for Social Media Applications: A Comparison of Support Vector Machines and Naïve Bayes,” Int. J. Informatics Vis., vol. 9, no. 2, pp. 796–807, 2025, doi: 10.62527/joiv.9.2.2905.

[20] M. Graff, D. Moctezuma, and E. S. Téllez, “Bag-of-Word approach is not dead: A performance analysis on a myriad of text classification challenges,” Nat. Lang. Process. J., vol. 11, p. 100154, Jun. 2025, doi: 10.1016/J.NLP.2025.100154.

[21] T. R. Widodo, I. N. Fajri, and B. W. Sari, “Sentiment Analysis of the Film ‘JUMBO’ on Twitter Using the Naive Bayes Method and Support Vector Machine (SVM) with a Text Mining Approach,” J. Appl. Informatics Comput., vol. 9, no. 5, pp. 2861–2868, 2025, doi: 10.30871/jaic.v9i5.10557.

[22] N. Wayan, I. Juliandewi, A. S. Kusuma, K. Martina, D. Putri, and I. Gusti, “Comparison of Naïve Bayes and Random Forest in Sentiment Analysis of State-Owned Banks Management by Danantara on X and YouTube,” vol. 6, no. 3, pp. 527–537, 2024.

[23] N. A. Maulana and D. Darwis, “Perbandingan Metode SVM dan Naïve Bayes untuk Analisis Sentimen pada Twitter tentang Obesitas di Kalangan Gen Z,” J. Pendidik. dan Teknol. Indones., vol. 5, no. 3, pp. 655–666, 2025, doi: 10.52436/1.jpti.691.

[24] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” Sci. Rep., vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-56706-x.

[25] Y. Mao, Q. Liu, and Y. Zhang, “Sentiment analysis methods, applications, and challenges: A systematic literature review,” J. King Saud Univ. - Comput. Inf. Sci., vol. 36, no. 4, p. 102048, 2024, doi: 10.1016/j.jksuci.2024.102048.

BIBLIOGRAPHY OF AUTHORS

Marsella Iriana Mamusung is an undergraduate student in Informatics Engineering at Universitas Papua, Manokwari, Indonesia, currently completing her bachelor’s thesis based on this research. Her research interests include sentiment analysis, text mining, social media analytics, and machine learning for natural language processing.

Mrs. Lorna Yertas Baisa, S.T., M.Kom., is a lecturer in the Department of Informatics Engineering, Faculty of Engineering, Universitas Papua, Manokwari, Indonesia. Her teaching focuses on data mining and related informatics courses. Her research interests include data mining and applied computing. She supervises undergraduate student projects and is involved in curriculum development as well as academic collaborations, including research grants and partnerships

Mr. Andreas Leonardo Sumendap, S.T., M.T., is a lecturer in the Informatics Engineering Department, Faculty of Engineering, Universitas Papua, Manokwari, Indonesia. His teaching and research interests include artificial intelligence, image processing, and computer networks. He supervises undergraduate student projects and is involved in academic collaborations that support teaching and research development.