Sentiment Analysis and Classification of Public Opinion on Prabowo Subianto Using Naïve Bayes on Twitter

Cucu Sulaiman[1], Ummy Gusti Salamah[2], Risna Oktavitati[3], Safira Faizah[4*]

[1,4*]Department Informatics Engineering, Faculty Engineering and Computer Science, Jakarta Global University, 16412, Indonesia
[2]Department Informatics Management, State Polytechnic of Sriwijaya, Palembang, 30139, Indonesia
[3]Department Informatics Engineering, State Polytechnic of Sriwijaya, Palembang, 30139, Indonesia

Article Info	ABSTRACT
	In recent years, microblogging platforms have emerged as highly popular communication tools among internet users. Microblogging refers to a form of social media that enables users to share short messages—typically limited to fewer than 200 characters—containing opinions, comments, or news. One widely used microblogging service is Twitter, which attracts users from diverse backgrounds. Following the Indonesian presidential election in February 2024, numerous trending tweets expressed various opinions about Prabowo Subianto, the 8th President of Indonesia. These tweets reflected a range of sentiments, including positive, negative, and neutral viewpoints. Based on this, a sentiment analysis system is needed to efficiently analyze and classify public responses on Twitter according to sentiment categories: positive, negative, or neutral. To perform this classification, an appropriate algorithm is required. One suitable method is the Naïve Bayes classifier, a probabilistic algorithm that categorizes data by computing probability values. This method is well-suited for text classification tasks such as sentiment analysis of tweets. In this study, a dataset comprising tweets was used, consisting of 96 entries for training and 150 entries for testing, totaling 246 data points. The Naïve Bayes classifier achieved an accuracy rate of 84% using this dataset.

Corresponding Author:

Safira Faizah
Department Informatics Engineering, Faculty Engineering and Computer Science, Jakarta Global University, Depok, 16412
Email: safirafaizah@jgu.ac.id

1. INTRODUCTION

Twitter (now rebranded as X) is a widely used microblogging platform that enables users to express thoughts, opinions, and responses in short text messages known as tweets. With a limit of 280 characters, tweets often reflect spontaneous and unfiltered public sentiment on current events. Due to its real-time nature and widespread user engagement, Twitter has become a rich and dynamic source of data for researchers across various disciplines. Numerous studies have exploited this data to gain insights into public attitudes and societal trends. For instance, Kamiński et al. (2024) [1] analyzed tweets involving fatphobia and body shaming to examine how discriminatory language affects engagement metrics like likes and retweets. Similarly, Lane et al. (2024)[2] explored how language in environmental-related tweets correlates with regional obesity rates, suggesting that areas with higher tweet activity on environmental issues tend to exhibit healthier behavior patterns. The vast volume and accessibility of tweet data allow researchers to apply machine learning and natural language processing (NLP) techniques to detect patterns, sentiments, sarcasm, and opinions. With millions of daily tweets covering topics like stock markets [3], sarcasm detection [4], and signed social network interactions [5], [6], Twitter has become an essential source for both real-time monitoring and retrospective analysis. Political discourse is another dominant theme on Twitter. Firdaus et al. (2024) [7] collected and analyzed 30,000 tweets to examine public sentiment toward presidential candidates in Indonesia's 2024 election, highlighting Twitter's role as a reflection of political engagement. As such, tweets serve not only as

57

J. of Glob. Engineering Res. & Science   *e*-ISSN: 2830-5418, *p*-ISSN: 2830-5906

individual expressions but also as aggregated indicators of collective sentiment, social awareness, and political alignment.

The analysis of sentiments expressed in tweets has gained prominence as a method to systematically interpret public emotions and opinions in textual data. This process involves the classification of tweets into predefined sentiment categories such as positive, negative, or neutral, using machine learning algorithms. Naïve Bayes classifiers, known for their simplicity and speed, have been frequently applied in this domain, with advancements such as the Naïve Bayes Enrichment Method (NBEM) enhancing recall and precision across diverse datasets [8]–[10]. Tan et al. (2024) [11] introduced a hybrid approach by combining Naïve Bayes with regularized logistic regression to improve accuracy, particularly in low-dimensional settings. Beyond traditional classifiers, novel models such as hybrid quantum neural networks [3] have emerged, integrating quantum computing with deep learning to improve predictive accuracy in complex tasks like stock price forecasting based on Twitter sentiment. In the context of elections, specialized models like ElecBERT [12] leverage transformer-based architectures fine-tuned on multi-lingual election-related tweet datasets, outperforming conventional models in accuracy and F1-score. These evolving methodologies demonstrate a broader trend: sentiment analysis is shifting toward more explainable, accurate, and context-aware systems. In parallel, the importance of rigorous threshold-setting and performance evaluation using confusion matrices, as highlighted by Phillips et al. (2024) [13], reinforces the need for careful balancing of classification accuracy, precision, and recall—particularly in skewed or imbalanced data scenarios often encountered in real-world sentiment datasets. This is particularly important in high-stakes environments such as political forecasting, public health monitoring, and financial market analysis. Firdaus et al. (2024) [7] emphasize that pre-election Twitter sentiment can serve as a baseline for analyzing political alignment shifts during campaign periods and post-election outcomes, showing how sentiment analysis contributes to a deeper understanding of public behavior over time.

While numerous studies have utilized Twitter data for sentiment analysis in diverse domains such as public health, environmental issues, financial prediction, and social behavior, there is limited research specifically focused on analyzing political sentiment in the context of Indonesian elections using classical machine learning approaches like Naïve Bayes [14], [15]. Most recent developments emphasize deep learning models or hybrid approaches, which, although powerful, often lack transparency and require extensive computational resources. Moreover, existing political sentiment studies on Twitter tend to focus on global or Western contexts, with relatively few addressing localized sociopolitical dynamics in Indonesia. Additionally, despite the proven effectiveness of Naïve Bayes in low-dimensional classification tasks and its ability to perform well with limited training data, its application for election-related sentiment classification in Indonesia remains underexplored. This study addresses this gap by applying and evaluating the performance of the Naïve Bayes classifier to classify sentiments toward Prabowo Subianto during the 2024 Indonesian presidential election, offering a more interpretable and resource-efficient alternative for political opinion mining.

2. METHOD

This study employs the Naïve Bayes classification method to perform sentiment analysis on tweets related to Prabowo Subianto during the 2024 Indonesian presidential election. The methodological steps begin with data collection using the Twitter API, where relevant tweets are gathered based on specific keywords associated with the candidate. The raw data undergo several preprocessing stages, including case folding, tokenization, stop-word removal, and stemming, to clean and standardize the text for analysis. After preprocessing, the dataset is divided into training and testing subsets, with 96 tweets used for training and 150 for testing. The Naïve Bayes algorithm is then applied to the training data to calculate the probability of each sentiment class such as positive, negative, or neutral based on the frequency of word occurrences. The trained model is evaluated using the test data to assess its classification accuracy. In this study, the Naïve Bayes classifier achieved an accuracy rate of 84%, demonstrating its effectiveness as a lightweight and interpretable approach for sentiment classification in political social media content.

3. RESULTS AND DISCUSSION

The results of this study present the performance of the Naïve Bayes classifier in analyzing and classifying public sentiment toward Prabowo Subianto based on Twitter data. The analysis focuses on the classification accuracy and distribution of sentiment categories, providing insights into public opinion during the 2024 Indonesian presidential election.

3.1. Naïve Bayes Classifier (NBC)

The Naïve Bayes classifier begins by calculating the probability of each sentiment class based on the entire training dataset. This involves determining how likely a tweet belongs to a particular class, namely positive, negative, or neutral, given the features present in the training data in table 1. The tweet data presented in Table 1 is displayed in the original Indonesian language to maintain authenticity.

Table 1. training data

No	Tweet Data	Class
1	Supporting Prabowo Subianto to win	Positive
2	Prabowo Subianto Is Very Good	Positive
3	Reject Prabowo Subianto as President	Negative
4	The most racist, the worst president	Negative
5	Why did Prabowo Subianto re-register as a presidential candidate?	Neutral
6	Prabowo Subianto will register again as a candidate for the presidency of the Republic of Indonesia	Neutral

The following is the probability calculation for each class, performed using the equation presented below

$$P\left(v_{positive}\right) = \frac{\Sigma positive}{\Sigma tweet\ data} \tag{1}$$

$$P\left(v_{negative}\right) = \frac{\Sigma negative}{\Sigma tweet\ data} \tag{2}$$

$$P\left(v_{neutral}\right) = \frac{\Sigma neutral}{\Sigma tweet\ data} \tag{3}$$

The probability of each class is 0.33, as there are two tweets in each category namely positive, negative, and neutral, out of a total of six tweet entries. Subsequently, the classification of new tweet data is performed using the previously calculated probability values for each class. The following is the tweet data to be classified using the Naïve Bayes classifier as test data in table 2 and preprocessing in table 3.

Table 2. test data

No	Tweet Data	Class
1	Let's support Prabowo Subianto so that he can win the presidential election of the Republic of Indonesia later!	?

Table 3. preprocessing data

No	Tweet Data	Class
1	Support win	?

Subsequently, the likelihood of each word appearing in the document for each class is calculated. The list of words retained after the preprocessing stage is presented in Table 4.

Table 4. List of words

No	word	class
1	support	Positive
2	win	Positive
3	good	Positive
4	reject	Negative
5	racist	Negative
6	list	Neutral
7	candidate	Neutral
8	nominator	Neutral

Next, the probability of the occurrence of the word w_i in a document belonging to class v_j is calculated.

$$P\left(w_i|v_j\right) = \frac{n_{ij}+1}{n_j+|word|} \tag{4}$$

In positive class,

$$P(support|positive) = \frac{1+1}{3+8} = 0,18$$

$$P(win|positive) = \frac{1+1}{3+8} = 0,18$$

In negative class

$$P(support|negative) = \frac{0+1}{2+8} = 0,1$$

$$P(win|negative) = \frac{0+1}{2+8} = 0,1$$

In neutral class,

$$P(support|neutral) = \frac{0+1}{3+8} = 0,09$$

$$P(win|netral) = \frac{0+1}{3+8} = 0,09$$

Subsequently, a calculation is performed to determine the class category that is most likely or closest to the given data.

$$V_{positive} = 0,33 * 0,18 * 0,18 = 0,010692$$

$$V_{negative} = 0,33 * 0,1 * 0,1 = 0,0033$$

$$V_{neutral} = 0,33 * 0,09 * 0,09 = 0,002673$$

Finally, the class is determined by identifying the highest probability value among the calculated results.

$$V_{MAP} = \underset{v_j \in V}{argmax} \ (V_{positive}|V_{negative}|V_{neutral}) \hspace{2cm} (5)$$

$$V_{MAP} = \underset{v_j \in V}{argmax} \ (0,010692|0,0033|0,002673)$$

The maximum value obtained is 0.010692, which corresponds to the V value for the positive class. Therefore, the test data is classified as belonging to the "Positive" class.

3.2. Accuracy, Precision, Recall, and Specificity Evaluation
From the test data, consisting of 12 entries (with 4 data points in each class), collected between March 16, 2023, and August 16, 2024, the following results were obtained:

Actually class	Prediction			
	Positive	**Negative**	**Neutral**	**Total**
Positive	4	0	0	4
Negative	1	3	0	4
Neutral	1	0	3	4
Total	6	3	3	12

Based on the confusion matrix evaluation, accuracy, precision, recall, and specificity can be calculated using the following formulas:

$$Accuracy = \frac{T_{positive}+T_{negative}+T_{neutral}}{T_{positive}+T_{negative}+T_{neutral}+F_{positive}+F_{negative}+F_{neutral}} \hspace{1cm} (6)$$

$$Accuracy = \frac{4+3+3}{4+3+3+0+1+1} = \frac{10}{12} = 0{,}83 = 83\%$$

Based on the calculation above, the accuracy of the sentiment classification model using the Naïve Bayes algorithm is 83%, meaning that 10 out of 12 test data points were correctly classified. The true positives for each class namely, positive, negative, and neutral, were 4, 3, and 3 respectively, indicating that the model was effective at identifying tweets that genuinely belonged to their respective sentiment categories. Meanwhile, the remaining 2 instances (1 false negative and 1 false neutral) represent misclassifications. This result suggests that the Naïve Bayes classifier is reasonably effective for sentiment analysis on Twitter data, particularly given the relatively small dataset size. An 83% accuracy demonstrates the model's potential to generalize well on unseen data, though there remains room for improvement, especially in reducing false classifications. Further optimization could be achieved by increasing the size of the training dataset, enhancing preprocessing techniques, or integrating additional linguistic features to improve classification robustness.

$$Precision_{positive} = \frac{T_{positive}}{T_{positive} + F_{positive}} = \frac{4}{4+2} = 0{,}67 = 67\%$$

$$Precision_{negative} = \frac{T_{negative}}{T_{negative} + F_{negative}} = \frac{3}{3+0} = 100\%$$

$$Precision_{neutral} = \frac{T_{neutral}}{T_{neutral} + F_{neutral}} = \frac{3}{3+0} = 100\%$$

$$Precision_{mean} = \frac{67 + 100 + 100}{3} = 89\%$$

The Naïve Bayes classifier demonstrates varying levels of precision across sentiment classes. The positive class achieved a precision of 67%, indicating that out of all tweets predicted as positive, 67% were correct. In contrast, both the negative and neutral classes achieved a perfect precision of 100%, meaning every tweet predicted as negative or neutral was classified correctly with no false positives. The mean precision across all three classes is calculated at 89%, reflecting strong overall performance of the model in terms of precision. These results suggest that while the classifier is highly reliable in predicting negative and neutral sentiments, it has some difficulty in accurately identifying positive sentiment, which may stem from overlapping linguistic features or limited training examples for that class. Improving feature representation or expanding the dataset could help balance the precision performance across all sentiment categories.

$$Recall_{positive} = \frac{T_{positive}}{T_{positive} + F_{negative} + F_{neutral}} = \frac{4}{4+0} = 1 = 100\%$$

$$Recall_{negative} = \frac{T_{negative}}{T_{negative} + F_{positive} + F_{neutral}} = \frac{3}{3+2} = 0.6 = 60\%$$

$$Recall_{neutral} = \frac{T_{neutral}}{T_{neutral} + F_{positive} + F_{negative}} = \frac{3}{3+2} = 0.6 = 60\%$$

$$Recall_{mean} = \frac{100 + 60 + 60}{3} = 73\%$$

The recall results illustrate the Naïve Bayes classifier's effectiveness in identifying actual instances of each sentiment class. The positive class achieved a perfect recall of 100%, meaning all tweets that truly belonged to the positive class were correctly identified. However, the negative and neutral classes both recorded a recall of 60%, indicating that 40% of the actual tweets in each of these classes were misclassified as another sentiment. The mean recall across all classes is 73%, which highlights that while the model is highly effective at detecting positive sentiments, it is less sensitive to detecting negative and neutral sentiments. This discrepancy could stem from overlapping features in tweet content, limited training data, or class imbalance.

To improve recall which is particularly for the negative and neutral classes, future work could explore enhanced feature extraction techniques, rebalance the dataset, or integrate ensemble methods to reduce misclassification rates.

$$Specificity_{positive} = \frac{T_{negative} + T_{neutral}}{T_{negative} + T_{neutral} + F_{positive}} = \frac{3+3}{3+3+2} = 0.75 = 75\%$$

$$Specificity_{negative} = \frac{T_{positive} + T_{neutral}}{T_{positive} + T_{neutral} + F_{negative}} = \frac{4+3}{4+3+0} = 1 = 100\%$$

$$Specificity_{neutral} = \frac{T_{positive} + T_{negative}}{T_{positive} + T_{negative} + F_{neutral}} = \frac{4+3}{4+3+0} = 1 = 100\%$$

$$Specificity_{mean} = \frac{75 + 100 + 100}{3} = 92\%$$

The specificity results presented indicate the Naïve Bayes classifier's ability to correctly identify tweets that do not belong to a particular sentiment class. The positive class achieved a specificity of 75%, meaning 75% of the tweets not belonging to the positive class were correctly excluded. In contrast, both the negative and neutral classes achieved perfect specificity scores of 100%, indicating no false positives were recorded for those classes. The mean specificity across all three classes is 92%, reflecting strong performance in correctly identifying non-class members. This high specificity suggests that the model is particularly effective at minimizing false positives, especially for the negative and neutral categories. However, the lower specificity for the positive class indicates some misclassification into this category, which may warrant further refinement of feature selection or training data balancing to improve overall classification accuracy.

4. CONCLUSION

The evaluation results show that the Naïve Bayes classifier is a promising method for sentiment analysis of public opinion on Twitter, particularly in the political context surrounding Prabowo Subianto. The model achieved a respectable accuracy of 83%, correctly classifying 10 out of 12 tweets. In terms of precision, the classifier performed best in the negative and neutral classes with perfect scores of 100%, while the positive class achieved 67%, resulting in a mean precision of 89%. This indicates the model is generally effective at minimizing false positives, particularly for negative and neutral sentiments. On the other hand, the recall scores varied, with the positive class achieving 100%, but the negative and neutral classes each scoring 60%, giving a mean recall of 73%. This suggests that while the model is good at identifying true positives in the positive class, it tends to miss some actual instances of negative and neutral sentiments. The specificity results are strong across the board, with 100% for negative and neutral classes and 75% for positive class, leading to a mean specificity of 92%, highlighting the model's robustness in excluding tweets that do not belong to a particular class.

In summary, the classifier shows strong overall performance, especially in avoiding misclassification, but still has room for improvement in capturing all relevant instances of certain sentiments. This points to the potential for refinement in areas such as data preprocessing, class balancing, or the integration of more advanced linguistic features. The findings validate the use of Naïve Bayes as a viable tool for sentiment classification in Indonesian political discourse on Twitter, with meaningful implications for political monitoring, campaign evaluation, and public opinion analysis.

ACKNOWLEDGEMENTS

REFERENCES

[1] M. Kamiński, T. Wieczorek, M. Kręgielska-Narożna, and P. Bogdański, 'Tweeting about fatphobia and body shaming: A retrospective infodemiological study', *Nutrition*, vol. 125, p. 112497, Sep. 2024, doi: 10.1016/j.nut.2024.112497.
[2] J. M. Lane *et al.*, 'Tweeting environmental pollution: Analyzing twitter language to uncover its correlation with county-level obesity rates in the United States', *Prev. Med. (Baltim).*, vol. 186, p. 108081, Sep. 2024, doi: 10.1016/j.ypmed.2024.108081.
[3] M. Gandhudi, A. P.J.A., U. Fiore, and G. G.R., 'Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction', *Comput. Electr. Eng.*, vol. 118, p. 109302, Aug. 2024, doi: 10.1016/j.compeleceng.2024.109302.
[4] Q. Abuein, R. M. Al-Khatib, A. Migdady, M. S. Jawarneh, and A. Al-Khateeb, 'ArSa-Tweets: A novel Arabic sarcasm detection system based on deep learning model', *Heliyon*, vol. 10, no. 17, p. e36892, Sep. 2024, doi: 10.1016/j.heliyon.2024.e36892.
[5] S.-Y. Liu, J. Xiao, and X.-K. Xu, 'Sign prediction by motif naive Bayes model in social networks', *Inf. Sci. (Ny).*, vol. 541, pp.

316–331, Dec. 2020, doi: 10.1016/j.ins.2020.05.128.

[6] Muhammad Aditya Wisnu Wardana et all., 'Development OF Canva Application Based Learning Media for E-Book Interactive', *J. Nalar Pendidik.*, vol. 10, pp. 71–79, 2023, doi: 10.26858/jnp.v10i1.

[7] A. A. Firdaus, A. Yudhana, I. Riadi, and Mahsun, 'Indonesian presidential election sentiment: Dataset of response public before 2024', *Data Br.*, vol. 52, p. 109993, Feb. 2024, doi: 10.1016/j.dib.2023.109993.

[8] O. Peretz, M. Koren, and O. Koren, 'Naive Bayes classifier – An ensemble procedure for recall and precision enrichment', *Eng. Appl. Artif. Intell.*, vol. 136, p. 108972, Oct. 2024, doi: 10.1016/j.engappai.2024.108972.

[9] Merinda Lestandy, Abdurrahim Abdurrahim, and Lailis Syafa'ah, 'Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent Neural Network dan Naïve Bayes', *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 802–808, 2021, doi: 10.29207/resti.v5i4.3308.

[10] D. Normawati and S. A. Prayogi, 'Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter', *J. Sains Komput. Inform. (J-SAKTI*, vol. 5, no. 2, pp. 697–711, 2021.

[11] Y. Tan, B. Sherwood, and P. P. Shenoy, 'A naïve Bayes regularized logistic regression estimator for low-dimensional classification', *Int. J. Approx. Reason.*, vol. 172, p. 109239, Sep. 2024, doi: 10.1016/j.ijar.2024.109239.

[12] A. Khan, H. Zhang, N. Boudjellal, A. Ahmad, and M. Khan, 'Improving Sentiment Analysis in Election-Based Conversations on Twitter with ElecBERT Language Model', *Comput. Mater. Contin.*, vol. 76, no. 3, pp. 3345–3361, 2023, doi: 10.32604/cmc.2023.041520.

[13] G. Phillips *et al.*, 'Setting nutrient boundaries to protect aquatic communities: The importance of comparing observed and predicted classifications using measures derived from a confusion matrix', *Sci. Total Environ.*, vol. 912, p. 168872, Feb. 2024, doi: 10.1016/j.scitotenv.2023.168872.

[14] P. Arsi, R. Wahyudi, and R. Waluyo, 'Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia', *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 231–237, 2021, doi: 10.29207/resti.v5i2.2698.

[15] R. Wahyudi and G. Kusumawardana, 'Sentiment Analysis on the Grab Application on Google Play Store Using Support Vector Machine', *J. Inform.*, vol. 8, no. 2, pp. 200–207, 2021.