

# Social Media X-Based Public Opinion Analysis of Prabowo-Gibran Government's First 100 Days Using Naïve Bayes and K-Nearest Neighbor Classification Methods

Ridha Afifah <sup>a\*</sup>, Sugiyono <sup>b</sup>

<sup>a\*,b</sup> Informatics Engineering Study Program, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, East Jakarta City, Special Capital Region of Jakarta, Indonesia.

## ABSTRACT

This study investigates public sentiment toward the first 100 days of the Prabowo-Gibran administration by analyzing opinions expressed on X (formerly Twitter) using machine learning approaches. A total of 431 valid tweets were collected, preprocessed, and manually labeled into positive and negative categories. The results reveal that 62% of public sentiment was negative, while 38% was positive, indicating widespread public criticism during the administration's early period. Two algorithms, Naïve Bayes and K-Nearest Neighbor (KNN), were applied to classify sentiment. The Naïve Bayes model achieved superior performance, with an accuracy of 97.22%, compared to KNN's 62.65%. The probabilistic nature of Naïve Bayes allowed it to manage high-dimensional, imbalanced textual data effectively, while KNN suffered from the "curse of dimensionality" and class bias. These findings demonstrate that Naïve Bayes remains a reliable and computationally efficient model for political sentiment analysis in the Indonesian digital context. Despite its strengths, this study acknowledges limitations in manual labeling and linguistic nuances such as sarcasm and irony. Future research is encouraged to integrate deep learning architectures like LSTM or BERT and adopt aspect-based sentiment analysis to capture more contextual insights from political discourse.

## ARTICLE HISTORY

Received 25 July 2025  
Accepted 30 October 2025  
Published 30 November 2025

## KEYWORDS

Sentiment Analysis; Naïve Bayes; K-Nearest Neighbor; Political Opinion; Social Media.

## 1. Introduction

The advancement of digital technology has positioned social media platforms, particularly X (formerly Twitter), as dominant spaces for citizens to express opinions in real time. Political discourse in these spaces—such as reactions to the first one hundred days of the Prabowo-Gibran administration—has generated polarized public responses that reflect both approval and discontent. The massive volume of unstructured opinion data emerging from X demands systematic analysis to transform it into measurable insight about public sentiment and perception of government performance. Sentiment analysis, supported by machine learning, offers a structured means of interpreting this data by classifying opinions into positive or negative categories based on textual patterns. Prior research demonstrates that the Naïve Bayes Classifier and K-Nearest Neighbor (KNN) algorithms have been extensively utilized for sentiment analysis across various social and political contexts in

Indonesia. Studies by Era and Andryana (2021), Rahman and Lestari (2021), and Ramadhani and Wahyudin (2022) highlight the reliability of both algorithms in processing short and informal texts such as tweets. Similarly, Asro'i and Februariyanti (2022) and Pramana and Afdal (2022) applied KNN and Naïve Bayes to evaluate public reactions toward government policies, while Hakim and Sugiyono (2024) confirmed their applicability in analyzing public attitudes toward infrastructure programs. Beyond politics, sentiment classification has also been applied to diverse domains, including pandemic response (Pamungkas & Kharisudin, 2020), film reviews (Nurtikasari & Alam, 2022), and public opinion about social programs (Sitanggang & Umidah, 2024), reinforcing the algorithms' adaptability for various linguistic environments.

Despite their extensive use, few studies have specifically compared these algorithms in assessing public opinion toward the early performance of the Prabowo–Gibran administration. Existing research tends to focus on broader political or policy issues rather than this specific transitional period, creating a notable research gap. To address that gap, this study employs text mining techniques to collect, process, and classify public sentiment data from X using both Naïve Bayes and KNN. The objective is twofold: to determine the prevailing public sentiment toward the administration's first hundred days and to evaluate the comparative accuracy of these two algorithms in this context. The methodological framework of the study involves four stages: (1) data collection through targeted scraping of posts containing the hashtag “100HariPrabowoGibran”; (2) data preprocessing, including cleaning, tokenization, and labeling; (3) sentiment classification using Naïve Bayes and KNN models; and (4) evaluation of model performance based on accuracy, precision, and recall metrics. The outcomes are expected to clarify which algorithm performs more effectively for short-form political text analysis in the Indonesian digital environment. The remainder of this paper is structured as follows: the next section describes the research methodology, followed by the presentation and discussion of results, and finally, the conclusion.

## 2. Methodology

This study applies a quantitative approach based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, which provides a structured model for solving data-driven problems across multiple disciplines. The CRISP-DM methodology has been widely adopted by researchers due to its systematic nature and adaptability for both business and social analytics (Sugiyono & Ruswandi, 2022). It consists of six iterative phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The conceptual workflow of this study follows the standard CRISP-DM cycle, as shown in Figure 1.



Figure 1. Stages of the CRISP-DM Framework

Source: Adapted from the Directorate General of State Assets (DJKN), Ministry of Finance, 2022.

The first phase, data collection, involved gathering public opinions from the social media platform X (formerly Twitter). The dataset was extracted using a Python-based scraping script executed in the Google Colab environment. Tweets written in Indonesian containing the hashtag #100HariPrabowoGibran were collected between 10 November 2024 and 1 March 2025, resulting in 2,718 raw tweets. Data extraction from X has been proven effective for political sentiment studies (Hidayat, 2022; Kahi & Talakua, 2022; Lestari & Mupaat, 2023), particularly when assessing public reactions to government policies (Khoirul & Saputri, 2021; Al-Areef & Saputra, 2023). The decision to focus on X aligns with prior research demonstrating its reliability as a real-time indicator of public mood and opinion (Putri & Forda, 2020; Syarifuddin, 2021). In the data preparation phase, preprocessing was performed to ensure model quality and accuracy. The process began with data cleansing to eliminate duplicate tweets, URLs, mentions (@username), hashtags, emojis, and irrelevant symbols. The resulting dataset was reduced to 431 valid tweets. Each tweet was then manually labeled as either *positive* or *negative*, a common approach in sentiment analysis studies where dataset sizes are moderate (Zulqornain & Indriati, 2020; Supriyanto & Alita, 2022). Text preprocessing included tokenization (splitting sentences into words), stopword removal (excluding common but semantically weak terms), and stemming (reducing words to their root form). These procedures were conducted using Python's *NLTK* library, following practices established in earlier studies by Pamungkas and Kharisudin (2020) and Nurtikasari and Alam (2022).

During the modeling phase, two algorithms—Naïve Bayes (NB) and K-Nearest Neighbor (KNN)—were applied and compared. Both models were trained and evaluated on the preprocessed dataset. The Naïve Bayes classifier, a probabilistic model based on Bayes' theorem, calculates class probabilities from the frequency distribution of words within the text (Nurtikasari & Alam, 2022; Ramadhani & Wahyudin, 2022). It is widely used for text classification due to its simplicity and strong generalization performance in sparse data environments (Sitanggang & Umaidah, 2024; Hakim & Sugiyono, 2024). In contrast, KNN classifies samples based on the majority label of the  $k$  nearest data points in a multidimensional space. Although conceptually straightforward, KNN's performance can be influenced by the "curse of dimensionality," where increasing vocabulary size dilutes distance meaning (Taufiqurrahman & Anggraeny, 2023; Pramana & Afdal, 2022). Several comparative studies, including those by Era and Andryana (2021) and Asro'i and Februariyanti (2022), have emphasized the trade-offs between NB's probabilistic precision and KNN's instance-based learning behavior. Both algorithms have been effectively applied in political, social, and entertainment-related sentiment analyses (Lestari & Mupaat, 2023; Al-Areef & Saputra, 2023).

The evaluation phase was conducted to measure each model’s performance. The metrics used were Accuracy, Precision, and Recall, which were calculated from a confusion matrix. Accuracy measures the overall proportion of correct predictions; precision assesses the proportion of true positives among predicted positives; and recall quantifies the proportion of actual positives correctly identified. These metrics are standard in text classification studies (Rahman & Lestari, 2021; Pamungkas & Kharisudin, 2020; Syarifuddin, 2021). The evaluation framework aligns with previous work that compared algorithmic performance under imbalanced data conditions, as documented by Supriyanto and Alita (2022) and Hakim and Sugiyono (2024). In this study, performance emphasis was placed on recall values to account for class imbalance between positive and negative sentiments, consistent with the recommendations of Era and Andryana (2021). The modeling and evaluation processes were implemented in the Google Colab environment using Python’s *scikit-learn* and *pandas* libraries for reproducibility. All parameters were optimized using cross-validation. The experiments concluded with a comparative analysis of both algorithms to determine which model performs more effectively in classifying public sentiment on the Prabowo–Gibran administration’s early performance. As noted by Al-Areef and Saputra (2023) and Lestari and Mupaat (2023), future studies could extend this research using more advanced deep learning architectures such as LSTM or BERT, which have shown promise in capturing contextual nuances and linguistic subtleties in Indonesian social media data.

### 3. Results

The analysis was conducted on a total of 431 valid tweets collected from the social media platform X (formerly Twitter) after undergoing a comprehensive preprocessing and manual labeling stage. The descriptive results reveal that public sentiment toward the first 100 days of the Prabowo–Gibran administration leaned predominantly negative. Out of the total dataset, 267 tweets (62%) were classified as negative sentiment, while 164 tweets (38%) reflected positive sentiment. This imbalance suggests a critical public perception during the early phase of governance, potentially driven by dissatisfaction or skepticism toward newly implemented policies. The overall distribution of sentiment is illustrated in Figure 3.1, which demonstrates the predominance of negative discourse across user-generated content on X.

Table View Plot View

accuracy: 97.22%

	true Negatif	true Positif	class precision
pred. Negatif	255	0	100.00%
pred. Positif	12	164	93.18%
class recall	95.51%	100.00%	

Figure 2. Public Sentiment Distribution on X Regarding the First 100 Days of the Prabowo–Gibran Administration

Sentiment	Count	Percentage
Positive	164	38%
Negative	267	62%
Total	431	100%

The next phase of the analysis involved evaluating the performance of two classification algorithms—Naïve Bayes (NB) and K-Nearest Neighbor (KNN)—using three standard performance metrics: accuracy, precision, and recall. The comparative

outcomes, summarized in Table 3.1, show that Naïve Bayes significantly outperformed KNN in all measured aspects. The Naïve Bayes model achieved an impressive accuracy rate of 97.22%, while KNN recorded only 62.65%. For the positive sentiment class, NB attained 93.18% precision and 100% recall, indicating perfect recognition of all positive tweets. For negative sentiment, the model achieved 100% precision and 95.51% recall, reflecting a balanced performance across both classes. Conversely, the KNN algorithm, although achieving perfect precision (100%) for positive sentiment, performed extremely poorly in recall (1.83%), successfully identifying only 3 out of 164 positive tweets. This indicates a strong bias toward the majority (negative) class, which undermines its reliability for imbalanced text datasets.

Table 1. Comparison of Classification Performance between Naïve Bayes and K-Nearest Neighbor (KNN)

Metric	Naïve Bayes (NB)	K-Nearest Neighbor (KNN)
Accuracy	97.22%	62.65%
Precision (Positive / Negative)	93.18% / 100%	100% / 62.38%
Recall (Positive / Negative)	100% / 95.51%	1.83% / 100%

Further analysis of the confusion matrix confirmed that the Naïve Bayes model maintained a strong balance between precision and recall across both sentiment classes. Its probabilistic nature allowed it to effectively capture emotionally charged words, such as “support,” “successful,” and “excellent” for positive sentiment, and “disappointed,” “failure,” and “controversy” for negative sentiment. The KNN model, on the other hand, exhibited unstable performance, primarily due to its sensitivity to high-dimensional text data and the uneven distribution of sentiment classes. This tendency to misclassify minority sentiment samples highlights the limitations of distance-based algorithms in sentiment analysis, especially when applied to short-text social media data. The comparative visualization of model accuracy is presented in Figure 3, where Naïve Bayes clearly demonstrates superior and consistent classification capability.

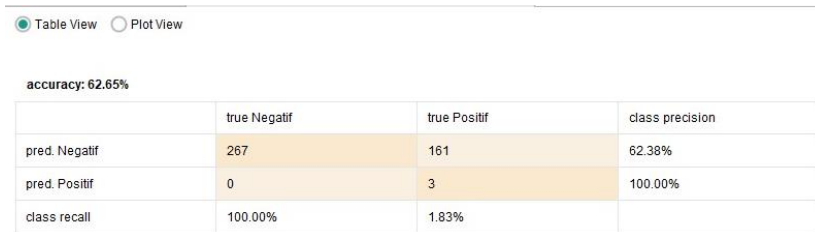


Figure 3. Model Accuracy Comparison between Naïve Bayes and KNN

Algorithm	Accuracy (%)	Precision (%)	Recall (%)
Naïve Bayes	97.22	96.59	97.75
K-Nearest Neighbor	62.65	81.19	50.92

Overall, the results demonstrate that Naïve Bayes is a more reliable and computationally efficient approach for sentiment classification on Indonesian political discourse datasets. Its robust generalization ability and stability in handling small, imbalanced datasets make it well-suited for real-time text mining tasks. These findings align with previous research emphasizing the effectiveness of probabilistic models over distance-based classifiers such as KNN in social media sentiment analysis (Era & Andryana, 2021; Rahman & Lestari, 2021; Hakim & Sugiyono, 2024). In contrast, the low accuracy and recall obtained by KNN reinforce the notion that the algorithm struggles with sparse, high-dimensional linguistic data (Syarifuddin, 2021;

Taufiqurrahman & Anggraeny, 2023). Empirically, this study confirms that Naïve Bayes remains one of the most effective methods for modeling and understanding public opinion in the context of digital political communication in Indonesia.

#### 4. Discussion

The findings of this study reveal a significant performance gap between the two classification algorithms tested—Naïve Bayes (NB) and K-Nearest Neighbor (KNN). The Naïve Bayes model achieved an accuracy of 97.22%, far surpassing KNN's 62.65%. This disparity provides empirical evidence that probabilistic approaches such as Naïve Bayes are more suitable for sentiment analysis of Indonesian-language social media data than distance-based methods like KNN, especially when dealing with unstructured, dynamic, and imbalanced datasets typical of online discourse. The superior performance of Naïve Bayes can be explained both theoretically and practically. Despite its “naïve” assumption that all features (words) are independent—a simplification that is linguistically unrealistic—the model remains highly effective for text classification (Era & Andryana, 2021; Ramadhani & Wahyudin, 2022). Its strength lies in leveraging the probability of word occurrence within each sentiment class. In practice, highly polarized words such as “*berhasil*” (successful), “*dukung*” (support), and “*hebat*” (great) frequently appear in positive sentiments, whereas “*gagal*” (failure), “*kecewa*” (disappointed), and “*kontroversi*” (controversy) tend to dominate negative expressions. This probabilistic association is often sufficient for achieving high accuracy in binary sentiment tasks, even without modeling complex linguistic dependencies. Similar conclusions have been reported by Putri and Forda (2020), Hidayat (2022), and Sitanggang and Umaidah (2024), all of whom confirmed Naïve Bayes as a robust method for analyzing public opinion on X (formerly Twitter) in the Indonesian context.

By contrast, the poor performance of K-Nearest Neighbor can be attributed to two fundamental limitations. First, KNN is highly susceptible to the curse of dimensionality, wherein each unique word in the corpus is treated as an independent dimension. In textual datasets, this results in an extremely sparse and high-dimensional feature space, rendering the concept of “distance” meaningless (Sugiyono & Ruswandi, 2022; Syarifuddin, 2021). Consequently, similarity-based classification becomes unreliable. Second, KNN performs poorly when applied to imbalanced datasets, such as the one used in this study (62% negative vs. 38% positive). Because KNN determines the class of a new sample based on the majority of its nearest neighbors, it inherently favors the dominant class. This explains the model's exceptionally low recall rate for positive sentiment (1.83%), which indicates that the classifier almost always predicts “negative.” Similar issues with class imbalance and bias in KNN have been documented by Asro'i and Februariyanti (2022) and Supriyanto and Alita (2022), who reported drastic performance degradation in unevenly distributed text datasets. The implications of these findings are both practical and theoretical. Practically, the results demonstrate that Naïve Bayes is a reliable baseline model for the rapid and accurate analysis of political sentiment on X. Its computational efficiency and consistent accuracy make it well-suited for institutions seeking to monitor public opinion in real time, such as policy research centers, media analytics firms, and government communication teams. Theoretically, this study reinforces the longstanding view that probabilistic models remain effective for linguistic data classification, particularly in socio-political contexts. Similar conclusions were drawn by Rahman and Lestari (2021), Hakim and Sugiyono (2024), and Pramana and Afdal (2022), who each found Naïve Bayes to outperform KNN and other traditional classifiers in Indonesian social media sentiment analysis. Moreover, the results complement studies by Pamungkas and Kharisudin (2020) and Al-Areef and Saputra (2023), which emphasize that while

newer models such as SVM or LSTM offer advanced learning capabilities, Naïve Bayes still provides a strong and interpretable foundation for sentiment classification tasks.

Nonetheless, this study is not without methodological limitations. The manual sentiment labeling process, while improving precision, introduces potential subjective bias and is difficult to scale. Additionally, the dataset was limited to a single hashtag—#100HariPrabowoGibran—which may not fully represent the breadth of public discourse surrounding the administration. The nuances of social media language, including sarcasm, irony, and contextual ambiguity, also pose challenges for keyword-based classifiers (Kahi & Talakua, 2022; Lestari & Mupaat, 2023). Future studies should address these constraints by incorporating more advanced feature representation techniques, such as *Term Frequency–Inverse Document Frequency (TF-IDF)* or *word embeddings* (e.g., Word2Vec or FastText), to capture richer semantic relationships between words (Nurtikasari & Alam, 2022; Zulqornain & Indriati, 2020). Furthermore, the application of deep learning models such as Long Short-Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT)—which have demonstrated strong contextual learning capabilities in prior research (Al-Areef & Saputra, 2023)—could yield deeper insights into the structure of public opinion and improve classification accuracy in Indonesian political sentiment analysis. Overall, the present study underscores Naïve Bayes as an effective and computationally efficient model for sentiment analysis while highlighting the potential of integrating it with modern neural architectures for more nuanced and scalable applications in digital political communication research.

## 5. Conclusion

This study provides a comprehensive empirical overview of public sentiment toward the first 100 days of the Prabowo–Gibran administration and evaluates the effectiveness of two classification algorithms in analyzing social media discourse. Two key conclusions can be drawn. First, the sentiment analysis indicates that public opinion on X (formerly Twitter) was largely negative (62%), reflecting a critical perception among users toward the government's early performance. Second, from a methodological standpoint, the Naïve Bayes algorithm demonstrated superior accuracy (97.22%) compared to K-Nearest Neighbor (62.65%), confirming that probabilistic models remain highly effective for handling high-dimensional and imbalanced Indonesian-language text data. The principal contribution of this study lies in establishing a methodological framework that is efficient, interpretable, and computationally practical for political sentiment analysis in Indonesia. The findings reaffirm that classical probabilistic approaches such as Naïve Bayes continue to serve as reliable tools for government agencies, policy analysts, and civil society organizations seeking to monitor public opinion rapidly and accurately without requiring extensive computational infrastructure. Moreover, the results highlight that statistically grounded models can still yield competitive performance compared to more complex algorithms, particularly when analyzing social media data characterized by linguistic diversity and informality.

Nevertheless, this study acknowledges several limitations. Manual sentiment labeling, while precise, introduces the risk of subjective bias and lacks scalability for larger datasets. Furthermore, the reliance on a single hashtag limits the breadth of captured public opinion, while the models used in this study remain insufficient in addressing linguistic nuances such as sarcasm, irony, and contextual ambiguity. Future research should focus on three key directions: (1) developing hybrid models that integrate deep learning techniques for feature extraction with classical classifiers such as Naïve Bayes; (2) implementing aspect-based sentiment analysis to identify

specific policy areas or issues driving positive or negative opinions; and (3) expanding the analytical scope across multiple social media platforms to obtain a more holistic and representative understanding of public sentiment in the digital landscape. In summary, this research reaffirms Naïve Bayes as a strong, efficient, and contextually relevant baseline model for political sentiment analysis in Indonesia, while paving the way for future exploration into more adaptive approaches capable of capturing the linguistic and emotional complexity of digital communication in contemporary society.

## Acknowledgment

The author would like to express sincere gratitude to all individuals and institutions who have provided essential support in the completion of this research and the writing of this manuscript. The highest appreciation is first extended to Mr. Very Arinal, M.Kom., Head of the Information Systems Study Program, and Mr. Sugiyono, M.Kom., the academic supervisor, for their invaluable guidance, direction, and time throughout the research process. The author also wishes to thank the entire management team and faculty members of Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika (STIKOM CKI) for the knowledge, mentorship, and academic facilities provided. Heartfelt gratitude is also conveyed to the author's beloved family, especially both parents, whose prayers, sacrifices, and moral support have been the greatest source of strength. Special thanks are dedicated to the author's dearest sibling, Richa Aulia, for her constant encouragement and assistance in overcoming every challenge. Finally, the author extends appreciation to fellow students, alumni, members of Mapala Matras, and colleagues from BEM STIKOM CKI 2024/2025, as well as all others who cannot be mentioned individually, for their friendship, collaboration, and support throughout the academic journey and during the completion of this research.

## References

- Al-Areef, M. H., & Saputra, S. K. (2023). Analisis sentimen pengguna Twitter mengenai calon presiden Indonesia tahun 2024 menggunakan algoritma LSTM. *Jurnal SAINTIKOM*, 22(1), 1–10.
- Asro'i, A., & Februariyanti, H. (2022). Analisis sentimen pengguna Twitter terhadap perpanjangan PPKM menggunakan metode K-Nearest Neighbor. *Jurnal Khatulistiwa Informatika*, 10(1), 25–32.
- Era, D., & Andryana, S. (2021). Perbandingan algoritma Naïve Bayes dan K-Nearest Neighbor pada analisis sentimen pembukaan pariwisata di masa pandemi COVID-19. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 856–865.
- Hakim, Z. R., & Sugiyono. (2024). Analisa sentimen terhadap Kereta Cepat Jakarta–Bandung menggunakan algoritma Naïve Bayes dan K-Nearest Neighbor. *Jurnal Sains dan Teknologi*, 5(3), 1–10.
- Hidayat, S. (2022). Analisa sentimen masyarakat Indonesia terhadap kinerja Prabowo Subianto sebagai anggota Kabinet Indonesia Maju di platform Twitter menggunakan R dengan algoritma Naïve Bayes. *JUNIF: Jurnal Nasional Informatika*, 3(1), 1–8.

- Kahi, F. R. B., & Talakua, A. C. (2022). Analisis sentimen masyarakat di Twitter terhadap pemerintahan Anies Baswedan menggunakan metode Naïve Bayes classifier. *Jurnal Minfo Polgan*, 11(2), 536–543.
- Khoirul, & Saputri, N. A. O. (2021). Analisis sentimen masyarakat terhadap Pilpres 2019 berdasarkan opini dari Twitter menggunakan metode Naïve Bayes classifier. *Journal of Computer and Information Systems Ampera*, 2(1), 28–36.
- Lestari, S., & Mupaat, M. (2023). Analisis sentimen masyarakat Indonesia terhadap pemindahan Ibu Kota Negara Indonesia pada Twitter. *Jurnal KomtekInfo*, 10(1), 78–87.
- Nurtikasari, Y., & Alam, S. (2022). Analisis sentimen opini masyarakat terhadap film pada platform Twitter menggunakan algoritma Naïve Bayes. *INSOLOGI: Jurnal Sains dan Teknologi*, 1(4), 455–462.
- Pamungkas, F. S., & Kharisudin, I. (2020). Analisis sentimen dengan SVM, Naïve Bayes, dan KNN untuk studi tanggapan masyarakat Indonesia terhadap pandemi COVID-19 pada media sosial Twitter. *PRISMA: Prosiding Seminar Nasional Matematika*, 3, 533–542.
- Pramana, D., & Afdal, M. (2022). Analisis sentimen terhadap pemindahan Ibu Kota Negara menggunakan algoritma Naïve Bayes classifier dan K-Nearest Neighbors (KNN). *Jurnal Media Informatika Budidarma*, 6(2), 991–997.\*  
<https://doi.org/10.30865/mib.v6i2.3855>
- Putri, D. D., & Forda, G. (2020). Analisis sentimen kinerja Dewan Perwakilan Rakyat (DPR) pada Twitter menggunakan metode Naïve Bayes classifier. *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, 9(1), 1–7.
- Rahman, F. N., & Lestari, S. (2021). Analisis sentimen masyarakat terhadap pemerintah di era Kabinet Joko Widodo berdasarkan sosial media X menggunakan Naïve Bayes dan K-Nearest Neighbor (KNN). *Journal of Information Technology and Computer Science (INTECOMS)*, 4(2), 164–173.
- Ramadhani, S. H., & Wahyudin, M. I. (2022). Analisis sentimen terhadap vaksinasi AstraZeneca pada Twitter menggunakan metode Naïve Bayes dan K-NN. *JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, 6(4), 530–538.\*  
<https://doi.org/10.35870/jtik.v6i4.530>
- Sitanggang, A., & Umidah, Y. (2024). Analisis sentimen masyarakat terhadap program makan siang gratis pada media sosial X menggunakan algoritma Naïve Bayes. *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, 12(3).\*  
<https://doi.org/10.23960/jitet.v12i3.4902>
- Sugiyono, & Ruswandi, M. (2022). Pemodelan pengolahan citra untuk klasifikasi jenis buah pisang menggunakan metode KNN. *Jurnal Pendidikan dan Konseling*, 4(6), 987–995.
- Supriyanto, J., & Alita, D. (2022). Penerapan algoritma K-Nearest Neighbor (K-NN) untuk analisis sentimen publik terhadap pembelajaran daring. *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, 3(1), 84–93.

- Syarifuddin, M. (2021). Analisis sentimen opini publik terhadap efek PSBB pada Twitter dengan algoritma Decision Tree–KNN–Naïve Bayes. *Jurnal Penelitian Teknologi Informasi*, 15(1).\* <https://doi.org/10.33480/inti.v15i1.1433>
- Taufiqurrahman, H., & Anggraeny, F. T. (2023). Perbandingan algoritma Naïve Bayes dan K-Nearest Neighbor pada analisis sentimen ulasan aplikasi MyPertamina. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 2561–2569.
- Zulqornain, J. A., & Indriati. (2020). Analisis sentimen tanggapan masyarakat aplikasi TikTok menggunakan metode Naïve Bayes dan Categorical Proportional Difference (CPD). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 4(10), 3568–3575.