

Research Horizon

ISSN: 2808-0696 (p), 2807-9531 (e)

Research Horizon

Volume: 05

Issue: 04

Year: 2025

Page: 1415–1424

Citation:

Febrilliani, J. S., & Wibowo, A. (2025). Twitter public sentiment towards vasectomy in Indonesia using SVM and Naïve Bayes. *Research Horizon*, 5(4), 1415–1424.

Article History:

Received: June 20, 2025

Revised: July 16, 2025

Accepted: July 29, 2025

Online since: August 30, 2025

Twitter Public Sentiment towards Vasectomy in Indonesia Using SVM and Naïve Bayes

Jihan Sastri Febrilliani^{1*}, Arief Wibowo¹

¹ Universitas Budi Luhur, Jakarta, Indonesia

* Corresponding author: Jihan Sastri Febrilliani (22411600147@student.budiluhur.co.id)

Abstract

Social media has experienced rapid growth in various sectors, including the health sector. Active users of social media or often called netizens, are quite active in the health community by sharing information and experiences. access to social media is the main means of finding information related to health. This study aims to find out public sentiment on the issue of vasectomy using the SVM and Naïve Bayes methods. This research collects data from social media platforms (X) or Twitter, using several keywords related to vasectomy which has recently become a hot topic in the Indonesian community. The results of this study showed that using the Support Vector Machine (SVM) algorithm and Naïve Bayes as a comparison obtained an accuracy performance of 71% for the support vector machine method and 51% for the naïve bayes method. The findings provide insights for policymakers and health advocates on addressing misconceptions and stigma surrounding vasectomy. The findings suggest to develop a more comprehensive Indonesian sentiment dictionary, use multi-platform data, and apply advanced NLP techniques to better capture cultural and contextual nuances.

Keywords

Sentiment Analysis, Support Vector Machine, Twitter, Social Media, Vasectomy.

1. Introduction

The vasectomy program is a key male contraception method aimed at controlling population growth and enhancing family life quality. In Indonesia, as in many countries, this program is actively promoted within the Family Planning (*Keluarga Berencana*/KB) strategy (Saputri et al., 2024). Public figures, such as the Governor of West Java, Dedi Mulyadi, have championed this initiative, leveraging their influence to raise awareness, despite their controversial approaches (Fahrudin et al., 2025). However, public acceptance of vasectomy remains hindered by challenges such as limited understanding, prevalent myths, and deep-rooted social stigma (Fahrudin et al., 2025). These barriers are particularly pronounced in Indonesia, where cultural and religious sensitivities often shape public perceptions of reproductive health interventions (Naik et al., 2025).

The rise of internet technology has transformed social media, particularly Twitter, into a primary platform for public discourse (Ryandi et al., 2025). Twitter's significance lies in its ability to facilitate real-time interactions and shape public opinion on various issues, including health-related topics (Aattouchi et al., 2021). With 540 million active users globally, Twitter serves as a critical space for campaigns and discussions on government programs in Indonesia, including health initiatives like vasectomy (Rodríguez-Ibáñez et al., 2023). Social media is no longer just an entertainment medium but a dynamic public space for exchanging ideas and information (Thelwall et al., 2012). In Indonesia, with 79.5% internet penetration as reported by Indonesian Internet Service Providers Association (*Asosiasi Penyelenggara Jasa Internet Indonesia*/APJII) chairman Muhammad Arif, social media platforms like Twitter have become vital for gauging public sentiment on sensitive issues such as male contraception (Wantika et al., 2021; Pradana, 2024).

Sentiment analysis, a process to determine whether a text expresses positive or negative opinions, is increasingly used to analyze public reactions to specific topics (Ananda et al., 2021; Romaito, 2021; Wankhade et al., 2022; Pertiwi et al., 2024). According to Kurniawan and Susanto (2019), sentiment analysis on Twitter can effectively capture public perceptions of controversial policies. Yet, few studies have applied this approach to reproductive health issues like vasectomy in the Indonesian context. This study addresses this research gap by analyzing Twitter opinions on vasectomy using advanced machine learning methods, specifically Support Vector Machine (SVM) and Naïve Bayes. While prior research, such as that by Darwis et al. (2020), has explored sentiment analysis for institutional performance, the application of these methods to male contraception policies remains underexplored, particularly in addressing cultural and social stigmas. SVM is employed for its robust classification capabilities, finding optimal boundaries to separate positive and negative sentiments, while Naïve Bayes serves as a comparative method using probabilistic word occurrence (Artanti et al., 2018; Sautomo et al., 2020; Samsir et al., 2021). Additionally, the integration of artificial intelligence in reproductive health, as highlighted by Medenica et al. (2022), underscores the potential of machine learning to enhance understanding of public attitudes toward interventions like vasectomy, yet such applications are scarce in Indonesia.

The objective of this research is to evaluate public sentiment on Twitter regarding vasectomy policies in Indonesia, particularly those promoted by Dedi Mulyadi, using SVM and Naïve Bayes methods. By collecting and analyzing 1008 tweets with the keyword "vasectomy," this study aims to provide a comprehensive understanding of public responses, categorized into positive, negative, and neutral sentiments (Buntoro, 2016; Widowati & Sadikin, 2020; Idris et al., 2023). The use of the Term Frequency-Inverse Document Frequency (TF-IDF) method for word weighting further enhances the analysis by prioritizing significant terms (Septian et al., 2019). This research seeks to acceptance and into public perceptions of vasectomy,

addressing the cultural and social barriers to its acceptance, and contributing to the broader discourse on leveraging social media for health policy evaluation. The findings are expected to inform policymakers and health advocates on strategies to improve public acceptance of male contraception in Indonesia.

2. Literature Review

2.1. Social Media and Public Sentiment Analysis

The rise of social media platforms, particularly Twitter, has transformed how public opinions are expressed and analyzed, making them valuable for sentiment analysis studies. Thelwall et al. (2012) highlighted that platforms like Twitter amplify emotional and polarized responses, enabling researchers to capture real-time public sentiments on various issues, including health policies. With 540 million active users globally, Twitter serves as a dynamic space for discourse, especially in Indonesia, where 79.5% internet penetration facilitates widespread engagement (Rodríguez-Ibáñez et al., 2023).

Sentiment analysis, as defined by Wankhade et al. (2022), involves determining whether a text expresses positive, negative, or neutral opinions, offering insights into public perceptions of controversial topics like vasectomy. In Indonesia, social media is a primary channel for health-related discussions, with netizens actively sharing experiences and information (Wantika et al., 2021). However, Kurniawan and Susanto (2019) noted that sentiment analysis on Twitter faces challenges due to non-standard language and slang, which complicates accurate classification in culturally diverse contexts like Indonesia. Their study on the 2019 presidential election demonstrated that Twitter data can effectively capture public reactions but requires robust preprocessing to handle linguistic variations. This underscores the need for tailored sentiment lexicons to improve analysis accuracy in Indonesian social media studies, particularly for sensitive health topics.

2.2. Machine Learning Methods for Sentiment Analysis

Machine learning techniques, such as Support Vector Machine (SVM) and Naïve Bayes, are widely used for sentiment classification due to their ability to process large datasets. Alhaq et al. (2021) applied SVM to analyze Twitter sentiments toward the Bukalapak marketplace, achieving an accuracy of 82%, which highlights SVM's strength in distinguishing sentiment classes by maximizing hyperplane margins. Similarly, Darwis et al. (2020) used SVM with Term Frequency-Inverse Document Frequency (TF-IDF) to classify sentiments about Indonesia's anti-corruption agency, reporting 77% negative sentiment, which suggests SVM's suitability for controversial policy analysis.

Naïve Bayes, while computationally efficient, often yields lower accuracy due to its reliance on probabilistic assumptions, as noted by Romaito et al. (2021), who compared SVM (80.7% accuracy) and Naïve Bayes (81.7%) in election sentiment analysis. Koto and Rahmaningtyas (2017) emphasized that lexicon-based approaches, like the InSet lexicon used in this study, enhance sentiment labeling but struggle with Indonesian slang, necessitating advanced preprocessing. These studies collectively indicate that SVM outperforms Naïve Bayes in complex datasets, but both methods require careful feature engineering to address cultural and linguistic nuances in Indonesian Twitter data.

2.3. Vasectomy and Public Perception in Indonesia

Vasectomy, as a male contraception method, faces significant cultural and social barriers in Indonesia, influencing public sentiment. Fahrudin et al. (2025) highlighted that public figures like Dedi Mulyadi promote vasectomy within Indonesia's Family Planning program, yet their controversial approaches often polarize opinions, as seen in Twitter discussions. Cultural stigmas and myths, such

as fears of reduced masculinity, hinder acceptance, as noted by Saputri et al. (2024), who identified determinants like limited understanding and religious sensitivities as key barriers.

Naik et al. (2025) argued that artificial intelligence, including sentiment analysis, can uncover public attitudes toward reproductive health interventions, yet such applications are scarce in Indonesia. Their scoping review emphasized the need for culturally sensitive approaches to address misconceptions. Social media’s role in shaping perceptions is critical, as Ryandi et al. (2025) found in their analysis of health ministry sentiments, where negative sentiments dominated due to cultural factors. These studies suggest that understanding public sentiment on vasectomy requires integrating machine learning with culturally tailored lexicons to capture Indonesia’s diverse linguistic and social landscape, providing a foundation for this study’s analysis.

3. Methods

This research was conducted through a systematic process to analyze public sentiment on Twitter regarding vasectomy, utilizing data collected from the platform. Data collection was performed by crawling tweets using the Python programming language on Google Colab, targeting the keyword “vasectomy” to gather 1,008 tweets. The process involved writing code with the assistance of the Node.js library in Python, incorporating an authentication token to access Twitter’s API, ensuring precise and relevant data retrieval based on the specified keyword. This approach enabled efficient extraction of real-time public opinions expressed on Twitter, forming the foundation for subsequent analysis.

Following data collection, text preprocessing was applied to clean and organize the data for analysis. This stage included several steps: removing special characters, replacing repetitive sequences, case folding, tokenizing, normalizing, removing stop words, and stemming to ensure the data was structured and free of noise. These preprocessing techniques transformed the raw tweet data into a format suitable for sentiment analysis, addressing issues such as non-standard language and slang commonly found in Indonesian tweets. The processed data was then labeled using the InSet lexicon approach, which facilitates automatic and rapid sentiment labeling by assigning weights to words based on a dictionary ranging from -5 to +5, covering positive and negative sentiments, emoticons, phrases, and booster words. This lexicon, containing 10,251 entries, was refined through stemming and synonym-based weighting to enhance labeling accuracy.

The sentiment analysis model was developed using the Support Vector Machine (SVM) method, employing k-fold cross-validation to eliminate the need for manual data splitting, ensuring robust testing of the model’s predictive accuracy. A data split ratio of 80:20 for training and testing was maintained to evaluate performance effectively.

$$Accuracy = \frac{TP + TN}{P + TN + FP + FN} \times 100\%$$

The model’s accuracy was assessed using a confusion matrix, which calculated metrics such as true positives, true negatives, false positives, and false negatives to determine precision, recall, and overall accuracy (García-Díaz et al., 2022). This evaluation provided a reliable measure of the model’s ability to classify sentiments correctly, with Naïve Bayes used as a comparative method to benchmark SVM’s performance. The integration of Term Frequency-Inverse Document Frequency (TF-IDF) for word weighting further enhanced the model by emphasizing significant terms within the dataset. This cohesive methodology ensured a

comprehensive analysis of Twitter sentiments toward vasectomy, capturing public perceptions with high precision.

4. Results

This study presents the findings from the sentiment analysis of Twitter data regarding public opinions on vasectomy in Indonesia, using Support Vector Machine (SVM) and Naïve Bayes methods. The results include the sentiment distribution, model performance metrics, and key terms identified in the dataset, derived from 1008 tweets collected using the keyword “vasectomy” included the non-standard language and slang commonly found in Indonesian tweets (Nugroho, 2016). All data were processed through text preprocessing, which involved multiple steps such as eliminating special characters, condensing repeated sequences, standardizing text case, tokenizing, normalizing, removing stop words, and applying stemming to ensure the dataset was well-structured and noise-free (Riyaddulloh et al., 2021). The tweets were then labeled using the InSet lexicon and analyzed with SVM and Naïve Bayes algorithms, with Term Frequency-Inverse Document Frequency (TF-IDF) applied for word weighting (Koto & Rahمانingtyas, 2017; Septian et al., 2019). To improve labeling accuracy, this lexicon, which contained 10,251 entries, was refined through stemming and synonym-based weighting (Abdillah et al., 2021). The findings are reported factually, with tables summarizing the sentiment distribution and performance metrics of the models.

Table 1. Results of Sentiment on Vasectomy

Sentiment	Percentage
Positive	42%
Negative	49.5%
Neutral	8.5%

The dataset of 1008 tweets was categorized into three sentiment groups: positive, negative, and neutral, based on the InSet lexicon labeling process. Table 1 presents the distribution of sentiments. The results showed that 424 tweets (42%) were classified as positive, 499 tweets (49.5%) as negative, and 86 tweets (8.5%) as neutral. The negative sentiment category had the highest representation, followed by positive, with neutral sentiments being the least frequent. The tweets were collected using Python-based crawling on Google Colab, targeting the keyword “vasectomy,” and preprocessed to remove noise such as special characters, non-standard language, and slang. The processed dataset served as the basis for subsequent classification using the SVM and Naïve Bayes models.

The SVM model was implemented with k-fold cross-validation to automate the division of training and testing data. The model achieved an overall accuracy of 71%, as determined by a confusion matrix evaluating True Positive, True Negative, False Positive, and False Negative predictions. Table 2 provides the detailed performance metrics for the SVM model across the three sentiment categories. For the positive sentiment category, the precision was 0.70, recall was 0.82, F1-score was 0.76, and support was 499. For the negative sentiment category, the precision was 0.85, recall was 0.07, F1-score was 0.12, and support was 86. For the neutral sentiment category, the precision was 0.71, recall was 0.71, F1-score was 0.71, and support was 424. These metrics indicate the model’s performance in classifying each sentiment category, with the support values reflecting the number of tweets in each group.

Table 2. The Results of Bayes' Naïve Calculations

Sentiment	Precision	Recall	F1 Score	Support
Negatif	0.52	0.92	0.67	90
Neutral	0.00	0.00	0.00	22
Positive	0.77	0.37	0.50	90
Accuracy	0.57			

The Naïve Bayes model, used as a comparative method, employed a 80:20 data split ratio for training and testing. The overall accuracy of the Naïve Bayes model was 51%, as calculated using a confusion matrix. Table 2 summarizes the performance metrics for the Naïve Bayes model. For the negative sentiment category, the precision was 0.52, the recall was 0.92, the F1-score was 0.67, and the support was 90. For the neutral sentiment category, the precision was 0.00, the recall was 0.00, the F1-score was 0.00, and the support was 22. For the positive sentiment category, the precision was 0.77, the recall was 0.37, the F1-score was 0.50, and the support was 90. The support values indicate the number of tweets classified in each sentiment category, with the neutral category having the smallest representation.

During the analysis, key terms frequently appearing in the tweets were identified through the TF-IDF weighting process. The most common terms included “vasectomy,” “talking,” “badung,” “keinget,” “same,” “people,” “friends,” “facebook,” “sharia,” and “hate.” These terms were extracted from the preprocessed dataset and reflect the vocabulary used in Twitter discussions about vasectomy. The TF-IDF method prioritized terms with higher significance in the context of the dataset, providing a basis for the sentiment classification performed by both models.

The data collection process involved crawling 1,008 tweets using the Node.js library in Python, along with an authentication token, to access Twitter’s API. Preprocessing steps included removing special characters, case folding, tokenizing, normalizing, removing stop words, and stemming. The InSet lexicon, containing 10,251 entries, was used to assign sentiment weights ranging from -5 to +5, covering positive and negative sentiments, emoticons, phrases, and booster words. The lexicon was refined through stemming and synonym-based weighting to ensure accurate labeling of the tweets.

The SVM model’s k-fold cross-validation approach eliminated the need for manual data splitting, ensuring robust testing across multiple folds. The 80:20 train-test split used for the Naïve Bayes model was applied consistently to evaluate its performance. Both models were assessed using precision, recall, F1-score, and accuracy metrics, with the confusion matrix providing a detailed breakdown of classification outcomes. The SVM model’s higher accuracy of 71% compared to the Naïve Bayes model’s 51% was derived from these evaluations.

The results also highlight the distribution of tweets across sentiment categories, with the negative sentiment category comprising the most significant share at 49.5%, followed by the positive sentiment category at 42%, and the neutral sentiment category at 8.5%, as shown in Table 1. The performance metrics in Table 2 provide a clear comparison of the two models, with SVM demonstrating higher precision, recall, and F1-scores across most categories, except for the negative sentiment category in the Naïve Bayes model, which had a higher recall of 0.92. The neutral category in the Naïve Bayes model showed zero precision, recall, and F1-score, indicating challenges in classifying this category.

The dataset’s composition and the models’ performance metrics provide a factual basis for understanding public sentiment toward vasectomy on Twitter. The identified key terms and their TF-IDF weights offer additional insight into the language used in these discussions.

5. Discussion

The analysis of Twitter sentiment toward vasectomy in Indonesia, as presented in the Results section, reveals a predominantly negative public perception, with 49.5% of the 1008 tweets classified as negative, 42% as positive, and 8.5% as neutral. This distribution suggests significant public skepticism, likely rooted in cultural and social factors, as vasectomy challenges traditional norms around male reproductive roles in Indonesia. The Support Vector Machine (SVM) model's 71% accuracy, compared to the Naïve Bayes model's 51%, indicates that SVM better captures the nuances of sentiment classification in this context. The frequent appearance of terms like "sharia" and "hate" in the TF-IDF analysis points to religious and emotional undertones in the discourse, reflecting the sensitivity of vasectomy as a topic in a culturally diverse society.

The higher accuracy of SVM aligns with findings by Alhaq et al. (2021), who reported an 82% accuracy using SVM for sentiment analysis of Twitter data on the Bukalapak marketplace, highlighting SVM's robustness in handling complex datasets. The low recall for negative sentiments in the SVM model and the zero precision for neutral sentiments in Naïve Bayes suggest challenges in classifying minority classes, possibly due to imbalanced data or the InSet lexicon's limitations in capturing Indonesian slang and cultural nuances. Koto and Rahmaningtyas (2017) noted that lexicon-based approaches like InSet may struggle with non-standard language, which is prevalent in Indonesian social media. The presence of slang terms like "keinget" and "badung" in the dataset underscores this issue, as these terms may not be adequately weighted in the lexicon, affecting classification accuracy.

The negative sentiment dominance (49.5%) resonates with Darwis et al. (2020), who found 77% negative sentiment in Twitter analysis of Indonesia's anti-corruption agency, suggesting that controversial public policies often elicit strong negative reactions online. The neutral sentiment's low representation (8.5%) may indicate polarized opinions, with users either strongly supporting or opposing vasectomy, leaving little room for ambivalence. This polarization could stem from cultural stigmas and myths surrounding vasectomy, as noted by Fahrudin et al. (2025), who highlighted the role of public figures like Dedi Mulyadi in shaping discourse, often in a controversial manner. The SVM model's ability to maximize hyperplane margins, as described by Laurensz and Sedyono (2021), likely contributed to its superior performance over Naïve Bayes, which relies on probabilistic assumptions less suited to complex sentiment patterns.

The findings also connect to broader theories of public opinion formation on social media. Thelwall et al. (2012) argued that platforms like Twitter amplify emotional and polarized responses, which may explain the high negative sentiment in this study. The low F1-score for neutral sentiments in Naïve Bayes (0.00, Table 3) further suggests that probabilistic models struggle with less distinct categories in emotionally charged datasets. The TF-IDF weighting's emphasis on terms like "hate" and "sharia" indicates that religious and cultural contexts heavily influence public sentiment, aligning with Naik et al. (2025) observation that cultural sensitivities shape reproductive health perceptions in Indonesia.

The implications of these findings are significant for health policymakers and advocates. The predominant negative sentiment suggests a need for targeted public education campaigns to address misconceptions and stigma surrounding vasectomy, particularly in culturally sensitive regions. By leveraging influential figures and social media platforms, as Fahrudin et al. (2025) suggest, policymakers can counteract negative narratives. Future research should focus on developing a more comprehensive Indonesian sentiment lexicon and incorporating multi-platform data to enhance the accuracy of sentiment analysis. Advanced natural language processing techniques, such as deep learning, can better capture contextual nuances,

thereby improving sentiment classification and informing strategies to enhance vasectomy acceptance in Indonesia.

6. Conclusion

This study analyzed public sentiment on Twitter regarding the vasectomy policy in Indonesia, particularly promoted by Dedi Mulyadi, using Support Vector Machine (SVM) and Naïve Bayes methods. The lexicon-based labeling revealed a predominantly negative sentiment, with 499 negative, 424 positive, and 86 neutral tweets out of 1,008 collected. The SVM model, combined with TF-IDF feature selection, achieved a promising accuracy of 71%, outperforming the Naïve Bayes model, which yielded an accuracy of 51%. These results indicate that SVM is more effective in classifying sentiments in this context, reflecting significant public skepticism toward vasectomy, likely influenced by cultural and social factors.

The findings offer valuable insights for policymakers and health advocates, highlighting the need to address public misconceptions and stigma surrounding vasectomy to improve its acceptance. However, the study faced limitations, including an imbalanced dataset, a constrained lexicon dictionary that may not fully capture Indonesia's diverse linguistic nuances, and a lack of cultural context in sentiment analysis. For future research, developing a more comprehensive sentiment dictionary tailored to Indonesian social media and incorporating multi-platform data could enhance accuracy and generalizability. Additionally, exploring advanced natural language processing techniques, such as deep learning, may provide a more accurate account of cultural and contextual factors in sentiment classification.

References

- Aattouchi, I., Elmendili, S., & Elmendili, F. (2021). Sentiment analysis of health care: Review. *Journal of Physics: Conference Series*, 1743(1), 120-135.
- Abdillah, W. F., Premana, A., & Bhakti, R. M. H. (2021). Analisis sentimen penanganan Covid-19 dengan Support Vector Machine: Evaluasi leksikon dan metode ekstraksi fitur. *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, 3(2), 160-170.
- Alhaq, Z., Mustopa, A., Mulyatun, S., & Santoso, J. D. (2021). Penerapan metode Support Vector Machine untuk analisis sentimen pengguna Twitter. *Journal of Information System Management (JOISM)*, 3(1), 16-21.
- Ananda, F. D., & Pristyanto, Y. (2021). Analisis sentimen pengguna Twitter terhadap layanan internet provider menggunakan algoritma Support Vector Machine. *Matrik: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 20(2), 407-416.
- Artanti, D. P., Muljono, M., Syukur, A., Prihandono, A., & Setiadi, D. R. M. (2018). Analisa sentimen untuk penilaian pelayanan situs belanja online menggunakan algoritma Naïve Bayes. *Konferensi Nasional Sistem Informasi (KNSI) 2018*, 8-9.
- Buntoro, G. A. (2016). Analisis sentimen hatespeech pada Twitter dengan metode Naïve Bayes Classifier dan Support Vector Machine. *Jurnal Dinamika Informatika*, 5(2), 1-12.
- Darwis, D., Pratiwi, E. S., & Pasaribu, A. F. O. (2020). Penerapan algoritma SVM untuk analisis sentimen pada data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. *Jurnal Ilmiah Educat: Pendidikan dan Informatika*, 7(1), 1-11.
- Fahrudin, A., Lisnarini, N., Dewi, G. K., Pangandaran, A. B. K., & Padjadjaran, U. (2025). Populisme kebijakan Gubernur Jawa Barat Dedi Mulyadi (studi analisis sentimen). *Massive: Jurnal Ilmu Komunikasi*, 5(1), 17-31.
- Fitriyana, V., Hakim, L., Novitasari, D. C. R., & Asyhar, A. H. (2023). Analisis sentimen ulasan aplikasi Jamsostek Mobile menggunakan metode Support Vector Machine. *Jurnal Buana Informatika*, 14(1), 40-49.
- García Díaz, P., Sánchez Berriel, I., Pontiel Martín, D., & González Ávila, J. L. (2022). A novel flexible feature extraction algorithm for Spanish tweet sentiment analysis based on the context of words. *Expert Systems with Applications*, 212(3), 808-817.

- Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis sentimen terhadap penggunaan aplikasi Shopee menggunakan algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi dan Komunikasi*, 5(1), 32–35.
- Koto, F., & Rahmaningtyas, G. Y. (2017). Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs. *Proceedings of the 2017 International Conference on Asian Language Processing, IALP 2017, 2018-January*, 391–394.
- Kurniawan, I., & Susanto, A. (2019). Implementasi metode K-Means dan Naïve Bayes Classifier untuk analisis sentimen pemilihan presiden (pilpres) 2019. *Jurnal Eksplorasi Informatika*, 9(1), 1–10.
- Laurensz, B., & Sedyono, E. (2021). Analisis sentimen masyarakat terhadap tindakan vaksinasi dalam upaya mengatasi pandemi Covid-19. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 10(2), 118–124.
- Medenica, S., Živanovic, D., Batkoska, L., Marinelli, S., Basile, G., Perino, A., ... & Zaami, S. (2022). The future is coming: Artificial intelligence in the treatment of infertility could improve assisted reproduction outcomes the value of regulatory frameworks. *Diagnostics*, 12(12), 2979–2996.
- Naik, N., Roth, B., & Lundy, S. D. (2025). Artificial intelligence for clinical management of male infertility, a scoping review. *Current Urology Reports*, 26(1), 17–30.
- Pertiwi, R. M., Sumahardi, D., & Indriani, F. (2024). The impact of banking security features on marketing performance: Gold Indonesia 2045. *Research Horizon*, 4(4), 281–286.
- Pradana, L. S. (2024). *Analisis sentimen masyarakat media sosial Twitter terhadap kinerja penjabat gubernur DKI Jakarta menggunakan model IndoBERT* (Bachelor's thesis, Jakarta: UIN Syarif Hidayatullah).
- Riyaddulloh, R., Romadhony, A., Informatika, F., & Telkom, U. (2021). Normalisasi teks bahasa Indonesia berbasis kamus slang studi kasus: Tweet produk gadget pada Twitter. *Jurnal Teknologi Informasi dan Komunikasi*, 8(4), 4216–4228.
- Rodríguez-Ibáñez, M., Casáñez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P. M. (2023). A review on sentiment analysis from social media platforms. *Expert Systems with Applications*, 223(3), 852–862.
- Romaito, E. S., Anam, M. K., & Rahmaddeni, A. N. U. (2021). Perbandingan algoritma SVM dan NBC dalam analisa sentimen pilkada pada Twitter. *CSRID Journal*, 13(3), 169–179.
- Ryandi, F. A., Pratiwi, D., & Sari, S. (2025). Analisis sentimen masyarakat di media sosial X terhadap Kemenkes dengan Naïve Bayes dan SVM. *Jurnal Sains dan Teknologi*, 7(1), 1–6.
- Samsir, A., Verawardina, U., Edi, F., & Watrionthos, R. (2021). Analisis sentimen pembelajaran daring pada Twitter di masa pandemi Covid-19 menggunakan metode Naïve Bayes. *Jurnal Media Informatika Budidarma*, 5(1), 157–163.
- Saputri, J. A., Safitri, N. J. C., Jaudah, H. A., & Herbawani, C. K. (2024). Determinan yang memengaruhi penerimaan metode kontrasepsi vasktomi pada pria di Indonesia. *Jurnal Penelitian Inovatif*, 4(3), 1469–1478.
- Sautomo, S., Hafidz, N., Achyani, Y. E., & Gata, W. (2020). Sentiment analysis due to “mudik” prohibited of Covid-19 through Twitter. *Jitk (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, 6(1), 7–12.
- Septian, J. A., Fahrudin, T. M., & Nugroho, A. (2019). Analisis sentimen pengguna Twitter terhadap polemik persepakbolaan Indonesia menggunakan pembobotan TF-IDF dan K-Nearest Neighbor. *Insyst: Journal of Intelligent System and Computation*, 1(1), 43–49.
- Sontayasara, T., Jariyapongpaiboon, S., Promjun, A., & Seelpipat, N. (2021). Twitter sentiment analysis of Bangkok tourism during Covid-19 pandemic using Support Vector Machine algorithm. *Journal of Disaster Research*, 16(1), 24–30.
- Thelwall, M., Sud, P., & Vis, F. (2012). Commenting on YouTube videos: From Guatemalan rock to el big bang. *Journal of the American Society for Information Science and Technology*, 63(3), 616–629.
- Tinaliah, T., & Elizabeth, T. (2022). Analisis sentimen ulasan aplikasi PrimaKu menggunakan metode Support Vector Machine. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(4), 3436–3442.
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780.
- Wantika, D. A., Sapari, Y., & Machsunah, U. (2021). Pemanfaatan Instagram sebagai media komunikasi pemasaran online (Analisis SWOT akun Instagram @olala.closets). *MASSIVE: Jurnal Ilmu Komunikasi*, 1(2), 37–45.

Widowati, T. T., & Sadikin, M. (2020). Analisis sentimen Twitter terhadap tokoh publik dengan algoritma Naïve Bayes dan Support Vector Machine. *Jurnal Simetris*, 11(2), 626–636.

Acknowledgment

We gratefully acknowledge the contributions of individuals who supported the completion of this article.

Funding Information

This research did not receive any funding.

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

Ethical Approval and Originality Statement

Ethical approval was obtained for this study. The manuscript represents original work and has not been previously published, nor is it under consideration by another journal.

Data Disclosure Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License (<https://creativecommons.org/licenses/by-sa/4.0/>).