BIMA
BULLETIN OF INTELLIGENT MACHINES AND ALGORITHMS
Vol. 1 No.3 March 2026
E-ISSN: 3132-5115
DOI:10.65780/bima.v1i3.18

Sentiment Analysis Model for the Free Lunch Program in Indonesia on Twitter (X) Based on Machine Learning

Amelia Tifany Dewi[1*], Nur Alamsyah[2], Arnold Ropen Sinaga[3]

Sistem Informasi[1,2,3]
Universitas Informatika dan Bisnis Indonesia, Bandung, Indonesia[1,2,3]
https://unibi.ac.id/[1,2,3]
amelia.td21@student.unibi.ac.id[1*], nuralamsyah@unibi.ac.id[2], arnoldropen@unibi.ac.id[3]

Abstract. Social media has become a primary platform for the public to voice their opinions on various public policies, including the free lunch program initiated by the Indonesian government. This study aims to analyze public sentiment toward this program through the Twitter (X) platform by utilizing machine learning algorithms. Data collection was conducted from January 2025 to June 2025, with a total of 2,045 comments successfully gathered. Sentiment labeling was performed manually, and only positive and negative sentiments were considered. The data, in the form of relevant comments, were pre-processed and classified into positive and negative sentiments. Three algorithms used in this study are Support Vector Machine (SVM), Naïve Bayes, and Random Forest. Evaluation was performed using data splitting schemes of 70:30 and 80:20, along with 5-fold cross-validation. Unlike previous studies, which primarily focused on sentiment analysis of general social issues or specific topics without emphasizing public policy, this study specifically investigates the public's sentiment regarding a government policy (the free lunch program) and compares the performance of different machine learning models. The results of the study show that the Random Forest model outperformed SVM and Naïve Bayes, achieving an accuracy of 89.41% with a standard deviation of 0.0138. Meanwhile, SVM achieved an accuracy of 88.96% and Naïve Bayes 88.72%. These findings suggest that Random Forest is the most optimal and consistent model for sentiment analysis of public policies on social media.

Keywords: Sentiment Analysis, Free Lunch Program, Random Forest, Support Vector Machine (SVM), Naïve Bayes

1. INTRODUCTION

Social media has become a primary platform for the public to share opinions, experiences, and views on various social issues and public policies. Among these platforms, Twitter (X) has been widely used as an open space for public communication and political discussion in Indonesia [7]. The interactions that occur on this platform generate valuable textual data that can be analyzed to understand public sentiment toward particular issues or policies [1], [2], [8], [16]. In this context, the free lunch program initiated by the Indonesian government has become a focal point of public attention and debate on social media [19], [22], [24], [27].

The free lunch program is intended to improve children's nutrition and support learning productivity in schools, thereby contributing to broader national development goals. As a large-scale public welfare policy, this program has generated a wide range of responses from the public, including support as well as criticism related to efficiency, food quality, budget allocation, and transparency [19], [22], [24], [26]. Through tweets and interactions among users on Twitter (X), the public continuously expresses opinions regarding the implementation and impact of this policy, creating an important source of data for understanding public sentiment [19], [24], [27].

It is important to examine how this program is perceived by the public and how its implementation is socially received. One effective way to do so is through sentiment analysis of public responses on social media. Sentiment analysis has been widely applied to Twitter data and other user-generated textual content to identify public opinion patterns on social, commercial, and governmental issues [1], [2], [6], [8], [18]. Therefore, this study aims to analyze public perception of the free lunch program through sentiment analysis applied to data collected from the Twitter (X) platform. Using machine learning algorithms such as Support Vector Machine (SVM), Naïve Bayes, and Random Forest, public opinions are classified into two main sentiment categories: positive and negative [6], [18], [21],

Vol. 1 No.3 March 2026
E-ISSN: 3132-5115
DOI:10.65780/bima.v1i3.18

[26]. This approach is expected not only to measure public responses objectively but also to identify the dominant issues influencing sentiment over time.

Research on sentiment analysis using machine learning algorithms has been widely conducted in various domains. For example, Ginabila and Fauzi analyzed sentiment toward Spotify using Naïve Bayes and SVM, and found that Naïve Bayes achieved better performance in their study [9]. Daryanti and Widodo compared SVM, Naïve Bayes, and Random Forest in analyzing Indonesian airline passenger reviews, and reported that Random Forest produced the best classification results [6]. Other studies have also examined sentiment in different contexts, such as Ruangguru application reviews [11], marketplace discussions on Twitter [13], responses to the COVID-19 pandemic [18], public service reviews [20], and public sentiment toward fuel price increase policies [21]. These studies demonstrate that machine learning methods are effective for sentiment classification across diverse topics.

However, although previous studies have shown the potential of sentiment analysis for public opinion mining, there is still a research gap in studies focusing specifically on large-scale social welfare policies in Indonesia. Several recent studies have examined public sentiment toward the free lunch program using methods such as SVM, Naïve Bayes, and general text mining approaches [19], [22], [24], [27]. In addition, Yuspita and Suryono compared various text classification methods for sentiment toward the free meal policy in Indonesia [26]. Nevertheless, many of these studies either focus on a single algorithm, compare limited methods, or do not specifically emphasize a direct comparison among SVM, Naïve Bayes, and Random Forest on the same dataset and sentiment categories. As a result, evidence remains limited regarding which of these three widely used machine learning algorithms performs most effectively and consistently for sentiment classification in the context of public welfare policy discourse in Indonesia.

The comparison of SVM, Naïve Bayes, and Random Forest is important because each algorithm has different strengths in handling textual data. SVM is well known for its ability to handle high-dimensional feature spaces and complex classification boundaries [1], [8], [14]. Naïve Bayes is computationally efficient and often performs well in text classification despite its simplifying independence assumption [9], [17], [25]. Meanwhile, Random Forest offers robustness through an ensemble mechanism and has shown strong performance in sentiment classification tasks [6], [15], [21]. Because sentiment data from Twitter are often noisy, dynamic, and context-dependent, comparing these three algorithms is relevant for determining which method is most suitable for analyzing public responses to the free lunch policy [6], [18], [21], [26].

This study is expected to contribute to the literature by providing a more focused comparison of SVM, Naïve Bayes, and Random Forest for sentiment analysis of a contemporary public welfare policy in Indonesia. The findings are expected to offer a clearer picture of public perception toward the free lunch program, identify the dominant issues shaping online opinion, and provide practical recommendations regarding the most effective machine learning algorithm for policy-related sentiment analysis. In addition, the results may serve as useful input for policymakers in designing more responsive, transparent, and socially acceptable public programs [19], [22], [24], [26].

2. RELATED WORK

Sentiment analysis is a technique in text mining used to identify opinions or attitudes expressed by users toward a particular topic based on textual data from social media. Twitter (X) is widely used as a data source because it allows users to express opinions openly regarding various social issues and public policies (Emeraldien et al., 2019) [7].

Several previous studies have applied machine learning algorithms for sentiment analysis. Rahutomo et al. (2018) [8] and Irfani (2020) [11] utilized the Support Vector Machine (SVM) algorithm to analyze sentiment in movie reviews and application reviews, showing that SVM can provide good classification performance for textual data. In addition, the Naïve Bayes algorithm is also widely used due to its simplicity and computational efficiency. Pamungkas and Kharisudin (2021) [18] and Ginabila and Fauzi (2023) [9] demonstrated that both Naïve Bayes and SVM can effectively classify public sentiment on social media.

Another method commonly used in sentiment analysis is Random Forest, an ensemble learning algorithm capable of improving classification accuracy through multiple decision trees. Basar et al. [4] and Mahmuda (2024) [15] showed that Random Forest performs well in classifying textual data and digital content.

Several studies have also conducted comparisons of classification algorithms to determine the most effective method. Mardiana et al. (2019) [16], Kurniawan et al. (2023) [13], and Samantri and Afiyati (2024) [21] compared several algorithms such as Naïve Bayes, SVM, and Random Forest in sentiment analysis on various public issues discussed on Twitter.

In the context of public policy, several recent studies have examined public opinion regarding the free lunch program on social media. Purwanti and Sugiyono (2024) [19], Saputra and Hasan (2024) [22], and Sitanggang et al. (2024) [24] used SVM and Naïve Bayes methods to classify public sentiment toward the program. Additionally, Yuspita and Suryono (2024) [26] compared various classification methods to analyze public sentiment regarding the free meal policy in Indonesia.

Based on previous studies, SVM, Naïve Bayes, and Random Forest have proven effective for sentiment analysis tasks. However, the performance of each algorithm may vary depending on the characteristics of the dataset and the data preprocessing techniques used. Therefore, this study evaluates and compares these classification methods to determine the

Vol. 1 No.3 March 2026
E-ISSN: 3132-5115
DOI:10.65780/bima.v1i3.18

most effective model for classifying public sentiment on social media.

3. METHODS

The research method used in this study is a quantitative research method, as shown in Figure 1. Quantitative research is an approach that focuses on the collection and analysis of numerical data to test hypotheses, identify patterns, or measure relationships between variables. This method uses structured instruments such as surveys, experiments, and statistical measurements to generate data that can be objectively measured and analyzed. The results are typically presented in the form of numbers, tables, graphs, or other statistical equations.

In this study, the we used the classification models Support Vector Machine (SVM), Naïve Bayes, and Random Forest. The research process is carried out in several sequential stages as shown in the following diagram.

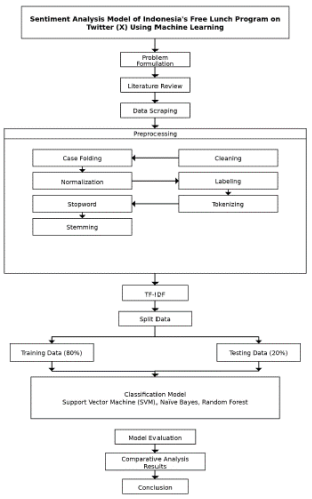

Fig.1. Research Process

Below is an explanation of the research stages:

1. Sentiment Analysis Model: This model analyzes the free lunch program in Indonesia on Twitter (X) using machine learning algorithms.
2. Problem Formulation: This stage defines the core problem to be investigated. In this study, the problem is how to analyze sentiment regarding the free lunch program on Twitter (X) using machine learning models, and which model is the best for this analysis.
3. Literature Review: This stage involves reviewing relevant written sources such as books, journals, articles, and theses to understand theories, concepts, and previous research.
4. Data Scraping: Data scraping is the process of automatically collecting data from websites or online sources using specific tools or scripts. In this study, the Tweet Harvest tool is used to scrape public comments from Twitter (X).
5. Pre-processing: This crucial first step cleans and prepares raw data for analysis. It includes:
   a. Cleaning: Removing unnecessary elements like special characters, punctuation, URLs, hashtags, and usernames.
   b. Case Folding: Standardizing text by converting all characters to lowercase.
   c. Normalization: Converting non-standard words and abbreviations to their correct forms.
   d. Labelling: Assigning sentiment labels (positive or negative) to text data.
   e. Tokenizing: Breaking text into smaller units like words or phrases.
   f. Stopword Removal: Removing common words that don't contribute meaningful information.
   i. Stemming: Reducing words to their base form.
6. TF-IDF: This method calculates the importance of words in a document based on their frequency (TF) and how unique they are across all documents (IDF).
7. Data Splitting: The dataset is divided into training and testing data. Training data is used to train the model, while testing data is used to evaluate its performance. Common splits include 80:20, 70:30, 90:10, and 60:40.
8. Classification Model: Machine learning models are used to classify data. This study tests three models for sentiment analysis:
   a. Support Vector Machine (SVM)
   b. Naïve Bayes
   c. Random Forest
9. Model Evaluation: Model evaluation focuses on accuracy using a confusion matrix, which calculates metrics like accuracy, precision, recall, and F1-score. Cross-validation is used to improve performance evaluation by splitting the training data into multiple subsets (folds).
10. Comparison Analysis: This research compares the performance of the three models by analyzing the results of the SVM, Naïve Bayes, and Random Forest algorithms.
11. Conclusion: The conclusion summarizes the findings and provides answers to the research questions based on the analysis.

4. RESULTS AND DISCUSSIONS

4.1. Data Scrapping

The initial step in the methodology of this study is the data collection phase, which is carried out through a crawling process on the Twitter (X) platform. The collected data consists of tweets that contain public opinions related to the research topic. Only tweet content is collected, without including media, links, or other additional data. Sentiment classification focuses only on positive and negative

Vol. 1 No.3 March 2026
E-ISSN: 3132-5115
DOI:10.65780/bima.v1i3.18

sentiments, and labeling is done manually by the researcher to ensure accuracy in sentiment classification. The data collection took place from January 2025 to June 2025, with a total of 2,045 comments successfully gathered.

4.2. Data Split.

The data split process is carried out by dividing the dataset into two parts: training data and testing data. After the data is represented using the TF-IDF method, the split is performed with two comparison scenarios, namely 70:30 and 80:20.

Next, during the model training phase, the k-fold cross-validation technique is used to evaluate the algorithm's performance. In this method, the training data is automatically split into several folds, with each fold alternating as both training and validation data. Thus, the model validation process is obtained directly from the previously performed training and testing data split, without the need to explicitly create an additional validation set. The goal of this approach is to maximize the use of the available data while ensuring a more objective and consistent model performance evaluation.

TABLE 1. Data Split

Description	Split 80:20	Split 70:30
Number of Training Data	1.398	1.223
Negative (Train)	777	680
Positive (Train)	621	543
Number of Test Data	350	525
Negative (Test)	195	292
Positive (Test)	155	233

4.3. Classifications Model.

SVM, Modeling is performed using several Python libraries, including pandas, scikit-learn, and TfidfVectorizer. Pandas is used to manage and display the evaluation results in tables, while train_test_split from sklearn.model_selection splits the dataset into training and testing data with ratios of 70:30 and 80:20. TfidfVectorizer from sklearn.feature_extraction.text is used to convert the pre-processed text into numerical vectors based on term frequency (TF) and inverse document frequency (IDF). The SVM shows fairly good performance in both data splitting scenarios, whether with a 70:30 or 80:20 ratio. In the 70:30 scheme, the SVM model achieved an accuracy of 87.81%, with a weighted recall of 87.81%, precision of 88.24%, and an F1-score of 87.73%. Meanwhile, in the 80:20 scheme, the model recorded a higher accuracy of 90.86%, with weighted recall and precision of 90.86% and 90.89%, respectively, and an F1-score of 90.82%.

Naïve Bayes, based on the evaluation results, the Naive Bayes algorithm model shows stable performance in both split schemes. In the 70:30 scheme, the model achieved an accuracy of 88.57%, with precision, recall, and F1-score all being 88.57%. Meanwhile, in the 80:20 scheme, the model recorded an accuracy of 89.43%, with precision of 89.50%, recall of 89.43%, and an F1-score of 89.45%. Although the improvement is not significant, these results indicate that Naive Bayes works fairly consistently and well in processing text-based sentiment classification.

Based on the evaluation results, the Random Forest model shows excellent and fairly stable performance in both data split schemes. In the 70:30 scheme, the model achieved an accuracy of 90.10%, with precision of 90.50%, recall of 90.10%, and an F1-score of 90.04%. Meanwhile, in the 80:20 scheme, the evaluation scores were slightly higher, with accuracy of 90.29%, precision of 90.28%, recall of 90.29%, and an F1-score of 90.28%. These results indicate that Random Forest works very effectively in classifying sentiment, especially when the proportion of training data is increased.

4.4 Cross Validation

As shown in Table 2, Based on the cross-validation results the SVM model shows a relatively high average accuracy in both data splits. In the 70:30 scheme, the mean accuracy reached 88.96% with a standard deviation of 1.89%, while in the 80:20 scheme, the mean accuracy was 88.70% with a standard deviation of 1.38%. The low standard deviation values indicate that the SVM model has good performance stability across each fold.

TABLE 3. Results of Cross Validation SVM

Split	Mean Accuracy	Std Dev
70.30	0.8896	0.0189
80.20	0.8870	0.0138

The Naive Bayes model also produces consistent accuracy in both data splits, as shown in Table 4. For the 70:30 split, the model achieved a mean accuracy of 88.72% with a standard deviation of 1.27%. Meanwhile, for the 80:20 split, the mean accuracy slightly decreased to 88.56% with a standard deviation of 1.70%. This indicates that the model's performance remains stable and does not experience large fluctuations across folds

TABLE 4. Results of Cross Validation Naïve Bayes

Split	Mean Accuracy	Std Dev
70:30	0.8872	0.0127
80.20	0.8856	0.0170

Random Forest shows comparable performance, even slightly higher in the 80:20 split, as shown in Table 5. With a mean accuracy of 88.80% in the 70:30 split and 89.41% in the 80:20 split, this model demonstrates good stability and effectiveness in handling the data. The standard deviation values remain low in both scenarios, 1.26% and 1.38%, indicating consistency in the prediction results across folds

TABLE 5. Results of Cross Validation Random Forest

Split	Mean Accuracy	Std Dev
70:30	0.8880	0.0126
80:20	0.8941	0.0138

Vol. 1 No.3 March 2026
E-ISSN: 3132-5115
DOI:10.65780/bima.v1i3.18

4.4. Comparison analysis

Performance evaluation results shown in the Table 6, the SVM model with an 80:20 data split achieved the highest accuracy of 90.9%. However, after performing cross-validation, the Random Forest model with an 80:20 data split showed the highest mean accuracy of 89.4%, outperforming the other models. This indicates that Random Forest has consistent and stable performance. Therefore, it can be concluded that Random Forest is the best model in this study for sentiment classification of public comments on the free lunch program.

TABLE 6. Results of the comparative analysis

Method	Data Split	Accuracy	Presisi	Recall	F1-Score	Mean Accuracy (Cross Validation)
Support Vector Machine (SVM)	70%:30%	87.8%	88.2%	87.8%	87.7%	88.9%
	80%:20%	90.9%	90.9%	90.9%	90.8%	88.7%
Naïve Bayes	70%:30%	88.6%	88.6%	88.6%	88.6%	88.7%
	80%:20%	89.4%	89.5%	89.4%	89.5%	88.6%
Random Forest	70%:30%	90.1%	90.5%	90.1%	90%	88.8%
	80%:20%	90.3%	90.3%	90.3%	90.3%	89.4%

5. CONCLUSIONS

Based on the evaluation of the classification models, it can be concluded that sentiment analysis of the free lunch program on the Twitter (X) platform was conducted through a text pre-processing stage, followed by data analysis using three machine learning algorithms: Support Vector Machine (SVM), Naive Bayes, and Random Forest. The performance of these models was evaluated based on data split schemes of 70:30 and 80:20, with validation using 5-fold cross-validation. The results showed that the Support Vector Machine (SVM) achieved the best performance on the 80%:20% data split, with an accuracy of 90.9%, followed by Random Forest with an accuracy of 90.3%, and Naive Bayes with 89.4%.

After performing cross-validation, the models showed the following average accuracies: SVM achieved the highest average accuracy of 88.9% on the 70%:30% data split, while Naive Bayes reached 88.7%, and Random Forest obtained the highest average accuracy of 89.4% on the 80%:20% data split. Despite SVM showing the highest accuracy on the 80%:20% data split, Random Forest demonstrated more consistent performance across cross-validation, with the highest mean accuracy of 89.4%. Given the evaluation and validation results, it is concluded that Random Forest is the most optimal model for sentiment analysis in this study.

REFERENCES

[1] Z. Alhaq, A. Mustopa, S. Mulyatun, and J. D. Santoso, "Penerapan metode support vector machine untuk analisis sentimen pengguna Twitter," Journal of Information System Management (JOISM), vol. 3, no. 2, pp. 44–49, 2021. doi: 10.24076/joism.2021v3i2.558.

[2] M. I. Amal, E. S. Rahmasita, E. Suryaputra, and N. A. Rakhmawati, "Analisis klasifikasi sentimen terhadap isu kebocoran data kartu identitas ponsel di Twitter," Jurnal Teknik Informatika dan Sistem Informasi, vol. 8, no. 3, 2022. doi: 10.28932/jutisi.v8i3.5483.

[3] I. Ardiansah and R. H. Permana, Memulai Python: Belajar Python dari Nol, CV. Cendekia Press, 2023.

[4] T. F. Basar, D. E. Ratnawati, and I. Arwani, "Analisis sentimen pengguna Twitter terhadap pembayaran cashless menggunakan Shopeepay dengan algoritma Random Forest," unpublished.

[5] H. Satria, "Cara mendapatkan data (crawl) Twitter X - Maret 2024," Helmi Satria, Jul. 2, 2024. [Online]. Available: https://helmisatria.com/blog/updated-crawl-data-twitter-x-maret-2024/

[6] D. Daryanti and T. Widodo, "Analisis sentimen ulasan penumpang maskapai penerbangan Indonesia menggunakan Support Vector Machine, Naive Bayes, dan Random Forest," The Indonesian Journal of Computer Science, vol. 13, no. 5, 2024. doi: 10.33022/ijcs.v13i5.4452.

[7] F. Z. Emeraldien, R. J. Sunarsono, and R. Alit, "Twitter sebagai platform komunikasi politik di Indonesia," Scan: Jurnal Teknologi Informasi Dan Komunikasi, vol. 14, no. 1, pp. 21–30, 2019. doi: 10.33005/scan.v14i1.1457.

[8] F. Rahutomo, P. Y. Saputra, and M. A. Fidyawan, "Implementasi Twitter sentiment analysis untuk review film menggunakan algoritma Support Vector Machine," Jurnal Informatika Polinema, vol. 4, no. 2, pp. 93, 2018. doi: 10.33795/jip.v4i2.152.

[9] G. Ginabila and A. Fauzi, "Analisis sentimen terhadap pemutar musik online Spotify dengan algoritma Naive Bayes dan Support Vector Machine," Jurnal Ilmiah ILKOMINFO - Ilmu Komputer & Informatika, vol. 6, no. 2, pp. 111–122, 2023. doi: 10.47324/ilkominfo.v6i2.180.

[10] A. A. Haikal and H. H. Anbiya, "Pengaruh program makan siang dan susu gratis Prabowo Gibran terhadap sektor industri manufaktur," unpublished.

[11] F. F. Irfani, "Analisis sentimen review aplikasi Ruangguru menggunakan algoritma Support Vector

Vol. 1 No.3 March 2026
E-ISSN: 3132-5115
DOI:10.65780/bima.v1i3.18

Machine," JBMI (Jurnal Bisnis, Manajemen, dan Informatika), vol. 16, no. 3, pp. 258–266, 2020. doi: 10.26487/jbmi.v16i3.8607.

[12] W. Irmayani, "Visualisasi data pada data mining menggunakan metode klasifikasi Naïve Bayes," Karimah Tauhid, vol. 2, no. 1, pp. 2, 2021.

[13] I. Kurniawan, A. L. Hananto, S. S. Hilabi, A. Hananto, and A. Y. Rahman, "Perbandingan algoritma Naive Bayes dan SVM dalam sentimen analisis marketplace pada Twitter," vol. 10, no. 1, 2023.

[14] K. Kusnawi and A. H. Pratama, Belajar Mudah dan Singkat Machine Learning: Panduan Praktis dengan Studi Kasus, Kode Program, dan Dataset, Penerbit Andi, 2024.

[15] S. Mahmuda, "Implementasi metode Random Forest pada kategori konten kanal Youtube," JURNAL JENDELA MATEMATIKA, vol. 2, no. 1, pp. 21–31, 2024. doi: 10.57008/jjm.v2i01.633.

[16] T. Mardiana, H. Syahreva, and T. Tuslaela, "Komparasi metode klasifikasi pada analisis sentimen usaha waralaba berdasarkan data Twitter," Jurnal Pilar Nusa Mandiri, vol. 15, no. 2, pp. 267–274, 2019. doi: 10.33480/pilar.v15i2.752.

[17] E. Martantoh and N. Yanih, "Implementasi metode Naïve Bayes untuk klasifikasi karakteristik kepribadian siswa di sekolah MTS Darussa'adah menggunakan PHP MySQL," Jurnal Teknologi Sistem Informasi, vol. 3, no. 2, pp. 166–175, 2022. doi: 10.35957/jtsi.v3i2.2896.

[18] F. S. Pamungkas and I. Kharisudin, "Analisis sentimen dengan SVM, Naïve Bayes dan KNN untuk studi tanggapan masyarakat Indonesia terhadap pandemi Covid-19 pada media sosial Twitter," vol. 4, 2021.

[19] Z. Purwanti and S. Sugiyono, "Pemodelan text mining untuk analisis sentimen terhadap program makan siang gratis di media sosial X menggunakan algoritma Support Vector Machine (SVM)," Jurnal Indonesia: Manajemen Informatika Dan Komunikasi, vol. 5, no. 3, Article 3, 2024. doi: 10.35870/jimik.v5i3.1001.

[20] E. Salim and M. Syafrullah, "Analisis sentimen pada ulasan pelayanan suku dinas kependudukan dan pencatatan sipil Kota Administrasi Jakarta Barat menggunakan algoritme K-Nearest Neighbor," Bit (Fakultas Teknologi Informasi Universitas Budi Luhur), vol. 20, no. 1, pp. 58, 2023. doi: 10.36080/bit.v20i1.2186.

[21] M. Samantri and N. Afiyati, "Perbandingan algoritma Support Vector Machine dan Random Forest untuk analisis sentimen terhadap kebijakan pemerintah Indonesia terkait kenaikan harga BBM tahun 2022," Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), vol. 8, no. 1, pp. 1–9, 2024. doi: 10.35870/jtik.v8i1.1202.

[22] R. Saputra and F. N. Hasan, "Analisis sentimen terhadap program makan siang & susu gratis menggunakan algoritma Naive Bayes," Jurnal Teknologi Dan Sistem Informasi Bisnis, vol. 6, no. 3, pp. 411–419, 2024. doi: 10.47233/jteksis.v6i3.1378.

[23] A. Y. Simanjuntak and I. S. S. Simatupang, "Implementasi data mining menggunakan metode Naïve Bayes classifier untuk data kenaikan pangkat dinas ketenagakerjaan Kota Medan," unpublished.

[24] A. Sitanggang, Y. Umaidah, and R. I. Adam, "Analisis sentimen masyarakat terhadap program makan siang gratis pada media sosial X menggunakan algoritma Naïve Bayes," Jurnal Informatika Dan Teknik Elektro Terapan, vol. 12, no. 3, Article 3, 2024. doi: 10.23960/jitet.v12i3.4902.

[25] T. Thoyyibah, F. Kurniawan, and T. Taryo, Dasar-Dasar Machine Learning pada Google Colabs, Eureka Media Aksara, 2024. [Online]. Available: https://repository.penerbiteureka.com/uk/publications/567741/

[26] E. Yuspita and R. R. Suryono, "Perbandingan berbagai metode klasifikasi teks untuk sentimen kebijakan makan gratis di Indonesia," The Indonesian Journal of Computer Science, vol. 13, no. 5, 2024. doi: 10.33022/ijcs.v13i5.4440.

[27] F. N. Zaman, M. A. Fadhilah, M. A. Ulinuha, and K. Umam, "Menganalisis respons netizen Twitter terhadap program makan siang gratis menerapkan NLP metode Naïve Bayes," Jurnal Sistem Informasi, vol. 14, no. 3, 2024.