Available online at https://jurnal.mifandimandiri.com/index.php/jphcs

Journal of Public Health and Clinical Studies

e-ISSN XXXX-XXXX

Vol. 01, No. 01, pp. 26-30, Juni 2026

Public Opinion Analysis on Free Nutritious Food through Random Forest-Based Sentiment Classification

Desmawati Zebua[1], Rivalri Kristianto Hondro[2*], Mawati Zalukhu[3]

[1]Universitas Prima Indonesia, Medan, Indonesia
[2]Univeristas Satya Terra Bhinneka, Medan, Indonesia
[3]Armari Family Four, Medan, Indonesia

**Corresponding author email: rivalryhondro@gmail.com*

Abstract

The Free Nutritious Meal Program is a social policy aimed at improving public access to nutritious food, particularly for vulnerable groups. The implementation of this program has generated various public responses that are widely expressed on social media. These opinions can serve as valuable information to understand public acceptance of the policy. This study aims to analyze public opinion toward the Free Nutritious Meal Program through sentiment classification using the Random Forest algorithm. The dataset consists of 1,000 Indonesian-language comments collected from social media and categorized into positive, negative, and neutral sentiments. The research stages include data collection, labeling, text preprocessing, feature extraction using TF-IDF, data splitting, Random Forest classification, and model evaluation. The results show that Random Forest can effectively classify public opinion by combining multiple decision trees to produce more stable predictions than a single-tree model. The model achieved an accuracy of 94%, precision of 0.93, recall of 0.92, and F1-score of 0.92. These findings indicate that Random Forest is a relevant method for sentiment analysis of public policy based on social media data.

Keywords: Public Opinion Analysis, Free Nutritious Meal Program, Sentiment Classification, Random Forest, TF-IDF

1. Introduction

Nutritional fulfillment is a crucial aspect of improving public health and productivity. The availability of nutritious food plays a significant role in supporting physical growth, cognitive development, and quality of life, especially for children, students, and economically disadvantaged groups. Therefore, policies aimed at increasing access to nutritious food are a crucial component of social development and public health (Brata et al., 2025). In Indonesia, free nutritious food programs, such as the free lunch program, have drawn public attention and often sparked a variety of reactions on social media.

The Free Nutritious Food Program (FBM) is a government intervention designed to help people access healthy food without the burden of cost. This program is expected to alleviate malnutrition, assist vulnerable groups, and support the improvement of human resources. However, during its implementation, the program has not only garnered support but also generated criticism from the public. This diverse public opinion highlights the importance of sentiment analysis to comprehensively understand public perceptions, particularly through social media platforms like Twitter (formerly Twitter), which serves as a primary forum for discussion (M. Azhari, 2024) and S. Rohman, 2025).

Various public responses to the program were widely found on social media. Some people considered the program beneficial because it helped meet nutritional needs and reduce the economic burden on families. On the other hand, there were also negative opinions highlighting issues with food quality, distribution equity, targeting accuracy, implementation readiness, and budget utilization. In this context, sentiment analysis is crucial for extracting valuable information from large volumes of data, enabling policymakers to understand the dynamics of public perception more accurately and adaptively (E. Yuspita, 2024). Furthermore, there are neutral comments that contain only information, questions, or responses without strong emotional tendencies.

The large number of public opinions on social media makes manual analysis inefficient. To address this, approaches based on text mining and machine learning are needed. These approaches enable the automatic identification of sentiment patterns from large amounts of text data, providing in-depth insights into public views on the free nutritious food program (T. Aziz et al., 2025) and (K. Y. Pratama et al., 2026). One approach that can be used is sentiment analysis, which is the process of identifying opinion tendencies in text and grouping them into certain categories, such as positive, negative, and neutral.

This study uses the Random Forest algorithm to classify public sentiment toward the Free Nutritious Food Program. Random Forest was chosen because it still has a conceptual relationship with Decision Tree, but works by building multiple decision trees. This model is known to have high predictive ability and is able to handle high-dimensional data and prevent overfitting (G. Ainur, et al., 2025) (J. Carberry, 2025). Each tree produces a prediction, then the final result is determined through a voting process. With this mechanism, Random Forest tends to be more stable, more robust against noise, and able to reduce the risk of overfitting compared to a single Decision Tree. Several previous studies have also shown the superiority of Random Forest in sentiment analysis, including its effectiveness in classifying public opinion toward similar programs on social media (P. M. Kadafi, et al., 2026) (I. A. N. Sabrina, 2025). This study aims to analyze public sentiment toward the Free Nutritious Meals policy in Indonesia based on social media data X, and to evaluate the performance of the Random Forest model in sentiment classification (M. M. A. Kadzim, 2025) (D. A. Sulistyo, 2025). Specifically, this study will measure the accuracy, precision, and recall of the Random Forest model to determine its effectiveness in identifying positive, negative, and neutral sentiment related to the program (F. M. Hidayat, 2025) (N. G. Ramadhan, 2025).

However, previous studies have shown that the performance of sentiment classification models can vary depending on the characteristics of the dataset and the preprocessing method used, as observed in studies using SVM and BERT (R. Munir, 2025), as well as IndoBERT, which demonstrated 60.27% accuracy with difficulty distinguishing the neutral class (M. Haikal, 2026). Previous studies have shown that Random Forest demonstrates superior performance compared to Linear Regression in sentiment analysis of the Free Nutritional Food Program, with an accuracy of 0.85 and an F1-Score of 0.90 (I. Hermansyah, 2025). This study aims to provide an empirical basis for policymakers in formulating communication strategies and program implementation that are more adaptive and responsive to community needs and concerns. Therefore, this research is expected to make a significant contribution to the development of more effective intervention strategies to address nutrition issues, particularly stunting, a serious challenge in Indonesia (A. F. Ningrum, 2024). The importance of addressing stunting is also emphasized by the Indonesian government, which targets a 14% reduction in stunting rates by 2024 through an acceleration program (A. F. Ningrum, 2024).

Based on this background, this study aims to analyze public opinion on the Free Nutritional Food Program through Random Forest-based sentiment classification. The results are expected to provide an objective picture of public opinion trends and serve as considerations in public policy evaluation.

2. Research Methods

2.1. Type of Research

This research uses a quantitative approach with an experimental method. This quantitative approach is used because text data is converted into numerical form so it can be processed and evaluated measurably. The experimental method was used to test the ability of the Random Forest algorithm to classify public sentiment based on social media comment data.

2.2. Data Sources

The research data consists of public comments on social media discussing the Free Nutritious Meals Program. Data were collected from platform X, namely Twitter, Instagram, and YouTube, using keywords such as "free nutritious food," "free meal program," and "MBG." The dataset used consisted of 1,000 comments in Indonesian. The collected comments were then filtered to remove duplicate data, irrelevant comments, and indications of automated comments or bots. After the filtering process, the data was used as the main material in sentiment analysis.

2.3. Data Labeling

The comment data was manually labeled into three sentiment classes: positive, negative, and neutral. Positive comments were comments that expressed support, appreciation, or good wishes for the Free Nutritious Meals Program. Negative comments were comments containing criticism, complaints, rejection, or concerns regarding the program's implementation. Neutral comments are informative, ask questions, or don't indicate a clear sentimental bias. Labeling is done by taking the context of the sentence into account so that the sentiment category is determined not only by specific words but also by the overall meaning of the comment.

2.4. Text Preprocessing

Preprocessing is performed to clean comment data before entering the classification stage. The preprocessing stages include:

1. Case folding, which converts all text to lowercase.
2. Cleaning, which removes URLs, numbers, punctuation, symbols, and unnecessary characters.
3. Tokenizing, which breaks sentences into words.
4. Stopword removal, which removes common words that don't significantly influence sentiment.
5. Stemming, which converts words with affixes to their base form.

This stage is performed to make the text data cleaner, more uniform, and easier to process by machine learning models.

2.5. TF-IDF Feature Extraction

After preprocessing, the text data is converted into numeric form using the Term Frequency–Inverse Document Frequency (TF-IDF) method. This method assigns weights to words based on their frequency of occurrence in a document and their importance within the overall dataset. Words that appear frequently in a particular comment, but are less common across the entire dataset, will receive higher weight. Thus, TF-IDF helps the model recognize relevant words to differentiate between positive, negative, and neutral sentiments.

2.6. Data Distribution

The dataset was divided into two parts: training data and test data. The training data was used to build the model, while the test data was used to evaluate the model's ability to predict new data. The data distribution was carried out with a ratio of 80% for training data and 20% for testing data. Of the total 1,000 comments, 800 comments were used as training data and 200 comments were used as testing data. The distribution was carried out using a stratified sampling technique to maintain a balanced proportion of each sentiment class in the training and testing data.

2.7. Random Forest Algorithm

Random Forest is an ensemble learning algorithm that builds multiple decision trees in the classification process. Each tree is trained using a randomly selected subset of data and features. The predictions from all trees are then combined through a voting mechanism to determine the final class. In this study, Random Forest was used to classify public comments into three sentiment classes. The main advantage of Random Forest is its ability to produce more stable predictions than a single Decision Tree. This is because Random Forest does not rely on a single decision tree, but rather on a combination of multiple trees.

2.8. Model Evaluation

Model evaluation was conducted using a confusion matrix and several measurement metrics, namely accuracy, precision, recall, and F1-score. Accuracy is used to determine the percentage of correct predictions based on all test data. Precision is used to measure the model's accuracy in predicting a class. Recall is used to measure the model's ability to identify all data belonging to a particular class. F1-score is used to assess the balance between precision and recall.

3. Results and Discussion

3.1. Preprocessing and TF-IDF Results

Preprocessing results show that social media comment data can be cleaned of irrelevant elements, such as links, symbols, numbers, punctuation, and common words. After going through case folding, cleaning, tokenizing, stopword removal, and stemming, the comments become more concise and ready for use in the feature extraction stage.

For example, a comment like "This MBG program is very helpful for school children, hopefully it's evenly distributed!" can be reduced to simpler stem word forms after preprocessing, such as "program," "mbg," "help," "child," "school," "hopefully," and "average." This form is easier for the model to process because unnecessary elements have been removed.

After preprocessing, TF-IDF is used to transform the text into a numeric representation. Words that are important to sentiment receive greater weight. Words like "help," "useful," and "good" tend to contribute to positive sentiment. Words like "bad," "uneven," "unworthy," and "wrong target" tend to lead to negative sentiment. Meanwhile, comments containing general information were more likely to fall into the neutral category.

3.2. Random Forest Classification Results

The Random Forest model was trained using 800 training data sets and tested using 200 test data sets. During the training process, the model built several decision trees, each learning different patterns from the data. The final classification result was determined based on the majority vote from these trees.

Based on the test results, Random Forest was able to effectively classify public opinion regarding the Free Nutritional Meals Program. The model was able to distinguish positive, negative, and neutral comments based on word patterns represented by TF-IDF.

In general, positive comments were characterized by words that indicated support for and benefits of the program. Negative comments were characterized by words related to criticism of the program's implementation. Neutral comments tended to contain information, questions, or statements without strong emotional overtones.

3.3. Model Performance Evaluation

The model was evaluated using 200 test data sets. The results of the Random Forest model evaluation are shown in the following table.

Table 1: Model Performance Evaluation

Evaluation Metrics	**Value**
Accuracy	94%
Precision	0,93
Recall	0,92
F1-score	0,92

An accuracy score of 94% indicates that most comments in the test data were correctly classified. A precision of 0.93 indicates that the model's predictions for each class are highly accurate. A recall of 0.92 indicates that the model was able to recognize most comments according to their actual class. An F1-score of 0.92 indicates that the model has a good balance between precision and recall.

These results indicate that Random Forest can be a more robust alternative to a single Decision Tree. In Decision Trees, the model relies on only one decision tree structure. Meanwhile, Random Forest combines multiple trees, making the final prediction more stable and less susceptible to specific data patterns.

3.4. Discussion of Public Sentiment

The classification results indicate that public opinion regarding the Free Nutritious Meals Program falls into three main categories. Positive sentiment reflects public support for the program. Comments in this category generally believe that the program can help meet nutritional needs, ease family burdens, and support the well-being of children or students.

Negative sentiment indicates criticism of aspects of the program's implementation. The criticisms relate to food quality, unequal distribution, potential mistargeting, and concerns about budget effectiveness. This criticism is important to note because it can provide input for policy improvement.

Neutral sentiment consists of comments that do not clearly indicate support or opposition. Comments in this category generally involve questions, information dissemination, or general responses to programs.

These findings suggest that social media can be a relevant data source for understanding public responses to public policies. By using Random Forest, large numbers of opinions can be classified more quickly, systematically, and measurably.

3.5. Model Strengths and Limitations

Random Forest has the advantage of mitigating the weaknesses of a single Decision Tree. By using multiple decision trees, the model becomes more stable and has better generalization capabilities. Furthermore, Random Forest is quite effective for text data that has been represented using TF-IDF.

However, this model also has limitations. Random Forest is more difficult to interpret than a single Decision Tree because the final decision comes from multiple trees. Furthermore, the model still has limitations in understanding comments containing sarcasm, irony, or informal language that often appear on social media. Therefore, the preprocessing and language normalization stages are important parts in improving the quality of classification.

4. Conclusion

This study analyzes public opinion on the Free Nutritious Food Program using Random Forest-based sentiment classification. The data used consisted of 1,000 Indonesian-language social media comments categorized into positive, negative, and neutral sentiments. The research stages included data collection, labeling, text preprocessing, TF-IDF feature extraction, data sharing, Random Forest model training, and model performance evaluation. The results showed that Random Forest was able to effectively classify public sentiment. The model achieved an accuracy of 94%, a precision of 0.93, a recall of 0.92, and an F1-score of 0.92. These results demonstrate that Random Forest can be an effective method for sentiment analysis because it combines multiple decision trees, resulting in more stable predictions. Positive opinions indicate public support for the benefits of the Free Nutritious Food Program. Negative opinions reflect criticism of program implementation, such as distribution, food quality, and targeting accuracy. Meanwhile, neutral opinions indicate comments that are informative or do not yet express a clear stance. Thus, Random Forest-based sentiment analysis can be used as an approach to understanding public perceptions of public policy. The results of this study can be used as input for the government in evaluating the implementation of the Free Nutritious Food Program and developing more appropriate communication strategies for the public.

References

Ainur, G., Latifah, N., & Nugraha, F. (2025). Analisis sentimen publik terhadap program makan bergizi gratis menggunakan metode Random Forest pada platform X. *SITECH Jurnal Sistem Informasi dan Teknologi, 8*(1), 75–84. https://doi.org/10.24176/sitech.v8i1.15628

Azhari, M., & Parjito, P. (2024). Analisis sentimen opini publik program makan siang gratis dengan Random Forest pada media. *Building of Informatics, Technology and Science (BITS), 6*(3), 1932–1942. https://doi.org/10.47065/bits.v6i3.6423

Aziz, T., Ismayadi, I., & Budiman, B. (2025). Analisis sentimen terhadap program makan siang gratis pada media sosial X menggunakan Logistic Regression dan SVM. *In Search, 24*(1), 18–28. https://doi.org/10.37278/insearch.v24i1.1238

Brata, D. W., Purnomo, W., Farisi, H., & Ratu, H. K. M. (2025). Menggali opini publik: Sentimen terhadap kebijakan makan siang gratis dengan supervised learning. *Jurnal Teknologi Informasi dan Ilmu Komputer, 12*(3), 533–538. https://doi.org/10.25126/jtiik.2025129565

Carberry, J., & Ikhsan, M. (2025). Sentiment analysis of “Free Lunch for Children” policy on social media X using Random Forest algorithm. *Journal of Information Systems and Informatics, 7*(1), 649–662. https://doi.org/10.51519/journalisi.v7i1.1039

Haikal, M., & Erfina, A. (2026). Analisis sentimen warganet terhadap program makan bergizi gratis (MBG) menggunakan model IndoBERT. *Jurnal Locus Penelitian dan Pengabdian, 5*(3), 1626–1636. https://doi.org/10.58344/locus.v5i3.5253

Hermansyah, I., & Hasibuan, M. A. (2025). Sentiment analysis of free nutritious meal programme on social media X using Linear Regression and Random Forest algorithms. *PIKSEL Penelitian Ilmu Komputer Sistem Embedded and Logic, 13*(1), 55–68. https://doi.org/10.33558/piksel.v13i1.10633

Hidayat, F. M., & Sanjaya, H. (2025). Analisis sentimen publik terhadap penjualan iPhone 16 dan kebijakan TKDN di Indonesia. *INFOTECH Journal, 11*(1), 74–80. https://doi.org/10.31949/infotech.v11i1.13159

Kadafi, P. M., Riyadi, H. Z., Raya, R. S. G., Kurniawati, I., Bismi, W., & Fahlapi, R. (2026). Penerapan machine learning untuk analisis sentimen program MBG pada platform X dan YouTube. *INFOTECH Journal*, *12*(1), 67–74. https://doi.org/10.31949/infotech.v12i1.17472

Kadzim, M. M. A. (2025). Analisis sentimen berbasis aspek terhadap program makan bergizi gratis di media sosial X menggunakan metode support vector machine [*Tesis, Universitas Islam Negeri Maulana Malik Ibrahim*].

Munir, R. (2025). Analisis sentimen cuitan di media sosial X tentang program makan bergizi gratis dengan metode NLP. *Jurnal Informatika dan Teknik Elektro Terapan, 13*(3). https://doi.org/10.23960/jitet.v13i3.6912

Ningrum, A. F. (2024). Sentiment analysis of public opinion on handling stunting in Indonesia using Random Forest. *Jurnal Statistika dan Aplikasinya, 8*(1), 31–40. https://doi.org/10.21009/jsa.08103

Ningrum, A. F., & Amri, I. F. (2024). Sentiment analysis of public opinion on handling stunting in Indonesia using Random Forest. *J Statistika Jurnal Ilmiah Teori dan Aplikasi Statistika, 17*(1), 645–654. https://doi.org/10.36456/jstat.vol17.no1.a9088

Pratama, K. Y., Achmadi, S., & Sari, K. A. (2026). Analisis sentimen terhadap makan bergizi gratis pada sosial media X menggunakan metode Decision Tree dengan pembanding Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika), 9*(6), 11061–11067. https://doi.org/10.36040/jati.v9i6.15525

Ramadhan, N. G., & Khoirunnisa, A. (2025). An integrated Random Forest for analyzing public sentiment on the “Makan Bergizi Gratis” program. *Journal of Information Systems and Informatics, 7*(3), 2314–2328. https://doi.org/10.51519/journalisi.v7i3.1184

Rohman, S., & Agung, I. W. P. (2025). Komparasi algoritma IndoBERT, SVM, dan Random Forest dalam analisis sentimen program makan bergizi gratis. *JATI (Jurnal Mahasiswa Teknik Informatika), 9*(6), 9464–9471. https://doi.org/10.36040/jati.v9i6.15611

Sabrina, I. A. N., Wuryandari, T. T., & Dapa, R. (2025). Implementation of Random Forest algorithm in classifying public sentiment towards free nutritious meal program. *Current Science Research Bulletin, 2*(5). https://doi.org/10.55677/csrb/02-v02i05y2025

Sulistyo, D. A., & Setiadi, E. (2025). Analisis sentimen kebijakan makan bergizi gratis menggunakan IndoBERT dan machine learning. *JURNAL FASILKOM, 15*(3), 507–513. https://doi.org/10.37859/jf.v15i3.10546

Yuspita, E., & Suryono, R. R. (2024). Perbandingan berbagai metode klasifikasi teks untuk sentimen kebijakan makan gratis di Indonesia. *Indonesian Journal of Computer Science, 13*(5). https://doi.org/10.33022/ijcs.v13i5.4440