Journal of Information Systems and Informatics

Vol. 7, No. 3, September 2025 e-ISSN: **2656-4882** p-ISSN: **2656-5935**

DOI: 10.51519/journalisi.v6i2.759 Published By DRPM-UBD

An Integrated Random Forest for Analyzing Public Sentiment on the "Makan Bergizi Gratis" Program

Nur Ghaniaviyanto Ramadhan[1,*], Azka Khoirunnisa[2]

[1,2]Department of Data Science, School of Computing, Telkom University, Bandung, Indonesia
Email: [1]nuruer@telkomuniversity.ac.id, [2]khoirunnisaazka@telkomuniversity.ac.id

Abstract

The "Makan Bergizi Gratis" (MBG) Program is a public policy aimed at improving the nutritional quality of the community, particularly vulnerable groups. However, the success of this program is heavily influenced by public sentiment and perception. This research analyzes public sentiment toward the MBG program thru the social media platform X using an ensemble-based machine learning approach. The proposed framework integrates the Random Forest algorithm and compares it with four other ensemble models: AdaBoost, XGBoost, Bagging, and Stacking. A total of 3,417 tweets were analyzed using the TF-IDF method, both with and without stemming. The Random Forest model showed the best performance with an accuracy of 91.15% and an ROC-AUC of 95.46% on the data without stemming, consistently outperforming the other models. Additionally, a visual analysis of word frequency provides a strong indication of public opinion. These findings demonstrate the effectiveness of Random Forest in managing unstructured sentiment data and provide valuable insights for policymakers to monitor public responses and improve program implementation with greater precision.

Keywords: Analysis Sentiment, Ensemble Learning, Random Forest, MBG, TF-IDF

1. INTRODUCTION

The "Makan Bergizi Gratis" (MBG) program is one of the public policy initiatives aimed at improving the nutritional quality of the community, especially for vulnerable groups such as students, children, and low-income families. This policy is expected to support the achievement of sustainable development goals (Sustainable Development Goals) in the aspects of health (Good Health and Well-being) and hunger eradication (Zero Hunger) [1]. However, every public policy implemented on a large scale is not free from public assessment and response. Public perception and sentiment play a crucial role in determining the success of program implementation, considering that the level of public acceptance can influence the effectiveness, sustainability, and legitimacy of the policy [2].

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

As social media and information technology have developed, the public is now more actively voicing their opinions, complaints, and support for various government programs. This massive public sentiment data becomes a valuable source of information for in-depth analysis [3]. However, processing text data from social media often faces challenges such as large data volumes, diverse language variations, and the presence of subjective and unstructured opinions [4]. Therefore, an analytical framework capable of effectively handling the complexity of such data is needed.

In this context, the Random Forest algorithm has become one of the popular machine learning methods due to its reliability in handling complex and diverse data. The advantage of Random Forest lies in its ability to perform classification with high accuracy, minimize overfitting, and provide interpretation through feature importance measurement [5]. However, the integrated application of Random Forest in public sentiment analysis still requires a systematic framework design, starting from the data preprocessing stage, feature selection, to model performance evaluation [6].

Some of the research conducted by Rahmatullah is using the Naive Bayes algorithm to conduct sentiment analysis on public reviews and comments about the free nutritious food program [7]. Triningsih, et al analyzed public sentiment towards the Free Nutritious Eating policy to understand public perception whether it is positive, neutral or negative towards the free nutritious meal program through a machine learning algorithm approach. To achieve this goal, this study compares the performance of two classification algorithms, namely Support Vector Machine and Random Forest in analyzing the sentiment of public opinion, especially on platform X [8]. Anggriyani analyzed public opinion related to various statements or comments about the free meal and free milk program [9].

Various previous studies have applied classification algorithms such as Naive Bayes and SVM for public policy sentiment analysis, including for the MBG program[7], [9]. However, there is still limited research that systematically explores the integration of ensemble learning methods, particularly Random Forest and its variants, within the context of policy sentiment analysis. This is even though ensemble learning is known to improve classification accuracy and handle high data complexity. Additionally, there haven't been many studies that directly link the results of this sentiment analysis to recommendations for policymakers.

Based on this background, this research aims to propose an Integrated Random Forest Framework designed to analyze public sentiment towards the MBG program through the social media platform X. This model is expected to provide accurate analysis results and support data-driven decision-making. This paper has the following structure: Part 1 discusses the background of the importance of

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

conducting an opinion analysis related to the MBG program and examples of the application of several related studies. Part 2 discusses the methods used, including an explanation of the theories used. Part 3 discusses the results obtained from the results of the experiment. Part 4 discusses the results in more depth. Part 5 is the conclusion and future work of this research.

2. METHODS

Figure 1 is a research diagram of this research. The research began with crawling tweet data through platform X. After obtaining the tweet data, the tweet data was cleaned and Exploratory Data Analytics (EDA). Furthermore, the TF-IDF process was carried out and then applied several ensemble-based evaluation models such as Random Forest, AdaBoost, XGBoost, Bagging, and Stacking.

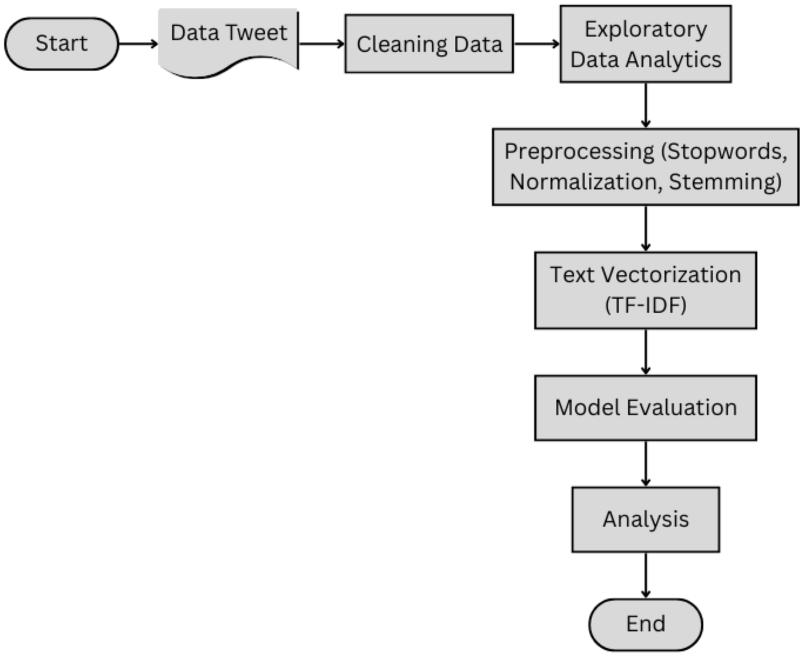

Figure 1. Proposed Research Diagram

2.1. Dataset

The dataset used in this study came from social media platform X in the form of tweets. The tweet data that was successfully crawled was 3417 tweets with a proportion of positive labels of 1,783 data and negative labels of 1,634 data. The data crawling period was carried out in the period from January to April 2025 were

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

adjusting the MBG program began to take effect. Table 1 is an example of the tweet data used.

Table 1. Data Tweet Sample

Tweet (Indonesia)	**Label**
@kumparan Klu saya tidak setuju bila MBG didanai dengan zakat karena klu MBG itu hanya untuk anak sekolah sedangkan zakat untuk kaum dhuafa misal besuk ganti presiden apa juga masih diprogramkan?	Negative
@03__nakula Kalau mbg nya makanan fresh sih tdk masalah.. Byk yg sdh jamuran atau sdh basi msh di distribusikan..	Negative
@LumowaS Bagaimana kalau anggaran MBG yg triliunan digunakan untuk menciptakan lap kerja atau memperbaiki sarpras sekolah/madrasah atau membangun jalan jembatan waduk listrik berapa banyak yg bisa terbantu tidak hanya sekedar jadi tai setelah makan	Negative
@BaseAnakFK MBG boleh dimulai pelan2 secara perlahan cth dimulai dari SD ya walaupun pasti nanti ada kecemburuan sosial tapi ya mau bagimana lagi jalanin aja kayak Sekolah Gratis/Program Wajib belajar (6th 9th 12th)	Positive
@hrdbacot secara umum tidak menyalahkan orang rusia yang buka loker di indonesia menurutku itu bagus buat indonesia juga jadi bisa kolaborasi transfer knowledge tapi ini kok tulisannya yg bisa gw baca kok cuma IT doang	Positive

2.2. Cleaning Data

In the data cleaning process, several important steps are taken to ensure the quality and reliability of the dataset used in sentiment classification. The first step is to remove missing values by checking if there are any rows in the dataset that have empty or undefined values in important columns such as tweet text or sentiment labels [10]. Data like this is removed because it doesn't provide useful information and can cause errors in later stages, such as tokenization, vectorization, or labelling during model training. The next step is to perform duplicate removal, which involves deleting identical or repeated entries. This duplication often originates from retweets or bot-generated content, which is quite common on social media platforms like X. If not removed, the presence of duplicates can cause the model to overfit, as it "remembers" the same patterns repeatedly too often [11], thus reducing the model's ability to generalize to new data. By removing the duplication, the dataset becomes cleaner and more representative of the actual distribution of public sentiment.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

2.3. Exploratory Data Analysis (EDA)

The EDA process is carried out to understand the distribution and characteristics of the data before entering the modelling stage [12]. The first step is a visualization of the distribution of sentiment labels that depicts the proportion of data labelled positive and negative. The aim is to identify whether there is a class imbalance in the dataset, which is important to know in sentiment analysis. Next, an analysis of the length of the text was carried out by counting the number of words in each entry using the len(x.split() function. The results were analysed using descriptive statistics to see the mean, minimum, maximum, and overall distribution. This distribution of text length is visualized with histograms and boxplots, which show the distribution and potential presence of outliers in the text data based on sentiment labels. In addition, EDA also includes visualizations of the words that appear most frequently in each of the sentiment classes. For each label (positive and negative), the 15 most popular words are counted based on the frequency with which they appear after the text cleanup process. These words are visualized in the form of horizontal bar charts to identify the dominant topics contained in public opinion of the MBG program.

2.4. Preprocessing Data

The data preprocessing step in this study was performed to prepare the text to be more structured and ready for use in the sentiment analysis stage [13], [14]. The initial step taken was to convert HTML characters, such as changing & to &, so that the original meaning of the text was preserved. Next, the system removed various elements that were irrelevant to sentiment analysis, including:

1) URL or link,
2) mention Twitter user account (for example @username),
3) hashtag (words that begin with the symbol #),
4) as well as non-alphabetic characters such as punctuation and symbols.

This cleaning is done using regular expressions to ensure that all these elements are automatically and consistently removed. After that, the entire text was converted to lowercase (case folding) to maintain consistency in word comparisons [15]. The next step is the tokenization process, which involves breaking down the text into words (tokens) [16], After that, normalization was carried out, which involved standardizing the variations in word spelling using a special dictionary. For example, the word "gk" was changed to "tidak". This process is important for unifying the meaning of various informal writing forms common on social media. Then, stop words were removed, which are common words that don't have significant meaning in sentiment analysis (such as "yang", "di", "ke"). However, some words that influence sentiment polarity, such as "tidak", "bukan", and "kurang", were retained because they play a role in forming negative meaning [17].

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

The next step is stemming, which involves changing words to their base form using an Indonesian stemmer. For example, the word "makanan" is changed to "makan" [18]. This process helps reduce text complexity and enrich the model's generalization in recognizing words with the same root. Finally, all the cleaned words are reassembled into sentences ready for use in the feature extraction and model training stages. This process produces two versions of the clean text, one with stemming and one without stemming, which are used for comparative analysis of model performance.

2.5. Text Vectorization

The TF-IDF (Term Frequency–Inverse Document Frequency) process is used as a method of representing text into numerical form so that it can be processed by machine learning algorithms [19]. TF-IDF works by giving weight to each word in a document based on two main components: term frequency (how often a word appears in a document) and inverse document frequency (how rarely the word appears throughout the document). Words that often appear in one document but rarely appear in another will be given higher weight, as they are considered more informative. The TF-IDF process is carried out using the TfidfVectorizer from the scikit-learn library. This feature is implemented twice, namely on text that has been cleaned with stemming (clean_text) and without stemming (clean_text_without_stemming). Each version of the text is converted into a high-dimensional numeric vector that represents the appearance of important words in the document. The result of this process is a dimensional matrix (number of documents x number of TF-IDF features), which is a representation of the main feature (X) in the classification model training process.

2.6. Model Evaluation

This study also conducted an experiment by comparing the performance of several ensemble learning-based models in classification tasks. The five models evaluated include Random Forest, AdaBoost, XGBoost, Bagging, and Stacking. The evaluation was conducted based on five key performance metrics, namely Accuracy, Precision, Recall, F1-Score, and ROC-AUC. Accuracy is used to measure the proportion of correct predictions to the total data. However, in sentiment analysis, this metric can be less representative when the amount of data for one class (e.g. neutral sentiment) is much more than the other, so this study did not use accuracy alone. Precision is important to ensure that predicted positive or negative sentiments are true, which is especially relevant in applications such as product reviews or public opinion [20]. Recall is used to find out the extent to which the model can capture all opinions with a particular sentiment. F1-Score was chosen because it provides a balance between precision and recall [21]. This is especially relevant on sentiment analysis tasks that often have an unbalanced class

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

distribution (e.g. more neutral reviews than positive or negative) [22]. The ROC-AUC is used to evaluate the model's ability to distinguish between different classes of sentiment in general at various thresholds.

3. RESULTS AND DISCUSSION

3.1. Experiment Performance

This study conducted a comparative analysis with and without stemming. The results of this experiment are shown in Table 2 and Table 3.

Table 2. Result of Experiment without Stemming

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Random Forest	**91.15**	**91.24**	**91.15**	**91.14**	**95.46**
AdaBoost	81.71	81.84	81.71	81.69	90.36
XGBoost	88.64	88.65	88.64	88.64	94.69
Bagging	87.90	87.93	87.90	87.9	94.21
Stacking	90.26	90.3	90.26	90.26	94.73

Based on the results of the experiment in Table 2, the Random Forest model achieved the highest score on all metrics. These results show that Random Forest can predict with high accuracy and consistency in distinguishing positive and negative classes in public sentiment. This is in line with the literature that states that Random Forest is effective in handling text data that has high dimensions and noise [22]. The stacking model ranks second with an accuracy of 90.26% and a ROC-AUC of 94.73%, close to Random Forest. This good performance indicates that combining the strengths of several basic models can improve the generalization of predictions in the analysis of public opinion.

The XGBoost model has a stable performance across all metrics, with an accuracy of 88.64% and a ROC-AUC of 94.69%. Suitable for more complex datasets, XGBoost can capture more subtle patterns in public opinion despite being slightly inferior to Random Forest and Stacking in terms of overall accuracy. Bagging model with an accuracy of 87.90% and an F1-Score of 87.90%, Bagging shows quite good performance but is still below XGBoost and Stacking. Its ability to reduce variance makes it stable but less than optimal in the context of diverse sentiment [23]. The AdaBoost model shows the lowest performance with an accuracy of 81.71% and a ROC-AUC of 90.36%. Although theoretically suitable for handling unbalanced data, in this study the model was less able to capture the diversity of opinions, likely due to its sensitivity to noise and data outliers.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 3. Experiment with Stemming

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Random Forest	**89.67**	**89.78**	**89.67**	**89.66**	**94.72**
AdaBoost	83.48	83.49	83.48	83.47	90
XGBoost	88.2	88.3	88.2	88.19	93.12
Bagging	87.46	87.46	87.46	87.46	93.52
Stacking	89.23	89.38	89.23	89.22	94.31

Based on the Table 3 the application of stemming, which is the process of changing words to their basic form influences the performance of the sentiment classification model. In this table, the five ensemble models are compared in terms of Accuracy, Precision, Recall, F1-Score, and ROC-AUC after the analyzed text is processed by stemming. The Random Forest model remains the best performing model overall, with an accuracy of 89.67% and the highest ROC-AUC of 94.72%. Despite a slight decline compared to no stemming (previously 91.15%), the model remains consistent and superior. This minor decline can occur because the stemming removes morphological context that may be relevant for the detection of sentimental nuances. The stacking model achieves an accuracy of 89.23% and a ROC-AUC of 94.31%, just slightly below Random Forest. The combination of several base models still shows stable and competitive performance even after the voting is done. This indicates that advanced ensemble approaches such as Stacking are quite resistant to variations in text representation.

The XGBoost model has strong performance with an accuracy of 88.20% and a ROC-AUC of 93.12%. This performance is slightly down from the no-stemming version, which suggests that the model is sensitive to the loss of subtle semantic information due to stemming. The Bagging model showed a stable performance (87.46% for all metrics) and a ROC-AUC of 93.52%. Interestingly, although it is simple, this model is quite resistant to changes in the representation of features such as stemming. The AdaBoost model remains the model with the lowest performance (Accuracy 83.48%, ROC-AUC 90.0%). Consistent with previous results, this lower performance can be attributed to AdaBoost's sensitivity to noise and the loss of important features due to stemming.

This study also visualized which words received the most positive and negative sentiments [24], [25]. Figure 2 and Figure 3 are the results of the visualization. Based on Figure 2, the words "gizi", "makan", and "gratis" show that the community appreciates the main essence of the MBG program. The words "dukung" and "sehat" reflect an affirmative attitude and positive expectations for the sustainability of the program. The appearance of the words "presiden" and "indonesia" indicates that most of the positive comments directly mention the

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

policy initiator and the context of his nationality. The words "anak", "papua", "banjir" also indicate that users associate the program with social impacts on vulnerable groups and specific regions, as well as the context of disasters. People with positive sentiments tend to highlight the substance of the MBG program and support for the government.

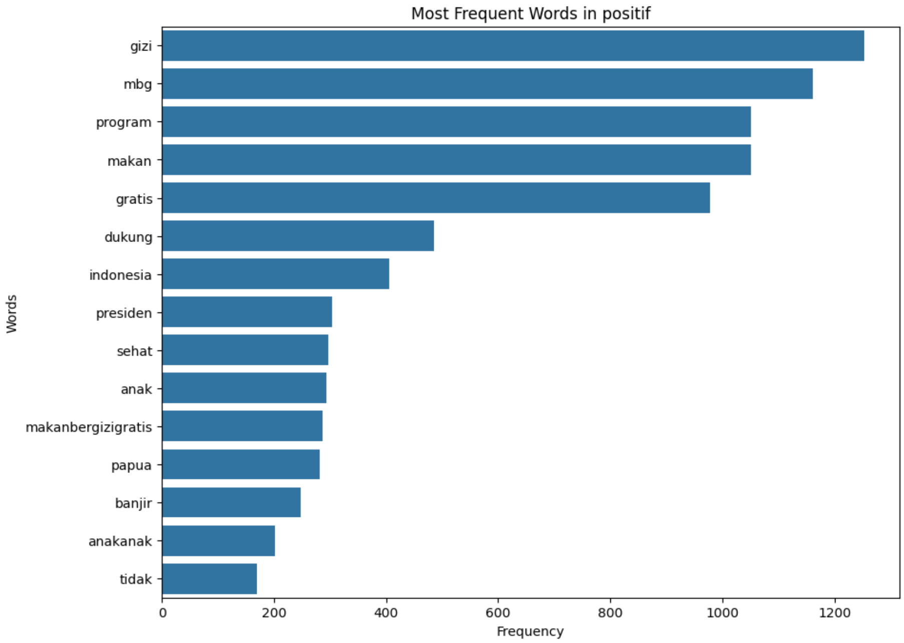

Figure 2. Most Frequent Words in Positive

Based on Figure 3, the words "tidak", "ya", "udah", and "nya" indicates a lot of informal comments, possibly in the form of complaints, sarcasm, or skepticism. The words "kerja", "rakyat", "pemerintah", and "sekolah" show a narrative of dissatisfaction with the program's implementation, affordability, or impact on the general public. Although the word "mbg", "makan", and "gizi" it also appears in negative sentiments, the context of which may be used sarcastically or as a criticism. Negative sentiment focuses more on doubts about the implementation of the program, such as not being on target, government performance, or unaffordability by the public. The words "sekolah", "anak", and "negara" showing concern that the parties who should have benefited have not felt it. This analysis shows that positive sentiment arises when the community responds to the ideal vision or benefits of the program, while negative sentiment arises when the realization on the ground is not as expected. These results reinforce the importance of aspect-based sentiment analysis approaches going forward, to distinguish sentiment based on specific aspects (e.g., food, distribution, school). Governments

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

and policymakers can leverage these insights to improve communication, benefit distribution, and transparency of program implementation.

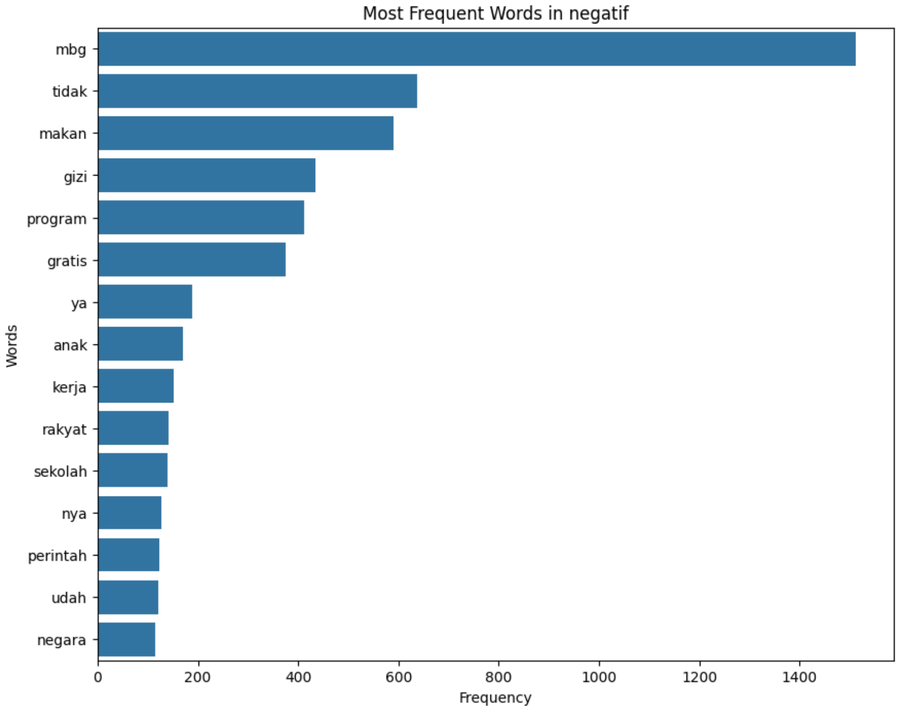

Figure 3. Most Frequent Words in Negative

3.2. Discussion

The results of this study conducted a comparative analysis of the performance of the ensemble model with and without the stemming process. The results of the experiment showed that Random Forest was consistently the best-performing model in the classification of public sentiment towards the MBG program. In Table 2 (without stemming), Random Forest recorded the highest score in all metrics, with an accuracy of 91.15%, an F1-Score of 91.14%, and the highest ROC-AUC of 95.46%. This shows that Random Forest can distinguish between positive and negative sentiment classes consistently and accurately. These results are in line with the literature that states that Random Forest effectively handles high-dimensional text data and is prone to noise [22].

The Stacking model ranks second with an accuracy of 90.26% and a ROC-AUC of 94.73%, indicating that the combination of several basic models can improve

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

the ability to generalize predictions in the analysis of public opinion. The XGBoost model also shows stable performance with an accuracy of 88.64% and a ROC-AUC of 94.69%, demonstrating its ability to capture fine patterns in the data despite being slightly lower than Random Forest and Stacking. Meanwhile, the Bagging model obtained an accuracy of 87.90%, with quite good performance but still below XGBoost and Stacking. The AdaBoost model has the lowest performance (accuracy of 81.71%, ROC-AUC of 90.36%) which is likely due to sensitivity to noise and outliers.

In Table 3 (with stemming), there is a decrease in performance in almost all models. Random Forest remained ahead with an accuracy of 89.67% and a ROC-AUC of 94.72%, although there was a slight decline compared to the non-voting period. This decline is suspected to be due to the stemming process eliminating the morphological context of the word which is important in capturing the nuances of sentiment. The Stacking model remained competitive with an accuracy of 89.23% and a ROC-AUC of 94.31%, indicating that the advanced ensemble method is quite resistant to variations in text representation. The XGBoost model also slightly decreased in performance (accuracy 88.20%, ROC-AUC 93.12%), indicating that the loss of semantic information due to stemming had an impact on its ability to understand context. The Bagging model is quite stable with an accuracy of 87.46% and the AdaBoost remains with the lowest performance (accuracy of 83.48%, ROC-AUC of 90.00%). Overall, the stemming process has a different impact on each model. Complex and flexible models like Random Forest and Stacking are more resistant to changes in text representation, while models like AdaBoost are more susceptible to the loss of important semantic features. This suggests that in the context of public opinion analysis of policy, a no-voting approach can provide a more informative representation.

Analysis of word frequency visualizations in positive and negative sentiment also supports these findings. In Figure 2, words such as *"gizi"*, *"makan"*, *"gratis"*, *"sehat"*, and *"dukung"* shows that the community appreciates the substance of the MBG program. Occurrence of the word *"presiden"*, *"indonesia"*, and *"anak"* show a sentiment of support for the government and concern for vulnerable groups. On the other hand, in Figure 3, words such as *"tidak"*, *"kerja"*, *"rakyat"*, *"pemerintah"*, and *"sekolah"* show doubts, criticism, and disappointment with the implementation of the MBG program in the field. Interestingly, words like *"mbg"*, *"makan"*, and *"gizi"* also appears in negative sentiments, possibly in the context of sarcasm or criticism of the implementation.

These results show that positive sentiment generally arises from expectations towards the program's ideal vision, while negative sentiment stems from a mismatch between expectations and realizations on the ground. Therefore, advanced approaches such as aspect-based sentiment analysis are highly

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

recommended to group sentiment based on specific aspects (e.g., food, distribution, education) [23]. This information can be used by the government and policy makers to develop communication strategies, benefit distribution, and transparency in program implementation.

The results of experiments showing the dominance of the Random Forest model's performance in the classification of public sentiment towards the MBG program have important implications, especially for policymakers. Word frequency analysis shows that positive sentiment generally arises from public expectations of the benefits of programs such as "gizi", "gratis", "sehat", and "anak". On the other hand, negative sentiment is more influenced by perceptions of the quality of implementation in the field, for example "tidak", "kerja", "rakyat", and "sekolah". his shows that aspects of food quality, equitable distribution, and accessibility of programs in disadvantaged areas are the most relevant dimensions to be followed up in follow-up policies.

Thus, the results of this analysis can be used to build a real-time public opinion monitoring system that helps stakeholders identify thematically dominant issues, such as dissatisfaction with distribution mechanisms, or concerns over food quality. In this context, further development in the form of aspect-based sentiment analysis (ABSA) is highly recommended in order to evaluate public perceptions of specific elements of the program, not just polarity classifications in general.mIn addition, compared to some previous studies such as [7] and [8], using Naive Bayes and SVM without ensemble learning integration, this study presents a more robust approach with systematic multi-model evaluation. This advantage lies in the use of five different ensemble methods as well as validation with two preprocessing conditions (with and without stemming), thus providing a more robust foundation in determining the most appropriate model for policy analysis based on public opinion.

To improve the validity of the results, a paired t-test was performed on the model's performance before and after the stemming process. The results showed that although there were no significant differences in the Accuracy, Precision, Recall, and F1-Score metrics ($p > 0.05$), there were significant differences in the ROC-AUC metrics ($p = 0.025 < 0.05$). This suggests that stemming can have a significant impact on the model's discriminating ability, although it does not affect the general accuracy of the classification consistently.

4. CONCLUSION

This study demonstrates that the Random Forest algorithm delivers the best performance in classifying public sentiment toward the *Makan Bergizi Gratis* (MBG) program compared to other ensemble models. With an accuracy of 91.15% and a

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

ROC-AUC of 95.46% on data without stemming, Random Forest has proven to be reliable in handling large and unstructured text data. Meanwhile, the Stacking and XGBoost methods also showed competitive performance, although there was a slight decrease when applied to data that had undergone stemming. On the other hand, the AdaBoost model exhibited the lowest performance, presumably due to its sensitivity to noise and the loss of semantic information caused by pre-processing. The visualization of the most frequently occurring words in each sentiment class shows that positive sentiment is dominated by support for the benefits of the program, while negative sentiment tends to focus more on criticism of its implementation in the field. These findings reinforce the importance of an aspect-based sentiment analysis (ABSA) approach to detect sentiment based on specific aspects of the policy, such as food quality, distribution, and accessibility.

As a recommendation for practical implementation, the proposed model can be directly applied by government agencies to monitor public opinion on various other social policy programs, such as cash assistance, free education programs, or healthcare subsidies. With minor adjustments to the training data and domain context, this approach can also be scaled to handle much larger datasets, such as nationwide social media reviews or multi-platform sources (Twitter, TikTok, Instagram). For future research, it is recommended to integrate this model with deep learning approaches such as IndoBERT to enhance deeper semantic understanding. In addition, techniques such as real-time streaming sentiment analysis could also be explored to support rapid decision-making needs by policymakers.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the generous support and funding provided by Telkom University through its internal research scheme, under contract number **186/LIT06/PPM-LIT/2025**. This support has been instrumental in enabling the completion of this research work.

REFERENCES

[1] N. K. Arora and I. Mishra, "United Nations Sustainable Development Goals 2030 and environmental sustainability: race against time," *Environ. Sustain.*, vol. 2, no. 4, pp. 339–342, Dec. 2019.
[2] J.-G. Shim, K.-H. Ryu, S. H. Lee, E.-A. Cho, Y. J. Lee, and J. H. Ahn, "Text mining approaches to analyze public sentiment changes regarding COVID-19 vaccines on social media in Korea," *Int. J. Environ. Res. Public Health*, vol. 18, no. 12, p. 6549, June 2021.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[3] M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, and U. R. Acharrya, "A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets," *Knowl. Based Syst.*, vol. 228, no. 107242, p. 107242, Sept. 2021.
[4] Z. Jianqiang, G. Xiaolin, and Z. Xuejun, "Deep convolution neural networks for twitter sentiment analysis," *IEEE Access*, vol. 6, pp. 23253–23260, 2018.
[5] N. G. Ramadhan, Adiwijaya, W. Maharani, and A. Akbar Gozali, "Chronic diseases prediction using machine learning with data preprocessing handling: a critical review," *IEEE Access*, vol. 12, pp. 80698–80730, 2024.
[6] A. Singh, N. Thakur, and A. Sharma, "A review of supervised machine learning algorithms," in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, pp. 1310–1315.
[7] B. Rahmatullah, S. A. Saputra, P. Budiono, and D. P. Wigandi, "Sentimen Analisis Makan Bergizi Gratis Menggunakan Algoritma Naive Bayes," *JIfoTech*, vol. 5, no. 1, Mar. 2025.
[8] Elsa Triningsih, "Analisis Sentimen Terhadap Program Makan Bergizi Gratis Menggunakan Algoritma Machine Learning Pada Sosial Media X," *BUILDING OF INFORMATICS, TECHNOLOGY AND SCIENCE (BITS)*, vol. 6, no. 4, Jan. 2025.
[9] W. Anggriyani, "Analisis Sentimen Program Makan Gratis Pada Media Sosial X Menggunakan Metode NLP," *JoSYC: Journal of Computer System and Informatics*, vol. 5, no. 4, 2024.
[10] A. Uddin, X. Tao, C.-C. Chou, and D. Yu, "Are missing values important for earnings forecast? a machine learning perspective," *Quant. Finance*, vol. 22, no. 6, pp. 1113–1132, Jan. 2022.
[11] Arpita, P. Kumar, and K. Garg, "Data cleaning of raw tweets for sentiment analysis," in *2020 Indo – Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN)*, Rajpura, Punjab, India, 2020, pp. 273–276.
[12] S. Aldera, A. Emam, M. Al-Qurishi, M. Alrubaian, and A. Alothaim, "Exploratory data analysis and classification of a new Arabic online extremism dataset," *IEEE Access*, vol. 9, pp. 161613–161626, 2021.
[13] H. Mohamed Zakir and S. Vinila Jinny, "A comparative study on data cleaning approaches in sentiment analysis," in *Lecture Notes in Electrical Engineering*, Singapore: Springer Singapore, 2020, pp. 421–431.
[14] H.-T. Duong and T.-A. Nguyen-Thi, "A review: preprocessing techniques and data augmentation for sentiment analysis," *Comput. Soc. Netw.*, vol. 8, no. 1, Dec. 2021.
[15] K. Afifah, I. N. Yulita, and I. Sarathan, "Sentiment analysis on telemedicine app reviews using XGBoost classifier," in *2021 International Conference on Artificial Intelligence and Big Data Analytics*, Bandung, Indonesia, 2021, pp. 22–27.
[16] W. Kim, "Comparative study of tokenizer based on learning for sentiment analysis," *Journal of the Korean Society for Quality Management*, vol. 48, Sept. 2020.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[17] I. Steinke, J. Wier, L. Simon, and R. I. Seetan, "Sentiment analysis of online movie reviews using machine learning," *Int. J. Adv. Comput. Sci. Appl.*, 2022.

[18] A. Jabbar, S. Iqbal, M. I. Tamimy, A. Rehman, S. A. Bahaj, and T. Saba, "An analytical analysis of text stemming methodologies in information retrieval and natural language processing systems," *IEEE Access*, vol. 11, pp. 133681–133702, 2023.

[19] A. Mee, E. Homapour, F. Chiclana, and O. Engel, "Sentiment analysis using TF–IDF weighting of UK MPs' tweets on Brexit," *Knowl. Based Syst.*, vol. 228, no. 107238, p. 107238, Sept. 2021.

[20] V. Balakrishnan, S. Khan, and H. R. Arabnia, "Improving cyberbullying detection using Twitter users' psychological features and machine learning," *Comput. Secur.*, vol. 90, no. 101710, p. 101710, Mar. 2020.

[21] N. G. Ramadhan and F. Adhinata, "Sentiment analysis on vaccine COVID-19 using word count and Gaussian Naïve Bayes," *Indones. J. Electr. Eng. Comput. Sci.*, June 2022.

[22] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022.

[23] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, "A survey on aspect-based sentiment analysis: Tasks, methods, and challenges," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11019–11038, Nov. 2023.

[24] S. Sharma, "SentiNet: A Word-Cloud based approach towards Social Media Sentiment Analysis," in *2024 IEEE International Conference on Contemporary Computing and Communications (InC4)*, Bangalore, India, 2024, vol. 821, pp. 1–6.

[25] A. Verma and R. Vashisth, "Sentiment analysis of Google play and app store reviews of threads," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, 2024, pp. 1–5.