Jurnal Matematika Thales (JMT)
Vol. 07, No. 01 (2025), pp. 35–53.

Public Sentiment Analysis and Distribution Optimization MBG

Akhmad Sultoni[1*], Dimaz Andhika Putra[2], Hidayah Noor Wahidah [3], Muhammad Mahathir Arief [4], and Pelangi Jingga Putria R.M. [5]

[1,2,3,4,5]**Department of Mathematics, University of Gadjah Mada, Yogyakarta, Indonesia**

[1]akhmad.s@mail.ugm.ac.id, [2]dimazandhikaputra@mail.ugm.ac.id,
[3]hidayah.noor.wahidah@mail.ugm.ac.id, [4]muhammadmahathirarief@mail.ugm.ac.id,
[5]pelangijinggaputriarisnamaulana@mail.ugm.ac.id

Abstract. This study analyzes public sentiment and regional prioritization regarding Indonesia's Makan Bergizi Gratis (MBG) program, a national initiative aimed at reducing stunting through the distribution of free meals to schoolchildren and pregnant women. Sentiment analysis was conducted on 47,803 posts from the social media platform X (formerly Twitter) using a lexicon-based labeling method and TF-IDF feature extraction. The results show that 22,504 posts (47.1%) expressed positive sentiment, 20,010 (41.9%) negative, and 5,289 (11.0%) neutral, indicating strong public support accompanied by considerable concerns. Eleven classification models were evaluated, with the Linear Support Vector Machine (SVM) achieving the highest accuracy (96.33%), and BERT-based models also demonstrating strong performance. Latent Dirichlet Allocation (LDA) topic modeling revealed five major themes in the negative sentiment, including transparency issues, maternal and child health, and inequality of access. Furthermore, provincial-level clustering using the K-Means algorithm grouped regions into three priority levels based on socioeconomic and health indicators. These findings provide critical insights for optimizing policy targeting and efficient resource allocation in the implementation of the MBG program.

Keywords: Sentiment Analysis, Makan Bergizi Gratis, Clustering, Stunting.

---

*Corresponding author
2020 Mathematics Subject Classification: 91C20
Received: 05-07-2025, accepted: 18-09-2025.

1. INTRODUCTION

Stunting is one of the chronic nutritional problems that has become a major concern for the Indonesian government. Based on the 2023 Indonesian Nutrition Status Survey, the government has targeted a reduction in the stunting rate to 14% by 2024, considering that its prevalence was still at 21.6% in 2022 [1]. In addition, malnutrition remains a serious public health issue in Indonesia, particularly among children. According to data from the Ministry of Health, approximately 3.8% of children under five in Indonesia experience malnutrition. This highlights a significant challenge in efforts to improve child nutrition and health. Various initiatives have been carried out to address this issue, one of which is the *Makan Bergizi Gratis* (MBG) program, formerly known as the Free Lunch Program. This program is an initiative of the President and Vice President elected in 2024, aimed at improving the nutritional intake of children and pregnant women by providing free lunches and milk at schools, pesantren (Islamic boarding schools), and for pregnant women, as part of the national stunting alleviation strategy [2].

Although the *Makan Bergizi Gratis* program offers potential benefits in efforts to combat stunting, it has also sparked controversy among the public, particularly on social media platforms such as X (formerly Twitter). One of the main issues under scrutiny is the substantial budget required, which is estimated to reach Rp450 trillion. Furthermore, the proposed funding plan—reportedly involving the use of the State Budget (APBN) for the education sector and School Operational Assistance (BOS) funds—has triggered opposition, including from the Indonesian Teachers Union Federation (FSGI) [3]. Concerns have also been raised regarding the potential impact of this policy on education costs, teacher welfare, and the overall financial stability of the country. Amid this public debate, doubts have emerged about the program's feasibility and long-term sustainability [4]. The growing polarization of public opinion highlights the need for sentiment analysis to better understand public perceptions of the policy, which could serve as valuable input for policymakers in evaluating and refining the program.

Sentiment analysis is a computational process aimed at evaluating individuals' opinions, feelings, and emotions toward an entity, event, or related attribute [5]. Its primary focus is to identify the polarity of a text—whether it is positive, negative, or neutral [6]. In today's digital era, people are increasingly active in using social media as a medium to express their views and emotions on various issues, including public policies. Therefore, it is important to analyze public sentiment and its changes over time to gain deeper insights into the public's responses to current issues [7]. In this study, data were collected from the social media platform X, which is widely used by users to discuss and respond to trending topics. In addition to serving as a space for expression, this platform is also considered a representative and reliable source of data for capturing public perception and opinion on ongoing policies and social phenomena, including *Makan Bergizi Gratis* [8].

In conducting sentiment classification, this study employs the Support Vector Machine (SVM) algorithm, which is known as one of the most effective and reliable

methods for text classification tasks [9]. SVM works by separating data into two or more classes through the identification of an optimal hyperplane that maximizes the margin between data points from each class [10]. In other words, the algorithm seeks the best decision boundary that minimizes classification errors. Its ability to handle high-dimensional data and its robustness against overfitting make SVM particularly suitable for text-based sentiment analysis, especially when dealing with social media data, which tends to be diverse and unstructured.

In addition to analyzing public sentiment, this study also aims to identify the level of need for the *Makan Bergizi Gratis* (MBG) program across various provinces in Indonesia. The identification process utilizes data from the Central Bureau of Statistics (BPS), which includes key indicators such as population size, stunting prevalence, poverty rate, average income, and education level. To classify the provinces based on their level of need, a clustering method is employed, allowing the division of regions into three categories: high, moderate, and low need. The results of this clustering are expected to provide a solid foundation for determining priority target areas, enabling the design of policies that are more focused, well-targeted, and efficient in accelerating the national effort to reduce stunting.

2. Related Work

Previous studies on sentiment analysis of users on platform X toward the *Makan Bergizi Gratis* (MBG) program include works by [11], [12], and [8]. These studies were limited to the use of three classification algorithms: Support Vector Machine (SVM), Random Forest, and Naive Bayes Classifier. In contrast, our study not only employs a broader range of machine learning models for sentiment classification, but also introduces a regional clustering analysis to identify variations in MBG-related needs across different provinces in Indonesia.

3. Methodology

3.1. **Data Collection.** This study utilizes two main datasets corresponding to the sentiment analysis and the clustering of MBG needs across provinces.

For the sentiment analysis, textual data were collected from the social media platform **X** (formerly Twitter). The data acquisition was conducted using keyword-based scraping with search terms such as *"Makan Bergizi Gratis," "Program Gizi," "Stunting,"* and related hashtags. The collected data consist of user-generated posts reflecting public opinion on the MBG program. The complete dataset can be accessed via the following link:
`https://drive.google.com/drive/folders/1OAb9G2avR0fv_BL82uLIxPeeXOVWG6qV`

For the MBG clustering analysis, we obtained structured quantitative data from the official portal of **Badan Pusat Statistik (BPS) Republik Indonesia**. The dataset includes a comprehensive set of socio-economic and health indicators at the provincial level, such as:

- Human Development Index (HDI)
- Gini Ratio
- Total Population
- Special Index for Stunting Management
- Stunting Prevalence among Children Under Five
- Open Unemployment Rate
- Number of Families at Risk of Stunting
- Percentage of Population Living Below the Poverty Line
- Poverty Depth Index
- Poverty Severity Index
- Average Hourly Wage
- Expenditure per Capita
- Prevalence of Inadequate Food Consumption

All indicators were compiled and normalized for clustering purposes. The processed dataset is available at:
`https://docs.google.com/spreadsheets/d/1nFo6YTkCv-it7_EHkw2T9_BkBnBzwMpJh_rysChgEGY`

3.2. **Sentiment Analysis.** Sentiment analysis is the process of analyzing textual data to determine its polarity, i.e., whether the opinion is positive, negative, or neutral. In this study, sentiment analysis is applied to X posts related to the *Makan Bergizi Gratis* program.

3.2.1. *Preprocessing.* The preprocessing steps are critical to ensure the quality of input data [5].

- **Cleaning:** Removing emojis, URLs, symbols, and punctuations.
- **Case folding:** Converting all characters to lowercase.
- **Normalization:** Replacing informal or slang words with their formal equivalents.
- **Tokenization:** Splitting text into individual words (tokens).
- **Stopword removal:** Eliminating common words that carry little semantic value.
- **Stemming:** Reducing words to their root forms using the Sastrawi stemmer.

3.2.2. *Labeling.* Sentiment labeling is conducted using the lexicon-based method with the InSet lexicon [13], which provides lists of positive and negative words. Each text is assigned a score:

- *Positive* if score > 0
- *Negative* if score < 0
- *Neutral* if score $= 0$

The score is calculated as:

$$\text{Sentiment Score} = \text{Positive Words} - \text{Negative Words} \tag{1}$$

3.2.3. *Feature Extraction.* Each document is transformed into a numerical vector using Term Frequency–Inverse Document Frequency (TF-IDF) [14]:

$$\text{TF-IDF}(t,d) = tf(t,d) \times \log\left(\frac{N}{df(t)}\right) \tag{2}$$

where $tf(t,d)$ represents the frequency of term t in document d; $df(t)$ denotes the number of documents in the corpus that contain the term t; and N is the total number of documents in the corpus. These components are used to compute the TF-IDF weight, which reflects how important a word is to a document in a given collection.

3.2.4. *Classification Modeling.* Eleven machine learning models are used:

- **Logistic Regression**
  Logistic Regression is a statistical method used for binary classification problems. It models the probability that a given input x belongs to a certain class $y = 1$ using the sigmoid (logistic) function:

$$P(y=1|x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\cdots+\beta_n x_n)}} \tag{3}$$

  where x_i denotes the i-th feature, β_i represents the coefficient for feature x_i, and β_0 is the intercept term. This method is interpretable and effective for linearly separable data [15].
- **Multinomial Naive Bayes**
  Multinomial Naive Bayes is a generative model based on Bayes' theorem, widely used for text classification. It assumes that features (typically word frequencies) are conditionally independent given the class. The classification rule is given by:

$$P(c|x) = \frac{P(c)\prod_{i=1}^{n} P(x_i|c)}{P(x)} \tag{4}$$

  where c is the class label, $x = (x_1, x_2, ..., x_n)$ is the feature vector representing word counts or frequencies, and $P(x_i|c)$ is the likelihood of word x_i given class c. It is simple, fast, and suitable for high-dimensional problems [16, 17, 18].
- **Support Vector Machine (SVM)**
  Support Vector Machine (SVM) is a powerful supervised learning model used for classification and regression tasks. It identifies the optimal hyperplane that maximally separates different class labels. For linear classification:

$$f(x) = \text{sign}(w \cdot x + b) \tag{5}$$

  where w is the weight vector, x is the input feature vector, and b is the bias term. For non-linear data, kernel functions such as the radial basis function (RBF) are employed:

$$K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right) \tag{6}$$

with γ controlling the influence of a single training example. SVM excels in both linearly and non-linearly separable data [19, 20, 21].

- **Random Forest**
  Random Forest is an ensemble learning method that constructs multiple decision trees and merges their outputs through majority voting for classification tasks:

$$\hat{y} = \text{majority\_vote}(h_1(x), h_2(x), \ldots, h_n(x)) \tag{7}$$

  where $h_i(x)$ is the prediction of the i-th decision tree for input x. This technique improves predictive accuracy and reduces overfitting [22, 23, 24].
- **AdaBoost**
  AdaBoost, or Adaptive Boosting, is an ensemble method that combines multiple weak learners in a sequential manner. Each subsequent model focuses on instances that were misclassified by previous ones. The weight of each learner is computed as:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \tag{8}$$

  where ϵ_t is the error rate of the t-th weak classifier, and α_t is its weight. Though sensitive to noisy data, AdaBoost can enhance model accuracy significantly [25].
- **XGBoost**
  XGBoost (Extreme Gradient Boosting) is an advanced implementation of gradient boosting algorithms. It incorporates regularization and second-order derivatives for enhanced performance. The loss function at iteration t is approximated by:

$$\mathcal{L}^{(t)} \approx \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t) \tag{9}$$

  where g_i and h_i are the first and second-order gradients of the loss with respect to predictions $f_t(x_i)$, and Ω is a regularization term. XGBoost is efficient, scalable, and robust against overfitting [26].
- **LightGBM**
  LightGBM is a fast and scalable gradient boosting framework that uses histogram-based algorithms and grows trees leaf-wise with depth constraints. It is optimized for memory usage and training time, making it suitable for large-scale data processing [27].
- **BERT (Bidirectional Encoder Representations from Transformers)**
  BERT is a transformer-based language model that captures the full context of a word by looking at both its left and right surroundings. The core mechanism is self-attention:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \tag{10}$$

where Q (queries), K (keys), and V (values) are matrices derived from the input, and d_k is the dimensionality of the keys. BERT excels in a variety of NLP tasks including sentiment analysis, question answering, and more [28].

- **DistilBERT**
  DistilBERT is a distilled version of BERT that retains most of its performance while being smaller and faster. It is trained using knowledge distillation to match the behavior of BERT with fewer parameters and computations, making it ideal for real-time or low-resource applications [29].
- **IndoBERTweet**
  IndoBERTweet is a variant of BERT pre-trained on Indonesian tweets. It is tailored to understand informal language, slang, abbreviations, and other characteristics unique to Indonesian social media. This makes it especially effective for sentiment analysis and opinion mining in Indonesian contexts [30].

Classical models are trained on TF-IDF vectors using `scikit-learn` [31]. Transformer-based models utilize pre-trained embeddings from Indonesian language models such as IndoBERTweet, powered by HuggingFace Transformers [32].

3.2.5. *Evaluation Metrics.* Model performance is evaluated using:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{12}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{13}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{14}$$

3.3. **Topic Modeling.** Topic modeling is an unsupervised learning technique used to discover hidden thematic structures in text corpora.

3.3.1. *Latent Dirichlet Allocation (LDA).* Topic analysis was conducted using the Latent Dirichlet Allocation (LDA) method, an unsupervised learning algorithm commonly used to uncover latent thematic structures within a collection of documents [33]. LDA operates under the assumption that each document is a mixture of multiple topics, and each topic is characterized by a specific distribution over words.

Prior to model training, data categorized as negative sentiment underwent a series of preprocessing steps, including case folding, removal of non-alphabetic characters, tokenization, and stopword removal using an Indonesian stopword dictionary. The resulting tokens were then converted into a bag-of-words representation and further processed into a corpus and dictionary using the `Gensim` library [34].

The probability of a topic $z_i = k$ for word w_i in document d_i is given by:

$$P(z_i = k | w_i = w, d_i = d) \propto \frac{n_{dk}^{-i} + \alpha}{n_d^{-i} + K\alpha} \cdot \frac{n_{kw}^{-i} + \beta}{n_k^{-i} + V\beta} \tag{15}$$

where n_{dk}^{-i} denotes the number of words in document d that are assigned to topic k, excluding the current word i; n_{kw}^{-i} represents the number of times word w is assigned to topic k, also excluding the current word. The parameters α and β are Dirichlet hyperparameters that control the sparsity of topic and word distributions, respectively. K refers to the total number of latent topics assumed in the model, while V denotes the vocabulary size, or the number of unique terms in the corpus. These parameters are used in the Gibbs sampling update equation for estimating the topic distribution in Latent Dirichlet Allocation.

The LDA model was trained using five topics and ten passes (iterations), producing outputs in the form of dominant keywords for each topic, along with topic distributions across the documents. Model visualization was carried out using the `pyLDAvis` library to enable interactive interpretation of inter-topic relationships.

3.4. **Clustering Analysis.**

3.4.1. *K-Means Clustering.* **Basic Concept of K-Means**

K-Means clustering is one of the most widely used unsupervised learning algorithms for data grouping [35]. Its main objective is to partition a set of n data points into k distinct clusters based on similarity. The algorithm aims to assign each observation to the cluster with the nearest mean, known as the *centroid*, thereby minimizing intra-cluster variance. The centroid is calculated as the average of all data points in a given cluster. K-Means thus seeks to minimize the total within-cluster variation, also known as the sum of squared errors (SSE), ensuring each cluster is as homogeneous as possible.

Algorithm Steps

The K-Means algorithm typically follows these steps [36, 37]:

(1) **Initialization:** Choose the number of clusters k to form.
(2) **Initial Centroids:** Randomly select k initial centroids from the dataset.
(3) **Assignment Step:** Assign each data point to the nearest centroid using a distance metric, commonly Euclidean distance.
(4) **Update Step:** Recalculate the centroid of each cluster by computing the mean of all points assigned to that cluster.
(5) **Convergence Check:** Repeat steps 3 and 4 until convergence, i.e., when there are no further changes in cluster assignments or centroids.

Mathematical Formulation

Euclidean Distance

The Euclidean distance between a data point $\mathbf{x}$ and a centroid $\boldsymbol{\mu}$ in d-dimensional

space is calculated as:

$$d(\mathbf{x}, \boldsymbol{\mu}) = \sqrt{\sum_{j=1}^{d} (x_j - \mu_j)^2} \tag{16}$$

Cluster Centroid Calculation
The centroid μ_k of cluster k is computed as the mean of all N_k points in that cluster [37]:

$$\mu_k = \frac{1}{N_k} \sum_{x \in S_k} x \tag{17}$$

where S_k is the set of data points in cluster k, and $N_k = |S_k|$ is the number of points in that cluster.
Objective Function (SSE)
The objective of K-Means is to minimize the total sum of squared distances between data points and their respective cluster centroids [36]:

$$J = \sum_{k=1}^{K} \sum_{x \in S_k} \|x - \mu_k\|^2 \tag{18}$$

This function, known as the within-cluster sum of squares (WCSS), quantifies the compactness of the clusters. The algorithm stops when J converges to a local minimum.

3.4.2. *K-Median Clustering.* **Basic Concept**
K-Median clustering is an alternative to K-Means that uses median values instead of means to define cluster centers. It is particularly useful when the dataset contains outliers or is not normally distributed. Unlike K-Means, which minimizes the sum of squared Euclidean distances, K-Median minimizes the sum of absolute distances (Manhattan distances) between data points and the cluster centroids (medians). This makes K-Median more robust to noise and extreme values [38].

Algorithm Steps
The K-Median algorithm proceeds as follows:

(1) **Initialize:** Choose the number of clusters k and randomly select k initial medians.
(2) **Assignment:** Assign each data point to the nearest median based on Manhattan distance.
(3) **Update:** For each cluster, update the median by computing the component-wise median of the assigned points.
(4) **Repeat:** Iterate steps 2 and 3 until cluster assignments no longer change or a convergence criterion is met.

Mathematical Formulation
Given a dataset $X = \{x_1, x_2, \ldots, x_n\}$ and a set of cluster centers $\{m_1, m_2, \ldots, m_k\}$,

the K-Median objective function is:

$$J = \sum_{i=1}^{n} \min_{j \in \{1,\ldots,k\}} \|x_i - m_j\|_1 \tag{19}$$

Here, $\|\cdot\|_1$ denotes the Manhattan (L1) norm. Each cluster center m_j is updated as the median of all data points assigned to cluster j:

$$m_j = \text{median}\{x_i \mid x_i \in C_j\} \tag{20}$$

Comparison with K-Means
While K-Means minimizes the sum of squared Euclidean distances and is sensitive to outliers, K-Median is more robust by minimizing the sum of absolute deviations. This makes K-Median particularly effective for applications involving skewed or noisy data.

3.4.3. *Clustering Analysis with Genetic Algorithm.* Genetic algorithm is a search and optimization method inspired by the principles of natural selection and biological genetics. This approach begins by generating a number of random solutions that form a population of chromosomes. Through evolutionary stages—including selection, crossover, and mutation—the algorithm aims to find a globally optimal solution. Selection is carried out by choosing chromosomes with the highest fitness values to form a new generation. The crossover process then combines genetic information from two parents to produce offspring with superior characteristics [39]. Genetic algorithm is applied to K-Means clustering to enhance clustering performance. The performance of K-Means clustering is known to be sensitive to suboptimal initial centroid selection. By using a genetic algorithm, the search for more representative cluster centers can be conducted more thoroughly.

In the implementation of genetic algorithm for K-Means clustering, the process begins by forming an initial population consisting of candidate solutions in the form of different centroid positions. Each chromosome represents a set of centroids as a potential solution. Evaluation of each chromosome is done by calculating its fitness value, typically using the within-cluster sum of squares, which measures how well the centroids divide the data. Selection then chooses the best-performing chromosomes to generate the next generation, followed by a crossover process that combines characteristics from two parent solutions to produce improved offspring. Random mutation is applied to maintain population diversity and prevent convergence to local optima. The processes of selection, crossover, and mutation are repeated over several generations until convergence is achieved or the best fitness solution is found, resulting in more optimal initial centroids for clustering.

3.5. **Experimental Setup.** This study consists of two main experimental components: sentiment analysis and provincial clustering.

3.5.1. *Sentiment Analysis.* To capture public response toward the Makan Bergizi Gratis (MBG) program, we collected textual data from social media platform **X**

(formerly Twitter) using relevant keywords and hashtags (e.g., #MBG, #MakanBergiziGratis, #ProgramGizi). The dataset was preprocessed using standard NLP techniques such as case-folding, tokenization, stopword removal, and stemming.

We implemented a supervised machine learning model for sentiment classification, categorizing each post as *positive*, *negative*, or *neutral*. We used the Python `scikit-learn` library, employing a TF-IDF vectorizer for feature extraction and a Logistic Regression classifier. Hyperparameters were tuned using grid search with 5-fold cross-validation. Performance metrics such as accuracy, precision, recall, and F1-score were evaluated on a 20% hold-out test set.

3.5.2. *Clustering of Provincial MBG Needs.* In the second phase, we analyzed official provincial-level indicators from Badan Pusat Statistik (BPS), including stunting prevalence, poverty rate, population, average income, and education level. The data were normalized using Min-Max scaling.

We applied K-Means clustering to group provinces based on their relative need for the MBG program. The optimal number of clusters (k) was determined using the Elbow Method and Silhouette Coefficient. Provinces were then categorized into three priority levels: *high need*, *moderate need*, and *low need*.

3.5.3. *Computational Environment.* All experiments were conducted using Python on the **Google Colab** platform, which provides cloud-based computation and interactive visualization capabilities. The main libraries used include `pandas` for data manipulation, `scikit-learn` for machine learning modeling, and `matplotlib` and `seaborn` for data visualization.

Sentiment data was collected from the social media platform **X** (formerly Twitter) via its API using keywords and hashtags related to the MBG program. Meanwhile, provincial-level indicator data was obtained from the official website of Badan Pusat Statistik (BPS) and manually compiled into a structured dataset using Google Sheets. The dataset can be accessed at the following link:
`https://docs.google.com/spreadsheets/d/1nFo6YTkCv-it7_EHkw2T9_BkBnBzwMpJh_rysChgEGY/edit?usp=sharing`

All source code and experimental documentation are publicly available for reproducibility at the following links:

- **Sentiment Analysis**: `https://drive.google.com/drive/folders/1OAb9G2avR0fv_BL82uLIxPeeXOVWG6qV?usp=drive_link`
- **MBG Clustering by Province**: `https://colab.research.google.com/drive/1EJgCbXpF8VIppbvfCMVghH3ZrGQfYC6A?usp=sharing`

4. Results and Discussion

4.1. Sentiment Analysis.

4.1.1. *Data Preprocessing.* The preprocessing stage begins with data cleaning, which involves the removal of emojis, symbols, URLs, and irrelevant punctuation. This is followed by case folding to standardize all characters to lowercase and the normalization of informal or slang words into their formal equivalents. The next step is

tokenization, where sentences are segmented into individual words. Subsequently, stopword removal is applied to eliminate common words that carry little semantic weight in the context of the analysis. Finally, stemming is performed using the Sastrawi library to reduce words to their root forms, thereby enhancing the quality of textual features used in the classification model.

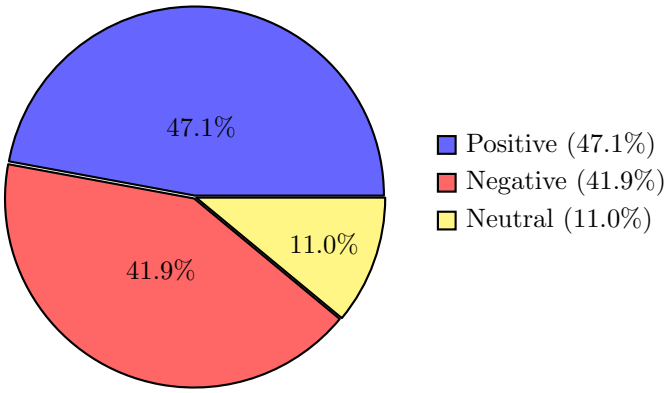

FIGURE 1. Sentiment distribution in the dataset. The legend shows each sentiment label and its corresponding percentage.

4.1.2. *Labeling.* Sentiment labeling was conducted using a lexicon-based approach, categorizing each text into positive, negative, or neutral sentiment. As shown in Figure (1), the dataset contains 22,504 positive samples (47.1%), 20,010 negative samples (41.9%), and 5,289 neutral samples (11.0%). This distribution indicates a generally positive trend in public sentiment, although a substantial portion also expresses negative opinions, suggesting ongoing debates or concerns regarding the topic.

4.1.3. *Feature Extraction.* Feature extraction was performed using the Term Frequency Inverse Document Frequency (TF-IDF) method, with a maximum of 5,000 features. This method generates a numerical representation of the text by quantifying the importance of each word in a document relative to the entire corpus.

4.1.4. *Modeling.* The dataset was divided into training and testing sets in an 80:20 ratio using stratified sampling to preserve the distribution of sentiment classes. A total of eleven classification models were employed in this study, comprising eight traditional machine learning models—Logistic Regression, Multinomial Naive Bayes, Support Vector Machines (with both Linear and RBF kernels), Random Forest, AdaBoost, XGBoost, and LightGBM—and four transformer-based models: BERT, DistilBERT, BERTweet, and IndoBERTweet.

4.1.5. *Model Evaluation.* Performance comparison among 11 sentiment classification models is presented in Table (1), covering both traditional machine learning and transformer-based approaches. The metrics evaluated include accuracy, precision, recall, and F1-score for both positive and negative classes.

TABLE 1. Performance Evaluation of Sentiment Classification Models

Model	Acc	P_{Pos}	R_{Pos}	$F1_{Pos}$	P_{Neg}	R_{Neg}	$F1_{Neg}$
Logistic Regression	0.9410	0.9496	0.9382	0.9439	0.9315	0.9440	0.9377
Multinomial NB	0.7976	0.8018	0.8205	0.8110	0.7927	0.7719	0.7821
SVM (Linear)	**0.9633**	**0.9653**	**0.9653**	**0.9653**	**0.9610**	**0.9610**	**0.9610**
SVM (RBF)	0.9467	0.9552	0.9436	0.9494	0.9374	0.9503	0.9438
Random Forest	0.8523	0.8603	0.8607	0.8605	0.8433	0.8428	0.8430
AdaBoost	0.6953	0.7870	0.5819	0.6691	0.6363	0.8228	0.7177
XGBoost	0.8578	0.8893	0.8354	0.8615	0.8267	0.8831	0.8539
LightGBM	0.8745	0.8978	0.8609	0.8790	0.8505	0.8898	0.8697
BERT	0.9153	0.9219	0.9178	0.9198	0.9080	0.9125	0.9103
DistilBERT	0.9066	0.9125	0.9109	0.9117	0.9000	0.9018	0.9009
BERTweet	0.9013	0.9148	0.8971	0.9059	0.8868	0.9060	0.8963
IndoBERTweet	0.8752	0.8914	0.8703	0.8807	0.8579	0.8808	0.8692

Among all models, Support Vector Machine with linear kernel (SVM Linear) achieved the highest overall performance, with an accuracy of 96.33%, and balanced precision, recall, and F1-score for both sentiment classes (all exceeding 96%). This indicates strong generalization and consistency in detecting sentiment polarity from social media text.

Transformer-based models also performed competitively. BERT achieved an accuracy of 91.53%, closely followed by DistilBERT and BERTweet, with F1-scores above 90% for both classes. These results affirm the effectiveness of pre-trained language models in capturing nuanced sentiment in informal and context-rich data.

In contrast, traditional models such as Multinomial Naive Bayes and AdaBoost demonstrated lower accuracy, at 79.76% and 69.53% respectively, with significantly imbalanced performance between positive and negative classes. This highlights their limitations in handling the complexity of social media text, particularly with regard to sarcasm, slang, and varying sentence structures.

Overall, the evaluation suggests that while classical models like SVM Linear remain highly effective with well-engineered features (e.g., TF-IDF), transformer-based models offer robust alternatives for future work, particularly when dealing with larger and more diverse datasets.

4.2. **Topic Modelling.**

4.2.1. *LDA Topic Modeling.* The results of LDA analysis indicate that negative sentiment toward the free meal program can be grouped into five major themes:

(1) **Children's Education and Nutrition**
Criticisms highlight disparities in access to child nutrition programs, particularly in remote regions such as Papua. Dominant keywords: *susu*, *anak*, *bantu*, *Papua*.

(2) **Maternal and Infant Welfare**
Complaints focus on the lack of government support for pregnant women

and infants, often accompanied by sarcastic remarks directed at public officials. Dominant keywords: *hamil*, *balita*, *sejahtera*, *hidup*.

(3) **Program and Funding Transparency**
Criticisms address the lack of clarity regarding the distribution of information and funds. Dominant keywords: *uang*, *hilang*, *informasi*, *masak*.

(4) **Budgeting and Public Policy Implementation**
Issues related to national budget (APBN) management, meal quality, and program implementation in schools. Dominant keywords: *menu*, *orang tua*, *budget*, *apbn*.

(5) **Public Distrust of Government**
Negative and often harsh expressions toward government programs, reflecting widespread public distrust. Dominant keywords: *tolol*, *tanggung*, *duit*, *gratis*.

4.3. **Clustering Analysis.** In this study, clustering analysis is applied to data related to health and nutrition, social and demographic conditions, as well as economic and employment indicators for each province in Indonesia. The data used is from 2023 and sourced from Badan Pusat Statistik Indonesia. The analysis focuses on data from 34 provinces in Indonesia, excluding the four new provinces established in 2022, as data for several variables in these provinces is not yet available. The purpose of the clustering analysis is to categorize regions based on their level of priority for receiving Makan Bergizi Gratis (MBG) program. This will enable the government to focus the implementation of the program on high-priority areas, ensuring more effective and targeted resource allocation. Priority levels are determined based on demographic factors, economic conditions, and the health and nutritional status of each region, with the goal of optimizing government spending for the program's implementation.

The clustering analysis is conducted using several algorithms, including K-Means, K-Median, and K-Means with Genetic Algorithm Optimization. The analysis aims to form three clusters, as predetermined by the researchers, representing high, medium, and low priority groups. The results from the three algorithms are then compared using the silhouette score evaluation metric to identify the most effective algorithm for clustering provinces based on their priority level. The clustering outcomes for each algorithm are presented in the following Table (2).

Table 2. Clustering Performance

Algorithm	**Silhoutte Score**
K-Means	**0.6047**
K-Median	0.4266
K-Means with GA	0.5729

Based on the clustering results from the three algorithms, the K-Means algorithm demonstrated the best performance, achieving the highest silhouette score compared to K-Median and K-Means with Genetic Algorithm Optimization. Therefore, the clustering results from K-Means will be used to determine the priority

levels of each region. Table (3) and Figure (2).presents the provincial clustering outcomes using the K-Means algorithm.

TABLE 3. Clustering Analysis Results

Cluster	**Province**
Cluster 1	Aceh, Sumatera Utara, Sumatera Barat, Jambi, Sumatera Selatan, Bengkulu, Lampung, Jawa Tengah, Jawa Timur, NTT, NTB, Kalimantan Barat, Kalimantan Selatan, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara
Cluster 2	Riau, Kep. Bangka Belitung, Kep. Riau, Jawa Barat, DI Yogyakarta, Banten, Bali, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Papua Barat, Papua
Cluster 3	DKI Jakarta

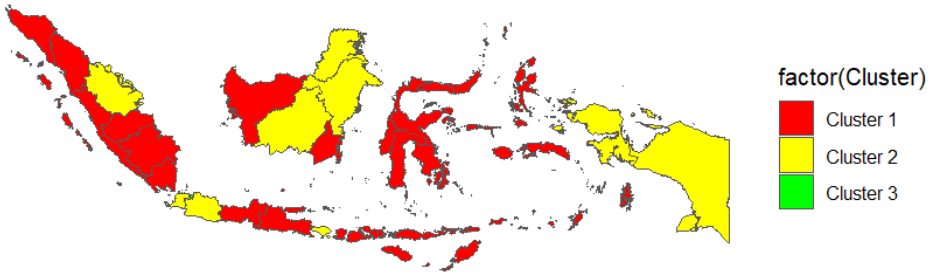

FIGURE 2. Provincial clustering results of the MBG program in Indonesia: Cluster 1 (red), Cluster 2 (yellow), and Cluster 3 (green).

Next, to identify the priority level of each cluster, it will be determined based on the characteristics of each cluster as observed from the centroid values of each variable within each cluster. The following are the centroids for each cluster in Table (4).

Provinces in Cluster 1 still show a quality of life that is not yet optimal, with the Human Development Index (HDI) categorized as moderate. The socio-economic conditions in this area face significant challenges, as evidenced by the high

prevalence of stunting and poverty that still threaten many families. Additionally, the community's purchasing power and wage levels are relatively low, while income inequality is at a moderate level. Therefore, this region requires special attention in social and public health development efforts to significantly improve the quality of life of its population.

Table 4. Centroids Each Cluster

Variables	**Cluster 1**	**Cluster 2**	**Cluster 3**
IPM	71.672	74.270	82.460
GR	0.338	0.349	0.431
Populasi	8301.162	7808.300	10672.100
IKPS	70.624	67.883	73.900
PBS	24.100	19.925	17.600
TPT	4.378	4.867	6.530
KRS	207.098	33.146	32.850
PPM	10.878	9.180	4.440
IKdM	1.804	1.695	0.690
IKpM	0.445	0.493	0.170
RRU	17127.905	22205.750	42354.000
PpK	1257437.670	1694760.080	2791716.000
PKK	11.247	12.654	2.570

Provinces in Cluster 2 show better development compared to Cluster 1, with HDI and quality of life relatively improved. Poverty levels and stunting prevalence in this area have started to decline, resulting in a much smaller number of families at risk of stunting. Purchasing power and wages have improved, although there are still challenges related to insufficient consumption among some groups. Then, cluster 3 represents regions that have reached a very advanced level of development, with a very high HDI reflecting excellent quality of education, health, and living standards. In this area, poverty and social risks are very low, while purchasing power is relatively high. The prevalence of stunting is also very low, indicating the success of various health and social programs implemented. Nevertheless, these regions still face challenges such as unemployment and income inequality that need to be managed well.

Based on the centroids of each cluster representing the economic conditions as well as the health and nutritional status of each region, priority levels for the implementation of the Makan Bergizi Gratis (MBG) program can be determined. Cluster 1 consists of provinces with a high priority for receiving Makan Bergizi Gratis (MBG) program. This is because these areas require special attention in social and public health development efforts to significantly improve the quality of life of their populations. Cluster 2 falls under medium priority for Makan Bergizi Gratis (MBG) program, where the government can target vulnerable groups such as poor families or children in certain schools. Meanwhile, Cluster 3 includes provinces with a low priority for Makan Bergizi Gratis (MBG) program, such as DKI Jakarta.

In this cluster, implementing Makan Bergizi Gratis (MBG) program is not yet an urgent priority; the government can focus more on lower-budget initiatives such as nutrition education. By referring to the results of this clustering analysis, the government can reassess the implementation of Makan Bergizi Gratis (MBG) program, allowing it to be carried out more targetedly and to save budget.

4.4. **Limitations.** This study has several limitations that should be acknowledged. First, the sentiment analysis was conducted solely on the social media platform **X** (formerly Twitter). While this platform provides timely and high-volume user-generated content, it does not capture sentiments from other widely used platforms such as Facebook, Instagram, or TikTok, which may reflect different user demographics and engagement patterns. As such, the sentiment findings may not be fully representative of the broader public opinion regarding the MBG program.

Second, the clustering analysis of MBG needs was performed at the provincial level due to data availability and granularity. While this provides a general overview of regional disparities, it lacks the precis

5. CONCLUSION

Sentiment analysis reveals that the majority of responses to the MBG program are positive (47.1%), followed by negative (41.9%) and neutral (11.0%), indicating strong public support, albeit with notable concerns. Among the 11 classification models evaluated, Linear SVM achieved the highest accuracy (96.33%) with balanced performance. Transformer-based models such as BERT and DistilBERT also performed well, effectively capturing the nuances of social media language. In contrast, traditional models like Naive Bayes and AdaBoost showed lower accuracy and less consistent performance across sentiment classes. Overall, transformer-based models are a strong choice for future analysis, particularly when dealing with complex and informal social media data.

The clustering analysis classified 34 provinces into three priority levels for the Makan Bergizi Gratis (MBG) program, with K-Means showing the best performance. Cluster 1 (high priority) includes provinces with lower development indicators, while Cluster 2 and Cluster 3 represent medium and low priority, respectively. These results enable more targeted and efficient program implementation. Future research is encouraged to use more granular data at the district or sub-district (kecamatan) level to improve policy targeting and resource allocation.

REFERENCES

[1] K. K. RI, "Prevalensi stunting di indonesia turun ke 21,6% dari 24,4%.," *Journal Name*, 2024.

[2] T. . Rachmawati, "Implementasi algoritma naive bayes untuk analisis sentimen terhadap program makan siang gratis," *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, vol. 5, no. 3, pp. 2925–2939, 2024.

[3] N. A. Rahmawati, S. A. Prasetyo, and M. W. Ramadhani, "Memetakan visi prabowo gibran pada masa kampanye dalam prespektif pembangunan:(analisis wacana kritis visi dan misi prabowo gibran dalam prespektif moderenisasi)," *WISSEN: Jurnal Ilmu Sosial dan Humaniora*, vol. 2, no. 3, pp. 97–120, 2024.

[4] P. A. Maharani, A. R. Namira, and T. V. Chairunnisa, "Peran makan siang gratis dalam janji kampanye prabowo gibran dan realisasinya," *Journal Of Law And Social Society*, vol. 1, no. 1, pp. 1–10, 2024.

[5] C. Steven and W. Wella, "The right sentiment analysis method of indonesian tourism in social media twitter," *IJNMT (International Journal of New Media Technology)*, vol. 7, no. 2, pp. 102–110, 2020.

[6] F. R. B. Kahi *et al.*, "Analisis sentimen masyarakat di twitter terhadap pemerintahan anies baswedan menggunakan metode naive bayes classifier," *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 324–336, 2024.

[7] M. Iqbal, M. Afdal, and R. Novita, "Implementasi algoritma support vector machine untuk analisa sentimen data ulasan aplikasi pinjaman online di google play store: Implementation of support vector machine algorithm for sentiment analysis of online loan application review data on google play store," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 4, pp. 1244–1252, 2024.

[8] 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, 2024.

[9] S. D. Wahyuni and R. H. Kusumodestoni, "Optimalisasi algoritma support vector machine (svm) dalam klasifikasi kejadian data stunting," *Bulletin of Information Technology (BIT)*, vol. 5, no. 2, pp. 56–64, 2024.

[10] A. M. Siregar, "Analisis sentimen pindah ibu kota negara (ikn) baru pada twitter menggunakan algoritma naive bayes dan support vector machine (svm)," *Faktor Exacta*, vol. 16, no. 3, 2023.

[11] E. 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, 2025.

[12] A. Ramadhani, I. Permana, M. Afdal, and M. Fronita, "Analisis sentimen tanggapan publik di twitter terkait program kerja makan siang gratis prabowo-gibran menggunakan algoritma naive bayes classifier dan support vector machine," *Building of Informatics, Technoogy and Science (BITS)*, vol. 6, no. 3, pp. 1509–1516, 2024.

[13] R. F. Fajri, M. Adriani, *et al.*, "Inset: A sentiment lexicon for indonesian social media," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.

[14] D. Jurafsky and J. H. Martin, *Speech and Language Processing (3rd ed. draft)*. Stanford University, 2021. Available at `https://web.stanford.edu/~jurafsky/slp3/`.

[15] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*. Wiley, 2013.

[16] A. McCallum and K. Nigam, "A comparison of event models for naive bayes text classification," *AAAI-98 workshop on learning for text categorization*, 1998.

[17] A. Sabrani and F. Bimantoro, "Multinomial naïve bayes untuk klasifikasi artikel online tentang gempa di indonesia," *Jurnal Teknologi Informasi, Komputer, Dan Aplikasinya (JTIKA)*, vol. 2, no. 1, pp. 89–100, 2020.

[18] A. W. Syaputri, E. Irwandi, and Mustakim, "Naive bayes algorithm for classification of student major's specialization," *Journal not specified*, vol. 1, no. 1, 2020.

[19] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[20] Saikin, S. Fadli, and M. Ashari, "Optimization of support vector machine method using feature selection to improve classification result," *Journal not specified*, vol. 4, no. 1, 2021.
[21] N. H. Ovirianti, M. Zarlis, and H. Mawengkang, "Support vector machine using a classification algorithm," *Journal not specified*, vol. 7, no. 3, 2022.
[22] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
[23] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *International Journal of Remote Sensing*, vol. 39, no. 9, pp. 2784–2817, 2018.
[24] Y. Religia, A. Nugroho, and W. Hadikristanto, "Analisis perbandingan algoritma optimasi pada random forest untuk klasifikasi data bank marketing," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 187–192, 2021.
[25] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.
[26] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
[27] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in *Advances in neural information processing systems*, pp. 3146–3154, 2017.
[28] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *NAACL*, 2019.
[29] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019.
[30] I. Alfina, R. Mulia, and M. I. Fanany, "Indobertweet: Pretrained transformer-based language model for indonesian twitter," *arXiv preprint arXiv:2109.05238*, 2021.
[31] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine learning in python," 2011. Journal of Machine Learning Research, 12, 2825–2830.
[32] T. Wolf, L. Debut, V. Sanh, *et al.*, "Transformers: State-of-the-art natural language processing." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 2020. HuggingFace Transformers library.
[33] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
[34] R. Řehůřek and P. Sojka, "Software framework for topic modelling with large corpora," 2010. Gensim Python library.
[35] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, University of California Press, 1967.
[36] S. P. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.
[37] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 264–323, 1999.
[38] V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit, "Local search heuristics for k-median and facility location problems," *SIAM Journal on Computing*, vol. 33, no. 3, pp. 544–562, 2004.
[39] M. Jahandideh-Tehrani, G. Jenkins, and F. Helfer, "A comparison of particle swarm optimization and genetic algorithm for daily rainfall-runoff modelling: a case study for southeast queensland, australia," *Optimization and Engineering*, vol. 22, no. 1, pp. 29–50, 2021.