Exacta: Journal of Pure and Fundamental Research
Vol 1 No 2 June 2026, Hal 21-31
ISSN: 3125-1447 (Print) ISSN: 3125-1439 (Electronic)
Open Access: https://sovereignresearch.org/exacta

Analysis of Public Sentiment Toward the Free Nutritious Meals Program Using the Naive Bayes Algorithm

Afdal Perangin Angin[1*], Indra Setiawan[2], Jeremia Revaldo Girsang[3], Aldo Jeremia Tumanggor[4], Andika Tioanta Ginting[5], Manuel Saputra Manalu[6], Desman Jaya Giawa[7]
[1-7] Universitas Mandiri Bina Prestasi, Indonesia
email: afdalperanginangin@email.com[1]

Article Info :
Received:
28-05-2026
Revised:
02-06-2026
Accepted:
10-06-2026

Abstract

The increasing use of digital platforms as spaces for expressing public opinion has created opportunities for applying natural language processing techniques to evaluate societal responses toward government policies. This study aims to analyze public sentiment toward the Program Makan Bergizi Gratis (MBG) using the Multinomial Naive Bayes algorithm. The research applies an empirical machine learning approach involving data preprocessing, lexicon-based sentiment labeling, feature extraction using Count Vectorizer with unigram and bigram representation, and model evaluation through accuracy, precision, recall, F1-score, and confusion matrix analysis. The dataset consists of 3,300 public comments processed through a structured computational pipeline and divided using stratified sampling into training and testing data. The findings indicate that public sentiment is dominated by negative responses, followed by neutral and positive categories, reflecting critical public evaluation toward program implementation aspects. The classification model demonstrates reliable performance in identifying sentiment patterns and provides an analytical framework for understanding digital public perception. This study contributes to computational social analysis by integrating machine learning techniques with policy-oriented sentiment monitoring.

Keywords: *Sentiment Analysis, Naive Bayes, Machine Learning, Public Opinion, Digital Policy Monitoring.*

INTRODUCTION

The rapid expansion of digital communication ecosystems has fundamentally transformed the manner in which citizens interact with, evaluate, and respond to public policies, positioning social media platforms as critical sources of collective intelligence for understanding societal perceptions. In contemporary computational research, the integration of Natural Language Processing (NLP), machine learning, and sentiment analysis has become an increasingly important approach for extracting latent public opinions from large-scale unstructured textual data, particularly in the context of policy evaluation and governance decision-making (Jurafsky & Martin, 2023; Liu, 2020). The emergence of data-driven policy analysis reflects a broader global transition from conventional survey-based opinion assessment toward computational approaches capable of identifying emotional tendencies, ideological polarization, and public acceptance patterns in real time. Sentiment analysis has demonstrated substantial potential in transforming massive volumes of user-generated content into meaningful analytical information through automated classification mechanisms, enabling researchers and policymakers to examine public responses toward social programs, governmental interventions, and strategic initiatives more efficiently (Liu, 2020). Within this technological landscape, the Indonesian Government's Program Makan Bergizi Gratis (MBG) represents a significant policy intervention aimed at improving nutritional accessibility and supporting human development, while simultaneously generating diverse reactions across digital platforms. The complexity of public discourse surrounding MBG creates an empirical context in which computational sentiment analysis becomes essential for identifying the dominant patterns of acceptance, criticism, and uncertainty expressed by society.

Recent studies have increasingly explored the application of machine learning algorithms for analyzing public sentiment toward the MBG program, demonstrating the effectiveness of computational classification methods in processing social media discourse. Triningsih (2025) investigated public sentiment toward MBG on the X platform using machine learning approaches and highlighted the

capability of automated classification models in identifying public reactions toward government programs based on textual characteristics. Similar findings were presented by Rahmatullah et al. (2025), who applied the Naive Bayes algorithm for MBG sentiment analysis and demonstrated that probabilistic classification methods remain relevant for handling text-based datasets due to their computational efficiency and suitability for high-dimensional linguistic features. Hidayat and Ratnaningsih (2025) further expanded this discussion by comparing Random Forest and Naive Bayes algorithms, indicating that algorithmic performance differences are influenced by feature representation, dataset characteristics, and distribution patterns within sentiment classes. Research conducted by Iranda and Huda (2025) comparing Support Vector Machine and Naive Bayes also confirmed that machine learning-based approaches provide valuable analytical capabilities for categorizing public opinions regarding nutritional policy programs. These studies collectively indicate that Naive Bayes remains one of the dominant approaches in sentiment classification research because of its simplicity, scalability, and effectiveness in processing large textual datasets, although variations in preprocessing techniques, labeling strategies, and evaluation frameworks produce different levels of classification performance.

Despite the increasing number of studies examining MBG-related sentiment, existing literature reveals several conceptual and methodological limitations that require further investigation. Most previous studies primarily emphasize algorithm comparison and performance measurement without sufficiently examining the methodological implications of data preparation, automated labeling mechanisms, and the representation of linguistic patterns within public discourse. Ihdayahuningsih et al. (2026) applied Naive Bayes for analyzing public sentiment on social media X and confirmed the applicability of this algorithm; however, the study still reflects the broader limitation of many sentiment analysis studies that treat sentiment categories as static outputs without critically evaluating the complexity of public narratives surrounding policy implementation. Putri and Novianto (2026) compared Naive Bayes and Support Vector Machine approaches, yet the emphasis remained largely centered on model accuracy rather than understanding how computational models capture social perceptions embedded within policy discussions. Similarly, Sakina et al. (2025) evaluated KNN and Naive Bayes algorithms for MBG sentiment classification, demonstrating algorithmic differences but leaving methodological questions regarding the reliability of automated sentiment labeling unresolved. These limitations indicate that existing research has not fully addressed the relationship between computational accuracy, linguistic representation, and the interpretative meaning of public sentiment, creating a research gap in developing a more systematic framework for understanding digital public opinion toward government nutritional policies.

The unresolved methodological issues surrounding sentiment classification have significant scientific and practical implications, particularly because inaccurate interpretation of public sentiment may influence how policymakers understand social acceptance and potential implementation challenges of large-scale public programs. Public reactions toward MBG are not merely numerical classifications of positive, negative, or neutral expressions, but represent complex manifestations of concerns related to budget allocation, program effectiveness, distribution mechanisms, transparency, and operational readiness. Research by Rifna and Nurdin (2026) examining TikTok user sentiment toward MBG using Naive Bayes demonstrated that different social media environments generate distinctive patterns of public expression, suggesting that platform characteristics and communication styles influence the formation of digital opinions. Likewise, Siregar et al. (2025) applied Naive Bayes to analyze netizen sentiment toward the free lunch program on social media X and reinforced the importance of computational approaches in capturing public reactions toward government initiatives. However, variations in datasets, preprocessing pipelines, and classification strategies across previous studies indicate the need for a more structured analytical approach that integrates reliable text processing techniques with transparent machine learning procedures. The development of such an approach is increasingly important as governments require evidence-based mechanisms to monitor public responses and evaluate policy implementation effectiveness within rapidly evolving digital environments.

This research is positioned within the intersection between computational intelligence, natural language processing, and digital governance by focusing on the utilization of Multinomial Naive Bayes as an analytical instrument for classifying public sentiment toward the MBG program. Unlike previous studies that predominantly concentrate on algorithm comparison, this research emphasizes the construction of an integrated sentiment analysis framework consisting of text preprocessing, lexicon-based labeling, feature extraction using Count Vectorizer with unigram and bigram representations,

stratified dataset partitioning, and performance evaluation through accuracy, precision, recall, F1-score, and confusion matrix analysis. The selection of Multinomial Naive Bayes is theoretically grounded in its probabilistic capability to estimate class probabilities based on word occurrence patterns, making it particularly suitable for textual classification problems characterized by high-dimensional feature spaces. This methodological positioning contributes to the broader development of applied machine learning research by demonstrating how relatively lightweight computational models can be utilized for analyzing large-scale public discourse related to strategic government programs.

This study aims to develop and evaluate a sentiment classification model based on the Multinomial Naive Bayes algorithm to identify public perceptions toward the Program Makan Bergizi Gratis through digital textual data analysis. The research contributes theoretically by strengthening the understanding of machine learning implementation in public policy sentiment analysis, particularly regarding the relationship between text representation, classification mechanisms, and digital opinion interpretation. Methodologically, this study contributes an integrated analytical framework that combines preprocessing optimization, lexicon-based sentiment annotation, feature engineering, and model evaluation procedures to improve the reliability of automated sentiment classification. The findings are expected to provide empirical insights into public responses toward MBG while offering a computational foundation that can support evidence-based policy evaluation and the development of intelligent systems for monitoring societal reactions in digital communication environments.

RESEARCH METHODS

This study adopts an empirical research design as it involves the development, implementation, and experimental evaluation of a machine learning-based sentiment classification system using real-world textual data. The methodological framework is designed as an end-to-end computational pipeline consisting of data acquisition, text preprocessing, feature engineering, model development, and experimental testing. The dataset consists of 3,300 public comments related to the Program Makan Bergizi Gratis (MBG) collected from digital platforms and stored in a structured Excel dataset. The innovation development process begins with text normalization procedures, including case folding, URL removal, elimination of non-alphabetic characters, and whitespace optimization to enhance textual consistency. Sentiment annotation is performed through a lexicon-based approach by assigning polarity scores based on predefined positive, negative, and neutral linguistic indicators, including contextual phrases and keyword weighting mechanisms. The processed textual data are subsequently transformed into numerical feature representations using Count Vectorizer with unigram and bigram configurations (*ngram_range = (1,2)*) and a maximum feature limitation of 1,000 attributes. The experimental model is developed using the Multinomial Naive Bayes algorithm, which applies probabilistic classification principles based on Bayes' theorem to estimate sentiment categories from extracted textual features. The dataset is partitioned using stratified sampling into 70% training data and 30% testing data to maintain class distribution consistency and improve the reliability of model generalization.

The validation process is conducted through systematic performance evaluation using quantitative machine learning assessment metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis. Accuracy is utilized to measure the overall correctness of sentiment predictions, while precision, recall, and F1-score provide a more comprehensive evaluation of classification effectiveness across each sentiment category by identifying potential bias toward dominant classes. The confusion matrix is employed to examine the distribution of correct and incorrect predictions, allowing further assessment of false positive and false negative classifications among positive, negative, and neutral sentiments. The methodological robustness of this research is strengthened through the combination of automated sentiment labeling, controlled preprocessing procedures, feature optimization using n-gram representation, and stratified model testing, which collectively enhance reproducibility and analytical consistency. The uniqueness of this study lies in the integration of a lightweight yet effective probabilistic machine learning framework with public opinion mining for evaluating societal responses toward a strategic governmental nutrition program, providing both computational insights and practical evidence for digital policy monitoring.

.

RESULTS AND DISCUSSION

Data Distribution and Sentiment Pattern Analysis of Public Responses toward the MBG Program

The empirical analysis began by examining the distribution pattern of sentiment categories generated through the lexicon-based annotation process applied to 3,300 textual comments related to the Program Makan Bergizi Gratis (MBG). The classification process identified three dominant sentiment categories consisting of positive, negative, and neutral expressions, which represent different forms of public interpretation toward the implementation of the government nutritional initiative. The utilization of lexicon-based labeling provides an initial analytical foundation for understanding the polarity structure of public discourse before the implementation of the machine learning classification stage. Similar approaches have been applied in previous MBG sentiment studies, where automated sentiment categorization was considered effective for transforming social media narratives into measurable analytical indicators (Triningsih, 2025).

The distribution of sentiment categories indicates that public perception toward MBG is characterized by a strong tendency toward critical responses rather than supportive expressions. The dominance of negative sentiment reflects the existence of concerns expressed by digital users regarding various dimensions of policy implementation, including budget effectiveness, distribution mechanisms, program management, and institutional readiness. This pattern demonstrates that social media functions not merely as a communication channel but also as a space where citizens evaluate and negotiate the legitimacy of public policies through textual interaction. Previous research on MBG sentiment analysis using Naive Bayes similarly identified significant variations in public responses, emphasizing that computational analysis can reveal underlying social concerns embedded within online discussions (Rahmatullah et al., 2025).

The quantitative distribution obtained from the sentiment annotation process is presented in Table 1, which illustrates the comparative proportion of each sentiment category within the research dataset. The table demonstrates that negative expressions represent the largest proportion, followed by neutral and positive categories, indicating an asymmetrical distribution of public opinion toward the MBG program. This imbalance becomes an important analytical consideration because machine learning models trained on uneven sentiment distributions may experience classification bias toward dominant classes. Previous comparative studies on MBG sentiment classification also emphasized that dataset composition significantly influences algorithmic performance and interpretative reliability (Hidayat & Ratnaningsih, 2025).

Table 1. Distribution of Public Sentiment Categories toward the MBG Program

Sentiment Category	Number of Comments	Percentage
Negative	2,762	83.7%
Neutral	432	13.1%
Positive	106	3.2%
Total	3,300	100%

The findings presented in Table 1 reveal that negative sentiment constitutes the dominant discourse pattern within the analyzed dataset, reaching 83.7% of total comments. This numerical composition suggests that public discussions surrounding MBG are largely driven by evaluative criticism rather than explicit endorsement, which may indicate the existence of uncertainty regarding program implementation. The relatively limited proportion of positive sentiment does not necessarily represent complete rejection of the policy but may reflect the tendency of social media users to express concerns more actively than approval. Research examining sentiment toward government meal programs has shown similar characteristics, where negative expressions frequently emerge as responses to perceived implementation risks and unresolved operational issues (Iranda & Huda, 2025).

The concentration of negative sentiment requires interpretation beyond statistical classification because textual opinions represent complex interactions between policy expectations and public experiences. Negative classifications identified within the dataset primarily correspond with concerns related to financial allocation, transparency, logistical preparedness, and implementation challenges

encountered in practical contexts. Such findings indicate that sentiment analysis can provide an early detection mechanism for identifying potential policy vulnerabilities through digital communication patterns. Ihdayahuningsih et al. (2026) demonstrated that Naive Bayes-based sentiment analysis on platform X can effectively identify dominant public perceptions, although interpretation of sentiment outcomes requires consideration of the broader social context surrounding policy discussions.

The neutral sentiment category represented 13.1% of the dataset, demonstrating the presence of users who expressed informational statements, inquiries, or balanced opinions without strong emotional orientation. This category plays an important role in sentiment analysis because neutral expressions often contain contextual information regarding public awareness, expectations, and knowledge development about government programs. The existence of neutral discourse suggests that digital discussions regarding MBG are not exclusively polarized but also include informational exchanges among users. Studies comparing multiple classification algorithms have indicated that neutral sentiment recognition remains a challenging aspect because linguistic ambiguity frequently affects the boundary between objective statements and implicit evaluations (Putri & Novianto, 2026).

The positive sentiment category accounted for only 3.2% of analyzed comments, representing the smallest proportion among the three sentiment classes. This limited representation suggests that explicit appreciation toward MBG was less frequently expressed compared with criticism or informational discussion within the collected dataset. Nevertheless, positive expressions provide important evidence that certain segments of society perceive MBG as a beneficial initiative related to nutritional improvement and social welfare enhancement. Previous research applying Naive Bayes and other machine learning approaches demonstrated that minority sentiment classes require careful evaluation because their limited frequency may affect model learning capability and prediction stability (Sakina et al., 2025).

The observed sentiment imbalance also influenced the methodological considerations during model development, particularly in relation to training data distribution and evaluation interpretation. The implementation of stratified sampling in this research ensured that the training and testing datasets maintained proportional representation of each sentiment category, reducing the potential distortion caused by random data separation. This procedure aligns with machine learning best practices for classification tasks involving imbalanced datasets, where preservation of class distribution contributes to more reliable model evaluation. Research involving MBG sentiment classification through Naive Bayes has similarly emphasized the importance of maintaining data consistency between training and testing phases to achieve valid performance measurement (Rifna & Nurdin, 2026).

The sentiment distribution findings provide empirical evidence that public discourse toward MBG reflects a complex relationship between policy objectives and societal expectations. The dominance of negative sentiment should not be interpreted solely as rejection of the program but rather as an indicator of active public scrutiny toward government intervention in the nutritional sector. Computational sentiment analysis enables researchers to identify these patterns systematically by converting unstructured textual communication into measurable analytical outcomes. Studies utilizing social media sentiment analysis have confirmed that digital opinion mining contributes significantly to understanding public reactions toward government policies and social programs (Siregar et al., 2025).

The results of this analysis establish the initial empirical foundation for evaluating the effectiveness of the Multinomial Naive Bayes model in classifying MBG-related public opinions. The distribution pattern demonstrates that the dataset contains substantial linguistic diversity and class imbalance, creating a realistic testing environment for machine learning-based sentiment classification. The findings also strengthen the argument that sentiment analysis is not merely a computational exercise but represents an interdisciplinary approach combining data science, communication studies, and policy evaluation. Previous investigations involving Naive Bayes implementation for MBG sentiment analysis have highlighted the relevance of computational models in supporting evidence-based understanding of public perception toward strategic government programs (Tsabitah et al., 2025).

Evaluation Performance of Multinomial Naive Bayes Model

The evaluation phase examined the capability of the Multinomial Naive Bayes (MNB) algorithm in classifying public sentiment toward the Program Makan Bergizi Gratis (MBG) based on the processed textual features derived from the dataset. The experimental procedure involved training the model using 70% of the annotated dataset and testing its predictive performance using the remaining

30% through stratified sampling techniques. This approach ensured that each sentiment category maintained proportional representation during model validation, allowing the evaluation process to reflect the actual classification capability of the developed system. The utilization of Multinomial Naive Bayes remains relevant in text classification research because its probabilistic framework effectively handles discrete word-frequency features generated from high-dimensional textual data (Liu, 2020).

The implementation of Count Vectorizer with unigram and bigram representations produced a structured numerical feature space consisting of frequently occurring linguistic patterns within MBG-related discussions. The extraction of both single-word and two-word combinations enabled the model to capture not only isolated keywords but also contextual relationships between terms that potentially influence sentiment polarity. This feature engineering strategy strengthened the representation of textual information because sentiment expressions in social media frequently depend on phrase combinations rather than individual words. Previous research applying Naive Bayes for MBG sentiment analysis demonstrated that feature extraction techniques significantly influence classification outcomes because the quality of textual representation determines the ability of algorithms to distinguish between sentiment categories (Noviyanti & Putra, 2025).

The performance results generated from the Multinomial Naive Bayes classification process are summarized in Table 2, which presents the evaluation metrics obtained from the testing dataset. The table demonstrates the model's capability in identifying sentiment classes through measurements of accuracy, precision, recall, and F1-score. These indicators provide complementary perspectives regarding classification effectiveness because accuracy alone may not sufficiently represent performance when the dataset contains unequal sentiment distributions. Similar evaluation frameworks have been adopted in comparative MBG sentiment studies to assess the reliability of machine learning models under different classification conditions (Saleh & Imanda, 2025).

Table 2. Performance Evaluation Results of Multinomial Naive Bayes Classification Model

Evaluation Metric	Value
Accuracy	86.2%
Precision	84.9%
Recall	86.2%
F1-Score	85.4%

The evaluation results presented in Table 2 indicate that the Multinomial Naive Bayes model achieved an accuracy level of 86.2%, demonstrating a relatively strong capability in predicting sentiment categories from MBG-related textual data. The precision value of 84.9% indicates that most predicted sentiment labels corresponded correctly with the actual sentiment classes, while the recall value of 86.2% shows that the model successfully identified a substantial proportion of relevant sentiment instances. The balanced F1-score of 85.4% reflects an acceptable equilibrium between precision and recall, suggesting that the model was able to maintain classification consistency despite the presence of sentiment imbalance. Research comparing Naive Bayes with other machine learning approaches has similarly shown that Naive Bayes provides competitive classification performance, particularly in text-based problems characterized by sparse feature distributions (Pratama & Suryono, 2025).

The confusion matrix analysis provides a deeper interpretation of model behavior by examining the relationship between actual sentiment labels and predicted classifications. The analysis revealed that the model performed effectively in identifying dominant negative sentiment due to the substantial number of training examples representing this category. However, classification challenges appeared in distinguishing minority sentiment categories, particularly positive and neutral expressions, because limited data availability reduced the linguistic diversity learned by the model. This phenomenon corresponds with previous findings indicating that imbalanced sentiment datasets remain one of the primary challenges in social media classification tasks because machine learning models tend to optimize predictions toward dominant classes (Amelsan & Muntahanah, 2026).

.

The classification capability of Multinomial Naive Bayes can be explained through its probabilistic mechanism, where each sentiment category is determined based on the probability distribution of word occurrences within training data. The model assumes conditional independence among textual features, allowing efficient computation despite handling thousands of linguistic attributes generated through vectorization. Although this assumption simplifies natural language complexity, the algorithm remains effective for many sentiment classification scenarios because word-frequency patterns often provide sufficient statistical signals for polarity identification. Studies examining Naive Bayes implementation in public sentiment analysis have confirmed that the algorithm maintains strong applicability due to its computational efficiency and interpretability compared with more complex models (Prastiawan & Yuniarto, 2026).

The obtained performance results also demonstrate the influence of preprocessing procedures on improving model stability and reducing textual noise. The elimination of URLs, symbols, numerical characters, and inconsistent capitalization contributed to producing cleaner linguistic features that improved the consistency of sentiment representation. The lexicon-based labeling mechanism further supported the supervised learning process by providing structured sentiment categories before model training. Previous research combining preprocessing optimization with Naive Bayes classification reported that systematic text preparation significantly affects predictive accuracy because irrelevant linguistic variations may reduce the quality of extracted features (Dinta et al., 2026).

The findings indicate that Multinomial Naive Bayes provides an efficient methodological solution for analyzing public sentiment toward policy-related discussions within digital environments. The model demonstrates that relatively simple probabilistic algorithms can generate meaningful analytical outcomes when supported by appropriate preprocessing, feature engineering, and validation strategies. This finding is relevant within the field of computational social science because policy-oriented sentiment analysis requires models that balance predictive capability, computational efficiency, and interpretability. Research comparing Naive Bayes with Support Vector Machine and other algorithms has shown that algorithm selection should consider not only predictive accuracy but also dataset characteristics and analytical objectives (Jakak & Rohman, 2025).

The evaluation results further indicate that the developed system possesses practical value for monitoring public responses toward strategic government initiatives such as MBG. The model can function as an automated analytical instrument capable of processing large-scale textual opinions that would be difficult to examine manually within a limited period. The integration of machine learning classification with social media analysis provides opportunities for developing responsive digital monitoring systems supporting evidence-based decision-making. Previous studies examining sentiment toward MBG on different social media platforms have emphasized that computational approaches can enhance policymakers' understanding of public acceptance patterns and emerging concerns (Akmal et al., 2026).

The experimental evaluation confirms that the Multinomial Naive Bayes model successfully fulfills its role as a sentiment classification mechanism for MBG-related public discourse. The achieved performance demonstrates that the combination of lexicon-based annotation, Count Vectorizer feature representation, and probabilistic classification generates a reliable analytical framework for processing Indonesian social media text. The findings contribute to existing sentiment analysis research by demonstrating the applicability of lightweight machine learning techniques for evaluating public perceptions toward government nutritional programs. The methodological outcomes also provide a foundation for further improvement through advanced feature selection, hybrid algorithms, or deep learning approaches in future investigations (Pratama et al., 2025).

Interpretation of Public Sentiment and Policy Implications of MBG Digital Discourse

The interpretation stage focuses on understanding the broader meaning of sentiment classification outcomes by connecting computational results with the social and policy context surrounding the Program Makan Bergizi Gratis (MBG). The empirical findings demonstrate that digital discourse regarding MBG represents a dynamic interaction between public expectations, policy objectives, and perceived implementation challenges. Sentiment categories generated through the Multinomial Naive Bayes model provide quantitative representations of public attitudes, yet their significance requires further examination through contextual interpretation of the linguistic patterns underlying each category. Sentiment analysis research emphasizes that computational classification should not be

limited to predictive accuracy but should also contribute to understanding collective perceptions and behavioral tendencies expressed through digital communication channels (Liu, 2020).

The dominance of negative sentiment identified in the dataset reflects a critical dimension of public engagement toward government policy rather than merely indicating opposition to the MBG initiative. Negative expressions frequently represent concerns regarding operational readiness, distribution effectiveness, financial accountability, and sustainability of implementation mechanisms. This pattern suggests that social media functions as an evaluative arena where citizens actively assess the practical consequences of public programs after policy announcements are made. Previous studies analyzing MBG-related discourse have similarly identified that negative sentiment often emerges from concerns regarding implementation quality rather than complete rejection of the underlying policy objectives (Anggraeni et al., 2025).

The relationship between sentiment polarity and policy evaluation can be further examined through the distribution of thematic concerns identified from classified comments, as presented in Table 3. The thematic mapping demonstrates that negative sentiment was associated with several recurring issues, while positive sentiment was primarily linked to expectations regarding nutritional improvement and social benefits. The identification of these themes provides additional analytical depth because sentiment categories alone cannot fully explain the reasons behind public reactions toward MBG implementation. Previous research has highlighted that combining sentiment classification with contextual interpretation improves the ability of computational approaches to generate meaningful insights for policy analysis (Siregar et al., 2025).

Table 3. Dominant Discussion Themes Associated with MBG Sentiment Categories

Sentiment Category	Dominant Discussion Themes	Interpretation
Negative	Budget concerns, distribution problems, implementation readiness, transparency issues	Public criticism toward operational and managerial aspects
Neutral	Information requests, policy discussions, factual statements	Public awareness and information-seeking behavior
Positive	Nutritional benefits, child welfare, social assistance expectations	Public support toward policy objectives

The thematic patterns presented in Table 3 indicate that negative sentiment should be interpreted as a form of public monitoring rather than a simple indicator of policy failure. Citizens expressing criticism demonstrate engagement with policy outcomes by questioning whether implementation mechanisms align with the intended objectives of MBG. The presence of neutral discussions further indicates that some users remain in an information-acquisition stage, suggesting that public understanding of the program continues to develop through digital interactions. Research on social media sentiment toward government programs has shown that online discussions frequently combine evaluation, information exchange, and collective interpretation processes within the same communication environment (Rifna & Nurdin, 2026).

The analytical findings also reveal the importance of considering platform characteristics in interpreting digital sentiment patterns. Public discourse collected from online platforms represents spontaneous expressions shaped by communication norms, audience interaction, and platform-specific engagement mechanisms. Although the present study focuses on aggregated textual data, the sentiment patterns provide evidence that digital platforms have become significant sources for observing societal reactions toward governmental initiatives. Studies examining MBG discussions across X, TikTok, and Instagram have demonstrated that differences in platform environments influence the intensity and form of public expression, requiring careful interpretation of computational results (Akmal et al., 2026).

The application of Multinomial Naive Bayes in this research demonstrates that computational models can support policy-oriented analysis by identifying dominant patterns within large-scale public conversations. The model does not replace qualitative policy evaluation but provides an initial analytical mechanism for detecting emerging issues and measuring general sentiment tendencies. This capability

becomes particularly relevant when policymakers must process extensive digital feedback within limited decision-making periods. Previous investigations comparing Naive Bayes with Support Vector Machine and other classification algorithms have emphasized that machine learning approaches provide valuable support for extracting actionable information from complex textual environments (Saleh & Imanda, 2025).

The findings also provide implications regarding the importance of improving communication strategies surrounding MBG implementation. The prevalence of negative sentiment indicates that public acceptance is influenced not only by policy substance but also by the availability of transparent information regarding implementation procedures, resource allocation, and program outcomes. Effective communication mechanisms may reduce uncertainty by providing society with clearer explanations regarding the objectives and operational framework of the program. Research on public sentiment classification has demonstrated that digital opinion analysis can assist institutions in identifying communication gaps between policy intentions and public interpretation (Triningsih, 2025).

From a technological perspective, this study confirms the relevance of lightweight machine learning models for developing scalable public opinion monitoring systems. The combination of preprocessing techniques, lexicon-based annotation, and Multinomial Naive Bayes classification provides an efficient computational framework that can be adapted for continuous monitoring of policy-related discussions. The approach offers advantages in terms of interpretability and computational efficiency compared with highly complex models that often require greater computational resources and extensive training datasets. Previous studies involving Naive Bayes implementation for MBG sentiment analysis have similarly demonstrated the practicality of probabilistic approaches for social media-based policy evaluation (Rahmatullah et al., 2025).

The empirical interpretation of sentiment patterns contributes to the broader field of digital governance by demonstrating how computational analysis can strengthen evidence-based understanding of public perception. The MBG case illustrates that government programs operating within contemporary digital societies are continuously evaluated through online communication spaces where public approval and criticism develop simultaneously. The findings emphasize that sentiment analysis should be positioned as a complementary analytical instrument capable of revealing societal responses that may not be captured through conventional evaluation mechanisms. Studies applying comparative machine learning approaches to MBG sentiment classification support the argument that computational opinion mining has become an important component of modern policy assessment frameworks (Pratama & Suryono, 2025).

The final interpretation of this research establishes that sentiment analysis using Multinomial Naive Bayes provides both methodological and practical contributions for understanding digital public discourse surrounding MBG. The classification results reveal patterns of acceptance, criticism, and uncertainty that reflect the complex relationship between governmental policy implementation and societal expectations. The integration of machine learning with public opinion analysis creates opportunities for developing responsive monitoring systems capable of supporting adaptive policy management. The findings contribute to technological innovation in sentiment analysis by demonstrating how empirical computational methods can be utilized to analyze public perception toward strategic social programs in Indonesia (Amelsan & Muntahanah, 2026).

CONCLUSION

This study demonstrates that the implementation of Multinomial Naive Bayes for sentiment analysis of public responses toward the Program Makan Bergizi Gratis (MBG) provides an effective computational framework for identifying patterns of digital public perception through textual data processing. The sentiment distribution analysis revealed that public discourse was predominantly characterized by negative sentiment, indicating substantial critical attention toward aspects of program implementation, including operational readiness, transparency, and policy execution mechanisms, while positive and neutral expressions reflected public expectations and informational engagement. The experimental evaluation confirmed that the developed model achieved reliable classification performance through the integration of lexicon-based labeling, text preprocessing, Count Vectorizer feature extraction, and stratified dataset validation, demonstrating the suitability of probabilistic machine learning approaches for large-scale social media analysis. The interpretation of computational findings indicates that sentiment analysis extends beyond classification accuracy by providing strategic

insights into societal responses and potential communication challenges surrounding government initiatives. This research contributes to the advancement of applied artificial intelligence in digital governance by presenting an efficient and reproducible analytical approach for monitoring public opinion, while also providing methodological foundations for future studies involving enhanced feature optimization, hybrid algorithms, and more adaptive sentiment intelligence systems.

REFERENCES

Akmal, Z. Z., Setiana, N., Musridho, J., Alfayet, A. Y., Fajar, A., Pratama, A., ... & Baihaqi, A. (2026). Analisis Sentimen Masyarakat terhadap Program Makan Bergizi Gratis (MBG) di Media Sosial Instagram Menggunakan Metode Naive Bayes dan Support Vector Machine. *Culture education and technology research (Cetera)*, *3*(2), 313-325. https://doi.org/10.31004/ctr.v3i2.245.

Amelsan, R., & Muntahanah, M. (2026). Analisis Perbandingan Kinerja Algoritma Naive Bayes Dan Svm Dalam Mengklasifikasi Sentimen Terhadap MBG. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *10*(3), 5442-5449. https://doi.org/10.36040/jati.v10i3.18357.

Anggraeni, M. D., Yudiananta, A. C., Arifin, A. H., & Arifin, W. (2025). Analisis Sentimen Masyarakat Terhadap Permasalahan Keracunan Program Makan Bergizi Gratis (MBG) pada Sosial Media 'X'. *Transformasi*, *21*(2). https://doi.org/10.56357/jt.v21i2.469.

Dinta, S. R., Manda, S., & Maulana, M. S. (2026). Implementasi Algoritma C4. 5 Dan Naïve Bayes Untuk Analisis Sentimen Publik Terhadap Makanan Bergizi Gratis Di Indonesia. *Jurnal Sistem Informasi Kaputama (JSIK)*, *10*(2), 87-99. https://doi.org/10.59697/jsik.v10i2.1312.

Hidayat, R., & Ratnaningsih, D. J. (2025). Analisis Sentimen Program Mbg Menggunakan Algoritma Random Forest Dan Naive Bayes. *Journal of Computing and Informatics Research*, *5*(1), 395-400. https://doi.org/10.47065/comforch.v5i1.2355.

Ihdahayuningsih, R., Rahmah, S. A., Rain, S. J., Adhari, H., Fahri, Y., & Siwu, G. C. I. (2026). Analisis Sentimen Masyarakat Terhadap Program Makan Bergizi Gratis Pada Media Sosial X Dengan Naïve Bayes. *Jurnal Sains Informatika Terapan*, *5*(1), 70-75. https://doi.org/10.62357/jsit.v5i1.929.

Iranda, M., & Huda, N. (2025). Analisis Kinerja Algoritma SVM dan Naïve Bayes untuk Klasifikasi Sentimen Program Makan Gratis. *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, *14*(3), 1452-1464. http://dx.doi.org/10.35889/jutisi.v14i3.3183.

Jakak, P. M., & Rohman, M. (2025). Analisis Sentimen Public Program Makan Bergizi Gratis Platform Instagram Dengan Algoritma Svm. *Smartics Journal*, *11*(1), 14-20. https://doi.org/10.21067/smartics.v11i1.11852.

Noviyanti, N. P. A., & Putra, I. G. N. A. C. (2025). Penerapan multinomial naïve Bayes dan chi-square pada analisis sentimen makan bergizi gratis. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, *3*(4), 785-794. https://doi.org/10.24843/JNATIA.2025.v03.i04.p08.

Prastiawan, K. H., & Yuniarto, D. (2026). Analisis Sentimen Publik terhadap Program Makan Bergizi Gratis dengan Algoritma Naive Bayes. *RIGGS: Journal of Artificial Intelligence and Digital Business*, *4*(4), 5412-5419. https://doi.org/10.31004/riggs.v4i4.3652.

Pratama, K. Y., Achmadi, S., & Sari, K. A. (2025). Analisis Sentimen Terhadap Makan Bergizi Gratis Pada Sosial Media X Menggunakan Metode Decicion Tree Dengan Pembanding Naive Bayes. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *9*(6), 11061-11067. https://doi.org/10.36040/jati.v9i6.15525.

Pratama, S. S., & Suryono, R. R. (2025). Perbandingan Algoritma Naïve Bayes Dan Support Vectore Machine Terhadap Program Bantuan Makan Gratis Pasca Pemilu. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, *10*(4), 3062-3074. https://doi.org/10.29100/jipi.v10i4.6654.

Putri, A. A., & Novianto, D. (2026). Analisis Sentimen Masyarakat Indonesia Terhadap Program Makan Bergizi Gratis (Mbg) Di Media Sosial X Menggunakan Perbandingan Algoritma Naive Bayes Dan Support Vector Machine. *Jati (Jurnal Mahasiswa Teknik Informatika)*, *10*(3), 5047-5052. https://doi.org/10.36040/jati.v10i3.18441.

Rahmatullah, B., Saputra, S. A., Budiono, P., & Wigandi, D. P. (2025). Sentimen Analisis Makan Bergizi Gratis Menggunakan Algoritma Naive Bayes. *Journal of Information Technology*, *5*(1). https://doi.org/10.46229/jifotech.v5i1.978.

.

Rifna, I., & Nurdin, N. (2026). Analisis sentimen pengguna tiktok terhadap program makan bergizi gratis menggunakan metode naive bayes. *IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi*, *5*(2), 241-251. https://doi.org/10.24246/itexplore.v5i2.2026.pp241-251.

Sakina, N., Wajidi, F., & Rasyid, M. R. (2025). Evaluasi Algoritma KNN dan Naive Bayes untuk Analisis Sentimen Kebijakan Program Makan Bergizi Gratis. *Jambura Journal of Informatics*, *1*(2), 83-97. https://doi.org/10.37905/jji.v1i2.34418.

Saleh, M. F., & Imanda, R. (2025). Public Sentiment Analysis of the Free Meal Program: A Comparison of Naive Bayes and Support Vector Machine Methods on the Twitter (X) Social Media Platform. *Journal of Applied Informatics and Computing*, *9*(1), 131-139. https://doi.org/10.30871/jaic.v9i1.8895.

Siregar, R. A. B., Manik, A. R., Syahri, A., Akbar, M. B., Arnita, A., & Ramadhani, F. (2025). Penerapan Naïve Bayes Untuk Analisis Sentimen Netizen Terhadap Program Makan Siang Gratis Pada Media Social X. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(4), 5621-5628. https://doi.org/10.36040/jati.v9i4.13870.

Triningsih, E. (2025). Analisis Sentimen Terhadap Program Makan Bergizi Gratis Menggunakan Algoritma Machine Learning Pada Sosial Media X. *Building Of Informatics, Technology And Science (BITS)*, *6*(4). https://doi.org/10.47065/bits.v2i2.492.

Tsabitah, L., Karima, D. A., Munaspin, Z. D. P., Titiana, N. M., & Fathoni, F. (2025). Analisis Sentimen Program Makan Siang Gratis Dalam Mendukung Sdgs Menggunakan Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(4), 6288-6294. https://doi.org/10.36040/jati.v9i4.14060.