International Journal Software Engineering and Computer Science (IJSECS)
6 (3), 2026, 1046-1057
Published Online December 2026 in IJSECS (http://www.journal.lembagakita.com/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v6i3.7809.

RESEARCH ARTICLE **Open Access**

Sentiment Analysis of Food Poisoning Incidents in Indonesia's Free Nutritious Meal Program Using TF-IDF and SVM

Muhammad Lucky Hermanto [1*], Ari Wedhasmara [2], Ken Ditha Tania [3], Allsela Meiriza [4]

[1*,2,3,4] Sriwijaya University, Palembang City, South Sumatra Province, Indonesia.

Article History:

Received: July 1, 2026.
Revised: July 3, 2026.
Accepted: August 27, 2026.
Available online: September 20, 2026.
Published: December 1, 2026.

***Corresponding Author:**
naufalromi83@gmail.com

Abstract: Food poisoning incidents associated with the Free Nutritious Meal Program (MBG) have triggered significant public reactions on the social media platform X/Twitter. This study identified and classified public sentiment regarding food poisoning occurrences in the MBG program by applying a Support Vector Machine (SVM) classification framework. Data were crawled from X/Twitter via Tweet Harvest using keywords related to MBG food poisoning incidents between January 2025 and June 2026. A total of 6,929 posts were collected, and 6,902 Indonesian-language records were retained after duplicate removal. The analytical workflow comprised data selection, automated sentiment labeling with IndoBERTweet, text preprocessing, Term Frequency-Inverse Document Frequency (TF-IDF) feature weighting, and classification using the SVM algorithm. Model performance was evaluated through a confusion matrix using accuracy, precision, recall, and F1-score metrics. Based on an evaluation of 1,989 testing records, the SVM model achieved an overall accuracy of 76.42%, exhibiting its strongest performance in the negative sentiment class. However, positive sentiment recorded a low recall of 26.03%, reflecting the difficulty of detecting the minority class. Sentiment distribution analysis revealed that negative sentiment dominated public discourse at 63.78%, followed by neutral sentiment at 29.05% and positive sentiment at 7.17%. These findings indicate that public perceptions of food poisoning incidents in the MBG program were overwhelmingly negative, underscoring critical concerns regarding food safety and program oversight. Beyond contributing to machine learning-based social media analytics, this study provides actionable insights for policymakers to evaluate service delivery and reinforce food safety standards in the MBG program.

Keywords: Sentiment Analysis; Free Nutritious Meal Program; Food Poisoning; Support Vector Machine; X/Twitter.

1. Introduction

The Free Nutritious Meal Program (Makan Bergizi Gratis [MBG]) has become one of the Indonesian government's flagship initiatives to improve public nutritional status and support human resource development toward the Indonesia Emas 2045 vision. Since its initial implementation on January 6, 2025, coordinated by the National Nutrition Agency (Badan Gizi Nasional [BGN]), the program has been progressively deployed in accordance with the operational readiness of Nutrition Fulfillment Service Units (Satuan Pelayanan Pemenuhan Gizi [SPPG]) across various regions. MBG functions as a preventive intervention designed to enhance nutritional intake, cultivate balanced dietary practices, and support the health of school-age children and other vulnerable demographics (bgn.go.id, 2025). The initiative is legally anchored by Presidential Regulation No. 83 of 2024 and further reinforced by Presidential Regulation No. 115 of 2025, which govern implementation mechanisms, priority locations, and operational standards (setkab.go.id, 2024; peraturan.go.id, 2025). Despite its strategic goals, MBG implementation has drawn extensive public scrutiny following reported food poisoning incidents associated with school meal distribution. A prominent incident occurred in Cimahi City on February 25, 2026, which intensified concerns regarding food safety protocols, kitchen sanitation certification, hygiene standards, and supply chain integrity (cimahikota.go.id, 2026). Media coverage subsequently reported widespread demands for rigorous evaluation and temporary operational suspensions of catering partners implicated in these service failures (cnnindonesia.com, 2025). Public complaints regarding food palatability, distribution delays, and kitchen operations further indicate that the program encounters substantial operational hurdles in maintaining consistent quality assurance and meal safety (gesuri.id, 2026; jemberkab.go.id, 2026). Public policy success is fundamentally influenced by community reception and trust. Negative perceptions triggered by service failures can erode public confidence and hinder policy acceptance (Wiranti & Ariawantara, 2024; Riyoldi *et al.*, 2025). In the context of MBG, food poisoning occurrences have provoked widespread criticism, skepticism, and apprehension, highlighting the importance of evaluating public reactions across digital communication channels (ombudsman.go.id, 2025; ugm.ac.id, 2025). Social media, particularly the platform X, represents an informative repository of real-time public sentiment, as users frequently voice spontaneous appraisals and grievances regarding public sector delivery. The platform's microblogging format facilitates rapid discourse propagation through reposts, replies, and hashtags, making X an effective environment for tracking public perception shifts during acute food safety incidents (Xia *et al.*, 2022; Xu, 2023).

To capture public sentiment from social media, text mining techniques can transform unstructured textual data into actionable insights through structured preprocessing and feature extraction pipelines (Silaban & Gumay, 2025). Sentiment analysis provides a practical framework for identifying public stances and emotional polarities. However, text derived from social media typically exhibits class imbalance, informal vocabulary, non-standard abbreviations, and syntactic irregularities. Consequently, robust machine learning algorithms are necessary for reliable classification. In text classification tasks, Support Vector Machine (SVM) has consistently demonstrated strong capabilities in handling high-dimensional, sparse feature spaces while maintaining robust generalization performance across complex text domains (Airlangga, 2024; Saputra *et al.*, 2025). Prior research has applied machine learning techniques to evaluate public opinion regarding government initiatives. Specifically, Triningsih *et al.* (2025) investigated general public sentiment toward the MBG program on social media X using SVM and Random Forest, observing that SVM coupled with SMOTE yielded an accuracy of 85.74%. Nevertheless, existing studies have primarily evaluated public perception of the MBG program at an aggregate level, without isolating public reactions specifically related to adverse food poisoning incidents. Consequently, a research gap remains regarding the targeted classification of public discourse surrounding meal poisoning incidents, which represent one of the most critical operational risks affecting program credibility. To address this gap, this study analyzes public sentiment toward food poisoning occurrences within the MBG program using Indonesian-language posts from platform X, with the Support Vector Machine algorithm serving as the primary classification framework. The findings aim to contribute methodologically to social media text mining applications for public policy monitoring while offering empirical, data-driven insights to assist regulatory authorities in strengthening food quality controls, kitchen inspection protocols, and responsive communication strategies.

2. Related Work

This section synthesizes prior studies investigating public sentiment toward the Free Nutritious Meal (MBG) program and examines methodological frameworks pertinent to this research, notably Support Vector Machine (SVM) classification, automated annotation via IndoBERTweet, and Word Cloud lexical analysis. To contextualize the methodological landscape and highlight the empirical gap addressed by this work, Table 1 summarizes relevant benchmark literature regarding analytical scope, data sources, corpus sizes, classification methods, evaluative performance, and inherent limitations.

Table 1. Summary of Prior Studies on MBG Sentiment Analysis and Machine Learning Classifiers

Reference	Research Focus	Data Source	Dataset Size	Method	Performance	Research Limitation
Triningsih *et al.* (2025)	General sentiment toward MBG	X/Twitter	2,400 tweets	TF-IDF, SVM, Random Forest, SMOTE	SVM + SMOTE: 85.74% accuracy	Examined general MBG sentiment rather than public reactions to specific food poisoning incidents.
Ramadhan & Khoirunnisa (2025)	General sentiment toward MBG	X/Twitter	3,417 tweets	TF-IDF, Random Forest, ensemble models	Random Forest: 91.15% accuracy; 95.46% ROC-AUC	Focused on overall MBG sentiment and ensemble performance rather than incident-specific food safety concerns.
Azmie *et al.* (2026)	Cross-platform sentiment toward MBG	X/Twitter and TikTok	5,173 posts	TF-IDF, Multinomial Naïve Bayes, SMOTE	SMOTE improved recall and class balance	Examined cross-platform MBG discourse without isolating food poisoning incidents as a primary focus.
Wibowo & Gunawan (2026)	General sentiment toward MBG	Twitter and TikTok	7,341 posts after cleaning	Bi-LSTM, IndoBERT, automated labeling via IndoBERTweet	IndoBERT: 82% accuracy; 79% F1-score	Centered on general program discourse and model benchmarking rather than incident-specific food safety sentiment.
Airlangga (2024)	Real-time disaster tweet classification	Twitter	Not specified	Naïve Bayes, RF, LR, SVM, LSTM	SVM: 89% accuracy; 88% F1-score	Validated event-driven classification performance but restricted to natural disaster domains.
Haliza & Ikhsan (2025)	Public sentiment toward IKN	X/Twitter	4,000 collected; 3,608 after preprocessing	TF-IDF, SVM, K-NN	SVM: 76% accuracy; 79% F1-score	Addressed national capital relocation policy; did not evaluate public health or food intervention programs.
Munthe & Lubis (2025)	Sentiment regarding IKN Garuda design	X/Twitter	5,128 tweets collected	SVM, Naïve Bayes, SMOTE	SVM: 77% (84% with SMOTE); Naïve Bayes: 73% (77% with SMOTE)	Evaluated architectural perception; relevance is strictly methodological regarding SVM

						performance on social data.
Gozali *et al*. (2026)	Sentiment toward #KaburAjaDulu	X/Twitter	12,353 tweets	IndoBERTweet labeling, IndoBERT embeddings, SVM	95.91% accuracy; 0.9581 weighted F1	Addressed a distinct socio-cultural trend; methodologically validated IndoBERTweet annotation paired with SVM classification.

2.1 Sentiment Analysis on the Free Nutritious Meal Program

Sentiment analysis has been widely applied to examine public perceptions of governmental interventions across microblogging platforms such as X/Twitter. In the context of the MBG initiative, several empirical inquiries have evaluated public reception using diverse supervised learning paradigms. Triningsih *et al*. (2025) analyzed 2,400 tweets using SVM and Random Forest combined with TF-IDF feature extraction and SMOTE balancing techniques. Their analysis revealed that negative sentiment was prominent (46%), primarily centered on budgetary constraints, while the integration of SVM with SMOTE yielded an optimal classification accuracy of 85.74%, demonstrating SVM's efficacy in handling polarized policy discourse. Expanding upon single-algorithm implementations, Ramadhan and Khoirunnisa (2025) deployed ensemble architectures across 3,417 tweets, reporting that an unstemmed Random Forest pipeline achieved 91.15% accuracy and an ROC-AUC of 95.46%, reinforcing the suitability of tree-based and margin-based classifiers for Indonesian policy tracking. Taking a multi-channel perspective, Azmie *et al*. (2026) evaluated public reactions across both X/Twitter and TikTok using Multinomial Naïve Bayes, TF-IDF, and SMOTE. Their findings highlighted distinct cross-platform dynamics: TikTok posts exhibited more pronounced critical rhetoric regarding kitchen accountability, whereas Twitter demonstrated relatively balanced evaluative discussions. These investigations collectively confirm the feasibility of computational text mining for tracking public opinion on MBG, yet they leave incident-specific risk occurrences largely unexamined.

2.2 Support Vector Machine for Social Media Sentiment Classification

Support Vector Machine (SVM) is widely recognized in text classification for its ability to project textual features into high-dimensional hyperplanes, thereby minimizing structural risk and mitigating overfitting in sparse data environments. Airlangga (2024) benchmarked multiple classifiers on crisis-related tweets, observing that SVM outperformed Naïve Bayes, Random Forest, and Logistic Regression with an accuracy of 89% and an F1-score of 88%, validating its stability on noisy, unstructured social corpora. The classifier's effectiveness has also been demonstrated across public sector evaluations in Indonesia. Haliza and Ikhsan (2025) examined citizen feedback regarding the Nusantara Capital City (IKN) policy, reporting that SVM achieved 76% accuracy and an F1-score of 79%, consistently outperforming K-Nearest Neighbor (K-NN). In a related policy assessment, Munthe and Lubis (2025) contrasted SVM against Naïve Bayes in classifying sentiment toward the IKN Garuda architectural design; SVM achieved superior baseline performance (77%) and climbed to 84% following synthetic minority oversampling. Synthesizing these empirical benchmarks indicates that SVM offers consistent discriminative capability when applied to informal Indonesian microblog texts, establishing its suitability for classifying targeted reactions to public health and meal safety concerns.

2.3 Automated Labeling Using IndoBERTweet

The high volume and real-time generation of social media corpora have increasingly motivated the deployment of domain-specific pretrained transformer models for automated data labeling, substantially reducing manual annotation overhead while preserving contextual fidelity. Wibowo and Gunawan (2026) utilized IndoBERTweet to automatically assign sentiment polarity to cross-platform MBG discussions, demonstrating that automated weak supervision produced reliable ground-truth distributions that supported downstream Bi-LSTM and IndoBERT models, achieving 82% accuracy and an F1-score of 79%. Similarly, Gozali *et al*. (2026) implemented IndoBERTweet to annotate over 12,000 tweets concerning social mobility trends before training an SVM classifier; their pipeline attained an overall accuracy of 95.91%, confirming that contextualized representations effectively capture informal slang, elisions, and colloquial Indonesian syntax without requiring manual annotations. These studies confirm that fine-tuned IndoBERTweet variants serve as a reliable, context-aware foundation for labeling complex public sentiment datasets.

2.4 Word Cloud Visualization in Sentiment Analysis

Complementing quantitative classification metrics, Word Cloud visualization provides a descriptive exploratory mechanism to surface lexical salience and discern thematic distributions within categorical subsets. Triningsih *et al.* (2025) applied lexical frequency visualizations to explore sentiment distributions surrounding the MBG program, observing that positive sentiment clustered around developmental descriptors such as "dukung", "sehat", and "stunting", whereas negative sentiment centered predominantly on fiscal allocation and systemic administration. Correspondingly, Azmie *et al.* (2026) employed comparative lexical clouds to examine cross-platform discourse differences, revealing that critical TikTok interactions featured sharper, accusatory vocabularies (e.g., "racun", "korupsi", and "dapur") compared to more policy-oriented phrasing on Twitter. Consequently, word cloud visualization functions as an effective interpretive aid, corroborating statistical model predictions by illustrating qualitative lexical patterns across distinct sentiment classes.

2.5 Research Gap and Positioning of This Study

As summarized in Table 1, existing scholarship on MBG sentiment analysis has concentrated primarily on aggregate policy assessments, focusing on broad themes such as macro-budgetary debates, administrative readiness, and overarching public acceptance. While these inquiries demonstrate the viability of machine learning classifiers (including SVM, Random Forest, Naïve Bayes, and transformer-based architectures), they evaluate public perceptions at a broad, general level. None of the reviewed MBG studies specifically isolated adverse operational failures—such as acute food poisoning incidents—as an independent analytical focus. This gap leaves an empirical blind spot regarding how the public perceives acute operational breakdowns within government nutritional initiatives. Public reactions to food poisoning incidents involve acute safety anxieties, institutional accountability questions, and direct concerns for student welfare, which differ fundamentally from general political or budgetary discourse. Addressing this distinction is vital for understanding how specific service failures shape institutional credibility. To bridge this empirical divide, this study shifts the analytical lens from aggregate program sentiment to incident-specific food safety perceptions. By capturing Indonesian-language discourse on X/Twitter, the methodology integrates automated labeling via IndoBERTweet, TF-IDF feature weighting, Linear SVM classification, and lexical visualization to delineate sentiment distribution and core thematic concerns regarding meal safety incidents. Consequently, the primary contribution of this study lies in providing a granular, incident-level evaluation of service delivery risks within the MBG initiative, directly complementing the macro-level policy literature.

3. Methodology

This study investigates public sentiment regarding food poisoning occurrences associated with the Free Nutritious Meal Program (MBG) using social media records retrieved from platform X. The research pipeline comprises data collection and filtering, initial preprocessing, automated sentiment labeling, advanced text preprocessing, feature extraction, sentiment classification, model evaluation, and lexical visualization. Figure 1 illustrates the overall analytical workflow.

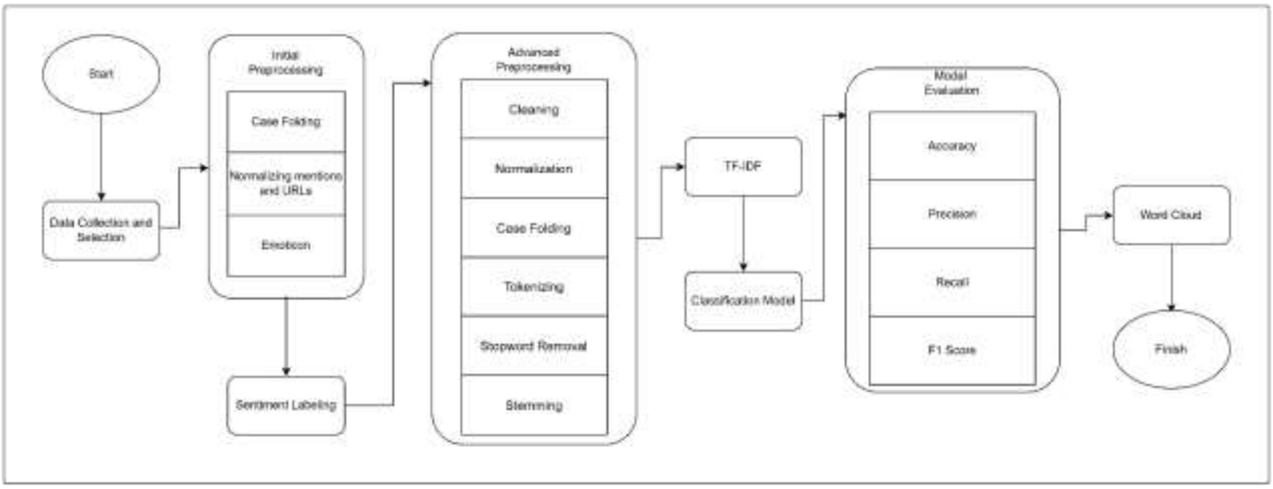

Figure 1. Research Framework

The corpus was gathered from X using Tweet Harvest. Data retrieval targeted Indonesian-language discussions related to food poisoning cases in the MBG program using key search queries, specifically "keracunan MBG" and "keracunan makan bergizi gratis", covering the period from January 2025 to June 2026. The collected data encompassed original tweets and direct user replies capturing public reactions, criticisms, concerns, and policy feedback. Following data extraction, duplicate entries were removed to eliminate informational redundancy and ensure corpus integrity, retaining only unique and topic-relevant records for subsequent analytical stages. Before performing sentiment labeling, the collected texts underwent initial preprocessing tailored to the architectural requirements of IndoBERTweet. Unlike conventional bag-of-words preprocessing, this stage preserved contextual syntax while standardizing social media elements. The procedure involved converting text to lowercase, replacing user mentions with the standardized token @USER, replacing web links with HTTPURL, and normalizing platform-specific emojis. This transformation strictly adheres to the input representation protocol established by Koto *et al.* (2021) for domain-specific IndoBERTweet modeling. Sentiment annotation was conducted automatically using a fine-tuned IndoBERTweet checkpoint, Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis (Ardiyanto, 2024).

This transformer architecture was fine-tuned specifically for three-way Indonesian sentiment classification: positive, negative, and neutral. In this pipeline, the model was executed in direct inference mode to categorize each tweet. Positive sentiment reflects support or approval of the MBG policy; negative sentiment captures criticism, anxiety, distress, or operational distrust concerning food poisoning incidents; and neutral sentiment comprises non-opinionated, informational reporting. To verify labeling fidelity, manual inspection was conducted on a randomized subset of the labeled corpus, confirming substantial alignment between predicted polarity and contextual tweet content. Following sentiment labeling, the dataset was processed through an advanced text-cleaning pipeline to prepare the textual features for machine learning classification. This stage encompassed punctuation and symbol removal, numerical stripping, case folding, tokenization, stopword removal based on formal and informal Indonesian lexicons, and morphological stemming. These transformations eliminated lexical noise, collapsed word inflections into base lemmas, and standardized morphological variants to enhance feature vector quality. To convert unstructured text into numerical inputs suitable for machine learning algorithms, Term Frequency–Inverse Document Frequency (TF-IDF) vectorization was applied. TF-IDF computes numerical weights by balancing the raw frequency of a term within an individual document against its inverse distribution across the entire corpus. This transformation maps the preprocessed texts into a high-dimensional feature space that emphasizes semantically discriminative terms while penalizing ubiquitous vocabulary. Sentiment classification was executed using Support Vector Machine (SVM) due to its established capability in handling high-dimensional, sparse text matrices generated by TF-IDF vectorization. The dataset was partitioned into training and testing subsets using a 70:30 stratified split, allocating 70% of the data to model training and the remaining 30% to unbiased performance evaluation. The classification model was instantiated using the LinearSVC module from Scikit-learn, which optimizes a linear decision boundary across high-dimensional feature spaces. The regularization parameter C was maintained at the default value of 1.0, and a random state of 42 was specified to ensure experimental reproducibility, without incorporating additional hyperparameter optimization. The classification performance of the SVM model was quantitatively evaluated through a multiclass confusion matrix and four benchmark performance metrics: accuracy, precision, recall, and F1-score. The evaluation metrics are formulated as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

$$\text{F}_1 = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \tag{4}$$

In these formulations, TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively, calculated across each sentiment category to measure the model's discriminative efficacy across negative, neutral, and positive classes. To corroborate quantitative model evaluations and explore dominant lexical patterns, Word Cloud visualizations were generated from the preprocessed corpus. This visualization highlights salient vocabulary by scaling word size proportionally to term frequency. Lexical clouds were rendered both globally and across stratified sentiment subsets to identify recurring thematic discourse and specific lexical markers associated with public reactions to MBG food poisoning incidents.

4. Results and Discussion

4.1 Results

The dataset was collected from platform X using the Tweet Harvest utility targeting discussions associated with food poisoning incidents within the Free Nutritious Meal Program (MBG). The scraping process yielded 6,929 raw records comprising 1,319 original tweets and 5,610 user comments and replies. Following data selection, 27 duplicate records were identified and removed, retaining 6,902 unique posts. During the advanced preprocessing stage, an additional 275 records were discarded due to the complete removal of non-alphanumeric noise, leaving no meaningful lexical tokens. Consequently, the finalized corpus comprised 6,627 cleaned records. Table 2 summarizes the corpus transformation at each pipeline stage.

Table 2.Summary of Collected and Preprocessed Dataset

Description	Count
Original tweets	1,319
Direct replies and comments	5,610
Total collected records	6,929
Duplicate records removed	27
Unique records after selection	6,902
Cleaned records after preprocessing	6,627

Automated sentiment annotation using the fine-tuned IndoBERTweet model revealed that public discourse was overwhelmingly polarized toward negative sentiment. As outlined in Table 3, negative sentiment accounted for 4,402 records (63.78%), whereas neutral sentiment comprised 2,005 records (29.05%), and positive sentiment constituted the smallest subset with 495 records (7.17%). This distribution demonstrates that online conversations surrounding MBG food poisoning occurrences were predominantly characterized by criticism, public apprehension, and dissatisfaction regarding food safety and meal handling, with supportive expressions remaining marginal.

Table 3. Distribution of Sentiment Classes

Sentiment Class	Count	Percentage
Negative	4,402	63.78%
Neutral	2,005	29.05%
Positive	495	7.17%

Following text filtering, the 6,627 preprocessed documents were partitioned into training (70%, n = 4,638) and testing (30%, n = 1,989) subsets. TF-IDF vectorization extracted 7,241 distinct unigram features from the training split. Table 4 presents the ten lexical terms exhibiting the highest mean TF-IDF weights across the corpus. The terms "mbg" (0.0347), "tidak" (0.0344), and "racun" (0.0313) emerged as the most discriminative textual features. The prominence of "racun" (poison) indicates that discussions maintained an acute focus on contamination cases, while the frequent appearance of "anak" (children) reflects public concern regarding the health and safety of school-age recipients.

Table 4. Top Ten Features Ranked by Mean TF-IDF Weight

Term	Mean TF-IDF
mbg	0.0347
tidak	0.0344
racun	0.0313
makan	0.0288
jadi	0.0204
gizi	0.0203
program	0.0196
gratis	0.0192
anak	0.0179
orang	0.0132

The Linear SVM model was trained on the TF-IDF feature representations and evaluated against the 1,989 unseen test instances. Overall, the classifier correctly predicted 1,520 instances while misclassifying 469 instances, attaining an overall classification accuracy of 76.42%. Figure 2 and Table 5 illustrate the complete multi-class confusion matrix. The model exhibited its highest classification sensitivity within the majority negative sentiment class, correctly identifying 1,174 out of 1,305 instances. Conversely, the model correctly classified 308 of 538 neutral instances (with 226 misclassified as negative) and 38 of 146 positive instances (with 80 misclassified as negative and 28 as neutral). This misclassification pattern indicates a notable classification bias toward the majority class, driven by the acute imbalance of the training corpus.

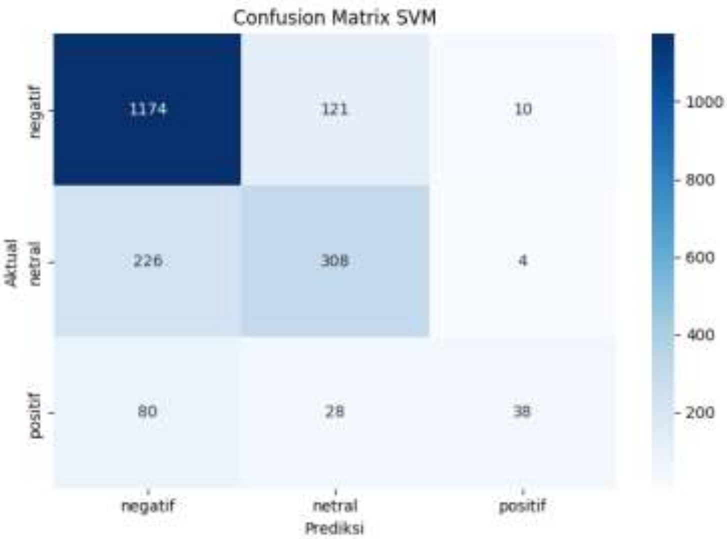

Figure 2. Confusion Matrix of the Linear SVM Model

Table 5. Confusion Matrix of Classification Predictions

Actual Sentiment	Predicted Negative	Predicted Neutral	Predicted Positive	Total
Negative	1,174	121	10	1,305
Neutral	226	308	4	538
Positive	80	28	38	146
Total	1,480	457	52	1,989

Detailed classification performance across individual classes is summarized in Table 6. Negative sentiment achieved the highest overall metrics, recording a precision of 0.7932, a recall of 0.8996, and an F1-score of 0.8431, confirming the model's reliability in identifying adverse public feedback. Neutral sentiment attained moderate performance with a precision of 0.6740, a recall of 0.5725, and an F1-score of 0.6191. In contrast, while the positive class achieved an acceptable precision of 0.7308, its recall dropped sharply to 0.2603, resulting in an F1-score of 0.3838. This indicates that while the model's positive predictions were relatively reliable, it failed to detect the majority of actual positive tweets, categorizing them instead under negative or neutral classes due to limited representative training instances.

Table 6. Classification Performance Metrics

Sentiment Class	Precision	Recall	F1-Score
Negative	0.7932	0.8996	0.8431
Neutral	0.6740	0.5725	0.6191
Positive	0.7308	0.2603	0.3838
Accuracy			0.7642

Word Cloud visualizations were constructed to examine qualitative discourse patterns across stratified sentiment categories, as illustrated in Figures 3, 4, and 5. For positive sentiment (Figure 3), discourse clustered around developmental and supportive terms, including "program", "gizi", "gratis", "anak", and "aman", emphasizing perceived nutritional benefits and child health improvements. Neutral discourse (Figure 4) was characterized by informational and reporting vocabulary such as "gizi", "makan", "mbg", and "program", reflecting non-evaluative news dissemination and policy updates. Conversely, negative sentiment (Figure 5) was dominated by acute risk terms such as "racun", "anak", "tidak", and "banyak", illustrating widespread public dissatisfaction, criticism regarding operational negligence, and anxieties surrounding student vulnerability during meal consumption.

Figure 3. Word Cloud Visualization for Positive Sentiment

Figure 4. Word Cloud Visualization for Neutral Sentiment

Figure 5. Word Cloud Visualization for Negative Sentiment

4.2 Discussion

The empirical findings reveal that public discussions concerning food poisoning incidents in the MBG program were overwhelmingly dominated by negative sentiment (63.78%). This substantial negative polarity demonstrates that operational meal safety incidents exert a critical influence on citizen evaluations of public health interventions. This finding aligns with Triningsih *et al.* (2025), who similarly identified negative sentiment as the prevailing response toward MBG on social media. However, the proportion of negative sentiment documented in this study is markedly higher than the 46% observed in their aggregate analysis. This contrast indicates that specific service breakdowns—particularly events threatening child safety—provoke substantially stronger critical reactions than general policy announcements or macroeconomic budgetary debates. The TF-IDF feature analysis and lexical visualizations further substantiate this dynamic. Salient terms such as "racun", "tidak", "anak", and "mbg" consistently registered the highest weights and frequencies across the corpus. The prominence of "anak" (children) underscores that public concern is intimately tied to the vulnerability of school-aged recipients, while the frequency of "racun" (poison) indicates acute public fixation on meal contamination risks. These lexical markers demonstrate that citizen dissatisfaction is driven not merely by bureaucratic disagreement, but by direct anxiety over child health, hygiene standards, and kitchen oversight. From a computational modeling perspective, the Linear SVM classifier achieved a robust accuracy of 76.42% and a weighted F1-score of 74.88%, demonstrating the viability of pairing TF-IDF representations with margin-based classifiers for Indonesian microblog texts. Nonetheless, performance disparities across sentiment classes were pronounced.

The model achieved superior predictive sensitivity for negative sentiment (recall: 0.8996), but experienced notable degradation when identifying positive sentiment (recall: 0.2603). This disparity is primarily attributable to severe class imbalance, as positive tweets constituted only 7.17% of the dataset compared to 63.78% for negative instances. With limited training exemplars, the classifier encountered difficulty establishing precise decision boundaries for the minority class, frequently misclassifying nuanced positive statements as negative. The Word Cloud analysis reinforces these quantitative distinctions by delineating thematic boundaries. While supportive interactions highlighted nutritional goals through terms like "gizi" and "aman", critical discourse was heavily anchored in operational grievances and contamination terminology. This demonstrates that while baseline appreciation for nutritional enhancement exists, acute operational incidents rapidly overshadow policy goodwill on public communication platforms. These findings carry direct practical implications for public administration and crisis management. The high concentration of negative sentiment underscores that food poisoning incidents represent a severe operational vulnerability capable of eroding trust in government initiatives. To mitigate reputational damage and protect public health, regulatory authorities must institute rigorous sanitation certification, real-time kitchen monitoring, and strict supply chain oversight. Furthermore, program administrators should deploy social media sentiment analytics as an active monitoring mechanism to detect emerging contamination reports rapidly. Transparent, immediate communication detailing corrective actions, kitchen sanctions, and clinical updates is essential to alleviate public anxiety and restore institutional credibility.

5. Conclusion and Recommendations

This study evaluated public sentiment surrounding food poisoning incidents within the Free Nutritious Meal Program (MBG) by implementing a machine learning classification framework on social media platform X. Based on the analysis of 6,902 unique posts annotated via IndoBERTweet, public discourse was heavily dominated by negative sentiment (63.78%), with neutral and positive expressions constituting 29.05% and 7.17%, respectively. These findings demonstrate that acute food safety incidents provoke widespread public criticism, apprehension, and operational skepticism, significantly outweighing positive sentiment regarding the nutritional objectives of the program. The classification pipeline combining TF-IDF feature representations with a Linear SVM classifier attained an overall accuracy of 76.42% and a weighted F1-score of 74.88%. While the classifier demonstrated robust predictive sensitivity in identifying majority negative feedback (recall: 0.8996), detection of positive sentiment was constrained by acute class imbalance (recall: 0.2603). Furthermore, lexical analysis substantiated that terms such as "mbg", "tidak", "racun", and "anak" anchored public conversations, illustrating that citizen concern centered primarily on food hygiene failures and the health vulnerability of school-aged recipients. These insights highlight that operational meal safety is a decisive determinant of institutional trust. Consequently, program authorities must enforce stringent kitchen hygiene certification, real-time sanitation inspections, transparent incident reporting mechanisms, and rapid corrective interventions to rebuild public confidence and ensure the long-term sustainability of the MBG initiative. Several limitations should be noted. This investigation relied exclusively on textual data retrieved from platform X, which may reflect specific demographic tendencies and exclude public opinions articulated across alternative digital platforms or offline communities. Additionally, sentiment annotations were generated through automated transformer-based inference; although manual spot-checking confirmed acceptable alignment, automated labeling remains susceptible to misinterpreting contextual ambiguity, irony, and complex colloquial syntax. Subsequent inquiries should address these constraints by capturing longitudinal datasets spanning multiple social networks, such as TikTok, Instagram, and Facebook, to achieve broader demographic representation. Methodologically, incorporating synthetic data balancing techniques, including SMOTE or adaptive thresholding, could mitigate performance degradation on minority classes. Finally, future work should explore comparative benchmarks between conventional margin-based algorithms and end-to-end fine-tuned transformer architectures, while integrating extensive multi-annotator manual validation to enhance label fidelity in public policy evaluation.

References

Airlangga, G. (2024). Comparative analysis of machine learning models for real-time disaster tweet classification: Enhancing emergency response with social media analytics. *Brilliance: Research of Artificial Intelligence*, *4*(1), 25–31. https://doi.org/10.47709/brilliance.v4i1.3669

Ardiyanto, M. (2024). *Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis* [Machine learning model]. Hugging Face. https://doi.org/10.57967/hf/5111

Azmie, F. U., Irawan, Y., & Setiawan, R. R. (2026). Sentiment analysis of Free Nutritious Meal programs using Naïve Bayes on platforms X and TikTok. *Jurnal Teknologi Informasi dan Pendidikan*, *19*(1), 1208–1227. https://doi.org/10.24036/jtip.v19i1.1112

Badan Gizi Nasional. (2025). *BGN akan memulai program MBG secara bertahap*. https://www.bgn.go.id/news/artikel/bgn-akan-memulai-program-mbg-secara-bertahap

CNN Indonesia. (2025, September 22). *Ribuan anak sekolah keracunan, KPAI usul MBG dihentikan sementara*. https://www.cnnindonesia.com/nasional/20250922155429-32-1276424/ribuan-anak-sekolah-keracunan-kpai-usul-mbg-dihentikan-sementara

Gesuri. (2026). *Reses, Sudin terima keluhan warga soal kualitas program makan bergizi gratis*. https://www.gesuri.id/kerakyatan/reses-sudin-terima-keluhan-warga-soal-kualitas-program-makan-bergizi-gratis-b2pNlZbdmd

Gozali, M. R. F., Tibyani, & Brata, D. W. (2026). Analisis sentimen pengguna Twitter/X terhadap fenomena #KaburAjaDulu menggunakan metode Support Vector Machine dan IndoBERT embedding. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *10*(2), 125–134. https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/16005

Haliza, D., & Ikhsan, M. (2025). Sentiment analysis on public perception of the Nusantara Capital on social media X using Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) methods. *Journal of Applied Informatics and Computing*, *9*(3), 716–723. https://doi.org/10.30871/jaic.v9i3.9459

Kementerian Hukum dan Hak Asasi Manusia Republik Indonesia. (2025). *Peraturan Presiden Nomor 115 Tahun 2025 tentang Tata Kelola Penyelenggaraan Program Makan Bergizi Gratis*. https://peraturan.go.id/id/perpres-no-115-tahun-2025

Koto, F., Lau, J. H., & Baldwin, T. (2021). INDOBERTWEET: A pretrained language model for Indonesian Twitter with effective domain-specific vocabulary initialization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 10660–10668). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.833

Munthe, M. H. R. P., & Lubis, A. H. (2025). Comparison of Support Vector Machine (SVM) and Naïve Bayes algorithm performance in analyzing Garuda bird design sentiment in IKN. *JITE (Journal of Informatics and Telecommunication Engineering)*, *8*(3 Special Issue), 1–8. https://doi.org/10.31289/jite.v8i3Spc.14830

Ombudsman Republik Indonesia. (2025). *Ombudsman RI temukan empat potensi maladministrasi dalam program makan bergizi gratis*. https://ombudsman.go.id/news/r/ombudsman-ri-temukan-empat-potensi-maladministrasi-dalam-program-makan-bergizi-gratis

Pemerintah Kabupaten Jember. (2026). *Tanggapi keluhan kualitas makan bergizi gratis, Pemkab Jember lakukan sidak dan peninjauan di beberapa SPPG*. https://www.jemberkab.go.id/tanggapi-keluhan-kualitas-makan-bergizi-gratis-pemkab-jember-lakukan-sidak-dan-peninjauan-di-beberapa-sppg/

Pemerintah Kota Cimahi. (2026). *Sejumlah siswa dan guru di Kota Cimahi diduga alami keracunan MBG*. https://cimahikota.go.id/berita/detail/83718-sejumlah-siswa-dan-guru-di-kota-cimahi-diduga-alami-keracunan-mbg

Ramadhan, N. G., & Khoirunnisa, A. (2025). An integrated Random Forest for analyzing public sentiment on the "Makan Bergizi Gratis" program. *Journal of Information Systems and Informatics*, *7*(3), 2314–2328. https://doi.org/10.51519/journalisi.v6i2.759

Riyoldi, R., Muhtar, E. A., & Karlina, N. (2025). Public perception analysis of digital population identity policy in Pontianak City using Technology Acceptance Model (TAM) approach. *International Journal of Social Science and Education*, *3*(1), 112–124. https://doi.org/10.59890/iijse.v3i1.6211

Saputra, A. N. A., Saputro, R. E., & Saputra, D. I. S. (2025). Enhancing sentiment analysis accuracy using SVM and slang word normalization on YouTube comments. *Sinkron*, *9*(2), 687–699. https://doi.org/10.33395/sinkron.v9i2.14613

Sekretariat Kabinet Republik Indonesia. (2024). *Inilah Perpres 83 Tahun 2024 tentang Badan Gizi Nasional*. https://setkab.go.id/inilah-perpres-83-tahun-2024-tentang-badan-gizi-nasional/

Silaban, N., & Gumay, M. G. (2025). Analisis sentimen pengguna aplikasi pinjaman online melalui media sosial X (Twitter) menggunakan metode Support Vector Machine. *Jurnal Minfo Polgan*, *14*(2), 2116–2130. https://doi.org/10.33395/jmp.v14i2.15300

Triningsih, E., Afdal, M., Permana, I., & Evrilyan Rozanda, N. (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), 2240–2250. https://doi.org/10.47065/bits.v6i4.6534

Universitas Gadjah Mada. (2025). *Keracunan MBG kembali terulang, guru besar UGM soroti lemahnya regulasi dan pengawasan*. https://ugm.ac.id/id/berita/keracunan-mbg-kembali-terulang-guru-besar-ugm-soroti-lemahnya-regulasi-dan-pengawasan/

Wibowo, A. P., & Gunawan, D. (2026). Analisis sentimen publik terhadap Makan Bergizi Gratis menggunakan Bi-LSTM dan IndoBERT. *Jurnal Sistem Komputer dan Informatika (JSON)*, *7*(3), 842–852. https://doi.org/10.30865/json.v7i3.9458

Wiranti, D. A., & Ariawantara, P. A. F. (2024). Assessing the role of Surabaya City Government's responsiveness toward public acceptance of COVID-19 mitigation policies. *Jurnal Borneo Administrator*, *20*(1), 85–100. https://doi.org/10.24258/jba.v20i1.1359

Xia, L., Chen, B., Hunt, K., Zhuang, J., & Song, C. (2022). Food safety awareness and opinions in China: A social network analysis approach. *Foods*, *11*(18), 2909. https://doi.org/10.3390/foods11182909

Xu, B. (2023). Monitoring and guidance of public opinions on food safety based on information retrieval and data mining: An empirical study of microblog public opinions on food safety in large farmers' markets. *Food Science*, *44*(7), 404–412. https://doi.org/10.7506/spkx1002-6630-20221125-297