



Comparison of Support Vector Machine and Extreme Learning Machine for Sentiment Analysis of Free Nutritious Meal Program on X Application

Rizqi Aulia Wijayanti*, Arief Agoestanto

Department of Mathematics, Faculty of Mathematics and Natural Science
Universitas Negeri Semarang, Semarang, 50229, Indonesia

Article Info

Article History:

Received 03-07-2025
Accepted 20-10-2025
Published 31-10-2025

Keywords:

Sentiment analysis, Support vector machine, Extreme learning machine, Free nutritious meal program

Abstract

The Free Nutritious Meal Program (MBG) has drawn wide opinion on social media, and the government needs a fast and reliable way to read it, but tweets are unstructured and numerous. This study compares Support Vector Machine (SVM), a strong classifier for high-dimensional TF-IDF text features, and Extreme Learning Machine (ELM), a fast, analytically trained single-hidden-layer network, for classifying the sentiment of MBG tweets. A total of 7,398 tweets (April 21–26, 2025) were cleaned, labeled positive or negative with a lexicon, and represented by TF-IDF with unigrams and bigrams. Four SVM kernels were tuned by grid search, and ELM was tested with four activation functions and four hidden-layer sizes; the models were compared by accuracy, precision, recall, F1-score, and the F1-gap between training and test data. The linear-kernel SVM was the best (accuracy 88.4%, F1-gap 5.82%), ahead of the linear-activation ELM (83.8%, F1-gap 2.54%), which generalized more tightly. Negative sentiment predominated (about 55%) and centered on food safety and distrust, whereas positive tweets referred to nutrition, health, and support. A linear SVM is recommended for monitoring sentiment, and the authorities are advised to strengthen food safety assurance and public communication.

How to cite:

Wijayanti, R. A., & Agoestanto, A. (2025). Comparison of support vector machine and extreme learning machine for sentiment analysis of free nutritious meal program on X application. *UNNES Journal of Mathematics*. 14(2): 49-57

1. Introduction

In the current digital era, data analysis has emerged as a critically important field, particularly in processing unstructured text data generated from various online interactions. Data mining, as one of the key data processing methods, serves to identify patterns and trends within datasets. Among data mining techniques, text mining plays a vital role—a specialized approach that extracts meaningful information from textual data by discovering underlying patterns (Alhaq et al., 2021). The capability of text mining to process and analyze large volumes of textual data makes it particularly valuable for sentiment analysis applications.

Sentiment analysis aims to classify opinions from social media comments or posts to understand public perspectives on specific topics (Indarbensyah & Rochmawati, 2021). This technique can be applied at various levels, from document-level to sentence-level analysis (Husada & Paramita, 2021). In understanding social media users' responses to policies, products, or phenomena, sentiment analysis plays a crucial role by classifying sentiments into positive, negative, and neutral categories while measuring classification accuracy. Among various methods for sentiment analysis, Support Vector Machine (SVM) and Extreme Learning Machine (ELM) algorithms have shown particular promise.

Social media has significantly transformed public behaviour, encouraging freer expression of opinions (Isnain et al., 2021). These platforms also serve as important channels for disseminating government program information and facilitating public discourse where citizens can voice their perspectives (Arifin et al., 2021). Application X (formerly Twitter) has become a prominent medium for public opinion expression. Recently, the Free Nutritious Meal Program has generated substantial discussion across various platforms, including Application X, where users share diverse opinions ranging from criticism to suggestions and praise.

The Free Nutritious Meal Program (MBG) represents a government initiative to provide complimentary nutritious meals for elementary through high school students, aiming to address nutritional deficiencies and support academic performance. Implemented gradually across various Indonesian regions, the program has generated numerous reviews from students and community members regarding its implementation. These public reviews enable the government to evaluate program effectiveness and to determine future directions for the program's continuation.

The Free Nutritious Meal Program (MBG) is a large national policy, and its early implementation has been accompanied by both public criticism, including complaints about food quality and safety, and support for its goals. The concrete problem studied here is that the government needs a fast and systematic way to read these opinions, but the tweets are unstructured, informal, and numerous (7,398 tweets were collected within six days), so that manual reading is not feasible, and it is not known which classifier labels them most accurately and most reliably. The urgency is that the program is being implemented in stages, so sentiment information is needed now to guide evaluation and communication, and a classifier that has been compared and validated is needed to produce it.

The Support Vector Machine (SVM) algorithm is primarily used as a binary classifier that separates data into two classes using a hyperplane (Arifin et al., 2021). SVM demonstrates effectiveness and stability in handling high-dimensional data, offers easy training processes, and maintains memory efficiency through kernel mapping to feature space (Iriananda et al., 2024). While real-world problems often involve non-linear separability, SVM can be extended to nonlinear classification through the implementation of the kernel trick (Fitriyah et al., 2020). SVM's kernel technique transforms data into a higher-dimensional feature space, converting nonlinear problems in the original space into linearly separable ones without significant computational overhead (Maggioni & Spinelli, 2024). Various kernel types, including linear, polynomial, radial basis function (RBF), and sigmoid, can be employed depending on data characteristics.

Extreme Learning Machine (ELM) represents a fast-learning algorithm that achieves good generalisation by randomly selecting hidden nodes and analytically determining output weights in single-hidden-layer feedforward networks (SLFNs) (F. Wang et al., 2017). ELM's core principle lies in the random initialisation of weights and biases between the input and hidden layers, followed by the analytical computation of output weights from the hidden layer (Ozyildirim & Coban, 2018). In its architecture, ELM employs fixed random initialisation for hidden layer weights and biases during training, where system nonlinearity is achieved through activation functions while maintaining linear model properties, as only output weights require optimisation (J. Wang et al., 2022). The ELM algorithm excels in computational efficiency by eliminating iterative bias and weight determination processes, enabling rapid approximation of target functions and effective data classification (Rausanfitra et al., 2018).

SVM and ELM were chosen for complementary reasons. SVM (Cortes & Vapnik, 1995) is a standard and strong classifier for sparse, high-dimensional TF-IDF text features, and its kernels allow linear and nonlinear class boundaries to be compared. ELM (Huang et al., 2006) trains a single-hidden-layer network in one analytical step, which makes it fast and allows the number of hidden neurons and the activation function to be compared quickly. Comparing them shows whether the stronger but heavier SVM or the faster ELM is preferable for this kind of data. TF-IDF with unigrams and bigrams was used because it is a common and simple representation for both classifiers (Pang et al., 2002; Liu, 2012).

Previous research by Purwanti & Sugiyono (2024) demonstrated SVM's effectiveness in processing data with high accuracy (94%), while analysis revealed predominantly positive public perception toward the free lunch program. Muttaqin & Kharisudin (2021) found that linear-kernel SVM achieved superior performance with 87.98% accuracy, 88.55% precision, and 95.43% recall. Riskiyah et al. (2024) showed ELM's capability in classifying Twitter user sentiments toward GOJEK services with 72% accuracy.

These studies examined SVM and ELM separately, on different data sets and with different preprocessing, so that their accuracies cannot be compared directly. ELM was also applied to Twitter sentiment by Ozyildirim and Coban (2018) and to Samsung Indonesia by Rausanfitra et al. (2018). To the authors' knowledge, no study has compared SVM kernels and ELM activation functions on the same tweets about the MBG program.

The novelty of this study is threefold: (1) SVM with four kernels and ELM with four activation functions and four hidden-layer sizes are compared on one data set of MBG tweets with the same preprocessing and labeling; (2) generalization is evaluated through the F1-gap between training and test data in addition to accuracy, precision, recall, and F1-score; and (3) the best classifier is used together with word clouds to interpret the content of positive and negative tweets and to derive recommendations for the program.

This study aims to compare the performance of SVM and ELM in classifying the sentiment of tweets about the MBG program and to identify public perceptions of the program. The research questions are: (1) which SVM kernel and which ELM configuration give the best classification; (2) which of the two methods performs better in accuracy, precision, recall, F1-score, and generalization (F1-gap); and (3) what do the positive and negative tweets reveal about public perception of the program, and what should be done in response?

2. Method

The data used in this study were obtained from the social media platform X (formerly known as Twitter). The dataset consists of tweets discussing the Free Nutritious Meal Program. Data collection was performed using a web scraping technique assisted by the Python tweet-harvest library, which enables the extraction of tweets based on relevant keywords. The acquired dataset includes several key variables: the date of tweet publication, tweet text content, and username.

The dataset analyzed in this study comprises three primary variables collected from the social media platform X (formerly Twitter). The first variable, *Timestamp*, captures the exact posting time of each tweet. The second variable, *Text Content*, contains the complete unstructured message of each tweet, including relevant linguistic features such as emojis, hashtags, and mentions related to the discussion of the Free Nutritious Meal Program. The third variable, *User Identifier*, refers to anonymized account usernames (hashed for privacy protection) to ensure data confidentiality while maintaining dataset integrity in accordance with ethical research standards.

In the data analysis stage, all data were processed using statistical techniques and machine learning algorithms to uncover meaningful patterns, relationships, and actionable insights that support evidence-based decision-making. The data analysis was conducted through the following stages.

(1). Data Preprocessing

Data were processed using Python through stages of cleaning special characters, punctuation, numbers, and irrelevant symbols, followed by text normalization, tokenization into word units, stopword removal, and stemming to return words to their base forms.

(2). Feature Extraction Using TF-IDF

Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method to weight each word based on its relative frequency, with unigram and bigram approaches to capture context and phrase patterns, as well as relative frequency distributions across documents.

(3). Labeling

Data labeling was performed automatically using Python with a lexicon-based approach to classify text into positive or negative sentiments. This process was accelerated by importing labeled data from external sources, and labeling consistency was essential to ensure classification model accuracy.

(4). Dataset Splitting

The processed dataset was divided into training and testing data using Python in a 70:30 or 80:20 ratio. This division was intended to prevent overfitting and evaluate model performance.

(5). SVM Model Building and Training

The model was developed and trained using the Support Vector Machine (SVM) algorithm with the kernel trick, selected for its ability to efficiently handle high-dimensional text data. Training was conducted using Python and the scikit-learn library on the extracted features to classify text into positive and negative sentiments. Mathematically, SVM works by forming a separating hyperplane in the feature space defined by the equation:

$$(w \cdot x) + b = 0 \quad (1)$$

where w denotes the weight vector, x is the feature vector, and b is the bias term. The SVM objective is to maximize the margin separating the two classes: positive sentiment ($y_i = 1$) and negative sentiment ($y_i = -1$), subject to the following constraints:

$$y_i(w \cdot x_i + b) \geq 1, \quad \forall i \quad (2)$$

The margin is defined as $\frac{2}{\|w\|}$, and the optimization objective becomes minimizing $\frac{1}{2} \|w\|^2$.

This constrained problem is solved using Lagrange multipliers (α_i):

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i(w \cdot x_i + b) - 1] \quad (3)$$

(6). ELM Model Building and Training

The model was developed using the Extreme Learning Machine (ELM) algorithm with a single hidden-layer feedforward neural network approach, which analytically determines the output weights while keeping the input weights randomly assigned. Training was performed using the extracted features and nonlinear activation functions, implemented in Python using libraries such as HPELM or scikit-learn-based ELM modules. Mathematically, for an input vector $x_i \in \mathbb{R}^n$ with n features, the nonlinear transformation to the hidden layer is defined as:

$$z_{ij} = g(w_j^T x_i + b_j) \quad (4)$$

where $w_j \in \mathbb{R}^n$ is the input weight vector to the j -th neuron, $b_j \in \mathbb{R}$ is the bias term of the j -th neuron, and $g(\cdot)$ denotes the activation function.

The hidden layer output for N training samples and L neurons is represented in matrix $H \in \mathbb{R}^{N \times L}$ with elements:

$$H_{ij} = g(w_i^T x_j + b_j) \quad (5)$$

The output weights $\beta \in \mathbb{R}^L$ are computed through least-squares solutions:

- Without regularization (Moore–Penrose Pseudo-Inverse):

$$\beta = (H^T H)^{-1} H^T t \quad (6)$$

- With ridge regularization:

$$\beta = (H^T H + \lambda I)^{-1} H^T t \quad (7)$$

where λ is the regularization parameter and I is the identity matrix.

(7). Model Evaluation

Model evaluation is a critical step in assessing the classification performance of both SVM and ELM. Key metrics, including accuracy, precision, recall, and F1-score, were computed on the test dataset to evaluate the models' generalization performance. Evaluation was performed using Python's scikit-learn library, which offers built-in functions for metric computation and visualization to facilitate result interpretation.

The metrics are accuracy = $(TP + TN) / (TP + TN + FP + FN)$, precision = $TP / (TP + FP)$, recall = $TP / (TP + FN)$, and F1-score = $2 \times \text{precision} \times \text{recall} / (\text{precision} + \text{recall})$, where TP, TN, FP, and FN are the numbers of true positives, true negatives, false positives, and false negatives. Generalization was measured with the F1-gap, defined as the F1-score on the training data minus the F1-score on the test data; a small gap indicates good generalization and a large gap indicates overfitting.

(8). Word Cloud Visualization of Text Data

Text data were visualized using word clouds, where word size corresponds to frequency of occurrence. This approach facilitates identification of key terms in the corpus and supports sentiment distribution analysis. The implementation utilized Python libraries such as wordcloud, matplotlib, and NLTK for visualization generation and customization. The flowchart of the research process stages is presented below.

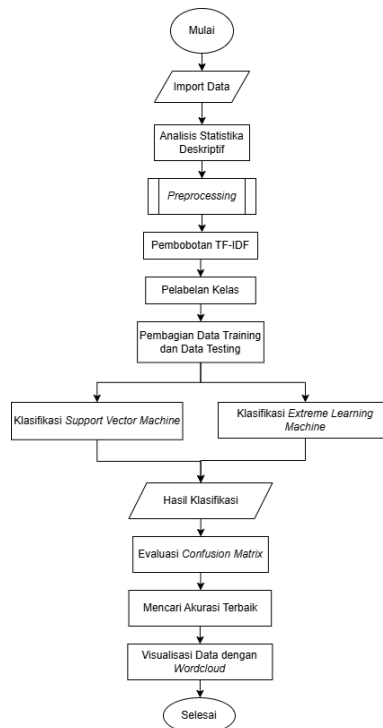


Figure 1. Research Flow Chart

3. Results and Discussion

3.1. Data Collection

The research dataset was collected through web crawling techniques from platform X (<https://x.com/>) using the tweet harvest library. The data collection process began with retrieving the auth_token via the browser's inspect element tool, which was then integrated into the crawling code to access X's official API. A total of 7,398 tweets were collected over two collection periods (April 21–23 and April 23–26, 2025) to overcome session-based data retrieval limits. Dividing the collection into separate periods ensured sufficient data coverage for comprehensive sentiment analysis.

3.2. Text Preprocessing

Prior to analysis, the raw crawled data underwent essential preprocessing stages. This transformation process converted unstructured, noisy data—containing special characters, duplicates, and irrelevant content—into a cleaner, structured format suitable for subsequent analysis. The implemented preprocessing pipeline included: case folding, cleansing, tokenisation, spelling normalisation, stopwords removal, stemming, and reconstruction into a single string.

Table 1. Input and Output Preprocessing Phase

Before	After
melihat banyaknya permasalahan yg timbul akibat program mbg ketua dpr ri @puanmaharani_ri mendesak pemerintah untuk segera melakukan evaluasi secara menyeluruh dan transparan terhadap program tersebut #puanmaharani https://t.co/twh2opkbv	banyak maslaah timbul akibat program mbg ketua dpr ri desak pemerintah evaluasi transparan program

3.3. Sentiment Labelling

In this stage, sentiment labelling was performed on the preprocessed dataset using a lexicon-based approach, employing a CSV file containing a list of words or phrases with predefined sentiment scores indicating positive or negative polarity. This dictionary enabled the automated identification and quantification of sentiment within the text. The lexicon-based classification results revealed the following sentiment distribution among platform x user reviews: 3,935 tweets were labeled as positive and 3,204 as negative.

3.4. TF-IDF Weighting

The processed textual data required numerical transformation through term weighting using the tf-idf method, implemented in python within the google colab environment. The resulting weights served as input features for the sentiment classification models. Analysis revealed the three most frequent terms: “mbg” (weight = 0.090879), “program” (weight = 0.157883), and “makan” (weight = 0.186094), highlighting the dominant lexical components of the dataset.

3.5. Support Vector Machine Modeling

The labeled (positive/negative) dataset was split into training and testing subsets, with a larger portion allocated to training for model development and optimization. The testing subset was used to evaluate model performance on unseen data. This study employed two split ratios (80:20 and 70:30) to examine their impact on model outcomes. For the svm implementation, four kernel types—linear, radial basis function (rbf), sigmoid, and polynomial—were evaluated to determine optimal classification performance. A grid search strategy was applied to identify the best combination of hyperparameters. The classification results, selected based on the highest accuracy achieved for each kernel type, are presented below.

Table 2. Results of SVM Model Test

Model	Accuracy	Precision	Recall	F1-score	F1-gap
Svm kernel linier	88.4%	88.4%	88.4%	88.4%	5.82%
Svm kernel rbf	87.3%	87.3%	87.3%	87.3%	4.3%
Svm kernel sigmoid	79.9%	80.09%	79.9%	79.9%	1.81%
Svm kernel polynomial	87.3%	87.3%	87.3%	87.3%	1.6%

Table 2 shows that the linear kernel gave the highest accuracy (88.4%), followed by the RBF and polynomial kernels (87.3%) and the sigmoid kernel (79.9%). The difference between the linear and the RBF kernel (1.1 points) is about one standard error (0.9 points for 1,428 test tweets), so the two are practically equal, whereas the sigmoid kernel is clearly worse. The F1-gap is largest for the linear kernel (5.82%) and smallest for the polynomial (1.6%) and sigmoid (1.81%) kernels, so the most accurate kernel also fits the training data most closely (a training F1-score of about 94.2%). A linear boundary is expected to work well for sparse TF-IDF features with unigrams and bigrams, where the number of features is large relative to the number of tweets and the classes are close to linearly separable, and this agrees with Muttaqin and Kharisudin (2021), who also found the linear kernel best.

3.6. Extreme Learning Machine Modeling

This stage involved text classification using the extreme learning machine (elm) algorithm applied to tf-idf-transformed and normalized feature vectors. The experiment evaluated four hidden layer configurations, with 100, 500, 1000, and 1500 neurons, and four activation functions were tested—linear, sigmoid, rbf, sin—each exhibiting distinct input transformation characteristics (Rausanfitra et al., 2018). The selection of both the number of hidden neurons and the type of activation function significantly influenced the model's ability to capture complex nonlinear patterns in the data. Experimental results identified the three best-performing configurations based on the highest classification accuracy.

Table 3. Results of ELM Model Test

Model	Hidden Neuron	Accuracy	Precision	Recall	F1-score	F1-gap
ELM Linear	500	83.8%	83.8%	83.8%	83.8%	2.54%

ELM Linear	500	82.8%	82.8%	82.8%	82.8%	4.75%
ELM Sigmoid	500	82.5%	82.5%	82.5%	82.5%	2.73%

For ELM, the linear activation with 500 hidden neurons gave the highest accuracy (83.8%), followed by the sigmoid activation (82.5%). That a linear activation is best means that the random nonlinear transformation of the hidden layer adds little for these TF-IDF features, which parallels the result for the SVM kernels. The ELM has a smaller F1-gap (2.54%) than the best SVM, and its training F1-score is about 86.3%, so it fits the training data less tightly and generalizes more evenly but at a lower level of accuracy.

3.7. Classification Comparison

Table 2 and Table 3 show that the SVM with a linear kernel is the best model, with an accuracy of 88.4%, a precision of 88.4%, and an F1-score of 88.4%, against 83.8% for the best ELM. With 1,428 test tweets, the approximate 95% intervals are 86.7–90.1% for the SVM and 81.9–85.7% for the ELM, which do not overlap, so the 4.6-point advantage of the SVM is unlikely to be due to chance if both models were evaluated on the same test set. The F1-gap, which measures the difference between the performance on the training and the test data, is 5.82% for the SVM and 2.54% for the ELM. A gap of about six points is moderate and does not indicate severe overfitting, but it is larger than that of the ELM and of the polynomial and sigmoid SVM, so the SVM gains accuracy at the cost of a larger gap. The confusion matrix of the SVM is presented below.

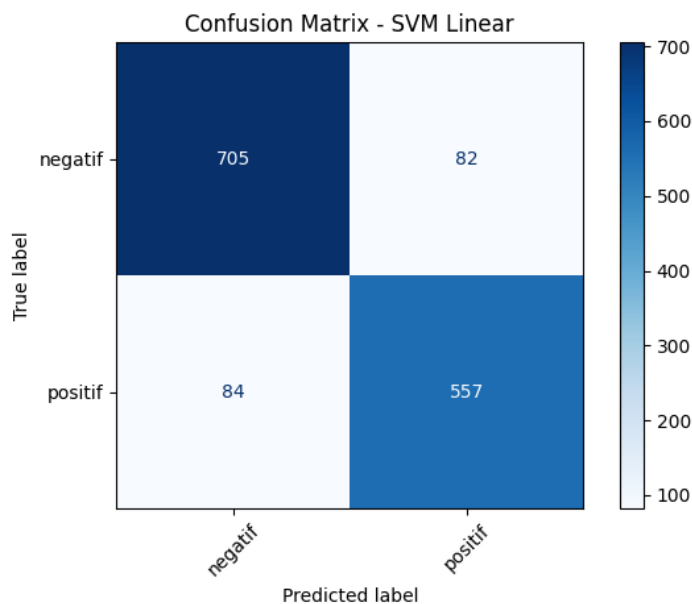


Figure 2. Confusion Matrix of SVM with Linear Kernel

The confusion matrix shows that of the 787 negative test tweets, 705 were classified correctly and 82 were classified as positive, and of the 641 positive test tweets, 557 were classified correctly and 84 were classified as negative. In total the model classified 1,262 of the 1,428 test tweets correctly (88.4%) and made 166 errors. The recall is 89.6% for the negative class and 86.9% for the positive class, and the precision is 89.4% and 87.2%, so the classifier is only slightly better on the majority (negative) class and is not biased toward one class. Its performance is sufficient for monitoring public perception of government programs.

3.8. Visualisation

The word cloud presents a visual representation of word frequency within the text, where the size of each word corresponds to its frequency of occurrence. Word clouds are utilized to identify dominant terms and general linguistic patterns in textual data. The larger a word appears, the more frequently it occurs in the corpus. Sentiment-based visualisations are categorised into two groups: positive sentiment and negative sentiment, as presented below.



Figure 3. Wordclouds of Negative Sentiments

The wordcloud visualisation and term frequency analysis identify dominant terms in negative sentiment regarding the free nutritious meal program. Keywords such as "racun" (poison), "tidak" (not), "mbg", and "gagal" (failed) consistently appear. This reflects public concerns regarding food safety, distrust of nutritional quality, and complaints about program implementation. The high frequency of the term "racun" indicates public perception associating the program with health risks. The analysis reveals declining public trust in the government program, particularly regarding food quality assurance and the competence of the organisers. These concerns underscore the need for prompt action by relevant authorities to provide clear explanations and effective solutions.



Figure 4. Wordclouds of positive Sentiments

The wordcloud visualisation highlights dominant terms in positive sentiment regarding the free nutritious meal program. Keywords such as "mbg", "program", "makan" (food), and "gizi" (nutrition) indicate public recognition of the program's benefits. The presence of "prabowo" and "presiden" (president) reflects appreciation for the policy initiators, while terms like "sehat" (healthy), "dukung" (support), and "lanjut" (continue) demonstrate positive perceptions of health impacts and program sustainability. The inclusion of "papua" and "desa" (village) further suggests the program's reach across diverse regions, including remote areas, reinforcing its positive perception as a tangible government effort to improve public welfare. This word cloud analysis illustrates public acceptance and support for the free nutritious meal program implementation.

3.9. Discussion

The results show that, for TF-IDF features of MBG tweets, the choice between a linear model and a nonlinear one matters less than the choice between the two learning methods: both the SVM and the ELM are best with a linear form (linear kernel, linear activation), and the SVM is more accurate. The maximum-margin training of the SVM with grid-searched hyperparameters appears to use the high-dimensional features more effectively than the ELM, whose hidden weights are random and are not tuned. The ELM is expected to be faster, which is the main reason to choose it, but training time was not measured in this study, so its efficiency advantage remains an expectation.

The SVM accuracy of 88.4% is close to the 87.98% of the linear kernel reported by Muttaqin and Kharisudin (2021) for GOJEK reviews, but lower than the 94% of Purwanti and Sugiyono (2024) for tweets on the free lunch program; the difference may come from the data, the labeling, and the preprocessing, which differ among the studies. The ELM accuracy of 83.8% is higher than the 72% reported by Riskiyah et al. (2024) for GOJEK. As in these studies, the SVM is at least as accurate as the ELM, and the results of the present study add that this ordering holds on the same data with the same preprocessing, and that the ELM generalizes more tightly. Because the data sets differ, the comparison with earlier studies is limited to the direction of the results.

The word clouds show what drives the sentiment. Negative tweets, which form the majority (about 55% of the test tweets), center on food safety and distrust (“racun”, “gagal”) and on the competence of the organizers, whereas positive tweets refer to nutrition and health (“gizi”, “sehat”), support and continuation (“dukung”, “lanjut”), and the reach of the program to regions such as Papua and villages. The novelty of the findings is therefore that the public discussion is split between a safety-and-quality concern and recognition of the program’s purpose, and that a simple linear SVM can separate them reliably.

These findings imply the following recommendations. For the implementers and relevant authorities: (1) strengthen food safety and quality assurance and give clear public explanations of incidents, because food safety is the main theme of negative sentiment; (2) communicate the nutritional benefits and the reach of the program, because the positive tweets show that these are valued; and (3) monitor public sentiment routinely with a linear SVM classifier and check its labels periodically against manual judgement. For researchers: use the linear SVM as the baseline for similar data, report training time when the speed of the ELM is a reason for its use, and repeat the ELM with several random seeds.

Several limitations should be noted. The labels come from a lexicon and not from human annotators, so the accuracy measures agreement with the lexicon and may overstate the agreement with human judgement, because the classifiers learn the same words that the lexicon uses. The classification is binary, the data cover six days from one platform and may include bot or buzzer accounts, and the hyperparameters, the ELM random initialization, and the class distribution need to be documented as indicated above.

4. Conclusion

This study compared SVM and ELM for classifying the sentiment of tweets about the Free Nutritious Meal Program, and its findings answer the three research questions. First, the best SVM uses the linear kernel (accuracy 88.4%, F1-gap 5.82%), and the best ELM uses the linear activation with 500 hidden neurons (accuracy 83.8%, F1-gap 2.54%). Second, the SVM is the better method in accuracy, precision, recall, and F1-score, with an advantage of 4.6 points, whereas the ELM has the smaller F1-gap and therefore generalizes more evenly; both methods favor a linear form for TF-IDF text features.

Third, negative sentiment predominates (about 55% of the labeled tweets) and centers on food safety, distrust, and the performance of the organizers, while the positive tweets refer to nutrition, health, support, and the reach of the program. The authorities are advised to strengthen food safety assurance and public communication and to monitor sentiment routinely with a linear SVM classifier, with periodic manual validation.

This study has several limitations that present opportunities for further development. First, filtering out potential buzzer or bot accounts is necessary to ensure that the collected tweets genuinely reflect public opinion. Second, expanding the sentiment classification framework from binary to three classes (positive, negative, and neutral) could provide more nuanced and accurate sentiment representation. Third, future experiments could explore alternative classification approaches such as Extreme Gradient Boosting (XGBoost), Deep Extreme Learning Machines (DELm), or hybrid models to potentially enhance classification accuracy. Fourth, diversifying data sources by incorporating additional social media platforms like Instagram and Facebook, as well as online discussion forums, would enable more comprehensive public sentiment analysis by capturing a wider spectrum of public discourse.

References

- Alhaq, Z., Mustopa, A., Mulyatun, S., & Santoso, J. D. (2021). Penerapan metode support vector machine untuk analisis sentimen pengguna Twitter. *Journal of Information System Management (JOISM)*, 3(2), 44–49.
- Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan algoritma support vector machine (SVM) dengan TF-IDF N-gram untuk text classification. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 6(2), 129.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Fitriyah, N., Warsito, B., & Di Asih, I. M. (2020). Analisis sentimen Gojek pada media sosial Twitter dengan klasifikasi support vector machine (SVM). *Jurnal Gaussian*, 9(3), 376–390.
- Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1–3), 489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>
- Husada, H. C., & Paramita, A. S. (2021). Analisis sentimen pada maskapai penerbangan di platform Twitter menggunakan algoritma support vector machine (SVM). *Teknika*, 10(1), 18–26.
- Indarbansyah, P. P. E., & Rochmawati, N. (2021). Penerapan N-gram menggunakan algoritma random forest dan naïve Bayes classifier pada analisis sentimen kebijakan PPKM 2021. *Journal of Informatics and Computer Science (JINACS)*, 2(4), 235–244.

- Iriananda, S. W., Budiawan, R. W., Rahman, A. Y., & Istiadi, I. (2024). Optimasi klasifikasi sentimen komentar pengguna game bergerak menggunakan SVM, grid search dan kombinasi N-gram. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(4), 743–752.
- Isnain, A. R., Sakti, A. I., Alita, D., & Marga, N. S. (2021). Sentimen analisis publik terhadap kebijakan lockdown pemerintah Jakarta menggunakan algoritma SVM. *Jurnal Data Mining dan Sistem Informasi*, 2(1), 31.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Maggioni, F., & Spinelli, A. (2024). *A novel robust optimization model for nonlinear support vector machine*. arXiv. <https://doi.org/10.48550/arXiv.2306.06223>
- Muttaqin, M. N., & Kharisudin, I. (2021). Analisis sentimen aplikasi Gojek menggunakan support vector machine dan K nearest neighbor. *UNNES Journal of Mathematics*, 10(2), 22–27.
- Ozyildirim, B. M., & Coban, O. (2018). An empirical study of the extreme learning machine for Twitter sentiment analysis. *International Journal of Intelligent Systems and Applications in Engineering*, 3(6), 178–184.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)* (pp. 79–86). Association for Computational Linguistics. <https://doi.org/10.3115/1118693.1118704>
- Purwanti, Z., & Sugiyono. (2024). Pemodelan text mining untuk analisis sentimen terhadap program makan siang gratis di media sosial X menggunakan algoritma support vector machine (SVM). *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 5(3), 3065–3079.
- Rausanfitra, A., Adikara, P. P., & Adinugroho, S. (2018). Analisis sentimen Twitter menggunakan ensemble feature dan metode extreme learning machine (ELM) (studi kasus: Samsung Indonesia). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(12), 6409–6417.
- Riskiyah, A., Fahrudin, T. M., & Hindrayani, K. M. (2024). Analisis sentimen kepuasan pelayanan transportasi online Gojek menggunakan algoritma extreme learning machine. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 5(2), 1273–1285.
- Wang, F., Zhang, Y., Rao, Q., Li, K., & Zhang, H. (2017). Exploring mutual information-based sentimental analysis with kernel-based extreme learning machine for stock prediction. *Soft Computing*, 21(12), 3193–3205.
- Wang, J., Lu, S., Wang, S.-H., & Zhang, Y.-D. (2022). A review on extreme learning machine. *Multimedia Tools and Applications*, 81(29), 41611–41660. <https://doi.org/10.1007/s11042-021-11007-7>