Journal of Applied Business and Technology

Comparison of Support Vector Machine Kernel Performance in Sentiment Analysis of The Free Lunch Program

Emanuel Zega [a*], Erlin Erlin [a]

[a] Faculty of Computer Science, Institut Bisnis dan Teknologi Pelita Indonesia, Indonesia

Article History
Received
31 October 2025
Received in revised form
31 January 2026
Accepted
1 May 2026
Published Online
31 May 2026

*Corresponding author
emanuel.zega@student.pelitaindonesia.ac.id

Abstract
The Free Lunch Program is a government policy that has generated various opinions on social media, ranging from support to criticism regarding its implementation. This study aims to analyze public sentiment before and after the implementation of the Free Lunch Program using the Support Vector Machine (SVM) algorithm and to compare the performance of several kernels, namely Linear, Polynomial, RBF, and Sigmoid. The data were collected from the X (Twitter) platform and classified into three sentiment categories: positive, negative, and neutral. The experimental results show that the RBF kernel achieved the best performance, with an accuracy of 97.04% on pre-implementation data and 96.32% on post-implementation data. Moreover, the sentiment analysis results indicate that the proportion of negative sentiment increased after the program's implementation, rising from 54.5% to 57.7%. These findings suggest that the RBF kernel is the most effective in classifying public opinions regarding the Free Lunch Program policy with a high level of accuracy, while also illustrating that public perception tends to be more critical after the program was implemented.

Keywords: Sentiment Analysis; Support Vector Machine; Free Lunch; Social Media
DOI: https://doi.org/10.35145/gtv1sr05
SDGs: Zero Hunger (2); Good Health and Well-being (3); Quality Education (4); Industry, Innovation and Infrastructure (9); Peace, Justice and Strong Institutions (16)

1.0 INTRODUCTION

The Free Lunch Program is a government policy that provides nutritious meals free of charge to students in Indonesia. This program aims to offer free lunches to schoolchildren with the hope of improving nutritional quality, reducing stunting rates, and supporting the education of children from underprivileged families. (Eliza et al., 2024).

Some people view this program as an effort to support education, as well-nourished children are expected to be more focused in their studies. In addition, questions have arisen regarding the distribution mechanism and transparency of its implementation, as well as whether the program will truly reach its intended targets. (Ernawati, 2024). This debate between supporters and critics is widely discussed on social media, especially on platforms such as X. The public actively expresses their opinions and responses, allowing for the exploration and analysis of sentiment patterns. (Erlin et al., 2021).

Sentiment analysis is a method used to identify and classify opinions, emotions, or attitudes contained in text, such as tweets, comments, or reviews. (Ipmawati et al., 2017). One of the most effective algorithms for sentiment analysis is the Support Vector Machine (SVM) (Ananda & Pristyanto, 2021). Previous research has also discussed the implementation of the SVM algorithm in various domains, highlighting its effectiveness in text classification, opinion mining, and sentiment analysis across different social media platforms (Junaedi et al., 2024), (Eldo et al., 2024), (Ahmad, 2023). Several previous studies (Armadianti et al., 2024), (Anggraini & Alita, 2024), (Rabbani et al., 2023) have shown that the Support Vector Machine (SVM) algorithm produces excellent results in sentiment analysis tasks, demonstrating high accuracy and reliable performance across various datasets.

The objective of this study is to classify public opinions and sentiments toward the Free Lunch Program into positive, negative, and neutral categories, and to compare the results of different Support Vector Machine (SVM) kernels. This research also compares the sentiment classification results of the Free Lunch Program before and after its implementation.

2.0 LITERATURE REVIEW

The Free Lunch Program

The Free Lunch Program is one of the government's initiatives that will begin implementation in January 2025. This program aims to improve the welfare of Indonesian children, particularly students and Islamic boarding school students, by providing nutritious meals free of charge. The government has allocated approximately IDR 71 trillion in the 2025 Draft State Budget (RAPBN 2025) to support the implementation of this program. Each child benefiting from the program will receive an allocation of around IDR 15,000 per day for their lunch (Ardelia Maharani et al., 2024).

Support Vector Machine

Support Vector Machine is a learning system that utilizes a hypothesis space consisting of linear functions within a high-dimensional feature space and implements a learning bias derived from statistical learning theory, trained using a learning algorithm (Isnain et al., 2021). The theory underlying SVM has been developed since the 1960s but was first introduced by Vapnik, Boser, and Guyon in 1992. According to (Lodhi et al., 2002) , the following are some commonly used kernel functions:

a. Kernel Linear :

$$K\ (xi, x) =\ xi^T x \quad (1)$$

b. Kernel *Polynomial* :

$$K(xi.x) = (y.xi^T x + r)\ p, y > 0 \quad (2)$$

c. Kernel *Radial Basis Function* (RBF) :

$$K\ (xi, x) = exp(-y\ |\ xi - x|\hat{}2\ ), y > 0 \quad (3)$$

d. Kernel *Sigmoid* :

$$(xi, x) =\ tanh(yxi\hat{}T + r) \quad (4)$$

Confusion Matrix

According to (Susanto, 2023), a Confusion Matrix is a table used to evaluate the performance of a classification model by comparing the model's predictions with the actual values. TP and TN describe cases where the classification process makes correct predictions, while FP and FN explain cases where the classification process makes incorrect predictions. he following is the formula of the confusion matrix used to measure the accuracy level, as shown in the equation.

$$Accuracy = \frac{TP+FN}{TP} \quad (5)$$

$$Precission = \frac{TP}{TP+FP} \quad (6)$$

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

$$F1\text{-}Score = \frac{2\ x\ Recall+Precission}{Recall+Precission} \quad (8)$$

Keterangan :

TP : *True Positive*
TN : *True Negative*
FP : *False Negative*
FN : *False Positive*

3.0 METHODOLOGY

This study employs the Support Vector Machine (SVM) method to perform sentiment analysis on the Free Lunch Program based on data collected from the X application. The process begins with data collection using a web crawler, followed by data cleaning, word weighting, model building, testing, and evaluation of the SVM model, and concludes with the final stage of deploying the developed model.

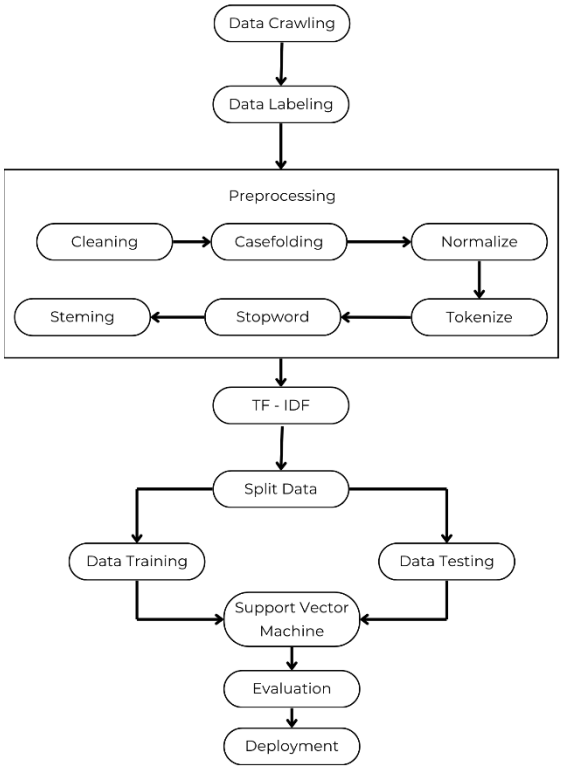

Figure 1. Research Stages

1. Data Crawling

In this study, data related to the Free Lunch Program were collected by crawling data from the X application using Google Colab. The obtained data include various public opinions based on keywords such as "free lunch program" and "#freelunchprogram." The analyzed data were taken from two periods: before the program implementation (December 12, 2023 – January 31, 2024, during the first presidential debate, and November 11, 2024 – December 31, 2024, when discussions about the program increased) and after the program implementation (January 6, 2025 – March 5, 2025).

2. Data Labeling

The collected tweets were manually classified into three sentiment categories: positive, negative, and neutral. This process was carried out manually rather than using an automated system to ensure more accurate and context-appropriate results (Homepage et al., 2024), (Kurniawan et al., 2023). In this labeling process, texts expressing support, appreciation, or praise for the Free Lunch Program were labeled as positive. Conversely, texts containing criticism, rejection, or concern about the program were labeled as negative. Meanwhile, texts that were purely informative, did not express a specific opinion, or merely questioned without taking a stance were labeled as neutral.

3. Preprocessing

At this stage, the data undergo several preprocessing steps, including cleaning, case folding, normalization, tokenization, stopword removal, and stemming (Khairunnisa et al., 2021). This process aims to improve data quality by removing irrelevant elements such as duplicate and ambiguous words, as well as converting text into its root form (Gunawan, 2016). Through these steps, the data become cleaner and ready for further analysis in the sentiment analysis process.

4. TF-IDF Feature Extraction

TF-IDF (Term Frequency – Inverse Document Frequency) feature extraction is a text processing technique that transforms text into a numerical representation based on how frequently a word appears in a document (Term Frequency) and how unique that word is across the entire set of documents (Inverse Document Frequency) (Asro'i Arief & Februariyanti Herny, 2022). The TF-IDF value can be determined using the following equation.

$$W_{if} = tf_{if} + f(x) = \{log \frac{d}{dfi}\} \quad (9)$$

Keterangan:

D = number of documents/tweets

df_i = occurrence of a word from D

W_{if} = weight of word t_j in document/tweet d_i

tf_{if} = number of occurrences of word t_j in document/tweet d_i

5. SVM Modeling

During the model-building stage, the k-fold cross-validation technique was used to achieve good model accuracy, and several kernels—Linear, Polynomial, RBF, and Sigmoid—were tested to obtain optimal performance (Feta, 2022). The linear kernel is used when the data can be separated by a straight line or hyperplane, whereas nonlinear kernels are applied when the data can only be separated by a curved line or a surface in a higher-dimensional space (Pratiwi & Setyawan, 2021).

6. Evaluation

At the evaluation stage, the final results of the model will be assessed to determine the extent to which the research objectives have been achieved (Nurhidayat & Dewi, 2023). This process aims to evaluate the model's performance in performing classification based on several evaluation metrics, such as accuracy, precision, recall, and F1-score, which are calculated based on the confusion matrix.

Deplyoment

The model deployment stage is the process of placing a trained and tested machine learning model into a production environment so that it can be used in real-world applications by users or other systems (Paramudita et al., 2024).

This section contains a complete and detailed description of the steps undertaken in conducting of research. In addition, the research step also needs to be shown in the form of flowchart of research or framework step and detailed including reflected algorithm, rule, modeling, design and others related to system design aspect.

All tables should be numbered with Calibri Light numerals. Every table should have a caption. Headings should be placed above tables and justified. Only horizontal lines should be used within a table, to distinguish the column headings from the body of the table, and immediately above and below the table. Tables must be embedded into the text and not supplied separately. Below is an example which the authors may find useful.

4.0 RESULTS AND DISCUSSION

Dataset

After performing data crawling, a total of 1,502 tweet data were obtained before the program implementation and 1,495 tweet data after the program implementation. Both datasets were stored in separate files and will be processed to compare the sentiment classification results before and after the implementation of the free lunch program.

Data Labeling

After the data labeling process, sentiment distribution can be analyzed. The sentiment distribution in the data before the free lunch program implementation is as follows: Negative 808 (53.8%), Positive 561 (37.4%), and Neutral 133 (8.9%). Meanwhile, the sentiment distribution in the data after the implementation of the free lunch program is: Negative 846 (56.6%), Positive 478 (32.0%), and Neutral 171 (11.4%).

Data Preprocessing

Before model construction, the data must go through a preprocessing stage. This stage aims to clean and normalize the data to make it easier to process by the model or algorithm. Tweet data must be cleaned of symbols, username tags, emojis, and irrelevant words. The preprocessing process consists of five stages: Cleaning, Case Folding, Normalization, Tokenizing, Filtering, and Stemming.

Table 1. Data Preprocessing

Tweet	Data Cleaning	Case Folding	Normalize	Tokenizing	Stopword	Steming
@WaktunyaID Maju Bagi2 bansos aja nggak merata dan sesuai datanya. Apalagi mau bagi2 susu dan makan siang	Bagi bansos aja nggak merata dan sesuai datanya Apalagi mau bagi susu dan makan siang gratis	bagi bansos aja nggak merata dan sesuai datanya apalagi mau bagi susu dan makan siang gratis	bagi bansos saja tidak merata dan sesuai datanya apalagi mau bagi susu dan makan siang gratis	['bagi', 'bansos', 'saja', 'tidak', 'merata', 'dan', 'sesuai', 'datanya', 'apalagi', 'mau', 'bagi', 'susu', 'dan', 'makan',	['bansos', 'merata', 'sesuai', 'datanya', 'susu', 'makan', 'siang', 'gratis']	bansos rata sesuai data susu makan siang gratis

gratis. #makansianggratis	'siang', 'gratis']

Resampling

One of the resampling methods is upsampling, which involves increasing the number of samples in the minority class by randomly duplicating data so that the class distribution becomes more balanced. After applying the resampling technique, the data becomes balanced, with 2,361 data samples before implementation and 2,313 data samples after implementation.

SVM Classification

1. Training and Testing Data Split

   After performing the data preprocessing and weighting process using Tf-Idf, the next step is classification. This study uses an 80:20 ratio for training and testing data.

Table 2. Data Split

Data Type	Before Implementation	After Implementation
Training Data	1889	1850
Testing Data	473	463
Total	2361	2313

2. K-Fold Cross Validation

   Based on the evaluation results using the 10-Fold Cross Validation method, the model used in the sentiment analysis of the free lunch program demonstrated consistent and high performance, with an average accuracy of 96.82% on the data before implementation and an average accuracy of 95.24% on the data after implementation.

3. SVM Kernel

   a. SVM Kernel on data before the implementation of the free lunch program (still in the planning stage)

Table 3. SVM Kernel Results Before Implementation

Kernel	Linear	Polynomial	RBF	Sigmoid
Accuracy	95.13%	90,69%	95,77%	93,86%
Recall	95.13%	90,69%	95,77%	93,86%
Precision	95.18%	92,21%	95,78%	93,90%
F1-Score	95.13%	90,52%	95,77%	93,86%

   Based on the testing results of the four types of kernels in the Support Vector Machine (SVM) algorithm, the RBF (Radial Basis Function) kernel demonstrated the best performance, achieving the highest accuracy of 95.77%.

   b. SVM Kernel on Data After the Implementation of the Free Lunch Program

Table 4. SVM Kernel Results After Implementation

Kernel	Linear	Polynomial	RBF	Sigmoid
Accuracy	95.46%	91,57%	96,32%	92,44%
Recall	95.46%	91,57%	96,32%	92,44%
Precision	95.47%	92,67%	96,32%	92,54%
F1-Score	95.44%	91,38%	96,32%	92,37%

   Based on the evaluation results of the four types of kernels in the Support Vector Machine (SVM) algorithm, the RBF (Radial Basis Function) kernel once again demonstrated the best performance, achieving the highest scores across all evaluation metrics, with an accuracy of 96.32%.

4. Testing on Data Split Variations

Based on the testing results of the four types of kernels in the Support Vector Machine (SVM) algorithm, the RBF (Radial Basis Function) kernel showed the best performance with the highest accuracy of 95.77%. Therefore, the RBF kernel was selected for use in testing various data split ratios.

a. Results of Data Split Variations on Data Before the Implementation of the Free Lunch Program (Planning Stage)

Table 5. Results of Data Split Variations Before Implementation

Data variaton	Data 70:30	Data 80:20	Data 90:10
Accuracy	88,29%	95,77%	97.04%
Recall	88,29%	95,77%	97.04%
Precision	90,70%	95,78%	97.03%
F1-Score	88,10%	95,77%	97.03%

The test results using the RBF kernel indicate that the larger the portion of training data, the better the SVM model's performance. At a 70:30 ratio, the accuracy was 88.29%, increasing to 95.77% at 80:20, and reaching 97.04% at 90:10. In conclusion, the 90:10 ratio provides the best results because the model has more training data to recognize sentiment patterns.

b. SVM Kernel on Data After the Implementation of the Free Lunch Program

Table 6. Results of Data Split Variations After Implementation

Data Variation	Data 70:30	Data 80:20	Data 90:10
Accuracy	96,10%	96,32%	94.82%
Recall	96,10%	96,32%	94,82%
Precision	96,10%	96,32%	94,85%
F1-Score	96,10%	96,32%	94,83%

Based on the test results, it can be seen that the data split variation affects the model's performance. At a 70:30 ratio, the accuracy, recall, precision, and F1-score values were all 96.10%. At an 80:20 ratio, all metrics slightly increased to 96.32%, indicating the best performance. Meanwhile, at a 90:10 ratio, the model's performance decreased, with an accuracy of 94.82%, recall of 94.82%, precision of 94.85%, and F1-score of 94.83%. This shows that the 80:20 split ratio is more optimal compared to the other variations.

Model Evaluation

1. Classification Report on Data Before the Implementation of the Free Lunch Program (Planning Stage)

```
                precision    recall  f1-score   support

      Negative       0.97      0.97      0.97        87
       Neutral       0.99      1.00      0.99        75
      Positive       0.96      0.95      0.95        75

      accuracy                           0.97       237
     macro avg       0.97      0.97      0.97       237
  weighted avg       0.97      0.97      0.97       237

[[84  0  3]
 [ 0 75  0]
 [ 3  1 71]]
Accuracy Score:  0.9704641350210971
```

Figure 2. Classification Report Before the Implementation of the Program

Based on the classification report Figure 2, the classification model demonstrates very good performance. The majority of the data were correctly classified, as indicated by the high values along the main diagonal:

160 Negative data, 151 Neutral data, and 142 Positive data were accurately classified. Misclassifications were relatively minor, with only a few cases of incorrect predictions between classes.

2. Classification Report on Data After the Implementation of the Free Lunch Program

```
                precision    recall  f1-score   support

      Negative       0.95      0.95      0.95       170
       Neutral       0.98      0.99      0.98       162
      Positive       0.95      0.95      0.95       131

      accuracy                           0.96       463
     macro avg       0.96      0.96      0.96       463
  weighted avg       0.96      0.96      0.96       463

[[161   3   6]
 [  2 160   0]
 [  6   0 125]]
Accuracy Score:  0.9632829373650108
```

Figure 3. Classification Report After the Implementation of the Program

Based on the displayed classification report, the model shows excellent performance. This can be seen from the number of correct predictions in each class: 161 out of 170 Negative class data were correctly predicted, 160 out of 162 Neutral class data were accurately classified, and 125 out of 131 Positive class data were also predicted correctly.

Deployment Model

(a)

```
def classify(tweet):
    pred = rbf.predict(Tfidf_vect.transform([tweet]))[0]  # Ambil nilai
    prediksi pertama

    if pred == 0:
        return "Negatif"
    elif pred == 1:
        return "Netral"
    elif pred == 2:
        return "Positif"
    else:
        return "Kategori Tidak Dikenal"  # Untuk menangani nilai yang tidak
        terduga
```

```
classify('Sampai akhirnya gw setuju makan siang gratis nya om Wowo....')
```

```
'Positif'
```

```
classify('Segini Anggaran Makan Siang Gratis Tuk Anak Sekolah di Indonesia
Program Pemerintah Era Prabowo')
```

```
'Netral'
```

```
classify('Duh kalau dikelola sama Pemda siap2 aja deh anggaran 400T utk program
makan siang gratis akan dijadikan para orang2 Pemda seperti PNS Polisi Dinsos
Tentara dll sebagai sarana utk korupsi')
```

```
'Negatif'
```

(b)

```
def classify(tweet):
    pred = rbf.predict(Tfidf_vect.transform([tweet]))[0]  # Ambil nilai
    prediksi pertama

    if pred == 0:
        return "Negatif"
    elif pred == 1:
        return "Netral"
    elif pred == 2:
        return "Positif"
    else:
        return "Kategori Tidak Dikenal"  # Untuk menangani nilai yang tidak
        terduga
```

```
classify('respek efisien banget makan siang gratis mantap gizi')
```

```
'Positif'
```

```
classify('tebak tebak menu makan siang gratis')
```

```
'Netral'
```

```
classify('tolong program makan siang gratis hapus apaapa pangkas program')
```

```
'Negatif'
```

Figure 4. (a) Before Implementation (b) After Implementation

Sentiment Classification Results

1. Sentiment Classification Distribution Results on Data Before the Implementation of the Free Lunch Program (Planning Stage)

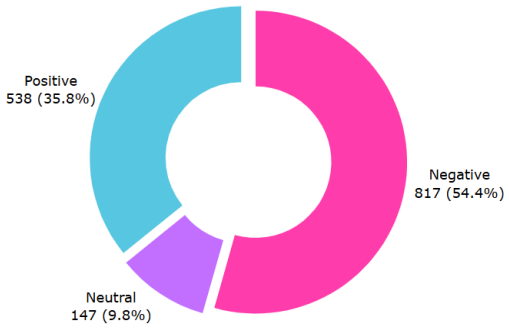

Figure 5. Sentiment Classification Results of Data Before Implementation

Based on the sentiment analysis before the implementation of the free lunch program, the majority of public responses tended to be negative, with a percentage of 54.4%, followed by positive sentiment at 35.8%, and neutral sentiment at 9.8%. These results indicate that before the program was implemented, many people responded to it with skepticism or criticism.

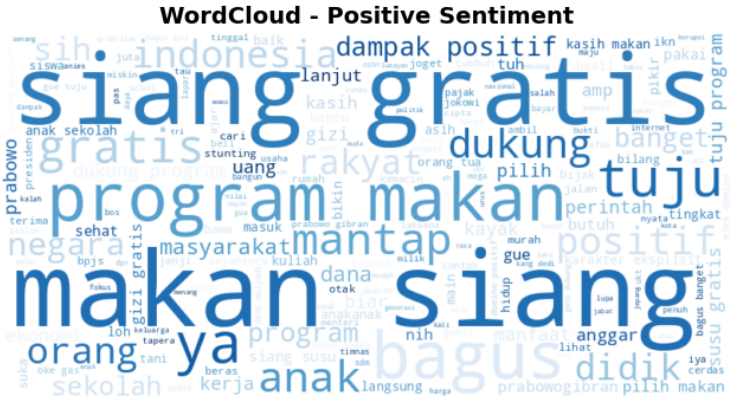

Figure 6. Positive Sentiment Before the Implementation of the Program

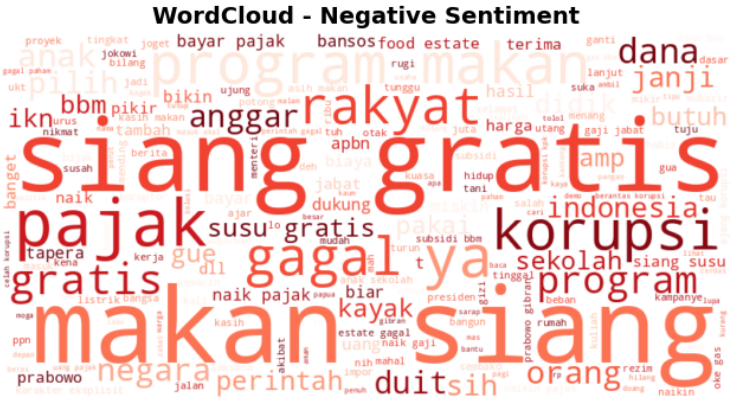

Figure 7. Negative Sentiment Before the Implementation of the Program

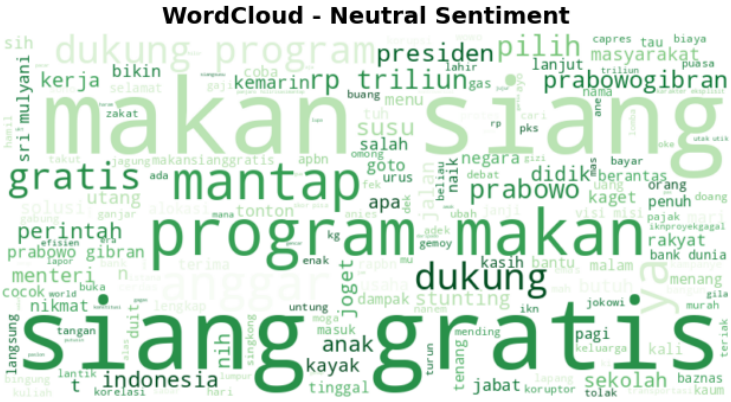

Figure 8. Neutral Sentiment Before the Implementation of the Program

2. Sentiment Classification Distribution Results on Data After the Implementation of the Free Lunch Program

Based on Figure 9, after the implementation of the free lunch program, negative sentiment still dominated public responses with a proportion of 57.7% (862 data). Meanwhile, positive sentiment was recorded at 467 data (31.2%) and neutral sentiment at 11.1% (166 data). This indicates that even though the program has been implemented, many people still responded critically, possibly concerning its implementation effectiveness, budget, or impact. However, the increase in positive sentiment also reflects that some members of the public have begun to recognize the benefits of the program.

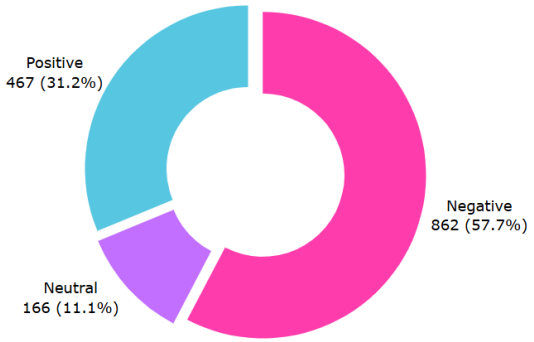

Figure 9. Sentiment Classification Results of Data After Implementation

Figure 10. Positive Sentiment After the Implementation of the Program

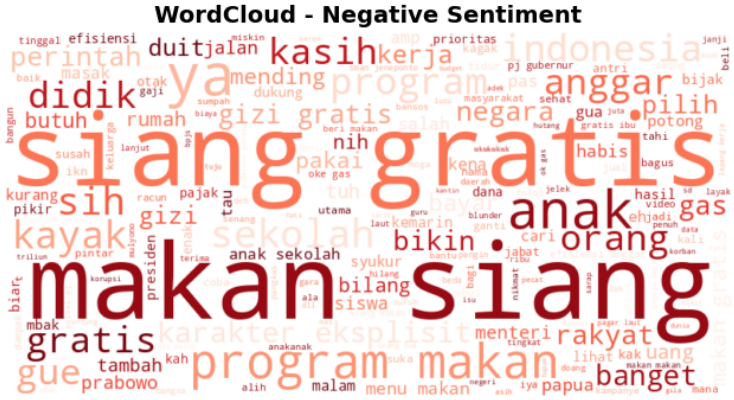

Figure 11. Negative Sentiment After the Implementation of the Program

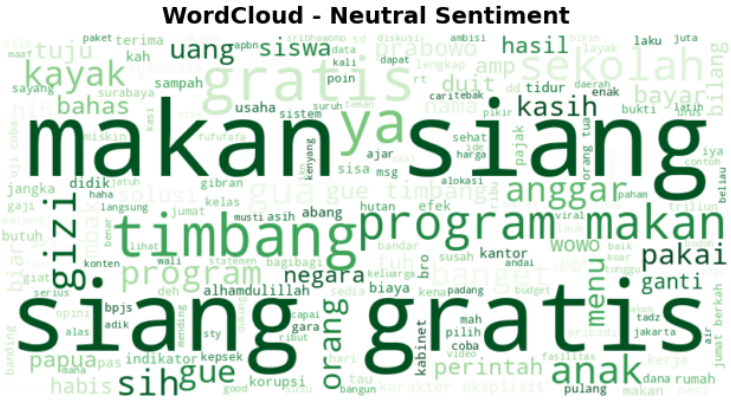

Figure 12. Neutral Sentiment After the Implementation of the Program

5.0 CONCLUSION

Based on the sentiment analysis of the free lunch program using the Support Vector Machine algorithm, several conclusions can be drawn as follows:

1. The best model for sentiment analysis was obtained using the SVM algorithm with the RBF kernel, achieving high accuracy of 95.77% on data before implementation and 96.32% on data after implementation.
2. In the data before the implementation of the free lunch program, negative sentiment dominated at 54.5%, followed by positive sentiment at 35.8%, and neutral sentiment at 9.8%. In the data after implementation, negative sentiment still dominated at 57.7%, followed by positive sentiment at 31.2%, and neutral sentiment at 11.1%.
3. Words such as “makan” (eat), “siang” (lunch), “program,” and “gratis” (free) consistently appeared across all sentiment categories. Negative sentiment included words like “korupsi” (corruption), “pajak” (tax), “janji” (promise), and “duit” (money), reflecting public concern about transparency and funding. In contrast, positive sentiment featured words like “bagus” (good), “setuju” (agree), “anak” (child), and “gizi” (nutrition), indicating hope and support, while neutral sentiment was dominated by descriptive words without specific opinions.
4. The comparison of sentiments before and after implementation shows that the free lunch program has not yet succeeded in improving public perception overall, as there was even an increase in negative sentiment.

References

Ahmad, K. (2023). Tahapan Pra-pemrosesan Teks (Text Preprocessing). *Journal of Mandalika Literature*, *4*(4), 323–332.

Ananda, F. D., & Pristyanto, Y. (2021). Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, *20*(2), 407–416. https://doi.org/10.30812/matrik.v20i2.1130

Anggraini, J., & Alita, D. (2024). Implementasi Metode SVM Pada Sentimen Analisis Terhadap Pemilihan Presiden (Pilpres) 2024 Di Twitter. *Jurnal Informatika: Jurnal Pengembangan IT*, *9*(2), 102–111. https://doi.org/10.30591/jpit.v9i2.6560

Armadianti, W., Syeh, A., Lastono, B., Putra, F. R., Kamil, I., Ghozi, A., & Rakhmawati, N. A. (2024). Analisis Sentimen Netizen Terhadap Personal Branding Elon Musk Pada Platform X Dengan Pendekatan Analisis Support Vector Machine. *Fountain of Informatics Journal*, *9*(1).

Asro’i Arief, & Februariyanti Herny. (2022). Analisis Sentimen Pengguna Twitter terhadap Perpanjangan PPKM Menggunakan Metode K-Nearest Neighbor. *Jurnal Khatulistiwa Informatika*, *10*(1), 17–24.

Eldo, H., Ayuliana, A., Suryadi, D., Chrisnawati, G., & Judijanto, L. (2024). Penggunaan Algoritma Support Vector Machine (SVM) Untuk Deteksi Penipuan pada Transaksi Online. *Jurnal Minfo Polgan*, *13*(2), 1627–1632. https://doi.org/10.33395/jmp.v13i2.14186

Eliza, F., Gistituati, N., Rusdinal, R., & Fadli, R. (2024). Analisis SWOT Kebijakan Makan Siang Gratis di Sekolah Menengah Kejuruan. *Juwara Jurnal Wawasan Dan Aksara*, *4*(1), 121–129. https://doi.org/10.58740/juwara.v4i1.91

Erlin, Josef Sianturi, Alyauma Hajjah, & Agustin. (2021). Analisis Sentimen Prosesor AMD Ryzen menggunakan Metode Support Vector Machine. *SATIN - Sains Dan Teknologi Informasi*, *7*(2), 129–141.

https://doi.org/10.33372/stn.v7i2.804

Ernawati. (2024). Jurnal Pendidikan Islam Anak Usia Dini Al-Amin. *Jurnal Pendidikan Islam Anak Usia Dini*, *2*(1), 30–37.

Feta, N. R. (2022). *Komparasi Fungsi Kernel Metode Support Vector Machine Untuk Comparison of the Kernel Function of Support Vector Machine Method for Modeling Classification of*. *July 2019*.

Gunawan, D. (2016). khazanah informatika Jurnal Ilmu Komputer dan Informatika Evaluasi Performa Pemecahan Database dengan Metode Klasifikasi pada Data Preprocessing Data Mining. *Khazanah Informatika: Jurnal Ilmiah Komputer Dan Informatika*, *1*(1), 1–4.

Homepage, J., Ningsih, W., Alfianda, B., & Wulandari, D. (2024). *MALCOM: Indonesian Journal of Machine Learning and Computer Science Comparison of Naive Bayes and SVM Algorithms in Twitter Sentiment Analysis on Electric Car Use in Indonesia Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada*. *4*(2), 556–562.

Ipmawati, J., Kusrini, & Taufiq Luthfi, E. (2017). 1444-1653-1-Sm. *Indonesian Journal on Networking and Security*, *6*(1), 28–36.

Junaedi, Hendra Gunawan, A., Kuswanto, V., & Jonathan. (2024). Tinjauan Support Vector Machine dalam Text-Mining untuk Analisis Sentimen di Sektor Pariwisata. *Bit-Tech*, *7*(2), 323–330. https://doi.org/10.32877/bt.v7i2.1810

Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *Jurnal Media Informatika Budidarma*, *5*(2), 406. https://doi.org/10.30865/mib.v5i2.2835

Kurniawan, I., Lia Hananto, A., Shofia Hilabi, S., Hananto, A., Priyatna, B., & Yuniar Rahman, A. (2023). Perbandingan Algoritma Naive Bayes Dan SVM Dalam Sentimen Analisis Marketplace Pada Twitter. *Jurnal Teknik Informatika Dan Sistem Informasi*, *10*(1), 731–740. http://jurnal.mdp.ac.id

Lodhi, H., Saunders, C., Shawe-Taylor, J., Cristianini, N., & Watkins, C. (2002). Text Classification using String Kernels. *Journal of Machine Learning Research*, *2*(3), 419–444. https://doi.org/10.1162/153244302760200687

Nurhidayat, R., & Dewi, K. E. (2023). KOMPUTA : Jurnal Ilmiah Komputer dan Informatika PENERAPAN ALGORITMA K-NEAREST NEIGHBOR DAN FITUR EKSTRAKSI N-GRAM DALAM ANALISIS SENTIMEN BERBASIS ASPEK. *Komputa : Jurnal Ilmiah Komputer Dan Informatika*, *12*(1), 91–100. https://www.kaggle.com/datasets/hafidahmusthaanah/skincare-review?select=00.+Review.csv.

Paramudita, F., Zulfa, M. I., & Taryana, A. (2024). Implementasi Devops Pada Pengembangan Aplikasi Android Pendeteksi Kualitas Beras Berbasis Machine Learning. *Transmisi: Jurnal Ilmiah Teknik Elektro*, *26*(3), 105–113. https://doi.org/10.14710/transmisi.26.3.105-113

Pratiwi, N., & Setyawan, Y. (2021). Analisis Akurasi Dari Perbedaan Fungsi Kernel Dan Cost Pada Support Vector Machine Studi Kasus Klasifikasi Curah Hujan Di Jakarta. *Journal of Fundamental Mathematics and Applications (JFMA)*, *4*(2), 203–212. https://doi.org/10.14710/jfma.v4i2.11691

Rabbani, S., Safitri, D., Rahmadhani, N., Sani, A. A. F., & Anam, M. K. (2023). Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *3*(2), 153–160. https://doi.org/10.57152/malcom.v3i2.897

Susanto, L. A. (2023). Pemilihan Hyperparameter Pada Alexnet Cnn Untuk Klasifikasi Citra Penyakit Kedelai. *Indexia*, *5*(02), 113. https://doi.org/10.30587/indexia.v5i02.5508