Jurnal Mandiri IT
Vol. 14 No. 1, July (2025), pp. 129-138

ISSN 2301-8984 (Print), 2809-1884 (Online)
Published by Institute of Computer Science

Analyzing public sentiment on youtube comments regarding the free lunch policy using the Support Vector Machine (SVM) algorithm

Linda Septiana[1], Verdy Yasin[2], Anton Zulkarnain Sianipar[3]
[1,2,3]Informatics Engineering, Sekolah Tinggi Teknologi Informasi dan Manajemen Komputer Jayakarta, Indonesia

ARTICLE INFO

Article history:

Received Jun 13, 2025
Revised Jul 17, 2025
Accepted Jul 24, 2025

Keywords:

Lunch Program;
Sentiment Analysis;
Support Vector Machine (SVM);
TF-IDF;
Youtube.

ABSTRACT

The advancement of information technology and social media has reshaped how individuals express their opinions on public policies. YouTube has emerged as a major platform where public sentiment is openly shared, including reactions to the government's Free Lunch Program for elementary school students. This study aims to analyze public sentiment toward the policy using the Support Vector Machine (SVM) algorithm with both linear and Radial Basis Function (RBF) kernels. A total of 1,883 YouTube comments were collected and manually labeled into three sentiment categories: positive, negative, and neutral. The preprocessing steps included cleansing, case folding, normalization, tokenization, stopword removal, and stemming, followed by TF-IDF transformation. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics, and validated using 10-Fold Cross Validation to ensure result consistency. The findings indicate that the SVM model with RBF kernel and 10-fold cross-validation achieved the highest accuracy at 81.46%. However, the linear kernel model provided a more balanced performance with superior precision, recall, and F1-score. These results highlight the importance of choosing the right kernel and validation strategy in developing sentiment analysis models, especially when dealing with imbalanced social media data.

Corresponding Author:

Linda Septiana,
Informatics Engineering,
Sekolah Tinggi Teknologi Informasi dan Manajemen Komputer Jayakarta,
Jl. Akasia Dalam, Mangga Dua Selatan, Sawah Besar, Kota Jakarta Pusat, 10250, Indonesia
Email: 21570010@stmik.jayakarta.ac.id

1. INTRODUCTION

In the last twenty years, rapid developments in digital technology have greatly influenced how people express their views and participate in discussions related to public policy issues. Among social media platforms, YouTube has emerged as a prominent medium enabling people to express their views openly and in real time regarding government initiatives. One of the policies that has attracted substantial public attention is the Free Lunch Program for elementary school students, introduced by the government to enhance children's nutritional standards and overall well-being in Indonesia (Palupi et al., 2020). Despite its positive social intent, this policy has also raised public concerns about funding availability, infrastructure adequacy, and the effectiveness of its implementation monitoring (Gunadi et al., 2025).

To explore public opinion dynamics, sentiment analysis offers a relevant and powerful method (Sebastian et al., 2024). As a branch of natural language processing (NLP), sentiment analysis seeks to categorize opinions into positive, negative, or neutral groups (Hermawan et al., 2023). The Support Vector Machine (SVM) algorithm is widely used for sentiment classification due to its capability to effectively distinguish between data classes, even in high-dimensional and unstructured content like social media comments. Several previous studies have confirmed SVM's robust performance. For example, one study reported that SVM achieved 82% accuracy, outperforming Naive Bayes at 80% when dealing with complex, non-linear data (Rakarahayu Putri & Cahyono, 2024). Another study showed that the linear kernel produced the highest accuracy (67.10%) compared to the sigmoid (64.73%) and RBF (62.43%) kernels in sentiment classification of product reviews (Wicaksono et al., 2023).

This study chose YouTube as the primary data source due to its unique features, which allow users to express public sentiment in a spontaneous, varied, and context-rich manner, surpassing many other social media platforms. YouTube's comment section serves as an interactive space where users directly respond to public policy-related content, making it a valuable resource for real-time and in-depth sentiment analysis (Safra & Zuliarso, 2025).

Although interest in sentiment analysis continues to grow, there are still limited studies that specifically examine public sentiment toward government policies by utilizing YouTube comments. In previous studies, such as those conducted by (Harun et al., 2023), the focus was on analyzing sentiment toward fuel price hike policies using data from the Twitter platform, where public sentiment tended to show dissatisfaction and harsh criticism of the economic burden caused. In contrast, this study focuses on the Free Lunch Program policy by analyzing YouTube user comments from six channels specifically selected for discussing the latest free lunch topics. The selection of these channels was based on the number of views and the relevance of their content in discussing the free lunch policy from various perspectives.

This study applies the Support Vector Machine (SVM) algorithm using both linear and Radial Basis Function (RBF) kernels to classify public sentiment. The model's effectiveness is assessed through K-Fold Cross Validation and key evaluation metrics, including accuracy, precision, recall, and F1-score, with special attention given to the challenge of imbalanced datasets. The findings are intended to enhance sentiment analysis accuracy and provide meaningful insights that support data-driven public policy decisions.

2. RESEARCH METHOD

The methodology adopted in this research follows a structured sequence consisting of several essential stages, such as data acquisition, preprocessing, feature transformation, implementation of the classification algorithm, and model performance assessment. A visual overview of the complete methodological process is presented in Figure 1.

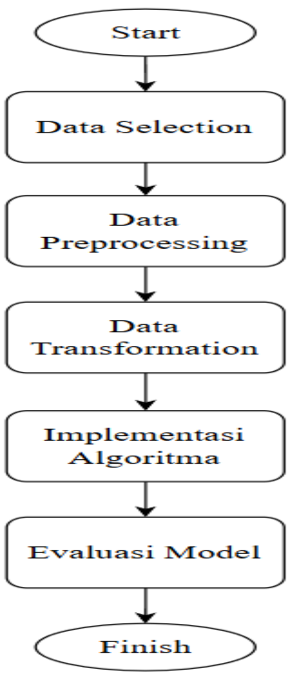

Figure 1. Research stages

Data Selection

During this phase of the research, data collection was carried out using scraping techniques by utilizing the YouTube Data API v3 to extract user comments from several YouTube channels discussing free nutritious meals for elementary school students. The research objects included user-generated comments found in videos from channels such as #FERI, #TEMPO, #PANDJI, #SEKERTARIAT NEGARA, #LEON and #TVONE. A total of 1,883 comments were collected, selected based on keywords and discussion topics relevant to the free nutritious meal policy.

After data collection, a selection process was conducted to ensure the validity and relevance of the dataset. This process involved assessing the suitability of comment content with the research topic and eliminating duplicate entries containing identical text. After the selection process was completed, the remaining comments were manually labeled into three sentiment categories: positive, negative, and neutral. The labeling process was carried out directly by researchers through an assessment of the contextual meaning and emotional tone of each comment. Labeling was done independently without involving language experts, and no inter-annotator agreement tests were conducted. This approach was taken in consideration of resource limitations, while still striving to maintain consistency in classification based on contextual understanding.

Data Preprocessing

Data preprocessing was carried out in six stages: a) Cleansing: cleaning the text of irrelevant elements or reducing noise in the collected tweet data, such as symbols, URL links, hashtags "#," and "@" when mentioning usernames; b) Case Folding: In this stage, words with uppercase letters are converted to lowercase letters. The purpose is to eliminate data redundancy by standardizing the form of words that are actually the same but written with different capitalization, such as "Makan" and "makan," which should be treated as the same word; c) Normalization: This is the process of correcting non-standard words, slang, abbreviations, or words in non-standard forms into standard word forms. Comments on social media often use informal language such as "gk" becoming "tidak," "bgus" becoming "bagus," or "udh" becoming "sudah."; d) Tokenization involves splitting sentences in text based on the word that makes them up. Each word is separated by a space and single quotation marks. For example, the comment "The free lunch program is very helpful" would be split into tokens: ["program," "lunch," "free," "very," "helpful"]; e) Stopword Removal: This step is performed to remove words considered unnecessary in the tweet data. Common stopwords include "from," "to," "that," "in," "and" and so on. These words frequently appear in almost all sentences but do not provide specific information about sentiment; f) Stemming: At this stage, words found in the comment data that contain affixes are reduced to their root or base forms. For instance, terms such as "repair," "repairs," and "repaired" are normalized to a common base form like "repair" or "repairing."

Data Transformation

At the data transformation stage, the preprocessed textual comments are transformed into numerical vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) technique. This step is intended to convert text into a structured format that can be efficiently utilized by classification algorithms. TF-IDF operates by assigning weights to words based on their frequency within individual documents and their rarity across the overall dataset, thereby highlighting terms that are most relevant for analysis.

Implementasi Algoritma

At this point in the process, sentiment classification is carried out using the Support Vector Machine (SVM) algorithm. The dataset, which has previously been transformed into numerical vector form through the Term Frequency–Inverse Document Frequency (TF-IDF) technique, serves as input for the classification task. This research utilizes two kernel types linear and Radial Basis Function (RBF) to compare their performance in categorizing public comments into three sentiment classes: positive, negative, and neutral.

To thoroughly assess the model's reliability and performance, this research contrasts the classification results obtained through the K-Fold Cross Validation method (with k set to 10) against those produced by a model trained using a conventional random train-test split without any cross-validation procedure. This strategy aims to examine the influence of K-Fold Cross Validation on model stability and accuracy, and to determine the degree to which this validation method helps in minimizing performance variability across different data partitions.

Evaluasi Model

The final phase of this study involves assessing the performance of the sentiment classification model developed using the Support Vector Machine (SVM) algorithm. This assessment is carried out through the use of a confusion matrix, which allows for evaluating how accurately the model predicts sentiment categories. Based on the confusion matrix, several key performance metrics are derived—namely accuracy, precision, recall, and F1-score—which serve to measure the model's effectiveness in terms of prediction accuracy, completeness, and its ability to maintain balanced classification across the three sentiment categories: positive, negative, and neutral.

3. RESULTS AND DISCUSSIONS

This research implements the Python programming language to develop a sentiment analysis model using the Support Vector Machine (SVM) algorithm with both linear and Radial Basis Function (RBF) kernels. The data was obtained from YouTube user comments discussing the Free Lunch program. Classification was performed into three sentiment categories: positive, negative, and neutral. The model's consistency and overall performance were evaluated using the K-Fold Cross Validation method with k set to 10.

Data Selection

Table 1 presents the raw dataset collected through the data scraping method.

Table 1. Results of initial data collection

No	Video_ID	Author	Date	Text
1	EEmfTnEhVJ0	@agungsawolo4918	2025-05-18 05:10	MBG ADALAH PROGRAM SANGAT BERMANFAAT BAGI RAKYAT NKRI ♥♥♥♥
2	a2WXt0aW76g	@EsSembilan-b9f	2025-06-21 06:17	Saya sangat mendukung bocor alus yang berani kritik kebijakan pemerintah karena pengelola uang negara petugas urus negara hrs diawasi
3	rrTT_3KwcSs	@nebulaharis	2025-03-08 07:39	terimakasih cerita ini membuat saya masih tetap optimis dalam menjalani hidup 🎧

The initial dataset consists of several attributes, and Table 2 contains explanations of each attribute.

Table 2. Attributes the initial dataset from data scraping results

No	Atribut	Description
1	Video_ID	This is the unique identifier of the YouTube video where the comment originated. This attribute serves to link each comment to a specific video, allowing researchers to understand the context of the comment based on the video content being discussed.
2	Author	The name of the YouTube user account that posted the comment.
3	Date	Information about when a tweet was created and uploaded. The date attribute includes the year, month, day, and time.
4	Text	The content of the comment itself in text form, which is the primary focus of sentiment analysis.

After completing the data selection stage, the comments that had passed the curation stage were then manually labeled with sentiment by the researchers themselves without involving language experts. The labeling was done based on an understanding of the Indonesian language context and the content of the comments relevant to the public policy topic being analyzed. This process aimed to identify sentiment trends in each comment, which were then grouped into three types: positive, neutral, and negative.

As indicated by the outcomes of manual labeling of all comment data obtained from the YouTube platform, the following sentiment distribution was obtained: 736 comments were categorized as positive, 1097 comments were categorized as negative, and 50 comments were categorized as neutral. Additional information on the classification outcomes can be found in Table 3.

Table 3. Sample data results label

No	Komentar	Label
1	MBG ADALAH PROGRAM SANGAT BERMANFAAT BAGI RAKYAT NKRI ♥♥♥♥	Positive
2	"Jangan sampai kita bodoh/di drive sama bias itu" Hal yang ga bisa di lakuin sma orang ga punya pendirian/krng berani bersuara. Mnurut gua	Negative
3	BANG, TOLONG BUATKAN VIDEO TENTANG KONTEN ANAK-ANAK YANG SANGAT MERESAHKAN: KONTEN MENGERIKAN, HAMIL BOHONGAN, HANTU-HANTUAN. SUDAH SANGAT MERESAHKAN. SUDAH DILAPORKAN, DIBLOK, DIKOMENTARIN TAPI BANYAK BANGET CHANNEL-CHANNEL YANG UNGGAH KONTEN SEPERTI ITU.	Neutral
4	Kalo idenya pak ahok aku setuju, berkaca sama KJP yg tiap bulan bisa dibelanjaain daging sapi sama susu telor sama ortunya. Dan bener adeknya yg belum sekolah juga bisa ikut makan. Tapi balik lagi mekanisme belanjanya kek gimana, harus ada kontrol juga biar duitnya bener2 buat belanja makanan bergizi.	Positive

Data Preprocessing

Preprocessing is done to ensure that the data becomes more organized, structured, and consistent. This is important because the initial data obtained still contains a lot of noise and words that do not have significant information value. Therefore, several preprocessing stages were applied to clean and refine the data before entering the modeling stage. A sample output from the preprocessing stage is presented in Table 4 below.

Table 4. Preprocessing result example

No	Proses	Hasil
1	Komentar Asli	MBG ADALAH PROGRAM SANGAT BERMANFAAT BAGI RAKYAT NKRI ♥♥♥♥
2	Cleansing	MBG adalah PROGRAM SANGAT BERMANFAAT BAGI RAKYAT NKRI
3	Case Folding	mbg adalah program sangat bermanfaat bagi rakyat nkri
4	Normalization	['mbg', 'adalah', 'program', 'sangat', 'bermanfaat', 'bagi', 'rakyat', 'nkri']
5	Tokenizing	['mbg', 'adalah', 'program', 'sangat', 'bermanfaat', 'bagi', 'rakyat', 'nkri']
6	Stopword Removal	['mbg', 'program', 'bermanfaat', 'rakyat', 'nkri']
7	Stemming	['mbg', 'program', 'manfaat', 'rakyat', 'nkri']

Data Transformation

After completing the preprocessing steps including stemming, the data is converted into numerical values using Term Frequency–Inverse Document Frequency (TF-IDF).

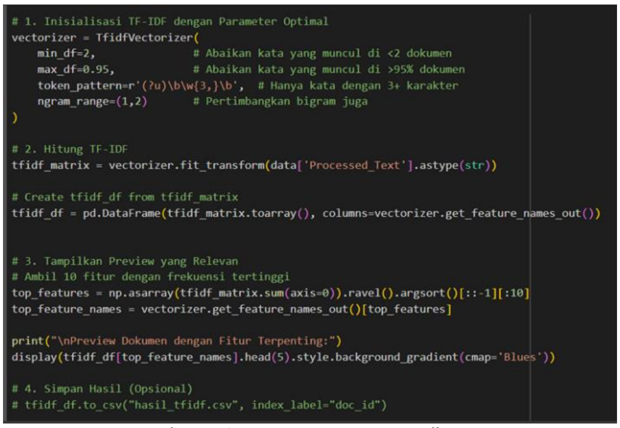

Figure 2. TF-IDF process coding

Previously processed data up to The text data transformation process was carried out using the TF-IDF method with TfidfVectorizer, which was configured to consider unigrams and bigrams, and ignore words that were too rare or too common. The transformation produced a TF-IDF matrix, from which the 10 most important features were selected based on the highest weights. The first five documents were displayed as a visual preview of the feature weights, as shown in the following image.

Preview Dokumen dengan Fitur Terpenting:

	yang	ini	makan	tidak	dan	ada	bang	pak	gratis	itu
0	0.000000	0.000000	0.000000	0.000000	0.000000	0.169883	0.000000	0.000000	0.000000	0.000000
1	0.087728	0.000000	0.000000	0.212610	0.000000	0.000000	0.000000	0.000000	0.000000	0.125764
2	0.121572	0.000000	0.000000	0.000000	0.000000	0.000000	0.090655	0.000000	0.000000	0.087141
3	0.111231	0.000000	0.141053	0.000000	0.071988	0.074189	0.000000	0.079729	0.000000	0.000000
4	0.062060	0.000000	0.078699	0.000000	0.080330	0.000000	0.000000	0.000000	0.087355	0.088968

Figure 3. TF-IDF results

Implementasi Algoritma

This process is carried out so that sentiment labels in text form can be converted into a numerical form that can be recognized by the classification algorithm. Meanwhile, the input features are the result of transforming the comment text using the TF-IDF method so that each document is represented as a high-dimensional numerical vector.

The data is then split into 80% training data and 20% test data using the train_test_split() function from scikit-learn, with the parameter stratify=y to maintain a balanced class distribution ratio in both subsets. This division results in 1,506 training data points and 377 test data points, which are subsequently used in the model training and evaluation process.

Table 5. Test scenario

Scenario	Model	Data Used	Description
1	SVM (Linear)	80 train, 20 test	Train SVM with linear kernel, evaluate on fixed test set
2	SVM (RBF)	80 train, 20 test	Train SVM with RBF kernel, evaluate on fixed test set
3	SVM (Linear)	80 train (K-fold)	10-fold CV: Each fold uses 72 for train, 8 for validation; results averaged across 10 folds
4	SVM (RBF)	80 train (K-fold)	10-fold CV: Each fold uses 72 for train, 8 for validation; results averaged across 10 folds

Evaluasi Model

The goal of the model evaluation stage is to determine the effectiveness of the implemented classification algorithm in assigning YouTube comments to one of three sentiment categories: positive, negative, or neutral. This evaluation utilizes a confusion matrix to generate values for True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). These values serve as the basis for calculating essential performance metrics such as accuracy, precision, recall, and F1-score. A summary of the evaluation outcomes for each testing approach is provided in Table 6 below.

Table 6. Evaluasi model

Model	Evaluasi	Accuracy	Precision	Recall	F1-Score
svm linear	train-test-split	75.60%	51.23	49.73	49.71
svm rbf	train-test-split	76.66%	73.85	57.07	60.77
svm linear	10-flod-cross validaion	80.13%	65.45	64.19	64.49
svm rbf	10-flod-cross validaion	81.46%	54.82	53.25	53.58

As presented in Table 6, the Support Vector Machine (SVM) algorithm was evaluated by applying two types of kernel functions—linear and Radial Basis Function (RBF)—alongside two distinct validation methods: the train-test split and the 10-fold cross-validation. The evaluation results demonstrate that the best accuracy, reaching 81.46%, was obtained when using the RBF kernel in combination with the 10-fold cross-validation strategy. However, model quality assessment cannot solely rely on accuracy parameters, especially when dealing with multi-class classification problems and dataset characteristics with imbalanced distributions, as encountered in this sentiment analysis study.

Upon closer examination, although it shows slightly lower accuracy (80.13%), the implementation of the linear kernel SVM with the 10-fold cross-validation approach was able to produce superior precision (65.45), recall (64.19), and F1-score (64.49) values compared to the RBF kernel SVM. This phenomenon indicates that the model with a linear kernel has better ability to classify data proportionally and stably for all existing categories, thereby demonstrating superior performance in terms of model generalization and reliability.

On the other hand, SVM with RBF kernel shows a tendency to be biased toward certain categories, which can be observed from the precision and recall values that tend to be lower even though the overall accuracy rate shows a higher number. This condition implies that the RBF model does indeed more often show “accuracy” in global predictions, but it may not necessarily provide fair and comprehensive classification for every label or category in the dataset.

Additionally, this performance difference can also be attributed to the inherent characteristics of each kernel. The linear kernel tends to be more effective in handling high-dimensional data with relatively simple structures, while the RBF kernel is more sensitive to parameter tuning and can experience overfitting on complex datasets. These results provide important insights that the selection of the appropriate kernel should consider not only accuracy but also performance balance across all evaluation metrics to ensure that the resulting model is reliable in practical applications.

4. CONCLUSION

Based on an in-depth analysis of sentiment analysis research on free meal programs using YouTube comment data, It may be inferred that the application of the Support Vector Machine (SVM) algorithm with various kernel configurations and validation methods produced significant findings. This study successfully processed 1,883 comments from various prominent YouTube channels such as #FERI, #TEMPO, #PANDJI, #SEKERTARIAT NEGARA, #LEON and #TVONE, which were then categorized into 736 positive comments, 1,097 negative comments, and 50 neutral comments, indicating a dominance of negative sentiment in public discussions regarding the free meal program policy for elementary school students.

The model evaluation results showed that although the SVM with the RBF kernel using 10-fold cross-validation achieved the highest accuracy of 81.46%, the model did not provide the best overall performance. Conversely, SVM with a linear kernel validated using 10-fold cross-validation proved to be superior, achieving higher precision (65.45), recall (64.19), and f1-score (64.49) values than the RBF kernel, despite its slightly lower accuracy (80.13%). This phenomenon indicates that the linear kernel has better capabilities in handling unbalanced data characteristics and is able to provide more proportional classification for all sentiment categories.

Another important finding is that the use of the 10-fold cross-validation technique consistently produces better performance than the simple train-test-split method, both for the linear and RBF kernels. This shows that cross-validation reduces variance and improves model stability, resulting in more robust and reliable evaluations. The significant performance difference between the two validation methods underscores the importance of selecting the appropriate validation technique in machine learning model development, especially for data with complex characteristics such as social media sentiment.

From a methodological perspective, this study demonstrates the significance of applying comprehensive and systematic data preprocessing steps, which include a series of data cleansing, case folding, text normalization, tokenization, stopword removal, and stemming processes. These entire stages have been proven to make a substantial contribution to improving the quality and consistency of the input data to be used in the sentiment classification process.

The data transformation process, which implements the Term Frequency-Inverse Document Frequency (TF-IDF) method configured to analyze unigram and bigram patterns, has demonstrated high effectiveness in converting textual data into a numerical representation format that can be understood and processed optimally by machine learning algorithms. The integration of comprehensive preprocessing stages with appropriate transformation techniques has established a solid and robust methodological foundation for developing sentiment classification models with high accuracy and reliability, particularly in the context of evaluating public opinion on government policies through the analysis of digital social media platform content.

Thus, the SVM approach and cross-validation play a significant role in sentiment analysis practices regarding public policy, particularly in terms of enhancing classification stability, addressing data imbalance, and providing reliable results for data-driven decision-making. Future development directions could focus on strategies for handling imbalanced data, further experiments with various hyperparameter combinations, and expanding the scope of the dataset—both in terms of the number of YouTube channels and other public policy topics—to enable the model to have better generalization capabilities in capturing public opinion more broadly across digital platforms.

REFERENCES

Ansori, Y., & Holle, K. F. H. (2022). Perbandingan Metode Machine Learning dalam Analisis Sentimen Twitter. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, *10*(4), 429. https://doi.org/10.26418/justin.v10i4.51784

Efraim, D. A. (2023). *Analisis Sentimen Pada Sosial Media Instagram Menggunakan Algoritma Naive Bayes (Studi Kasus : Timnas Futsal Indonesia)*. *April 2012*, 498–509.

Gunadi, G. A., Raharjo, J. S., & Setianingsih, S. (2025). *Analisis Kemanfaatan Kebijakan Program Makan Siang Gratis bagi Peserta Didik dan Pemerintahan*. *9*(1), 7403–7411.

Harun, R., Ishak, R., & Panna, S. (2023). Analisis Sentimen Opini Publik Pengguna Twitter Terhadap Kenaikan Harga BBM Menggunakan Algoritma Naïve Bayes. *Jurnal Ilmiah Ilmu Komputer Banthayo Lo Komputer*, *2*(1), 26–33. https://doi.org/10.37195/balok.v2i1.414

Hermawan, A., Jowensen, I., Junaedi, J., & Edy. (2023). Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine. *JST (Jurnal Sains Dan Teknologi)*, *12*(1), 129–137. https://doi.org/10.23887/jstundiksha.v12i1.52358

Imaddudin, S., Astuti, I., & Ruhama, S. (2025). *Studi Sentimen Masyarakat terhadap PSSI di Era Erick Thohir menggunakan Algoritma Support Vector Machine (SVM) pada Media Sosial X*. *1*(8), 1003–1013.

Liu, H., Chen, X., & Liu, X. (2022). A Study of the Application of Weight Distributing Method Combining Sentiment Dictionary and TF-IDF for Text Sentiment Analysis. *IEEE Access*, *10*, 32280–32289. https://doi.org/10.1109/ACCESS.2022.3160172

Mufid, A. (2023). *Google Colab: Pengertian, Keuntungan, dan Cara Menggunakan*. http://blog.rumahweb.com/google-colab-adalah/

Muttaqin, Wahyu Wijaya Widiyanto, M. M., Green Ferry Mandias, Stenly Richard Pungus, A. W., Wiranti Kusuma Hapsari, S. A. H., Aslam Fatkhudin, Pasnur, E. F. B., & Mochammad Anshori, Suryani, N. S. (2023). *Pengenalan Data Mining* (Issue July).

Muzaki, A., Febriana, V., & Cholifah, W. N. (2024). Analisis Sentimen Pada Ulasan Produk di E-Commerce dengan Metode Naive Bayes. *Jurnal Riset Dan Aplikasi Mahasiswa Informatika (JRAMI)*, *5*(4), 758–765. https://doi.org/10.30998/jrami.v5i4.9647

Negi, M., Vishwakarma, K., & Badhani, P. (2021). *Kajian Analisis Sentimen Twitter Menggunakan Algoritma Machine Learning di Studi Analisis Sentimen Twitter Menggunakan Mesin Belajar Algoritma di Python*. https://doi.org/10.5120/ijca20

Nurkholis, A., Alita, D., & Munandar, A. (2022). Comparison of Kernel Support Vector Machine Multi-Class in PPKM Sentiment Analysis on Twitter. *Jurnal RESTI*, *6*(2), 227–233. https://doi.org/10.29207/resti.v6i2.3906

Palupi, I. R., Rachmawati, V. N., & Prawiningdyah, Y. (2020). 632 HIGEIA 4 (4) (2020) HIGEIA JOURNAL OF PUBLIC HEALTH RESEARCH AND DEVELOPMENT Pemenuhan Gizi dari Penyelenggaraan Makan Siang Sekolah dan Konsentrasi Siswa Sekolah Dasar. *HIGEIA Journal of Public Health Reseacrch and Development*, *4*(4), 632–644.

Permana, A. A., Fahrezi, M. F., Kristiyanti, D. A., & Sihotang, M. (2021). Sentimen Analisis Opini Masyarakat Pada Media Sosial Twitter Terhadap Vaksin Berbayar Menggunakan Metode Naïve Bayes Classifier (Nbc). *Jurnal Teknik*, *10*(2), 84–92. https://doi.org/10.31000/jt.v10i2.5471

Pratama, J. A., Suprijadi, Y., & Zulhanif, Z. (2017). The Analisis Sentimen Sosial Media Twitter Dengan Algoritma Machine Learning Menggunakan Software R. *Jurnal Fourier*, *6*(2), 85. https://doi.org/10.14421/fourier.2017.62.85-89

Rahardi, M., Aminuddin, A., Abdulloh, F. F., & Nugroho, R. A. (2022). Sentiment Analysis of Covid-19 Vaccination using Support Vector Machine in Indonesia. *International Journal of Advanced Computer Science and Applications*, *13*(6), 534–539. https://doi.org/10.14569/IJACSA.2022.0130665

Rakarahayu Putri, R., & Cahyono, N. (2024). Analisis Sentimen Komentar Masyarakat Terhadap Pelayanan Publik Pemerintah Dki Jakarta Dengan Algoritma Super Vector Machine Dan Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 2363–2371. https://doi.org/10.36040/jati.v8i2.9472

Retnoningsih, E., & Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python. *Bina Insani Ict Journal*, *7*(2), 156. https://doi.org/10.51211/biict.v7i2.1422

Safitri, T., Umaidah, Y., & Maulana, I. (2023). Analisis Sentimen Pengguna Twitter Terhadap Grup Musik BTS Menggunakan Algoritma Support Vector Machine. *Journal of Applied Informatics and Computing*, *7*(1), 28–35. https://doi.org/10.30871/jaic.v7i1.5039

Safra, K. K., & Zuliarso, E. (2025). Analisis Sentimen Terhadap Pelaksanaan Pilkada 2024 Pada Media Sosial Youtube Menggunakan Metode Decision Tree. *Jurnal Informatika Teknologi Dan Sains (Jinteks)*, *7*(1), 117–126. https://doi.org/10.51401/jinteks.v7i1.5295

Sebastian, D. F., Sulistiani, H., & Isnain, A. R. (2024). *Sentiment Analysis of Public Opinion on the Right of Inquiry in Indonesia in 2024 Using the Support Vector Machine (Svm) Method Analisis Sentimen Opini Masyarakat Mengenai Hak Angket Di Indonesia Tahun 2024 Menggunakan Metode Support Vector Machine (Sv*. *5*(4), 1025–1034.

Suryadevara, C. K., & Services, R. S. (2023). *INSIGHTS FROM THE TUBE : ANALYZING YOUTUBE COMMENTS WITH International Engineering Journal for Research & Development Issue 5 INSIGHTS*

FROM THE TUBE : ANALYZING YOUTUBE COMMENTS WITH E-ISSN NO : 2349-0721. *September 2016*.

Wicaksono, M. H., Purbolaksono, M. D., & Faraby, S. Al. (2023). Perbandingan Algoritma Machine Learning untuk Analisis Sentimen Berbasis Aspek pada Review Female Daily. *EProceedings of Engineering*, *10*(3), 3591–3600.

Widayani, W., & Harliana, H. (2021). Analisis Support Vector Machine Untuk Pemberian Rekomendasi Penundaan Biaya Kuliah Mahasiswa. *Jurnal Sains Dan Informatika*, *7*(1), 20–27. https://doi.org/10.34128/jsi.v7i1.268

Zain, H. H., Awannga, R. M., & Rahayu, W. I. (2023). Perbandingan Model Svm, Knn Dan Naïve Bayes Untuk Analisis Sentiment Pada Data Twitter: Studi Kasus Calon Presiden 2024. *JIMPS: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, *8*(3), 2083–2093. https://jim.usk.ac.id/sejarah