Smart Techno (Smart Technology, Informatic, and Technopreneurship)
Volume 8, Number 2, 2026 pp. 291-302
E-ISSN : 2541-0679
Open Access: https://lppm.primakara.ac.id/jurnal/index.php/smart-techno/



Sentiment Analysis of Public Opinions on Electric Vehicles on Social Media Using the Support Vector Machine (SVM) Algorithm with Optimized Preprocessing

Aurora Aulia[1], Fernando Yunianda[2], Bima Sanggraloka[3], Jovi Satria[4], Eko Prasetyo[5] Rahmawati Febrifyaning Tias[6]
[1,2,3,4,5,6] Informatics Engineering, Universitas Bhayangkara Surabaya, Surabaya, Indonesia

ARTICLE INFO
Article history:
Initial submission
15-06-2026
Received in revised form
07-08-2026
Accepted 13-08-2026
Available online 30-09-2026

Keywords:
User-Centered Design, User Interface, Usability Testing, System Usability Scale, Academic Information System

DOI:
https://doi.org/10.59356/smart-techno.v8i2.229

ABSTRACT

The development of regulations and trends in electric vehicle adoption in Indonesia has generated diverse public responses and opinions on social media platforms such as YouTube. However, manually evaluating large volumes of unstructured review data tends to be inefficient and less effective. This study applies the Support Vector Machine (SVM) algorithm with a Radial Basis Function (RBF) kernel, utilizing Term Frequency-Inverse Document Frequency (TF-IDF) weighting and optimized preprocessing techniques to automatically classify public opinion sentiment. The dataset consists of 1,517 Indonesian-language YouTube comments that underwent a series of preprocessing stages, including case folding, cleaning, tokenization, stopword removal, normalization, and stemming. The resulting cleaned dataset was then transformed into a numerical matrix containing 1,180 word features and divided into training and testing sets using a 70:30 ratio. The initial labeling results showed that public opinions were dominated by negative sentiment, accounting for 57.3% (869 comments), followed by positive sentiment at 33.2% (504 comments) and neutral sentiment at 9.5% (144 comments). The performance evaluation of the SVM classification model using a confusion matrix achieved an overall accuracy of 78.1%. These findings demonstrate that the application of optimized preprocessing techniques combined with a nonlinear SVM kernel provides reliable performance in accurately identifying and mapping the characteristics of digital public opinion.

1. INTRODUCTION

Electric vehicles have currently become a major focus of the global automotive industry and in Indonesia as one of the concrete efforts to reduce negative environmental impacts (Dzaky Candy Fahrezy & Fandy Setyo Utomo, 2024). The continuous increase in the use of fossil fuel-powered motor vehicles has caused Indonesia's fuel reserves to decline gradually over the years (Erfina & Lestari, 2023). Therefore, the Indonesian government strongly supports the transition from conventional vehicles to electric vehicles to reduce carbon emissions and greenhouse gas emissions in a measurable manner (Farhani, 2024). Nevertheless, the presence and adoption of electric vehicles continue to generate both support and opposition and remain a subject of public debate (Sri Widagdo Adika Nuresa Qodri et al., 2023).

Research Gap

Although electric vehicles are widely considered an environmentally friendly solution, public understanding of this technology in Indonesia remains relatively limited (Luthfi Krisna Bayu & Wurijanto, 2024). Public opinion analysis indicates that limited public literacy is an issue that directly contributes to the emergence of negative opinions (Yulianto et al., 2025). Previous studies have extensively employed standard algorithms such as Naïve Bayes, K-

* *Corresponding Author: Eko Prasetyo: eko@ubhara.ac.id*

Nearest Neighbor (KNN), and Decision Tree; however, the classification performance of these models has varied and still requires improvement (Sehabudin et al., 2025). As a distinction and key advantage over twelve previous studies, most of which analyzed short reviews from Twitter without adequately addressing the quality of raw text, this study specifically utilizes YouTube comments, which generally contain longer and more complex sentence structures (Sehabudin et al., 2025). This study introduces novelty through rigorous preprocessing optimization, including the conversion of informal automotive language commonly used by netizens before the data are used for model training (Aryanti & Santoso, 2023). The combination of effective text processing at the initial stage and the nonlinear separation capability of the Support Vector Machine (SVM) algorithm using an RBF kernel is specifically designed to address the complexity of informal language (Pramana et al., 2025). This approach is proposed to demonstrate that optimization during the preprocessing phase can significantly improve the model's performance and classification accuracy compared with approaches that focus solely on algorithmic tuning (Ningsih et al., 2024).

Problem Statement

Based on the aforementioned background and research gap, the primary issue concerns how Indonesian society responds to and expresses sentiment toward the adoption of electric vehicles (Aryanti & Santoso, 2023). Furthermore, it is important to identify how public opinion specifically responds to the electric vehicle incentive policies implemented by the government (Pramana et al., 2025). The subsequent question concerns which machine learning method or algorithm provides the most optimal performance in classifying public support for and opposition to the transition from conventional vehicles to electric vehicles (Ningsih et al., 2024).

Proposed Solution

To address these research problems, sentiment analysis is proposed as the primary method for classifying public opinion texts into positive, negative, or neutral sentiment categories (Ningsih et al., 2024). Public opinion data will be collected specifically from Indonesian-language comments on YouTube, which is frequently used as a platform for public discussion (Erfina & Lestari, 2023). The raw text data will then undergo advanced preprocessing and be transformed into numerical representations using the Term Frequency-Inverse Document Frequency (TF-IDF) weighting method (Pramana et al., 2025). Finally, the sentiment classification model will be evaluated and optimized using the nonlinear approach of the Support Vector Machine (SVM) algorithm with an RBF kernel to obtain more precise measurements of public perceptions (Daniati et al., 2026).

2. LITERATURE REVIEW

Research on public opinion toward electric vehicles using the Naïve Bayes algorithm has shown that this decision-making method can achieve competitive accuracy in text classification (Yulianto et al., 2025). Sentiment analysis of electric vehicle incentive policies using a combination of TF-IDF weighting and Naïve Bayes has effectively mapped public perceptions (Pramana et al., 2025). A comparative study of the Naïve Bayes and Support Vector Machine (SVM) algorithms on digital data revealed that SVM offers greater stability in processing opinion data related to electric vehicles (Ningsih et al., 2024). Sentiment evaluation based on public platforms regarding the development trends of electric vehicles in Indonesia has also confirmed the computational efficiency of these methods in handling large-scale opinion datasets (Aryanti & Santoso, 2023).

Meanwhile, opinion analysis of YouTube comments concerning electric vehicle content using the SVM algorithm has demonstrated its reliability in generating an optimal decision boundary or hyperplane (Sehabudin et al., 2025). Evaluation of deep learning methods such as Long Short-Term Memory (LSTM) has shown strong performance in handling the contextual

complexity of informal sentences in electric vehicle reviews (Sri Widagdo Adika Nuresa Qodri et al., 2023). Other studies exploring public opinions across various digital platforms have concluded that automated sentiment mapping can significantly assist in overcoming the limitations of manual analysis (Dzaky Candy Fahrezy & Fandy Setyo Utomo, 2024). The integration of appropriate feature extraction techniques has also been shown to improve evaluation metrics in sentiment classification models for electric vehicles in Indonesia (Farhani, 2024). Supporting studies on electric vehicles further emphasize that the non-standard language characteristics commonly used by netizens require accurate preprocessing to prevent degradation in model performance and classification accuracy (Erfina & Lestari, 2023). Ultimately, sentiment classification using machine learning approaches can provide an objective overview for the government in evaluating public responses to energy transition programs (Luthfi Krisna Bayu & Wurijanto, 2024).

Social Media as a Source of Public Opinion

YouTube has increasingly evolved into a highly dynamic digital public discussion space for society (Dzaky Candy Fahrezy & Fandy Setyo Utomo, 2024). Comments and video posts on the platform provide a broad source of opinion data that can be continuously collected and analyzed in near real time (Ningsih et al., 2024). The high level of interaction among digital users makes YouTube a valuable source of data for capturing public perspectives in an objective manner (Pramana et al., 2025).

Electric Vehicle Policies and Public Opinion

The implementation of electric vehicle policies in Indonesia has frequently triggered emotional responses and diverse debates among various audiences (Dzaky Candy Fahrezy & Fandy Setyo Utomo, 2024). Public interactions in digital spaces reflect communication in which support and opposition are expressed through diverse forms of sentiment (Aryanti & Santoso, 2023). Analyzing these public opinions is important for the government to evaluate the effectiveness and impact of implemented regulations (Farhani, 2024).

Sentiment Analysis

Sentiment analysis is a text-processing technique used to automatically identify and process information to extract public opinions from textual data (Erfina & Lestari, 2023). This process utilizes natural language processing and machine learning techniques to classify opinions from large-scale datasets into specific categories, such as positive and negative sentiments (Sehabudin et al., 2025). Automated sentiment analysis is essential because the large volume of available data makes manual opinion evaluation inefficient and impractical (Luthfi Krisna Bayu & Wurijanto, 2024).

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a machine learning algorithm that is widely used and reliable for both binary and multiclass classification tasks (Daniati et al., 2026). The method operates by identifying an optimal decision boundary or hyperplane that separates different sentiment classes while maximizing the margin between them (Ningsih et al., 2024). SVM has consistently demonstrated effectiveness, stability, and reliability in handling high-dimensional text classification tasks, including large-scale datasets (Dzaky Candy Fahrezy & Fandy Setyo Utomo, 2024).

3. METHOD

This study employs a structured framework to extract information from a dataset of public opinion texts contained in YouTube comments concerning electric vehicles (Luthfi Krisna Bayu & Wurijanto, 2024). The research methodology is systematically designed, with a primary focus on optimizing data preprocessing to improve the classification accuracy of the Support Vector Machine (SVM) algorithm (Sehabudin et al., 2025).

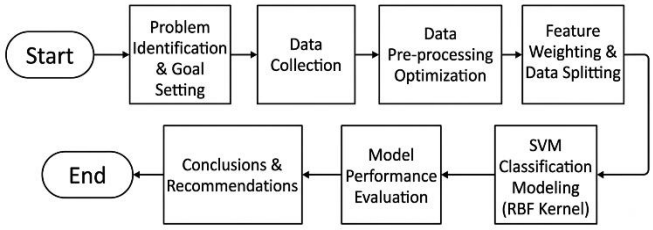


Figure 1. Flowchart of the Sentiment Analysis Research Stages

Data Preprocessing

Preprocessing optimization is an essential stage for cleaning unstructured textual data to generate high-quality features for the SVM predictive model [8]. The preprocessing stages are performed sequentially as follows:

1. Case Folding: All characters in the text dataset are converted to lowercase to ensure consistency in the alphabetic format (Pramana et al., 2025).
2. Cleaning: This stage removes noise from the data using Regular Expressions (Regex) to eliminate numbers, punctuation marks, special characters, and empty or NaN comments (Ningsih et al., 2024).
3. Tokenization: This step divides complete review sentences into individual word tokens, making them easier to process and enabling the calculation of word frequency distributions (Aryanti & Santoso, 2023).
4. Stopword Removal: This stage removes conjunctions and other common words that do not provide meaningful sentiment information using a standard stopword list (Farhani, 2024).
5. Normalization: The remaining text is normalized by converting abbreviations and informal language commonly used by Indonesian netizens into standardized words based on a customized conversion dictionary (Yulianto et al., 2025).
6. Stemming: The final stage removes prefixes, suffixes, and other affixes to reduce words to their root forms, thereby ensuring consistency in the vocabulary used for classification (Yulianto et al., 2025).

Feature Weighting and Data Splitting

Before the feature weighting stage, the preprocessed text data are first divided into training and testing sets using the train_test_split function with a 70:30 ratio. Stratified sampling is applied to maintain the proportion of the three sentiment classes (positive, negative, and neutral) across both subsets. Of the initial 1,517 comments, 2 comments were removed, for example, because they were empty or became duplicates after preprocessing. Consequently, 1,515 comments were used for the data-splitting process. The resulting dataset consisted of 1,060 training samples (70%) and 455 testing samples (30%) (Luthfi Krisna Bayu & Wurijanto, 2024).

After the data-splitting process, the text data are transformed into numerical vectors using the Term Frequency-Inverse Document Frequency (TF-IDF) weighting method. The structured word tokens are converted into numerical vectors so that they can be processed computationally by the machine learning model (Daniati et al., 2026). It is important to emphasize that the TF-IDF vocabulary is constructed exclusively from the training data, while the testing data are transformed using the vocabulary and inverse document frequency (IDF) weights learned from the training data, without contributing to the formation of a new vocabulary. This procedure is essential to ensure that the testing data genuinely represent unseen data that have not been accessed by the model in any form, including during the feature extraction stage. Therefore, the final evaluation results are protected from data leakage between the training and testing sets.

The transformation relies on TF-IDF calculations to assign a specific weight to each word based on the following formula:

$$Wt,d = TFt,d \times \log(\frac{N}{DFt}) \quad (1)$$

Where Wt,d denotes the weight of term t in documnet d, N represents the total number of documents, and DFt denotes the number of documents containing termt (Prasetyo et al., 2025).

Classification Using Support Vector Machine (SVM)

Because the dataset exhibits nonlinear characteristics, the decision function was implemented using the Radial Basis Function (RBF) kernel. The regularization parameter (C) and kernel parameter (γ) were not arbitrarily determined but optimized using GridSearchCV with a 5-fold cross-validation scheme. The search evaluated combinations of C={0.01,0.1,1,5,10,50,100,1000}, γ={scale,0.1,0.01,0.001,0.0001}, and class_weight = {None, balanced}, resulting in a total of 80 parameter combinations and 400 model-training iterations. The macro-F1 score was used as the selection metric because it provides a more representative evaluation than accuracy when dealing with imbalanced class distributions.

The grid search results indicated that the best parameter combination was C=10, γ=0.1, and class_weight = 'balanced', achieving a cross-validation macro-F1 score of 0.574. The value C=1000 was explicitly excluded because diagnostic analysis showed that increasing C from 10 to 100 improved the test accuracy by only 2%, while the training accuracy approached 100%, indicating a potential overfitting issue. Therefore, larger C values did not provide sufficient generalization benefits to justify the increased risk of overfitting.

To address class imbalance (see Table 1: Negative 57.3%, Positive 33.2%, and Neutral 9.5%), two approaches were compared: (1) using class_weight='balanced' in the SVM model and (2) applying Synthetic Minority Oversampling Technique (SMOTE) exclusively to the training data after the data-splitting process, ensuring that the testing set remained unchanged and consisted entirely of original data. SMOTE balanced the training data across the three classes to 608 samples per class, from the original distribution of 608:353:99. The C and γ parameter search was also repeated specifically for the SMOTE-based approach using imblearn.pipeline.Pipeline. This ensured that the resampling process was performed independently within each cross-validation fold and did not leak information into the validation data. The SMOTE-based approach produced an optimal combination of C=100 and γ=0.01, achieving a cross-validation macro-F1 score of 0.588. A comparison of the performance of the two approaches for handling class imbalance is presented in Table 6 (see Revised Section 4.7).

The SVM approach performs classification by identifying the optimal hyperplane in a high-dimensional feature space to clearly separate opinions among different sentiment classes (Sehabudin et al., 2025). The RBF kernel is used to transform the feature space so that nonlinear data can be optimally separated by the model (Ningsih et al., 2024). The mathematical formulation of the SVM decision function is given as follows:

$$f(\mathbf{x}) = sign\,(\sum_{i}^{n} \alpha iyiK(\mathbf{x}i,\mathbf{x}) + b) \quad (2)$$

The RBF kernel function used to transform the feature space is formulated as follows:

$$K(\mathbf{x}i,\mathbf{x}) = \exp(-\gamma||\mathbf{x}i - \mathbf{x}||^2 \quad (3)$$

Dimana :
x = Test data
xi = Support vector obtained from the training process
αi = Lagrange coefficient of the support vector
yi = Support vector class label (+1 or −1)
b = Bias or intercept
γ = RBF kernel parameter controlling the influence of each support vector
K(xi,x) = Similarity value between the test data and the support vector
sign\ \left(.\right) = Class decision function

Model Performance Evaluation

The reliability of the algorithm in classifying public opinions regarding electric vehicles was evaluated through mathematical calculations based on the Confusion Matrix (Prasetyo et al., 2025). The evaluation metrics were conventionally defined using the following formulas:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

Where:

- TP (True Positive) = Number of positive samples correctly classified as positive.
- TN (True Negative) = Number of negative samples correctly classified as negative.
- FP (False Positive) = Number of negative samples incorrectly classified as positive.
- FN (False Negative) = Number of positive samples incorrectly classified as negative.
- Accuracy is the most basic metric for measuring the proportion of total reviews whose sentiment is correctly classified by the model relative to the total number of reviews.
- Precision measures the accuracy of positive predictions among all texts classified by the model as having positive sentiment.
- Recall measures the sensitivity of the model in correctly identifying samples that actually belong to the positive sentiment class.
- F1-Score is the harmonic mean of precision and recall. This metric is used to achieve an optimal balance between the two measures.

In addition to these four metrics, this study also reports the macro-averaged F1-Score (F1-macro), weighted-averaged F1-Score (F1-weighted), and Cohen's Kappa. F1-macro calculates the average F1-Score across the three sentiment classes with equal weighting, thereby preventing the evaluation from being biased toward the majority class. This metric is

more representative than accuracy for imbalanced datasets such as the one used in this study. Cohen's Kappa measures the level of agreement between the model predictions and the actual labels while accounting for the possibility of agreement occurring by chance. Its value ranges from -1 (no agreement) to 1 (perfect agreement).

4. RESULT AND DISCUSSION

Results of Data Collection

The data collection process of public comments regarding electric vehicles in Indonesia successfully obtained 1,517 actual comment records from the YouTube platform. The collected data were then stored in a structured .csv format to facilitate processing in subsequent stages. The dataset contains diverse expressions of public opinion regarding environmentally friendly vehicle incentive policies. The size and diversity of the dataset provide an appropriate foundation for evaluating the classification performance of the machine learning algorithm.

Sentiment Labeling

The labeling stage classified the 1,517 comments into three sentiment polarity categories: negative, positive, and neutral. The labeling results showed that negative sentiment dominated the corpus, with a total of 869 comments (57.3%), followed by positive sentiment with 504 comments (33.2%), and neutral sentiment with 144 comments (9.5%). Accurate sentiment labeling is crucial because the resulting labels serve as the ground truth for the supervised learning model.

Table 1. Distribution of Sentiment Labels for Electric Vehicle Comments

No.	Sentiment Class	Number of Comments	Percentage (%)
1	Negative	869	57.3%
2	Positive	504	33.2%
3	Neutral	144	9.5%
Total	Overall	1,517	100.0%

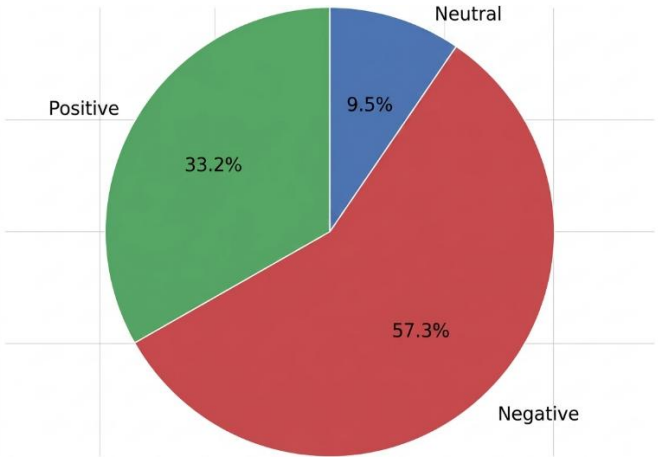


Figure 2. Percentage Distribution of Sentiments in Public Opinion

Case Folding

Case folding was applied as the initial text-cleaning step to convert all letters in the comment text into lowercase consistently. This standardization is necessary to prevent the system from treating identical words as different features due to variations in capitalization.

The resulting transformation ensures that all alphabetic characters follow a consistent standard format.

Table 2. Examples of Case Folding Results

No	Original Comment Text	Case-Folded Text
1	Model jelek Kwalitas buruk Harga Mahal	model jelek kwalitas buruk harga mahal
2	Iklan kudu Kencang	iklan kudu kencang
3	harga mahal croot	harga mahal croot
2	subsidi... tepat sasar!!	subsidi tepat sasar

Tokenization

Tokenization was performed to split the cleaned comment sentences into individual word tokens. This process facilitates the analysis of word occurrences within each review by the classification algorithm. The frequency of key terms identified through the tokenization process is visualized using a WordCloud.

Table 3. Examples of Data Tokenization Results

N	Cleaning Results	Tokenization Results
1	iklan kudu kencang	['iklan', 'kudu', 'kencang']
2	subsidi tepat sasar	['subsidi', 'tepat', 'sasar']

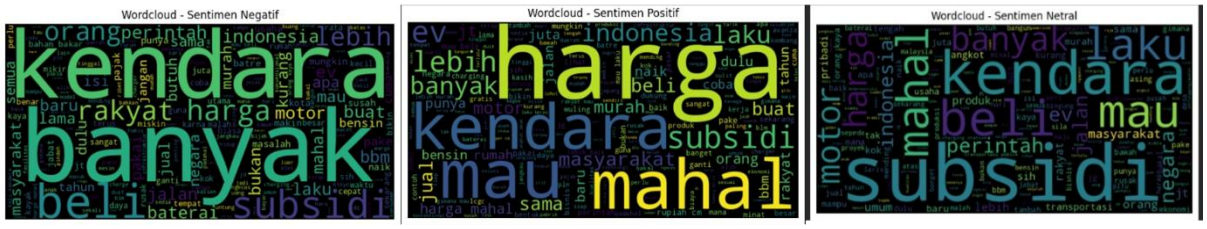


Figure 3. WordCloud Visualization of Sentiment Categories (Positive, Negative, and Neutral)

Feature Weighting (TF-IDF)

The preprocessed text data were transformed into a numerical matrix using the Term Frequency-Inverse Document Frequency (TF-IDF) weighting method. Because TF-IDF was fitted exclusively on the training data (Section 3.2), the resulting feature space consisted of 863 unique numerical features derived from the 1,060 training samples. This dimensionality is smaller than that obtained when TF-IDF is calculated from the entire corpus because the vocabulary was constructed solely from terms appearing in the 70% training data. The testing data (455 samples) were subsequently transformed using the same 863 features without generating a new vocabulary.

Results of SVM Classification and Model Evaluation

After hyperparameter optimization using GridSearchCV with a 5-fold cross-validation scheme, the optimal parameter combination was obtained as C=10, γ=0.1, and class_weight='balanced', with the highest cross-validation macro-F1 score of 0.574 on the original dataset without oversampling. These optimized parameters were then used as the basis for comparing the following four modeling schemes:

1. Baseline - SVM using the tuned C and γ values without any explicit class-imbalance handling.
2. SVM + Class Weighting - SVM using class_weight='balanced' to assign greater weights to minority classes during model training.

3. SVM + SMOTE - SVM trained on the oversampled training data using SMOTE while retaining the C and γ values obtained from hyperparameter tuning on the original data.
4. SVM + SMOTE (Retuned) - SVM trained on the SMOTE-resampled training data, with C and γ re-optimized specifically for this scheme. The optimal parameters were C=100 and γ=0.01, obtained using imblearn.pipeline.Pipeline. This approach ensured that the resampling process was independently performed within each cross-validation fold, thereby preventing data leakage into the validation data.

A summary of the hyperparameter search results for the two data schemes (original and SMOTE) is presented in Table 4.

Table 4. Summary of Hyperparameter Search Results (5-Fold Cross-Validation)

Data Scheme	C	y	class_weight	F1-macro CV
Original Data	10	0.1	balanced	0.574
SMOTE Data (Leakage-Free)	100	0.0.1	-	0.588

Among the four schemes described above, the performance of each model was subsequently evaluated on the unseen test set, which had not been accessed by the models during training or hyperparameter optimization. The evaluation results are summarized in Table 5.

Table 5. Performance Comparison of the Four Model Schemes on the Test Set

Model	Akurasi	Cohen's Kappa	F1-macro	F1-weighted
Baseliine	0.7225	0.4681	0.5256	0.6999
SVM + Classs Weighting	0.7077	0.4713	0.5833	0.7093
SVM + SMOTE	0.7275	0.4916	0.5977	0.7203
SVM + SMOTE (tuning ulang)	0.6967	0.4535	0.576	0.7001

Based on Table 5, the SVM + SMOTE model (C=10, γ=0.1, trained on the oversampled training data) demonstrated the best overall performance among the four evaluated schemes, achieving an F1-macro score of 0.598 and Cohen's Kappa of 0.492. Therefore, this model was selected as the final model in this study, achieving an accuracy of 72.75% on the test set. The detailed per-class classification report for the final model is presented in Table 6.

Table 6. Per-Class Classification Report of the Final Model (SVM + SMOTE)

Kelas	Precision	Recall	F1-Score	Support
Negatif (-1)	0.779	0.835	0.806	261
Netral (0)	0.344	0.256	0.293	43
Positif (1)	0.713	0.675	0.694	151
Macro avg	0.612	0.589	0.598	455
Weighted avg	0.716	0.727	0.72	455

Table 7. Confusion Matrix of SVM Classification Results

Aktual / Prediksi	Negatif	Netral	Positif
Negatif	218	14	29
Netral	20	11	12
Positif	42	7	102
Total Prediksi	**280**	**32**	**143**

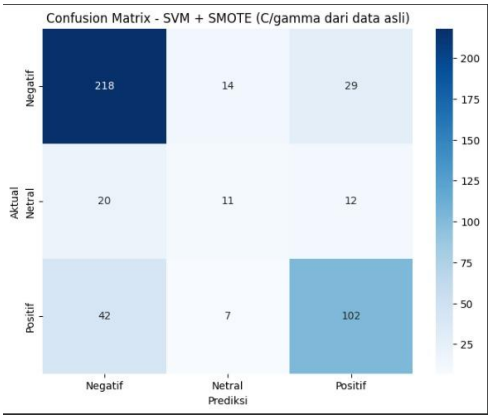


Figure 4. Heatmap of the Sentiment Classification Confusion Matrix

Compared with the results reported in the previous draft, in which the model achieved an accuracy of 78.1% (with TF-IDF calculated from the entire corpus before the data were split), the final model accuracy in this revised study is slightly lower at 72.75%. However, this decrease does not indicate that the model has become less effective. Rather, it is a consequence of the methodological improvements implemented in this study. In the previous draft, TF-IDF weighting was calculated using the entire dataset, including the test data, before the data were divided into training and testing sets. This resulted in data leakage, causing the evaluation score to be overly optimistic and not representative of the model's actual performance on unseen data. After the TF-IDF procedure was revised so that it was fitted exclusively on the training data (see Revised Section 3.2), the reported score became more reliable and accurately reflects the model's generalization capability when applied to new data.

The validity of this methodological improvement can also be observed from the close relationship between the cross-validation macro-F1 score (0.588) and the test-set macro-F1 score (0.576) for the retuned SMOTE scheme. The relatively small difference of approximately ±0.01 indicates that the model validation process was conducted without data leakage, thereby providing more reliable evaluation results.

Meanwhile, the Neutral class remained the lowest-performing class among the three sentiment categories, with an F1-score of only 0.293, substantially lower than the Negative class (0.806) and the Positive class (0.694). Nevertheless, a considerable improvement was observed compared with the Baseline scheme, which did not address class imbalance, where the F1-score for the Neutral class was only 0.075. This issue is discussed further in Sections 4.8 and 5.

Discussion

Based on the series of computational processes conducted in this study, public opinions regarding electric vehicles were successfully classified using a dataset of 1,517 comments collected from the YouTube platform. The sentiment analysis results indicate that public opinion was predominantly negative, accounting for 57.3% (869 comments), followed by positive sentiment at 33.2% (504 comments) and neutral sentiment at 9.5% (144 comments). The high proportion of negative sentiment is consistent with findings indicating that limited public literacy regarding electric vehicles may directly contribute to the emergence of negative perceptions and resistance toward their adoption. Overall, the initial implementation of the RBF-kernel-based SVM classification model, using a 70:30 training-to-testing data split, achieved an accuracy of 78.1%.

Methodological improvements were subsequently implemented in the hyperparameter selection and class-imbalance handling stages. The C and γ parameters of the RBF kernel were determined through a systematic search using GridSearchCV with a 5-fold cross-validation scheme rather than being selected arbitrarily. Thus, the selected parameter values were based on quantitative evidence and can be reproduced. Two class-imbalance handling

techniques were also directly compared, namely class weighting and SMOTE. Among the evaluated approaches, SMOTE demonstrated the best performance based on the F1-macro and Cohen's Kappa metrics. The selection of F1-macro as the primary evaluation metric, rather than relying solely on accuracy, was intended to ensure that the model's performance on the minority class (Neutral) was adequately considered rather than being dominated by its performance on the majority class (Negative).

The performance achieved by the SVM algorithm addresses the research gap identified in the previous sections. While many previous studies extracted short reviews from Twitter or X without placing particular emphasis on the quality and cleanliness of raw text, this study demonstrates the ability to handle the greater linguistic complexity of YouTube comments, which generally contain longer and more diverse sentence structures. Unlike previous approaches that employed standard algorithms such as Naïve Bayes, K-Nearest Neighbor (KNN), and Decision Tree, whose classification accuracies varied across studies, this research demonstrates that extensive preprocessing optimization, including the normalization of automotive-related slang and informal language, can be effectively combined with the nonlinear classification capability of the SVM algorithm. This comprehensive approach improves the stability and precision of the model in handling the linguistic complexity of informal user-generated content compared with approaches that focus solely on algorithmic parameter tuning.

5. CONCLUSION

Although the model achieved a relatively high overall prediction accuracy, evaluation using the confusion matrix revealed several limitations in the classification process. The most prominent limitation was observed in the Neutral sentiment class, for which the model achieved a precision of only 34.4% and a recall of 25.6%. The reduced performance in recognizing the Neutral class was strongly influenced by the class imbalance in the dataset, as the Neutral class represented the minority class.

This study evaluated two imbalance-mitigation techniques, namely class weighting and SMOTE. The results showed that SMOTE improved the performance of the Neutral class, with its F1-score increasing from 0.075 in the model without class-imbalance handling to 0.293. Nevertheless, the performance gap between the Neutral and majority classes was not completely eliminated. The class imbalance caused the model to become relatively biased toward the Negative class, which contained a substantially larger number of samples, making the classifier more sensitive to patterns associated with the majority class.

Future studies are therefore recommended to explore other approaches for addressing class imbalance, such as advanced cost-sensitive learning techniques or expanding the dataset with additional accurately labeled Neutral samples. These approaches may further improve the model's ability to recognize minority-class sentiments and reduce the performance gap among sentiment categories.

6. REFERENCES

Aryanti, P. G., & Santoso, I. (2023). Analisis Sentimen Pada Twitter Terhadap Mobil Listrik Menggunakan Algoritma Naive Bayes. IKRA-ITH Informatika : Jurnal Komputer Dan Informatika, 7(2), 133–137. https://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/2821

Daniati, E., Nugroho, A., Ristyawan, A., Utama, H., Studi, P., Informasi, S., Nusantara, U., Kedir, P., Gang, M., Kediri, K., & Timur, J. (2026). Perbandingan Kinerja Algoritma SVM , LSTM , dan Fine-tuned IndoBERT dalam Analisis Sentimen Opini Masyarakat Indonesia terhadap Mobil Listrik. IJCSR, 5(1), 1–13.

Dzaky Candy Fahrezy, & Fandy Setyo Utomo. (2024). Analisis Sentimen Opini Publik Tentang Kendaraan Listrik di Indonesia. Journal of Informatics and Interactive Technology, 1(2), 127–139. https://doi.org/10.63547/jiite.v1i2.37

Erfina, A., & Lestari, R. A. (2023). Sentiment Analysis of Electric Vehicles using the Naïve Bayes Algorithm. Sistemasi, 12(1), 178. https://doi.org/10.32520/stmsi.v12i1.2417

Farhani, A. (2024). Analisis Sentimen Terhadap Kendaraan Listrik di Indonesia Menggunakan Metode

Klasifikasi Naïve Bayes. JIMIK, 5(3), 2680–2690.

Luthfi Krisna Bayu, & Wurijanto, T. (2024). Analisis sentimen mobil listrik menggunakan metode Naïve Bayes Classifier. INFOTECH : Jurnal Informatika & Teknologi, 5(2), 328–335. https://doi.org/10.37373/infotech.v5i2.1465

Ningsih, W., Alfianda, B., Rahmaddeni, R., & Wulandari, D. (2024). Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 4(2), 556–562. https://doi.org/10.57152/malcom.v4i2.1253

Pramana, Y., Pradeva, Y., Widiartha, I. M., & Satwika, I. P. (2025). Analisis Sentimen Kebijakan Insentif Mobil Listrik Menggunakan TF-IDF dan Naïve Bayes. JNATIA, 4(1), 181–188.

Prasetyo, T., Waskita, A. A., & Taryo, T. (2025). Analisis Sentimen Pengguna Seputar Kendaraan Listrik di Twitter dengan Penerapan Algoritma Naïve Bayes, KNN, dan Decision Tree untuk Klasifikasi. Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan), VIII(2), 108–117.

Sehabudin, S., Irma Purnamasari, A., Bahtiar, A., & Wahyudin, E. (2025). Peningkatan Model Klasifikasi Sentimen Publik Di Youtube Cnbc Indonesia Terhadap Mobil Listrik Menggunakan Algoritma Support Vektor Machine. JATI (Jurnal Mahasiswa Teknik Informatika), 9(4), 6888–6896. https://doi.org/10.36040/jati.v9i4.14184

Sri Widagdo Adika Nuresa Qodri, Krisna Edi Nugroho, Fachruddin S, Akbar Rizky, & Nisrina P. (2023). Analisis Sentimen Mobil Listrik di Indonesia Menggunakan Long-Short Term Memory (LSTM). Jurnal Fasilkom, 13(3), 416–423.

Yulianto, M. F., Hana, F. M., & Prihandono, A. (2025). Penerapan Algoritma Naïve Bayes dalam Analisis Sentimen terhadap Mobil Listrik. Sainteks, 22(1), 109–115. https://doi.org/10.30595/sainteks.v22i1.26036