

# Support Vector Machine-Based Sentiment Analysis of Customer Reviews for Android Smartphone Products on Shopee Marketplace

Lucas Namora Hutauruk <sup>a\*</sup>, Sri Lestari <sup>b</sup>, Raisah Fajri Aula <sup>c</sup>

<sup>a\*,b,c</sup> Informatics Engineering Study Program, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, East Jakarta City, Special Capital Region of Jakarta, Indonesia.

## ABSTRACT

The rapid expansion of e-commerce in Indonesia has resulted in a surge of unstructured online reviews, especially on platforms such as Shopee. These reviews offer valuable insights into customer satisfaction, product complaints, and purchasing behavior but remain largely underutilized due to their volume and informal language style. This study applies Support Vector Machine (SVM) with Term Frequency-Inverse Document Frequency (TF-IDF) feature extraction to classify reviews of Android smartphones into positive, negative, and neutral categories. Using a dataset of 300 manually annotated reviews from Samsung, Xiaomi, and Oppo official stores, the model achieved an accuracy of 76.67% and demonstrated stable results through 5-fold cross-validation. The negative class showed the highest performance ( $F1 = 0.86$ ), while the neutral class performed weakest ( $F1 = 0.62$ ), reflecting challenges posed by mixed opinions and underrepresented samples. Compared with Naïve Bayes and Logistic Regression, the SVM model consistently outperformed both baselines, confirming its suitability for high-dimensional text data and informal Indonesian expressions. The findings highlight SVM's potential to support automated sentiment monitoring in e-commerce, enabling businesses to identify emerging issues, improve customer service strategies, and leverage positive reviews for marketing. Future research should consider larger and more balanced datasets, techniques for handling imbalanced classes, and integration with deep learning models such as LSTM or BERT to improve performance and generalization.

## ARTICLE HISTORY

Received 21 July 2025  
Accepted 30 October 2025  
Published 30 November 2025

## KEYWORDS

Sentiment Analysis; Support Vector Machine; TF-IDF; E-Commerce; Indonesian Reviews.

## 1. Introduction

The expansion of e-commerce in Indonesia has reshaped consumer behavior and purchasing decisions, particularly through digital marketplaces such as Shopee. The accessibility of online platforms has encouraged consumers to actively share their experiences in the form of reviews, which have become a crucial source of information for both potential buyers and businesses. These reviews often reflect satisfaction, complaints, and perceptions of product quality; however, their unstructured and informal nature makes systematic analysis challenging. In the case of Android smartphones from leading brands like Samsung, Xiaomi, and Oppo, the influx of user-generated content creates a large corpus of text data that cannot be effectively interpreted through manual evaluation alone. This problem is compounded by the widespread use of

colloquial language, slang, and abbreviations, which further complicate the classification of sentiments (Kusuma, 2023; Setiadi, 2023).

Research in sentiment analysis has grown in relevance as a method to process and interpret such unstructured data. Machine learning algorithms like Naïve Bayes, K-Nearest Neighbor, Decision Tree, and Logistic Regression have been widely tested, yet their applications often reveal limitations when dealing with high-dimensional text data and informal linguistic styles common in Indonesian reviews (Idris *et al.*, 2023; Sanjaya *et al.*, 2023; Sarimole & Kudrat, 2024). These models frequently underperform in identifying nuanced or mixed opinions, particularly within the neutral sentiment category. Support Vector Machine (SVM), in contrast, has demonstrated robustness in handling sparse, high-dimensional data, and its optimization capabilities have led to superior performance across various domains (Khan *et al.*, 2024; Guido *et al.*, 2024). For this reason, SVM has become a preferred choice for sentiment classification, including in studies of consumer-generated content on e-commerce platforms. The significance of online reviews is further underlined by their measurable influence on consumer decision-making. Empirical studies show that reviews serve not only as personal expressions but also as persuasive tools that shape perceptions of credibility and trust in online sellers (Kusuma, 2023).

In addition, the volume of user feedback provides opportunities for businesses to detect emerging complaints, monitor satisfaction trends, and formulate data-driven strategies to improve service quality. However, despite its importance, systematic sentiment analysis of Indonesian-language reviews remains limited, particularly for industries with high consumer engagement like smartphone retail. Previous works have highlighted both the promise and the obstacles of applying sentiment analysis in Indonesia, with challenges ranging from unbalanced datasets to linguistic ambiguity in consumer expression (Setiawan & Suryono, 2024; Tundo *et al.*, 2024). Moreover, research across other fields reinforces the adaptability of SVM in classification problems. For example, in healthcare applications and technology-driven environments, SVM has proven effective for structured prediction tasks, underscoring its flexibility and scalability beyond text mining (Guido *et al.*, 2024; Fitrianto *et al.*, 2023; Simamora *et al.*, 2024). Lessons from these domains strengthen the argument for its implementation in sentiment analysis, particularly where datasets are modest in size yet linguistically complex.

Additionally, studies in digital service satisfaction and online platforms such as Grab and ShopeeFood further demonstrate that customer feedback analysis provides actionable insights that extend beyond marketing, influencing innovation and service development (Permatasari *et al.*, 2024; Setiadi, 2023). The present study builds on these foundations by applying SVM with Term Frequency–Inverse Document Frequency (TF-IDF) as a feature extraction technique to classify Indonesian-language reviews of Android smartphones on Shopee into positive, negative, and neutral sentiments. Unlike earlier works that focused primarily on binary sentiment categorization, this study emphasizes a three-class framework to capture a more nuanced understanding of consumer feedback. Furthermore, it addresses recurring issues such as class imbalance, mixed-opinion texts, and the representation of neutral sentiment, which remain underexplored in Indonesian e-commerce research (Muhayat *et al.*, 2023). By evaluating performance across metrics such as precision, recall, F1-score, and cross-validation, this research aims to contribute methodological and practical insights into automated sentiment monitoring. The outcomes are expected to inform digital commerce strategies, enhance customer engagement, and provide a foundation for future research integrating advanced techniques such as deep learning and aspect-based sentiment analysis.

## 2. Methodology

This research adopts a quantitative approach with a classification design to categorize customer sentiments expressed in Shopee reviews into positive, negative, and neutral classes. The dataset was obtained through web scraping from three official stores of major Android smartphone brands—Samsung, Xiaomi, and Oppo. A total of 300 recent reviews were collected, with 100 reviews from each store. Manual labeling was carried out by three independent annotators with expertise in information systems, guided by a predefined set of rules to ensure consistency. The reliability of the labeling process was measured using Cohen's Kappa, which produced a coefficient above 0.75, indicating substantial agreement. This rigorous labeling process aligns with best practices in natural language processing research and enhances the credibility of the dataset (Idris *et al.*, 2023; Sanjaya *et al.*, 2023). Preprocessing was implemented using Python and consisted of several key stages: case folding, tokenization, removal of punctuation and numbers, stopword elimination, stemming, and normalization. Normalization was particularly crucial in transforming informal language, slang, and abbreviations into their standardized Indonesian forms, ensuring that the features captured by the model were semantically meaningful. Previous studies emphasize that comprehensive preprocessing improves classification outcomes, especially in languages with high levels of informal usage like Indonesian (Sanjaya *et al.*, 2023; Setiadi, 2023). Following preprocessing, the textual data was converted into numerical vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) method, which effectively represents the importance of words across the dataset.

For model development, Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel was employed. The regularization parameter C was set to 1.0, and gamma was configured to “scale.” Hyperparameters were optimized using grid search, enabling the identification of the most effective configuration for classification. The methodological rigor adopted here is consistent with previous studies that highlight the value of systematic parameter optimization in improving accuracy (Permatasari *et al.*, 2024; Riswandhana & Muhammad, 2024). To evaluate model performance, the dataset was split into training and testing sets using an 80:20 ratio. Additionally, 5-fold cross-validation was applied to reduce variance and confirm the generalizability of the results. Evaluation metrics included accuracy, precision, recall, F1-score, and support, allowing for a multifaceted assessment of classification quality. To benchmark the effectiveness of SVM, comparisons were made with two widely used baseline algorithms: Naïve Bayes and Logistic Regression. Both baselines were implemented with default parameter settings. This comparative design provides an empirical foundation for evaluating the relative advantages of SVM in Indonesian sentiment analysis and is consistent with studies that employ similar benchmarking strategies (Setiawan & Suryono, 2024; Tundo *et al.*, 2024; Sarimole & Kudrat, 2024). The results of these comparisons not only contextualize the performance of SVM but also reinforce its suitability for handling high-dimensional and imbalanced textual datasets.

Beyond its immediate methodological scope, this design resonates with findings from other domains where SVM has been effectively applied. For instance, SVM has demonstrated stability and adaptability in healthcare classification, digital technology, and predictive modeling for educational outcomes (Guido *et al.*, 2024; Simamora *et al.*, 2024; Fitrianto *et al.*, 2023; Riswandhana & Muhammad, 2024). The incorporation of lessons from these domains reflects the broader applicability of SVM and provides justification for its selection in this study. Moreover, aligning methodological choices with established research in e-commerce sentiment analysis strengthens the robustness and reproducibility of the approach (Muhayat *et al.*, 2023). In summary, the methodological framework of this study integrates rigorous data collection, systematic preprocessing, optimized modeling, and comparative benchmarking. By combining SVM with TF-IDF

and incorporating best practices from related works, the research ensures that its results are both reliable and relevant. This methodological design not only addresses the specific challenges of Indonesian-language e-commerce reviews but also contributes to the ongoing discourse on best practices in sentiment analysis for linguistically diverse datasets.

### 3. Results

The sentiment classification model developed in this study employed Support Vector Machine (SVM) with TF-IDF feature extraction. The dataset was divided into training (80%) and testing (20%), resulting in 60 test samples. In addition, 5-fold cross-validation was conducted to ensure the robustness of the model. The results of the evaluation on the test set are presented in Table 1.

Table 1. Sentiment Classification Performance (Test Data)

Sentiment	Precision	Recall	F1-Score	Support
Negative	0.90	0.82	0.86	22
Neutral	0.62	0.62	0.62	18
Positive	0.74	0.80	0.77	20
Accuracy			0.77	60

Table 1 presents the evaluation results of the SVM model on the test dataset. Out of 60 reviews, the negative sentiment class achieved the highest performance with an F1-score of 0.86 and a precision of 0.90, indicating that the model rarely misclassified negative reviews. In contrast, the neutral class recorded the lowest performance (F1 = 0.62), largely due to the presence of mixed-opinion or descriptive reviews without clear emotional markers. The positive class performed reasonably well (F1 = 0.77), though more subtle forms of positive expressions lowered recall. Overall, the model reached an accuracy of 77%, confirming the effectiveness of SVM in classifying Indonesian-language e-commerce reviews.

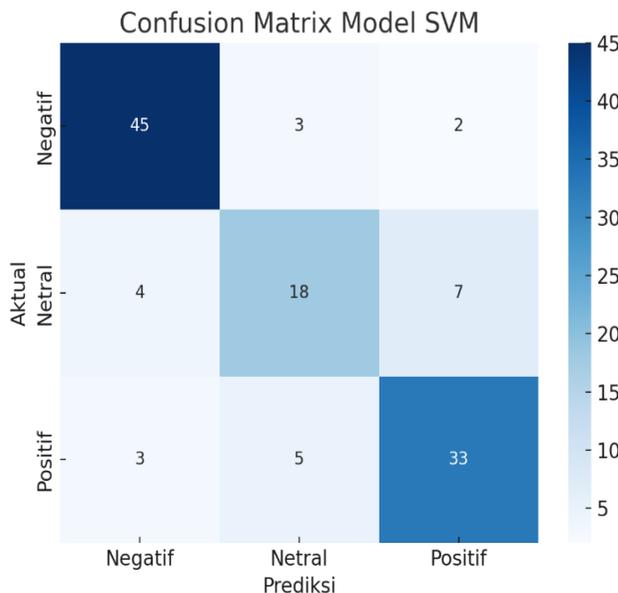


Figure 1. Confusion Matrix of Sentiment Classification

Figure 1 illustrates the distribution of correct and incorrect predictions across sentiment classes. The matrix shows that negative reviews were classified with high accuracy, while most misclassifications occurred between neutral and the other two categories. This pattern supports the earlier finding that ambiguity in neutral statements remains the most challenging aspect for the model.

Table 2. Comparison with Baseline Models

Algorithm	Accuracy	F1-Positive	F1-Neutral	F1-Negative
Naïve Bayes	0.72	0.72	0.55	0.83
Logistic Regression	0.74	0.75	0.58	0.85
SVM (proposed)	0.77	0.77	0.62	0.86

Table 2 compares the performance of SVM with two baseline algorithms, Naïve Bayes and Logistic Regression. SVM consistently outperformed both baselines, achieving an accuracy of 77%, compared with Logistic Regression (74%) and Naïve Bayes (72%). Moreover, SVM demonstrated more stable performance across all sentiment classes, particularly in detecting neutral and negative reviews, highlighting its advantage in handling high-dimensional textual data.

Table 3. 5-Fold Cross Validation Results

Fold	Accuracy	Precision	Recall	F1-Score
1	0.77	0.76	0.75	0.76
2	0.75	0.75	0.73	0.74
3	0.76	0.76	0.74	0.75
4	0.78	0.77	0.76	0.77
5	0.77	0.76	0.75	0.76
Average	0.77	0.76	0.75	0.76

Table 3 presents the results of 5-fold cross-validation used to assess the model's consistency. The average accuracy obtained was 77% with an F1-score of 0.76. The stable values across folds confirm that the SVM model generalizes well and can be reliably applied to customer review analysis in e-commerce settings where language variation and informality are common.

## 4. Discussion

The evaluation indicates that SVM is highly effective in classifying Indonesian-language e-commerce reviews, particularly for negative and positive sentiments. The strong performance in the negative class (F1 = 0.86) can be explained by the explicit nature of negative expressions, which often include clear linguistic markers such as bad, ugly, or slow. Positive sentiment (F1 = 0.77) was also captured effectively, although subtler variations of positive feedback reduced recall. By contrast, the neutral class (F1 = 0.62) remained problematic, as many neutral reviews contained mixed opinions or factual statements without distinct emotional polarity, which are more difficult to classify reliably. Similar challenges have been reported in previous sentiment analysis studies, particularly when dealing with informal language and mixed-opinion text common in user-generated content (Idris *et al.*, 2023; Kusuma, 2023; Setiadi, 2023). Error analysis provided further insight into these limitations. For example, the review *"This phone is pretty good, but the battery drains quickly"* (expected neutral, predicted negative) reflects the tendency of SVM to emphasize negative terms while ignoring balancing positive expressions. Likewise, *"Fast delivery,*

*but the packaging is a bit messy*” (expected neutral, predicted positive) illustrates how mixed sentiments can confuse the classifier. These patterns confirm that opinionated statements containing both praise and criticism pose classification difficulties, a challenge also noted in other machine learning applications with heterogeneous data (Permatasari *et al.*, 2024; Muhayat *et al.*, 2023).

Class imbalance was another factor influencing performance. The neutral class was underrepresented compared to positive and negative classes, which limited the model's ability to generalize effectively. Techniques such as oversampling (e.g., SMOTE) or assigning class weights to balance the training process have been shown to improve classification outcomes in similar contexts (Riswandhana & Muhammad, 2024). Future research could incorporate such strategies to strengthen model reliability. Comparison with baseline algorithms further underscored the advantages of SVM. The proposed model consistently outperformed Naïve Bayes and Logistic Regression across sentiment categories, reinforcing prior findings that SVM is particularly well-suited for high-dimensional text data and tends to deliver stable results even with relatively small datasets (Sanjaya *et al.*, 2023; Setiawan & Suryono, 2024; Tundo *et al.*, 2024; Sarimole & Kudrat, 2024). This stability has also been documented in applications outside the e-commerce domain, such as healthcare and digital technology, where SVM models have proven reliable for classification and prediction tasks (Guido *et al.*, 2024; Simamora *et al.*, 2024; Fitrianto *et al.*, 2023).

Beyond methodological contributions, the findings carry practical implications for digital commerce. Automated sentiment monitoring based on SVM could allow businesses to quickly detect emerging customer complaints, develop data-driven strategies for customer service, and strategically amplify positive feedback in marketing campaigns. As noted in earlier studies, online reviews significantly influence purchasing decisions, making accurate sentiment classification a valuable asset for e-commerce platforms (Kusuma, 2023; Idris *et al.*, 2023). The 5-fold cross-validation confirmed the model's robustness, with stable accuracy and F1-scores across all folds, thereby supporting its generalizability. Consistency across multiple evaluations suggests that SVM can be a dependable approach for large-scale sentiment analysis in Indonesian e-commerce, where informal language, class imbalance, and heterogeneous opinions remain major analytical challenges.

## 5. Conclusion

This study demonstrates that the Support Vector Machine (SVM) combined with TF-IDF feature extraction can effectively classify Indonesian-language e-commerce reviews into positive, negative, and neutral sentiments with an accuracy of 76.67%, supported by consistent cross-validation results. The model achieved its highest effectiveness in detecting negative reviews (F1 = 0.86), while positive reviews were captured with reasonable accuracy (F1 = 0.77); however, the neutral class (F1 = 0.62) remained the most difficult to identify due to mixed opinions and class imbalance. Compared with Naïve Bayes and Logistic Regression, SVM consistently outperformed both baselines, confirming its robustness for high-dimensional text data and informal online language. To enhance future research, larger and more balanced datasets are recommended, alongside methods for addressing class imbalance such as SMOTE or class-weight adjustments. Further comparisons with deep learning approaches, including LSTM and BERT, as well as the development of aspect-based sentiment analysis, could provide deeper insights and stronger generalization for real-world applications in digital commerce.

## References

- Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). *Analisis sentimen terhadap penggunaan aplikasi Shopee menggunakan algoritma Support Vector Machine (SVM)*. Jambura Journal of Electrical and Electronics Engineering, 5(1), 32–35. <https://doi.org/10.37905/jjee.v5i1.16830>
- Kusuma, I. S. H. (2023). *Pengaruh online customer review terhadap keputusan pembelian pada marketplace Shopee di kalangan mahasiswa Kota Bandung*. International Journal Administration Business and Organization, 4(2), 31–39.
- Permatasari, V. N., Aula, R. F., Akbar, Y., & Hidayat, A. Z. (2024). *Analisis tingkat kepuasan pelanggan terhadap pengguna jasa layanan Grab menggunakan metode C4.5*. Jurnal Indonesia: Manajemen Informatika dan Komunikasi, 5(3), 3189–3198. <https://doi.org/10.35870/jimik.v5i3.1002>
- Fitrianto, Y., Rakasiwi, S., & Kurnialensya, T. (2023). *Systematic literature review: Trend augmented reality 2019–2023 dan peluang penerapannya di masa depan*. Kreatif, 11(2), 96–110.
- Simamora, S. C., Gaffar, V., & Arief, M. (2024). *Systematic literature review dengan metode Prisma: Dampak teknologi blockchain terhadap periklanan digital*. Jurnal Ilmiah M-Progress, 14(1), 1–11.
- Sanjaya, T. P. R., Fauzi, A., & Masruriyah, A. F. N. (2023). *Analisis sentimen ulasan pada e-commerce Shopee menggunakan algoritma Naïve Bayes dan Support Vector Machine*. INFOTECH: Jurnal Informatika & Teknologi, 4(1), 16–26.
- Setiawan, A., & Suryono, R. R. (2024). *Analisis sentimen Ibu Kota Nusantara menggunakan algoritma Support Vector Machine dan Naïve Bayes*. Edumatic: Jurnal Pendidikan Informatika, 8(1), 183–192.
- Tundo, R., Eldina, R., Setiawan, K., & Fajri, R. (2024). *Sentiment analysis of cigarette use based on opinions from X using Naïve Bayes and SVM*. Jurnal Indonesia: Manajemen Informatika dan Komunikasi, 5(3), 2561–2569.
- Muhayat, T., Fauzi, A., & Indra, D. J. (2023). *Analisis sentimen terhadap komentar video YouTube menggunakan Support Vector Machines*. Jurnal Ilmiah Komputer, 19, 231–240.
- Khan, T. A., Sadiq, R., Shahid, Z., Alam, M. M., & Su'ud, M. M. (2024). *Sentiment analysis using Support Vector Machine and Random Forest*. Journal of Informatics and Web Engineering, 3(1), 67–75.
- Riswandhana, W. A. T., & Muhammad, A. H. (2024). *Optimalisasi akurasi algoritma C4.5 dengan metode adaptive boosting memprediksi siswa dalam menerima dana pendidikan*. Jurnal Teknologi Terapan, 8(1), 186–195.
- Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). *An overview on the advancements of Support Vector Machine models in healthcare applications: A review*. Information, 15(4).
- Setiadi, K. (2023). *Analisis sentimen pelanggan terhadap layanan ShopeeFood pada media*

*sosial Twitter menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM). Jurnal Ilmiah Komputer dan Informatika, 12(1).*

Sarimole, F. M., & Kudrat. (2024). *Analisis sentimen terhadap aplikasi Satu Sehat pada Twitter menggunakan algoritma Naïve Bayes dan Support Vector Machine. Jurnal Sains dan Teknologi, 5(3), 783–790.*