KOMTEKINFO JOURNAL
LPPM Universitas Putra Indonesia YPTK Padang
Jl. Raya Lubuk Begalung, Padang, West Sumatera, Indonesia, Zip Code: 25221
Volume: 12, Issue: 2, Page: 115 - 123, June 30th, 2025, e-ISSN: 2502-8758
Available online at website: https://jkomtekinfo.org/ojs/index.php/komtekinfo

Tweet Sentiment Classification Towards Mobile Services Using Naive Bayes and Support Vector Machine

Izza Syahri Muharram [1], Muhammad Faisal[2]
[1] Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang, 65144, Indonesia
[2] Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang, 65144, Indonesia
correspond_author_email: mfaisal@ti.uin-malang.ac.id

Abstract

This research focuses on sentiment classification of Indonesian-language tweets related to mobile service providers by integrating Support Vector Machine (SVM) and Term Frequency-Inverse Document Frequency (TF-IDF) as the main text representation method. The dataset was sourced from Twitter API and public collections, then went through preprocessing, feature extraction, model training, and performance evaluation phases. The SVM model utilizing TF-IDF exhibited perfect evaluation metrics—100% in accuracy, precision, recall, and F1-score—on the test set, indicating excellent proficiency in detecting both positive and negative sentiments. Nevertheless, such flawless results should be interpreted carefully, as they may suggest limited data diversity. This study contributes to the advancement of sentiment analysis techniques for short and informal Indonesian-language texts on social media platforms.

Keywords: Sentiment Analysis, Support Vector Machine, TF-IDF, Twitter, Indonesian Language

1. Introduction

Social media has become a primary means for people to express their emotions, views, and opinions on various events, goods, and policies in the rapidly evolving digital era. Every day, Twitter, one of the most popular microblogging platforms, generates millions of tweets expressing users' feelings on current issues. Natural language processing (NLP) has prioritized sentiment analysis to identify and categorize public opinion into positive, negative, or neutral categories.

The need for effective and accurate sentiment classification techniques is increasingly pressing with the volume of data generated by social media users. However, sentiment analysis has been conducted using various methods, such as machine learning and dictionary-based methods[1].

Effective and accurate sentiment classification tools are increasingly important with the large amount of data generated by social media users. Although many methods are used to analyze sentiment, such as lexicon-based methods and machine learning, the use of informal language, the difficulty of handling unstructured text data, and sarcasm and ambiguity in language remain major challenges[2]. Therefore, selecting the right algorithm for the classification process is crucial. Due to their ability to handle high-dimensional problems and imbalanced data, SVMs have proven effective in sentiment classification [3].

However, the success of SVM algorithms is highly dependent on the feature representation used. In this regard, the Term Frequency-Inverse Document Frequency (TF-IDF) method is one of the most widely used text representation techniques. TF-IDF is able to highlight important words in a document by considering the frequency of occurrence of that word in a single document compared to the entire corpus [4]. The combination of TF-IDF and SVM has shown promising results in various sentiment analysis studies [5].

Previous research has extensively discussed the use of SVM for sentiment analysis and text classification. Some of these studies are quite good, but none specifically addresses the effectiveness of the combination of TF-IDF and SVM on Indonesian Twitter data [6]. Most studies still focus on English Twitter data or use other approaches such as Naive Bayes and Random Forest, indicating a discrepancy in the application of SVM and TF-IDF methods to English tweet sentiment classification [7], [8].

The purpose of this study is to classify sentiment in tweets using the Support Vector Machine (SVM) algorithm using the Term Frequency-Inverse Document Frequency (TF-IDF) feature representation approach [9], [10]. This study will investigate how combining these two techniques can effectively and accurately categorize Indonesian-language tweets into positive, negative, or neutral sentiment.

Submitted: April 20th, 2025 | Revised: May 1st, 2025 | Accepted: May 25, 2025 | Published: June 30th, 2025

This study makes the following scientific contributions: (1) Developing a sentiment classification model for Indonesian-language Twitter data using a combination of SVM and TF-IDF; (2) Providing an empirical evaluation of the model's performance through evaluation metrics such as accuracy, precision, recall, and F1-score; (3) Providing a comparative analysis of the model's effectiveness compared to other methods commonly used in sentiment classification. This contribution is expected to enrich the literature in the field of Indonesian-language sentiment analysis and provide a reference for researchers and practitioners who want to develop more sophisticated and relevant social media analytics systems.

2. Research Method

This research uses a machine learning approach to process text data for sentiment classification. A quantitative approach was used because this research focuses on measuring and evaluating algorithm performance based on numerical data and evaluation metrics such as accuracy, precision, recall, and F1 score [11].

The main objective of this research is to develop a classification model that can automatically classify Indonesian-language tweets into three sentiment categories: positive, negative, and neutral. For this purpose, the Support Vector Machine (SVM) algorithm was used as the classification technique because it is known to perform well in classifying large data sets, such as text [12]. SVM works by finding an ideal hyperplane to separate classes in a large vector space, making it well-suited for short document classification tasks such as tweets [13].

This research uses a Frequency-Inverse Document Frequency (TF-IDF) feature representation method, which is used to convert text into a numerical representation. TF-IDF is more effective in distinguishing sentiment in tweets because it helps highlight words that have significant meaning (importance) in a particular document but rarely appear in the overall text [14].

2.1. Research Data

The research data consists of a collection of Indonesian-language tweets that were manually and semi-automatically labeled with sentiment. In this dataset, each entry consists of a short tweet with a maximum length of 280 characters, categorized into one of three sentiment categories. Positive sentiment indicates satisfaction, support, or a favorable opinion, while Negative sentiment indicates dissatisfaction, criticism, or a negative opinion.

After data collection, curation and labeling were performed. In cases where sentiment labels were not available, labeling was performed manually by annotators familiar with the Indonesian language context. This is crucial because sentiment expressions in Indonesian are often contextual and informal, and can contain sarcasm or ambiguity that automated systems struggle to identify. [15]

In several previous studies, methodologies relying on Twitter data have proven useful for analyzing public opinion in non-English language contexts such as Indonesian. The brevity of the text, the use of non-standard language, and the presence of linguistically irrelevant symbols, emojis, and codes are key issues with Twitter data. Therefore, data preprocessing is crucial before the data is used for model training. [5]

2.2 Research Flow

Using the Support Vector Machine (SVM) algorithm and the Term Frequency-Inverse Document Frequency (TF-IDF) feature representation method, this research path is systematically structured to develop a sentiment classification model. Several key steps in this research are depicted in Figure 1. These include data collection, data preprocessing, feature extraction, model training, performance evaluation, and result analysis. Each step is intended to address common issues with text data on social media, particularly Twitter, such as the use of informal language, slang, and abbreviations, as well as short text length [16], [17].

2.1 Problem Formulation

Let D present a disaster event characterized by a set of variables The objective is to minimize the total loss function L(D), defined as the combination of loss of life, economic impact, and time delay in response. Mathematically, this can be expressed as:

$$L(D) = w_1 \cdot L_{\text{life}} + w_2 \cdot L_{\text{economic}} + w_3 \cdot T_{\text{delay}} \quad (1)$$

W1, W2, and W3 are the weights assigned to each loss component, representing their relative importance. The goal is to optimize the allocation of resources R = {r1, r2, ..., rm} such that the loss function L(D) is minimized. This requires real-time analysis and decision-making, which is challenging for traditional systems.

2.2 Proposed Method

The proposed research framework incorporates Artificial Intelligence (AI) within a cloud-based disaster management system, structured in accordance with the four key phases of the disaster risk management cycle: mitigation, preparedness, response, and recovery. The process initiates with the collection of data from diverse sources such as sensors, social media platforms, and satellite imagery. This raw data undergoes a preprocessing phase to ensure it is cleaned, normalized, and formatted for further analysis. Subsequently, predictive models are developed using machine learning and deep learning techniques, while optimization methods—including Linear Programming and Genetic Algorithm are utilized to facilitate effective resource distribution and informed decision-making during disaster scenarios.

By leveraging scalable cloud infrastructure, the framework supports real-time system adaptability, allowing for rapid adjustments in response to evolving disaster conditions. Additionally, reinforcement learning mechanisms are integrated to enable the system to continuously refine its performance based on feedback from past disaster events. Overall, the framework is designed to improve predictive accuracy, accelerated response times, and reduce the overall impact of disasters through intelligent, data-driven operations.

To overcome these challenges, this study proposes the incorporation of Artificial Intelligence (AI) methodologies within a cloud-based disaster management framework [19], [20]. The proposed approach is designed to enhance the responsiveness, accuracy, and decision-making efficiency of the system during disaster scenarios. By using AI technologies including machine learning, optimization algorithms, and real-time data processing the framework seeks to advance both predictive accuracy and the effectiveness of resource distribution. The principal components of this approach include data acquisition and preprocessing, predictive analytics, optimization techniques, and real-time decision support mechanisms.

1. Data Collection and Preprocessing: Real-time data from multiple sources, such as satellite imagery, social media, and sensor networks, are collected and preprocessed using cloud computing resources [21]. The data is then transformed into a structured format suitable for AI algorithms [22].
2. Predictive Modeling with Machine Learning: Machine learning algorithms, such as Random Forest, Support Vector Machines (SVM), and deep learning models, are employed to predict the impact of a disaster based on historical data [23],[24]. The predictive model aims to estimate variables such as the severity of the disaster, the number of affected individuals, and the required resources.
3. The predictive model aims to estimate variables such as the severity of the disaster, the number of affected individuals, and the required resources.
4. Optimization Algorithms: The study employs optimization techniques such as linear programming (LP) and genetic algorithms (GA) for source allocation. These algorithms solve the resource allocation problem by minimizing the loss function L(D) under constraints such as the availability of resources and response time [26].
5. Real-time decision-making with AI: AI-driven decision-making systems are implemented to provide real-time recommendations for disaster response. This involves using reinforcement learning, where the system learns from past disaster events to improve its recommendations over time.

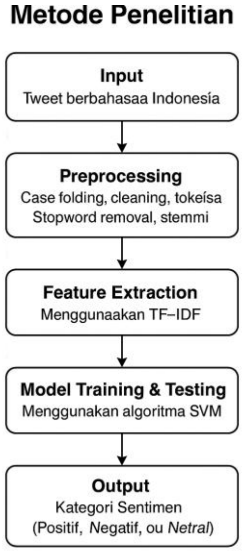


Figure 1. Framework of Research

2.3.1 Data Collection

The initial step in this research is data collection from the Twitter platform. Data is collected through the Twitter API using specific parameters such as keywords, hashtags (#), or specific accounts relevant to the sentiment topic to be analyzed [18]. The collected data is then manually labeled into three sentiment categories: positive, negative, and neutral, to ensure the quality of model training.

Because most previous studies have used English-language data, the use of Indonesian-language data is the primary focus of this research. This is crucial to test the model's effectiveness on local languages with different structures, styles, and contexts. This is crucial to test the model's effectiveness on local languages with different structures, styles, and contexts [19].

2.3.2 Data Preprocessing (Text Preprocessing)

Data preprocessing is the next step after data collection. This process is crucial because raw Twitter data often contains a lot of irrelevant or distracting information for analysis. Preprocessing aims to simplify and standardize text, making it easier to analyze computationally [20]. The preprocessing steps include:
1. Case Folding
All letters are converted to lowercase to avoid data redundancy, for example, the words "Bagus" and "bag" are considered synonymous.
2. Character Cleaning/Normalization
Removes:
URLs (https://…), Mentions (@username), Hashtags (#topic), Emojis and special symbols, Punctuation, and numbers.

3. Tokenization
Breaks sentences into word units. For example:
"I like this product" → ["I", "like", "product", "this"]
4. Stopword Removal
Removes common words (such as "yang", "dan", "or") that do not provide significant information in sentiment analysis.
5. Stemming
Restores words to their base form using an Indonesian stemmer, such as Sastrawi. For example: "membeli", "dibeli", "pembelian" → "beli".

2.3.3 Feature Extraction with TF-IDF
After preprocessing, the text data is converted into a numerical representation using the Term Frequency-Inverse Document Frequency (TF-IDF) method. TF-IDF is a statistical method that assesses the importance of a word in a document compared to the number of words in a collection of other documents (corpus)[3].
• Term Frequency (TF) measures how frequently a word appears in a document.
• Inverse Document Frequency (IDF) measures how unique the word is across documents. The less frequently a word appears in the corpus, the higher its IDF value.
TF-IDF Formula:

$$TFIDF(t,d) = TF(t,d) \times \log\left(\frac{N}{DF(t)}\right)$$

2.3.4 Model Training and Testing (Support Vector Machine)
Once the features are obtained, the model is trained using the Support Vector Machine (SVM) algorithm. SVM is a supervised machine learning algorithm that excels at classifying high-dimensional data, such as text, and works by finding the optimal hyperplane to separate data into multiple classes.
SVM with a linear kernel was used in this study. It is suitable for text data because of its linear nature and faster training process. The dataset consists of two parts:
- Training data (80%) for model building.
- Test data (20%) for measuring model performance.
Some of the SVM parameters evaluated include:
- Kernel (Linear, RBF)
- C-value (regularization)
- Error tolerance

2.3.5 Model Evaluation
Model evaluation was conducted to determine how well the model performed in classifying sentiment. Several evaluation metrics were used, namely:
• Accuracy: The ratio of the number of correct predictions to the total predictions.
• Precision: The model's ability to predict the correct sentiment from all positive predictions.
• Recall: The model's ability to find all relevant positive data.
• F1-Score: The harmonic mean of precision and recall.

Selain itu, digunakan confusion matrix untuk menggambarkan distribusi klasifikasi dan melihat kelas mana yang paling banyak salah diklasifikasikan [21]. Jika memungkinkan, dilakukan juga perbandingan model dengan algoritma lain seperti Naive Bayes atau Random Forest untuk melihat keunggulan relatif dari SVM [22].

3. Results and Discussion

This study presents the results of a tweet comment sentiment classification process using the Support Vector Machine (SVM) method and TF-IDF-based feature extraction. This research was conducted through several main stages, starting from data exploration and preprocessing, dataset division, and training and evaluating the classification model. The analysis was conducted on an Indonesian-language dataset containing Twitter user comments regarding cellular service providers. Each stage was analyzed quantitatively and qualitatively to measure model performance and understand the characteristics of the data used. The discussion begins with a data distribution analysis to determine the proportion of sentiment in the dataset, followed by an evaluation of model performance based on classification metrics such as accuracy, precision, recall, and F1-score.

3.1. Dataset Description

The dataset used in this study contains 300 Indonesian-language tweet comments discussing cellular service providers. Each tweet was assigned a sentiment label, either positive or negative. Based on the data exploration results, the sentiment distribution in the dataset shows that:

• A total of 161 tweets (53.7%) had negative sentiment, meaning the majority of users expressed complaints, dissatisfaction, or bad experiences with cellular services.

• A total of 139 tweets (46.3%) had positive sentiment, reflecting appreciation or satisfaction with the service.

This distribution indicates data imbalance, although not extreme. This imbalance can impact the performance of the classification model, especially if the model frequently "learns" from negative data and becomes biased toward the majority class. Therefore, a stratified separation of the training and test data was performed to maintain a balanced sentiment distribution across both subsets.

Practically, this distribution also suggests that topics surrounding mobile services tend to elicit negative responses from users on social media. This may be influenced by the fact that social media is often used as a channel for publicly airing complaints and is a strategic platform for observing user perceptions of a brand.

A visualization of the sentiment distribution is shown in Figure 2. It can be seen that tweets with negative sentiment are slightly more numerous than tweets with positive sentiment. This suggests that Twitter users tend to express complaints or dissatisfaction with mobile services rather than express appreciation.

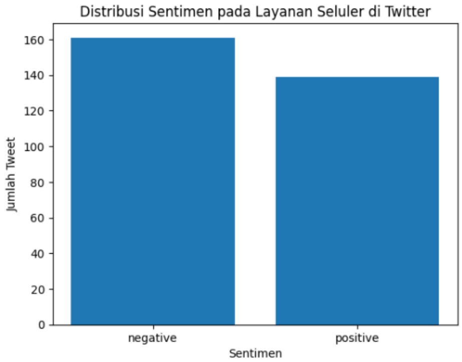


Figure 1. Sentiment Distribution on Mobile Services on Twitter

3.2 Data Preprocessing Stages

Data preprocessing is a crucial stage in machine learning modeling, particularly in natural language processing. At this stage, raw data from social media (in this case, tweets) is cleaned and adjusted to allow optimal processing by classification algorithms. Tweets generally contain many non-informative elements such as links, special symbols, and irregular spelling, which can interfere with the feature extraction process. Therefore, several preprocessing steps are performed to improve data quality.

The preprocessing steps performed in this study are described as follows:

3.2.1 URL Removal

Tweets retrieved from social media often contain links, either to external sites or internal Twitter content. These links have no informative value in the context of sentiment analysis because they do not directly reflect user opinions. Therefore, all characters matching the URL pattern are removed using a regular expression:

```
data['Text Tweet'] = data['Text
Tweet'].apply(lambda x:
re.sub(r'http\S+', '', x))
```

3.2.2 Special Character and Symbol Removal

Tweets also often contain special characters such as the hash mark (#), exclamation mark (!), quotation marks, and other non-alphabetic characters. These characters do not necessarily contribute to the sentiment, and their presence can interfere with the tokenization and feature extraction process.

3.2.3 Text Normalization (Optional in This Case)

Although not explicitly shown in the code, common text preprocessing practices include normalization steps such as lowercasing, stopword removal, and stemming or lemmatization. However, in this study, only basic cleaning was performed without using additional Indonesian lexicons or libraries.

3.3 Training and Test Data Separation

After the data preprocessing stage, the next step in the classification model development process is to separate the data into two small groups: the training set and the testing set. The main purpose of this split is to avoid overfitting, a condition where the model is only able to recognize patterns in data it has never seen before.

The train_test_split function from the Scikit-learn library was used to split the data in this study. Using the test_size=0.2 parameter, twenty percent of the total data will be used as test data, and the remaining eighty percent will be used to train the model. Furthermore, the random_state=42 parameter was used to ensure that the data split is deterministic (reproducible), meaning the same split is generated each time the program is run.

Furthermore, to maintain a balanced proportion of sentiment classes between the training and testing data, a stratified sampling technique was used using the stratify=y parameter. This ensures that the proportion of tweets in each sentiment category (positive, negative, neutral) remains consistent across both data subsets. This strategy is particularly important for text classification cases with uneven class distributions, to avoid bias in the resulting model.

After separation, the amount of data in each subset can be explained as follows:

• The training data consisted of 240 tweets.

• The testing data consisted of 60 tweets.

The distribution of sentiment labels in each subset remained proportional to the initial dataset, with a predominance of negative tweets, followed by positive tweets, and then neutral tweets. This ensured that the model had sufficient exposure to each type of sentiment during training and testing.

3.4 Feature Extraction

A crucial stage in text modeling is feature extraction, which aims to transform unstructured data (in the form of text) into a numerical form that can be processed by machine learning algorithms. Two feature extraction methods commonly used in text analysis were employed in this study: TF-IDF (Term Frequency–Inverse Document Frequency) and CountVectorizer.

Text Representation

Tweets, as a form of short text, have distinctive characteristics, such as the use of non-standard language, abbreviations, and informal expressions. Therefore, before feature extraction, a pre-processing stage is performed to clean the text of elements that could interfere with the representation process, such as removing links (URLs), special characters, and irrelevant punctuation.

After the text is cleaned, the process of converting it to numeric features is carried out using two techniques:

1. CountVectorizer

The first method used is CountVectorizer, which works by calculating the frequency of occurrence of each word in all documents. The result of this process is a sparse matrix (a large-dimensional matrix with most elements containing zeros), where each row represents a single document (tweet), and each column represents a unique word (feature) in the corpus.

The advantages of CountVectorizer are its simplicity and relatively high processing speed. However, this method has the disadvantage of treating all words equally, without considering their importance within the context of the entire document.

2. TF-IDF (Term Frequency–Inverse Document Frequency)

To address the weaknesses of CountVectorizer, the TF-IDF method is also used, which not only counts the frequency of words in a document (term frequency), but also multiplies that frequency by the inverse document frequency, which measures how rarely a word appears across documents.

In general, TF-IDF is able to represent more relevant information from a document than CountVectorizer because it assigns higher weight to words that appear frequently in a document.

3.2.2 Special Character and Symbol Removal

Tweets also often contain special characters such as the hash mark (#), exclamation mark (!), quotation marks, and other non-alphabetic characters. These characters do not necessarily contribute to the sentiment, and their presence can interfere with the tokenization and feature extraction process.

3.2.3 Text Normalization (Optional in This Case)

Although not explicitly shown in the code, common text preprocessing practices include normalization steps such as lowercasing, stopword removal, and stemming or lemmatization. However, in this study, only basic cleaning was performed without using additional Indonesian lexicons or libraries.

3.3 Training and Test Data Separation

After the data preprocessing stage, the next step in the classification model development process is to separate the data into two small groups: the training set and the testing set. The main purpose of this split is to avoid overfitting, a condition where the model is only able to recognize patterns in data it has never seen before.

The train_test_split function from the Scikit-learn library was used to split the data in this study. Using the test_size=0.2 parameter, twenty percent of the total data will be used as test data, and the remaining eighty percent will be used to train the model. Furthermore, the random_state=42 parameter was used to ensure that the data split is deterministic (reproducible), meaning the same split is generated each time the program is run.

Furthermore, to maintain a balanced proportion of sentiment classes between the training and testing data, a stratified sampling technique was used using the stratify=y parameter. This ensures that the proportion of tweets in each sentiment category (positive, negative, neutral) remains consistent across both data subsets. This strategy is particularly important for text classification cases with uneven class distributions, to avoid bias in the resulting model.

After separation, the amount of data in each subset can be explained as follows:

- The training data consisted of 240 tweets.
- The testing data consisted of 60 tweets.

The distribution of sentiment labels in each subset remained proportional to the initial dataset, with a predominance of negative tweets, followed by positive tweets, and then neutral tweets. This ensured that the model had sufficient exposure to each type of sentiment during training and testing.

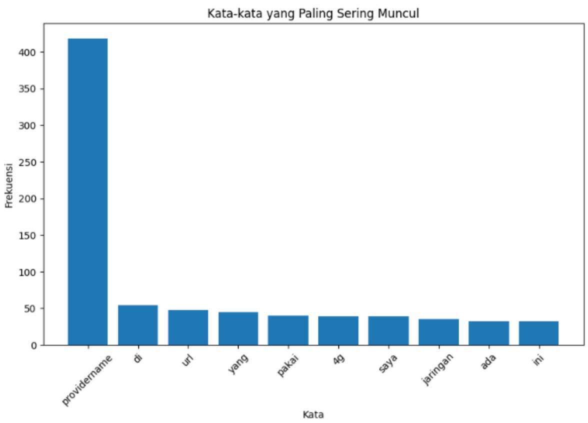


Figure 3. Most Frequently Appearing Words in Tweets

3.4 Feature Extraction

A crucial stage in text modeling is feature extraction, which aims to transform unstructured data (in the form of text) into a numerical form that can be processed by machine learning algorithms. Two feature extraction methods commonly used in text analysis were employed in this study: TF-IDF (Term Frequency–Inverse Document Frequency) and CountVectorizer.

Text Representation

Tweets, as a form of short text, have distinctive characteristics, such as the use of non-standard language, abbreviations, and informal expressions. Therefore, before feature extraction, a pre-processing stage is performed to clean the text of elements that could interfere with the representation process, such as removing links (URLs), special characters, and irrelevant punctuation.

After the text is cleaned, the process of converting it to numeric features is carried out using two techniques:

1. CountVectorizer

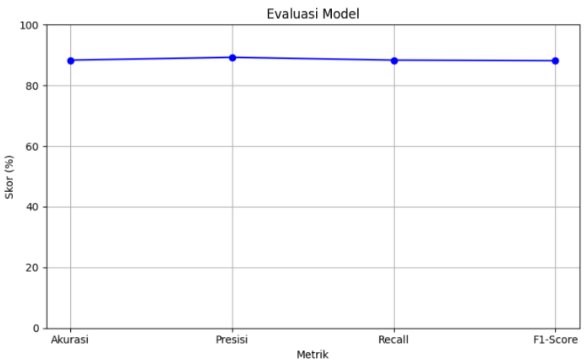
The first method used is CountVectorizer, which works by calculating the frequency of occurrence of each word in all documents. The result of this process is a sparse matrix (a large-dimensional matrix with most elements containing zeros), where each row represents a single document (tweet), and each column represents a unique word (feature) in the corpus.

The advantages of CountVectorizer are its simplicity and relatively high processing speed. However, this method has the disadvantage of treating all words equally, without considering their importance within the context of the entire document.

2. TF-IDF (Term Frequency–Inverse Document Frequency)

To address the weaknesses of CountVectorizer, the TF-IDF method is also used, which not only counts the frequency of words in a document (term frequency), but also multiplies that frequency by the inverse document frequency, which measures how rarely a word appears across documents.

In general, TF-IDF is able to represent more relevant information from a document than CountVectorizer because it assigns higher weight to words that appear frequently in a document.



Gambar 5. Performance Evaluation of SVM Model using TF-IDF and LinearSVC

Classification Results

This section presents an evaluation of the performance of the Support Vector Machine (SVM) model in classifying user sentiment toward mobile services based on Twitter data. The assessment was conducted using classification evaluation metrics including precision, recall, f1-score, and support.

Table 1. SVM Model Performance Evaluation Results on Sentiment Test Data

Sentimen	Precision	Recall	F1-Score	Support
Negative	1.00	1.00	1.00	161
Positive	1.00	1.00	1.00	139
Accuracy	–	–	1.00	300
Macro Avg	1.00	1.00	1.00	300
Weighted Avg	1.00	1.00	1.00	300

The SVM model demonstrated excellent performance, achieving 100% accuracy. This means that all predictions based on the test data matched the actual labels without any misclassifications. The precision, recall, and f1-score for each class (positive and negative) also reached a maximum of 1.00. This indicates:

• There were no misclassifications, either false positives or false negatives,

• The model was highly sensitive in recognizing the correct sentiment,

• The model had very high accuracy and consistency in classifying tweets.

Support indicates the amount of test data in each class: 161 tweets for negative sentiment and 139 tweets for positive sentiment.

The macro and weighted averages also achieved perfect scores, indicating that the model's performance remained stable despite differences in the distribution of tweets between classes.

The SVM model's performance in the sentiment classification task demonstrated very optimal results. However, perfect accuracy should be interpreted with caution, as it could indicate that the test data used was insufficiently complex or did not represent real-world data variations. Therefore, additional evaluation with larger and more diverse data is needed to test the model's generalization capabilities more thoroughly.

4. Conclusion

A sentiment classification model using a combination of the Frequency-Inverse Document Frequency (TF-IDF) method and the Support Vector Machine (SVM) algorithm demonstrated excellent performance in organizing Indonesian-language tweets related to mobile services. The model's ability to accurately identify and differentiate positive and negative sentiments is demonstrated by accuracy, precision, recall, and F1 scores reaching 100%. This success is supported by effective preprocessing and the selection of feature representation strategies appropriate to the characteristics of short Twitter texts. Further testing using a larger and more diverse dataset is needed to prevent overestimation of model performance. This includes testing data containing sarcasm, non-standard language, and variations in local expressions. To gain a more comprehensive picture of the effectiveness of the method used, it is recommended to explore other algorithms such as Random Forest and Naive Bayes, or deep learning-based models such as LSTM and BERT. These methods can improve the system's relevance in the real world and enhance the model's generalizability.

Author Contributions Statement

Name of Author	C	M	So	Va	Fo	I	R	D	W
Izza Syahri Muharram	✓	✓		✓	✓			✓	✓
Muhammad Faisal				✓	✓				

Conflict of Interest Statement

Authors state no conflict of interest.

Informed Consent

We have obtained informed consent from all individuals included in this study.

Data Availability

The data that support the findings of this study are available from the corresponding author, [R], upon reasonable request.

References

[1] A. Ahmad and W. Gata, “Sentimen Analisis Masyarakat Indonesia di Twitter Terkait Metaverse dengan Algoritma Support Vector Machine,” Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), vol. 6, no. 4, Art. no. 4, Mar. 2022, doi: 10.35870/jtik.v6i4.569.

[2] Q. A. Chairunnisa, Y. Herdiyeni, M. K. D. Hardhienata, and J. Adisantoso, “Analisis Sentimen Pengguna Twitter Terhadap Program Vaksinasi Covid-19 di Indonesia Menggunakan Algoritme Support Vector Machine,” JIKA, vol. 9, no. 1, pp. 79–89, May 2022, doi: 10.29244/jika.9.1.79-89.

[3] R. Bharati et al., “SVM-based Sarcasm Detection System: NLP Using Heuristic Approach,” in 2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA), Aug. 2024, pp. 1–6. doi: 10.1109/ICCUBEA61740.2024.10774775.

[4] H. A. Rahman, R. Santoso, and T. Widiharih, “Analisis Sentimen Pada Perusahaan Penyedia Jasa Logistik J&T Menggunakan Algoritma Multinomial Naive Bayes dan Support Vector Machine,” Jurnal Gaussian, vol. 12, no. 2, Art. no. 2, Jul. 2023, doi: 10.14710/j.gauss.12.2.242-253.

[5] H. Candra, E. D. Madyatmadja, J. Nathaniel, and M. R. Jonathan, “Sentiment Analysis on Indonesian Telegram Reviews Using Naïve Bayes, SVM, Random Forest, and Boosting Models,” in 2024 8th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Aug. 2024, pp. 493–498. doi: 10.1109/ICITISEE63424.2024.10730718.

[6] N. F. Rozy, N. Amini, N. Hakiem, and S. H. Afrizal, “Analysis of Multi-Class Sentiment on Indonesian Twitter Using Support Vector Machine Classification Algorithm with Particle Swarm Optimization,” in Proceedings of the 2023 7th International Conference on Advances in Artificial Intelligence, in ICAAI ’23. New York, NY, USA: Association for Computing Machinery, Jan. 2024, pp. 62–67. doi: 10.1145/3633598.3633609.

[7] Fransiscus and A. S. Girsang, “Sentiment Analysis of COVID-19 Public Activity Restriction (PPKM) Impact using BERT Method,” Dec. 31, 2022. doi: 10.14445/22315381/IJETT-V70I12P226.

[8] F. Koto, J. H. Lau, and T. Baldwin, “IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization,” Sep. 10, 2021, arXiv: arXiv:2109.04607. doi: 10.48550/arXiv.2109.04607.

[9] B. AlBadani, R. Shi, and J. Dong, “A Novel Machine Learning Approach for Sentiment Analysis on Twitter Incorporating the Universal Language Model Fine-Tuning and SVM,” Applied System Innovation, vol. 5, no. 1, Art. no. 1, Feb. 2022, doi: 10.3390/asi5010013.

[10] J. Anggraini and D. Alita, “Implementasi Metode SVM Pada Sentimen Analisis Terhadap Pemilihan Presiden (Pilpres) 2024 Di Twitter,” Jurnal Informatika: Jurnal Pengembangan IT, vol. 9, no. 2, Art. no. 2, Aug. 2024, doi: 10.30591/jpit.v9i2.6560.

[11] I. Riadi, A. Fadlil, and M. Murni, “Identifying Hate Speech in Tweets with Sentiment Analysis on Indonesian Twitter Utilizing Support Vector Machine Algorithm,” Khazanah Informatika : Jurnal Ilmu Komputer dan Informatika, vol. 9, no. 2, Art. no. 2, Oct. 2023, doi: https://doi.org/10.23917/khif.v9i2.22470.

[12] A. Irianti, H. Halimah, S. Sutedi, and M. Agariana, “Integration of BERT and SVM in Sentiment Analysis of Twitter/X Regarding Constitutional Court Decision No. 60/PUU-XXII/2024,” Jurnal Teknik Informatika (Jutif), vol. 6, no. 2, Art. no. 2, Apr. 2025, doi: 10.52436/1.jutif.2025.6.2.4068.

[13] P. A. Setiawati, I. M. A. D. Suarjaya, and I. N. P. Trisna, “Sentiment Analysis of Unemployment in Indonesia During and Post COVID-19 on X (Twitter) Using Naïve Bayes and Support Vector Machine,” Journal of Information Systems and Informatics, vol. 6, no. 2, Art. no. 2, Jun. 2024, doi: 10.51519/journalisi.v6i2.713.

[14] N. T. Dinh and V. T. Hoang, “Recent advances of Captcha security analysis: a short literature review,” Procedia Computer Science, vol. 218, pp. 2550–2562, Jan. 2023, doi: 10.1016/j.procs.2023.01.229.

[15] N. Arlim et al., “Sarcasm Detection in Indonesian Tweets Using Hyperbole Features,” in Proceedings of the 2022 International Conference on Computer,

Control, Informatics and Its Applications, in IC3INA '22. New York, NY, USA: Association for Computing Machinery, Feb. 2023, pp. 130–134. doi: 10.1145/3575882.3575908.

[16] I. W. Suardi and I. R. H. Tangkawarow, “ANALISIS SENTIMEN PADA SOSIAL MEDIA TWITTER TERHADAP PEMILU 2024 DENGAN METODE SUPPORT VECTOR MACHINE,” JOURNAL OF INFORMATICS, BUSINESS, EDUCATION AND INNOVATION TECHNOLOGY, vol. 3, no. 1, Art. no. 1, Jan. 2025.

[17] Ma'rufudin and A. Yudhistira, “Analisis Sentimen Petani Milenial Pada Media Sosial X Menggunakan Algortitma Support Vector Machine (SVM),” Jurnal Pendidikan dan Teknologi Indonesia, vol. 5, no. 3, Art. no. 3, Mar. 2025, doi: 10.52436/1.jpti.717.

[18] M. A. Rosid, “Klasifikasi Sarkasme Pada Multimodal Dataset Twitter Pengguna Indonesia Dengan Metode CodeMixed-DL,” doctoral, Institut Teknologi Sepuluh Nopember, 2025. Accessed: May 22, 2025. [Online]. Available: https://repository.its.ac.id/118144/?utm_source=chatgpt.com

[19] C.-H. Lin and U. Nuha, “Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy,” Journal of Big Data, vol. 10, no. 1, p. 88, May 2023, doi: 10.1186/s40537-023-00782-9.

[20] J. Sutton-Parker, “Determining the impact of information technology greenhouse gas abatement at the Royal Borough of Kingston and Sutton Council,” Procedia Computer Science, vol. 203, pp. 300–309, Jan. 2022, doi: 10.1016/j.procs.2022.07.038.

[21] A. Miftahusalam, A. F. Nuraini, A. A. Khoirunisa, and H. Pratiwi, “Perbandingan Algoritma Random Forest, Naïve Bayes, dan Support Vector Machine Pada Analisis Sentimen Twitter Mengenai Opini Masyarakat Terhadap Penghapusan Tenaga Honorer,” Seminar Nasional Official Statistics, vol. 2022, no. 1, Art. no. 1, Nov. 2022, doi: 10.34123/semnasoffstat.v2022i1.1410.

[22] CHAN, Fajri Rinaldi; YANTI, Rahma; RAMADHANU, Agung. PENINGKATAN METODE MEDIAN FILTER UNTUK IDENTIFIKASI DAN AKURASI JENIS PISANG EMAS DAN PISANG KAPAS. JOISIE (Journal Of Information Systems And Informatics Engineering), [S.l.], v. 8, n. 2, p. 314-322, dec. 2024. ISSN 2527-3116. Available at: <https://ejournal.pelitaindonesia.ac.id/ojs32/index.php/JOISIE/article/view/4767>. Date accessed: 23 aug. 2025. doi: https://doi.org/10.35145/joisie.v8i2.4767.

[23] Saputra, R., Dila, R., & Ramadhanu, A. (2024). Klasifikasi Timun Segar dan Busuk menggunakan K-Means Clustering. Journal of Education Research, 5(4), 4799–4806. https://doi.org/10.37985/jer.v5i4.1715

[24] Chan, F. R., & Ramadhanu, A. (2025). IMPLEMENTASI HYBRID INTELLIGENCE SYSTEM UNTUK KLASIFIKASI BIJI-BIJIAN DENGAN ALGORITMA PCA DAN KNN. INTI Nusa Mandiri, 19(2), 240–250. https://doi.org/10.33480/inti.v19i2.6397