



Social Media Sentiment Analyst on Danantara Using Support Vector Machine with TF-IDF Weighting

Faizal Risqi Andika¹, Muhamad Arielfa Rizky², Kharisma Ummi Khairunnisa³,
Hana Amirah⁴, Abdillah Alfazhari⁵, Hana Aulia⁶

^{1,2}Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

³Department of Dirasat Islamiyah, Faculty of Syari'ah Islamiyah,
University of Al Azhar, Egypt

⁴Department of Business Administration, Faculty of Business,
Karabuk University, Turkey

⁵Department of Ushul Fiqh, Faculty of Sharia, Mu'tah University, Jordan

⁶Bachelor Degree of Arabic Language and Multimedia, Faculty of Islamic Civilization Studies,
Selangor Islamic University, Malaysia

E-Mail: ¹12350311779@students.uin-suska.id, ²12350313074@students.uin-suska.id,
³ummikharisma07@gmail.com, ⁴hanaamirah2704@gmail.com,
⁵abdillahaal2006@gmail.com, ⁶hanaaulia29@gmail.com

Received Jul 13th 2026; Revised Aug 14th 2026; Accepted Sep 18th 2026; Available Online Sep 30th 2026

Corresponding Author: Faizal Risqi Andika

Copyright © 2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

The establishment of Badan Pengelola Investasi Daya Anagata Nusantara (Danantara) has generated extensive public discussion on social media, reflecting diverse perceptions regarding its governance, transparency, and potential contribution to national development. Despite growing public discourse, empirical studies examining public sentiment toward this newly established institution remain limited. This study addresses that gap by analyzing public sentiment toward Danantara using a Support Vector Machine (SVM) classifier with Term Frequency–Inverse Document Frequency (TF-IDF) feature weighting, applied to a multi-platform dataset spanning X (Twitter), Instagram, and YouTube an approach that captures a broader and more representative range of public opinion than single-platform studies. A dataset of 3,000 comments was collected and underwent preprocessing, including case folding, cleaning, normalization, tokenization, stopword removal, and stemming, followed by manual labeling into three sentiment classes: positive, negative, and neutral. The proposed LinearSVC model achieved an accuracy of 84.67%, a macro-average F1-score of 0.82, and a weighted-average F1-score of 0.85. Sentiment distribution analysis reveals that negative sentiment dominates public discourse (57.13%), followed by positive (24.37%) and neutral sentiments (18.50%). This study contributes both a novel empirical case being among the first sentiment analyses of public perception toward Danantara and evidence that SVM with TF-IDF offers an effective, robust approach for multiclass sentiment classification of Indonesian social media text, providing actionable insights for policymakers and institutional stakeholders.

Keywords: Danantara, Sentiment Analysis, Social Media, Support Vector Machine, TF-IDF

1. INTRODUCTION

The rapid growth of social media platforms, including X (formerly Twitter), Instagram, and YouTube, has provided Indonesian users with increasingly accessible spaces to express opinions and share information on various social, economic, and political issues. The vast amount of user-generated content on these platforms offers valuable opportunities to understand public sentiment in Indonesia [1]. However, the diversity of Indonesian language expressions, including informal language, slang, abbreviations, and regional linguistic variations, poses challenges for accurately identifying sentiment in social media texts [2]. Consequently, sentiment analysis using Natural Language Processing (NLP) and machine learning has emerged as an important approach for automatically classifying Indonesian-language opinions as positive, negative, or neutral, enabling more efficient and systematic analysis of public opinion at scale [3].

One of the most consequential national policy developments currently attracting public attention in Indonesia is the establishment of Badan Pengelola Investasi Daya Anagata Nusantara, commonly known as Danantara. Officially launched by President Prabowo Subianto, Danantara is a sovereign wealth fund designed



to optimize the assets of state-owned enterprises and accelerate national economic growth through strategic investment management [4]. As a newly established institution with far-reaching fiscal implications, Danantara has generated a broad spectrum of public responses, ranging from optimistic expectations regarding national development to critical concerns about governance transparency and financial risk. These concerns are consistent with studies showing that public trust in sovereign wealth funds is strongly influenced by governance legitimacy and transparency mechanisms [29], [30].

Sentiment analysis, a subfield of NLP, provides a systematic and automated framework for categorizing textual expressions into predefined polarity classes such as positive, negative, and neutral [5]. This technique has been applied extensively across diverse domains, including political discourse, product evaluation, and public policy monitoring [4]. In the context of government initiatives, sentiment analysis enables policymakers to gain empirical insights into citizens' perceptions and adapt communication strategies accordingly [3]. Recent studies also demonstrate that social media sentiment analysis has become an important instrument for monitoring public opinion and supporting evidence-based policy decisions [23], [24], [25]. Despite the expanding body of work on sentiment analysis applied to Indonesian policy issues, dedicated research examining public opinion toward Danantara remains relatively scarce.

Among the supervised machine learning algorithms commonly employed for text classification, the Support Vector Machine (SVM) has consistently demonstrated strong, reliable performance, particularly in high-dimensional feature spaces characteristic of textual data [5]. By identifying the optimal separating hyperplane in the feature space, SVM achieves robust generalization and is applicable to both binary and multiclass classification tasks [6]. Multiple empirical studies have confirmed the effectiveness of SVM for sentiment classification tasks across various social media contexts, reporting accuracy ranging from 80% to 90% depending on the preprocessing pipeline and feature representation employed [7].

For the purpose of transforming raw text into a numerical representation suitable for machine learning, this study employs Term Frequency–Inverse Document Frequency (TF-IDF) weighting. TF-IDF is a well-established statistical technique that assigns greater weight to terms that appear frequently within a given document but are relatively rare across the entire corpus, thereby capturing the discriminative power of individual words. When integrated with an SVM classifier, TF-IDF has been shown to yield effective, computationally efficient sentiment models across a wide range of text domains, including social media [8]. This combination is therefore well-suited to classifying opinionated Twitter data in Indonesian.

Building on the findings of previous studies, this research addresses the limited research on public sentiment toward Danantara, Indonesia's newly established sovereign wealth fund, by analyzing Indonesian-language social media. Although previous research has demonstrated the effectiveness of SVM and TF-IDF in sentiment classification, existing studies have primarily focused on other topics, such as application reviews, political discussions, and general public opinions. Consequently, the application and performance of the SVM–TF-IDF approach for analyzing public perceptions of Danantara within Indonesian public policy discourse remain insufficiently explored. To address this gap, this study develops and evaluates a sentiment analysis system that employs TF-IDF for feature weighting and SVM for sentiment classification of Indonesian-language social media data concerning Danantara. The novelty of this research lies in its empirical investigation of public sentiment toward Danantara through a focused analysis of Indonesian social media discourse, rather than in proposing a new classification algorithm. The study further contributes by examining the distribution of positive, negative, and neutral sentiments and evaluating the performance of the proposed SVM–TF-IDF approach using standard classification metrics. The findings are expected to provide data-driven insights into public perceptions of Danantara and contribute to the growing body of research on machine learning-based sentiment analysis in Indonesian public policy discourse.

2. MATERIAL AND METHOD

This study follows a systematic pipeline for sentiment analysis that encompasses data collection, preprocessing, feature extraction, model training, and evaluation. The overall research workflow is illustrated in Figure 1. Each stage is described in detail in the following subsections.

2.1. Data Collection

The dataset used in this study consists of public user-generated content collected from three major social media platforms, namely X (formerly Twitter), Instagram, and YouTube, related to Danantara, the Indonesian sovereign wealth fund officially launched in 2025. These platforms were selected because they represent the most widely used social media channels in Indonesia, each offering distinct modes of public expression: short-form posts on X, photo and video-based comments on Instagram, and video discussion threads on YouTube [9]. The use of multiple social media platforms is recommended because each platform exhibits different interaction patterns and sentiment expressions that can enrich public opinion analysis [15], [16], [28]. Data collection was conducted using keyword-based crawling with the YouTube Data API v3 for YouTube comments, the Instagram Graph API for Instagram comments, and the X API v2 for posts on X. Keywords used included “Danantara,” “investasi Danantara,” and “dana Danantara.” The collection period covered the

official announcement of Danantara through three months following its launch. A total of approximately 1,500 user comments were gathered across the three platforms. Only comments written in Indonesian were included; reposts, duplicate entries, spam, and posts containing only media without text were excluded to ensure data quality and linguistic consistency [10].

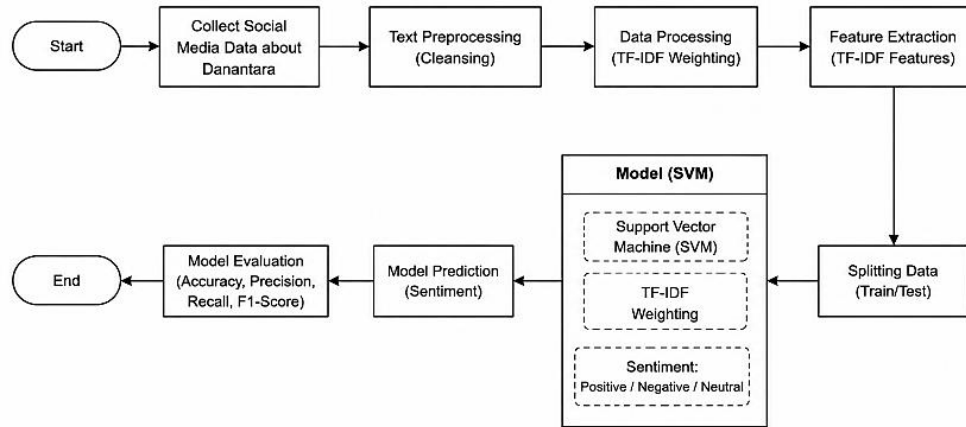


Figure 1. Research Methodology

2.2. Data Labeling

Each comment in the collected dataset was manually annotated into one of three sentiment categories: positive, negative, or neutral. Positive sentiment was assigned to comments expressing support, optimism, or approval toward Danantara. Negative sentiment was assigned to comments reflecting criticism, skepticism, or disapproval. Neutral sentiment was assigned to comments that conveyed factual information without explicit emotional polarity. Annotation was performed by two independent annotators with a background in Indonesian linguistics and social media analysis. Inter-annotator agreement was measured using Cohen's Kappa coefficient (κ), yielding a value of $\kappa = 0.81$, which indicates strong agreement [11]. In cases of disagreement between the two annotators, a third annotator served as the adjudicator to arrive at a final label. The resulting labeled dataset consists of three classes: positive, negative, and neutral.

2.3. Text Preprocessing

The raw social media comments collected from all three platforms underwent a multi-step preprocessing pipeline to reduce noise and improve the quality of feature extraction. The preprocessing stages were applied sequentially as follows: (1) case folding, in which all characters were converted to lowercase to ensure uniformity; (2) cleaning, which involved removing URLs, usernames (mentions), hashtags, emoji, punctuation marks, and numerical characters; (3) normalization, where informal and non-standard Indonesian words were mapped to their formal equivalents using a custom slang dictionary; (4) tokenization, which segmented each cleaned text into individual word tokens; (5) stopwords removal, eliminating high-frequency function words that carry minimal semantic value, using a standard Indonesian stopword list; and (6) stemming, in which words were reduced to their morphological root forms using the Sastrawi library for the Indonesian language [12]. The output of this pipeline is a clean set of tokens that serve as the basis for feature extraction. Previous studies on Indonesian social media sentiment analysis have shown that preprocessing procedures, particularly slang normalization and stemming, significantly improve classification performance by reducing lexical variation and noise [31], [32].

2.4. Feature Extraction Using TF-IDF

Term Frequency–Inverse Document Frequency (TF-IDF) weighting was employed to transform the preprocessed text tokens into numerical feature vectors suitable for input to the SVM classifier. TF-IDF assigns a weight to each term based on two components: the term frequency (TF), which measures how frequently a term appears in a given document, and the inverse document frequency (IDF), which penalizes terms that appear across many documents and thus carry lower discriminative value [6]. The TF-IDF weight for term t in document d within a corpus D is computed as equation 1.

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t) \quad (1)$$

$$\text{IDF}(t) = \log\left(\frac{|D|}{|\{d \in D: t \in d\}|}\right) \quad (2)$$

where $|D|$ denotes the total number of documents in the corpus, and $|\{d \in D : t \in d\}|$ denotes the number of documents containing term t . The resulting TF-IDF matrix represents each tweet as a high-dimensional sparse vector, with each dimension corresponding to a unique term in the vocabulary. Feature vectors with a maximum vocabulary size of 5,000 unique terms were retained for model training [10].

2.5. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm that identifies the optimal hyperplane that maximally separates data points belonging to different classes in a high-dimensional feature space [5]. For a binary classification problem, given a training dataset of labeled feature vectors $\{(x_i, y_i)\}$, where $y_i \in \{-1, +1\}$, the SVM minimizes the following objective function subject to constraints in equation 3.

$$\min_{w,b,\xi} \left(\frac{1}{2} |w|^2 + C \sum_{i=1}^n \xi_i \right) \quad (3)$$

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

where w is the weight vector defining the hyperplane, b is the bias term, C is the regularization parameter that controls the trade-off between margin maximization and misclassification, and ξ_i are slack variables that permit misclassification of training samples. Since the sentiment classification in this study involves three classes (positive, negative, neutral), the One-vs-Rest (OvR) strategy was adopted to extend the binary SVM to a multiclass setting [7]. A linear kernel was selected, given the high-dimensional and sparse nature of TF-IDF feature vectors, as linear kernels are well-suited for text classification tasks and are computationally efficient [6].

2.6. Model Training and Evaluation

The labeled dataset was divided into a training set (80%) and a test set (20%) using stratified random sampling to preserve the class distribution across both partitions. The SVM model was trained on the training set using TF-IDF feature vectors, with hyperparameter C tuned through 5-fold cross-validation. Model performance was assessed on the held-out test set using the following standard evaluation metrics: Accuracy, Precision, Recall, and F1-Score [5]. These metrics are defined as Equation 4-7.

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

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

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

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

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. For multiclass evaluation, the macro-average of Precision, Recall, and F1-Score was computed across all three sentiment classes to account for potential class imbalance. The use of macro-averaged metrics is particularly appropriate for imbalanced multiclass datasets because it evaluates classifier performance equally across all classes and avoids dominance by majority classes [21], [22]. The full experimental pipeline was implemented in Python using the scikit-learn library [13].

3. RESULTS AND DISCUSSION

3.1. Dataset Overview and Sentiment Distribution

The dataset used in this study consists of 3,000 comments collected from three social media platforms, YouTube, Instagram, and X (formerly Twitter), related to Danantara. After preprocessing and manual labeling, the dataset was distributed as follows: 1,714 comments (57.13%) were classified as Negative, 731 (24.37%) as Positive, and 555 (18.50%) as Neutral. This class imbalance reflects the predominantly critical nature of public discourse surrounding Danantara on social media, consistent with findings from similar studies examining public opinion toward newly established government institutions [4]. The predominance of negative sentiment may be attributed to widespread public skepticism about governance transparency, financial management, and Danantara's institutional mandate as a sovereign wealth fund [3]. Similar patterns of negative sentiment toward government policies have been reported in studies conducted in several countries, where concerns regarding transparency and accountability often dominate public discourse on social media [26], [27].

To address the class imbalance, the LinearSVC classifier was configured with the `class_weight='balanced'` parameter, which adjusts the penalty weights inversely proportional to class frequencies. This approach is recognized as an effective strategy for improving minority-class classification performance without requiring data resampling [11]. The dataset was then split using stratified sampling at an 80:20 ratio, yielding 2,400 training samples and 600 test samples while preserving the original class proportions across both subsets.

Tabel 1. Sentiment Distribution of the Dataset

Sentiment	Count	Percentage (%)
Negative	1,714	57.13
Positive	731	24.37
Neutral	555	18.50
Total	3,000	100.00

3.2. TF-IDF Feature Extraction Results

The TF-IDF vectorization was applied with a unigram and bigram configuration (`ngram_range=(1,2)`) and a minimum document frequency threshold of 2 (`min_df=2`) to filter rare terms. The resulting feature matrix consisted of 4,548 unique features. The training set comprised 2,400 samples represented as 4,548-dimensional sparse vectors, and the test set comprised 600 samples with the same dimensionality. The use of bigrams alongside unigrams enables the model to capture short-range contextual patterns, such as negation constructs (e.g., “tidak bagus”) and commonly co-occurring sentiment phrases [6]. This configuration aligns with established practices in Indonesian sentiment analysis, where multi-word expressions frequently carry significant sentiment value [12].

3.3. Model Performance Evaluation

The LinearSVC model trained on TF-IDF features achieved an overall accuracy of 84.67% on the held-out test set. The per-class performance metrics are summarized in Table 2. The Negative class achieved the highest F1-score of 0.89 (precision=0.87, recall=0.90), consistent with its majority representation in the training data. The Positive class obtained an F1-score of 0.83 (precision=0.82, recall=0.84), while the Neutral class, being the smallest class, achieved an F1-score of 0.74 (precision=0.79, recall=0.69). The macro-average F1-score across all three classes was 0.82, and the weighted-average F1-score was 0.85. These results demonstrate that the SVM model with TF-IDF features provides strong classification performance even under class-imbalanced conditions, consistent with prior research reporting SVM accuracy above 80% for multiclass sentiment tasks on social media text [7]. Comparable findings have also been reported in comparative studies between lexicon-based and machine-learning approaches, where SVM consistently demonstrates competitive performance for sentiment classification tasks on social media datasets [15], [17].

Tabel 2. Classification Report of SVM with TF-IDF

Class	Precision	Recall	F1-Score	Support
Negative	0.87	0.90	0.89	343
Neutral	0.79	0.69	0.74	111
Positive	0.82	0.84	0.83	146
Macro avg	0.83	0.81	0.82	600
Weighted avg	0.85	0.85	0.85	600

The relatively lower recall for the Neutral class (0.69) indicates that the model tends to misclassify some Neutral comments as either Positive or Negative. This is a known challenge in three-class sentiment analysis, as Neutral comments often contain mixed or ambiguous linguistic signals that overlap with the feature distributions of the other two classes [5]. The relatively lower representation of Neutral samples in the training set (18.50%) further limits the model’s capacity to learn discriminative features for this class. Despite this limitation, the balanced class weighting strategy employed in LinearSVC prevented the model from defaulting entirely to the majority class, thereby enabling meaningful recognition of all three sentiment categories [9].

3.4. Confusion Matrix Analysis

Figure 2 presents the confusion matrix of the SVM classifier on the held-out test set. The diagonal entries represent correctly classified samples for each sentiment class, while the off-diagonal entries indicate misclassifications. The Negative class achieved the highest number of correct predictions (309 out of 343 actual Negative samples), consistent with its dominant representation in the dataset. The Positive class was correctly identified in 122 out of 146 cases, while the Neutral class, with 77 correct predictions out of 111, exhibited the highest misclassification rate. Specifically, 23 Neutral samples were misclassified as Negative and 11 as Positive, reflecting the inherent ambiguity of Neutral comments that contain overlapping linguistic features

with the other two sentiment categories. These patterns corroborate the per-class recall values reported in Table 2 and confirm that the Neutral class remains the most challenging to classify in the three-class setting.

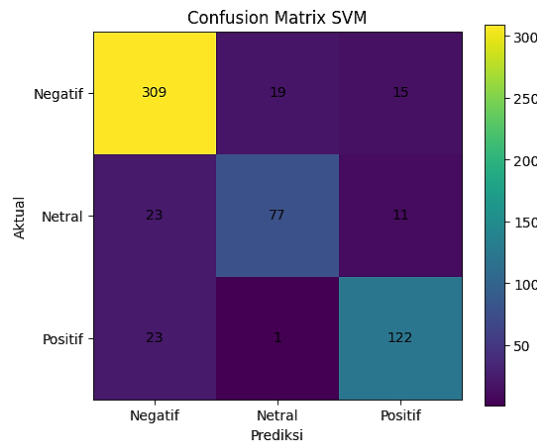


Figure 2. Confusion Matrix of SVM

3.5. Dominant Features per Sentiment Class

Analysis of the top 15 TF-IDF-weighted features per class reveals semantically coherent term clusters for each sentiment category. For the Negative class, the highest-weight terms include “suram” (gloomy), “parah” (severe/terrible), “korupsi” (corruption), “bubarkan” (disband), “oligarki” (oligarchy), and “hutang” (debt), reflecting strong public concerns about governance failures, corruption risks, and fiscal mismanagement. Studies on political communication have similarly found that terms associated with corruption, institutional trust, and economic risk frequently emerge as dominant indicators of negative sentiment in online public discourse [16], [23]. These terms are substantively aligned with common themes in critical public discourse toward state-linked institutions in Indonesia [3]. For the Positive class, dominant features include “bagus” (good), “mantap” (great), “keren” (cool), “efisiensi” (efficiency), “kesejahteraan” (welfare), and “profesional” (professional), suggesting that supportive commenters emphasize national development aspirations and institutional competence. For the Neutral class, terms such as “apakah” (whether), “gimana” (how), “perampangan” (streamlining), and “foreigners” suggest informational or inquisitive content without explicit emotional polarity, consistent with the definitional profile of neutral sentiment [5].

3.6. Cross-Validation and Model Stability

To further assess the generalizability and stability of the SVM model, a 5-fold cross-validation was conducted on the training set, using macro F1 Score as the evaluation metric. The per-fold results and overall mean are presented in Table 3. The mean macro F1-Score across the five folds was 0.8093, with individual fold scores ranging from 0.7989 (Fold 1) to 0.8219 (Fold 2). The standard deviation across folds was approximately 0.0082, indicating minimal variance between folds and confirming that the model generalizes consistently across different data partitions [14]. This stability is particularly important in the context of imbalanced datasets, where a single train-test split may yield misleading performance estimates [13].

Tabel 3. 5-Fold Cross-Validation Results (Macro F1-Score)

Fold	Macro F1-Score
1	0.7989
2	0.8219
3	0.8016
4	0.8151
5	0.8090
Mean	0.8093

The cross-validation macro F1-Score of 0.8093 closely approximates the single-split macro F1-Score of 0.82 reported in the classification report, suggesting that the model is not overfitting to a particular data partition and that the performance estimate is reliable. This level of consistency across folds is characteristic of well-regularized linear SVM classifiers when trained on sufficiently large TF-IDF feature spaces [6]. The results are comparable with studies that applied SVM and TF-IDF to Indonesian social media datasets, where macro F1-Scores typically fall in the range of 0.78 to 0.85 depending on dataset size, class distribution, and preprocessing quality [9][12]. Previous investigations also indicate that performance variability in text

classification models can be further influenced by sampling strategy, feature representation, and imbalance-handling techniques such as oversampling and class weighting [19], [20].

Taken together, the results of this study demonstrate that the combination of SVM with linear kernel and TF-IDF bigram features provides a robust and interpretable baseline for multiclass sentiment classification of Indonesian social media text. The predominantly negative public sentiment toward Danantara, as quantified by this model, offers empirical evidence that the institution faces significant communication challenges that policymakers should address. Future research may consider incorporating contextual embedding models, such as IndoBERT or RoBERTa-based approaches, to further improve Neutral class detection and to handle the semantic ambiguity inherent in informal Indonesian [8].

3.7. Discussion

The LinearSVC model achieved an accuracy of 84.67%, with a macro-average F1-score of 0.82 and a weighted-average F1-score of 0.85. The fivefold cross-validation yielded an average macro F1 Score of 0.8093, indicating relatively stable model performance. Negative sentiment dominated the dataset, accounting for 57.13% of all comments. This dominance may reflect public concerns, criticism, or uncertainty regarding Danantara. However, the dataset only represents opinions expressed on the selected social media platforms and may not reflect the views of the entire population.

The negative class achieved the highest F1-score (0.89), followed by the positive class (0.83). In contrast, the neutral class obtained the lowest F1-score (0.74) and recall (0.69). Neutral comments are more difficult to classify because they often contain factual statements, questions, or ambiguous expressions without clear sentiment indicators. The confusion matrix also showed that several neutral comments were classified as negative or positive.

These findings are consistent with previous sentiment analysis studies, indicating that TF-IDF combined with Support Vector Machines can produce effective results for text classification. However, neutral sentiment generally remains more challenging because TF-IDF does not fully capture context, sarcasm, negation, or implicit meaning. Future research may improve the results by using contextual models such as IndoBERT and increasing the number of neutral training samples.

4. CONCLUSION

This study developed a sentiment analysis model for Danantara-related social media comments using a Support Vector Machine with TF-IDF weighting. The model achieved an accuracy of 84.67%, with a macro-average F1-score of 0.82 and a weighted-average F1-score of 0.85. These results demonstrate that the proposed method can effectively classify public opinions expressed on social media platforms.

The sentiment distribution was dominated by negative comments (57.13%), followed by positive comments (24.37%) and neutral comments (18.50%). This distribution indicates that negative responses were more prominent in the collected dataset, potentially reflecting public concerns, criticism, or uncertainty regarding Danantara. These findings can help policymakers and related institutions identify public concerns, evaluate public responses, and develop more effective public communication strategies. Nevertheless, the model showed lower performance in classifying neutral sentiment, which obtained an F1-score of 0.74 and a recall of 0.69. This limitation suggests that neutral comments may contain ambiguous expressions, factual statements, or limited sentiment indicators that are difficult to distinguish using TF-IDF features.

Future research should improve the classification of neutral sentiment by increasing the number and diversity of neutral training samples, applying more advanced linguistic preprocessing, and testing contextual language models such as IndoBERT. Further studies should also collect data from additional social media platforms, use larger, more balanced datasets, and conduct periodic sentiment analysis to track changes in public opinion over time.

REFERENCES

- [1] D. K. J. Samuel, M. Swamy, P. Koparde, and S. N., "NLP - Powered Sentiment Analysis on the Twitter," *Saudi Journal of Engineering and Technology*, vol. 9, no. 01, pp. 1–11, Jan. 2024, doi: 10.36348/sjet.2024.v09i01.001.
- [2] L. Xiao, Q. Li, Q. Ma, J. Shen, Y. Yang, and D. Li, "Text classification algorithm of tourist attractions subcategories with modified TF-IDF and Word2Vec," *PLoS One*, vol. 19, no. 10 October, Oct. 2024, doi: 10.1371/journal.pone.0305095.
- [3] S. Gnanapriya and K. Anandan, "Sentiment Analysis in Political discourse: Understanding public opinion from social media," *Sage Journals*, Sep. 2025, doi: 10.1177/20427530251343205.
- [4] G. P. Rizky and W. Alrasyid, "AI-Based Sentiment Analysis of Social Media to Detect Public Opinion on Government Policies," 2025.
- [5] K. L. Tan, C. P. Lee, and K. M. Lim, "A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research," *Apr. 01, 2023, MDPI*. doi: 10.3390/app13074550.

- [6] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Enhancing machine learning-based sentiment analysis through feature extraction techniques," *PLoS One*, vol. 19, no. 2 February, Feb. 2024, doi: 10.1371/journal.pone.0294968.
- [7] K. Mathur, P. Jain, S. Gupta, and P. Mathur, "Sentiment Analysis Framework of Social Media Text by Feature Extraction and Machine Learning Model," *Indian J. Sci. Technol.*, vol. 16, no. 29, pp. 2233–2243, Aug. 2023, doi: 10.17485/IJST/v16i29.1537.
- [8] K. L. Tan, C. P. Lee, K. M. Lim, and K. S. M. Anbananthen, "Sentiment Analysis With Ensemble Hybrid Deep Learning Model," *IEEE Access*, vol. 10, pp. 103694–103704, 2022, doi: 10.1109/ACCESS.2022.3210182.
- [9] S. Rais Dwi Jayanto, "Sentiment Analysis on X, TikTok, and Instagram on Indonesian Capital relocation using Support Vector Machine," *Journal of Information System Research*, vol. 7, no. 2, pp. 349–359, 2026, doi: 10.47065/josh.v7i2.8839.
- [10] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, "Enhancing machine learning-based sentiment analysis through feature extraction techniques," *PLoS One*, Feb. 2024, doi: <https://doi.org/10.1371/journal.pone.0294968>.
- [11] A. N. A. Saputra, R. E. Saputro, and D. I. S. Saputra, "Enhancing Sentiment Analysis Accuracy Using SVM and Slang Word Normalization on YouTube Comments," *Jurnal dan Penelitian Teknik Informatika*, Apr. 2025, doi: <https://doi.org/10.33395/sinkron.v9i2.14513>.
- [12] Muhamad Agung Nulhakim, Yuliant Sibaroni, and Ku Muhammad Naim Ku Khalif, "Geospatial Sentiment Analysis Using Twitter Data on Natural Disasters in Indonesia with Support Vector Machine (SVM) Algorithm," *International Journal on Information and Communication Technology (IJoICT)*, vol. 10, no. 2, pp. 242–258, Jan. 2025, doi: 10.21108/ijoict.v10i2.1032.
- [13] S. K. Al Sarayrah, N. M. Alkudah, and H. Y. Al-kharabsheh, "Twitter Sentiment Analysis Using Machine Learning and Deep Learning Techniques," *Journal of Computer Science*, Mar. 2025, doi: 10.3844/jcssp.2025.1785.1794.
- [14] T. Abedin, H. Xu, and S. Uddin, "The impact of K selection in K-fold cross-validation on bias and variance in supervised learning models," *Sci. Rep.*, vol. 16, no. 1, Dec. 2026, doi: 10.1038/s41598-026-37247-x.
- [15] Y. Qi and Z. Shabrina, "Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach," *Soc. Netw. Anal. Min.*, vol. 13, no. 1, Art. 31, Feb. 2023, doi: 10.1007/s13278-023-01030-x.
- [16] P. A. Henríquez and F. Alessandri, "Analyzing digital societal interactions and sentiment classification in Twitter (X) during critical events in Chile," *Heliyon*, vol. 10, no. 12, Jun. 2024, doi: 10.1016/j.heliyon.2024.e32572.
- [17] F. Cao and X. Huang, "Performance analysis of aspect-level sentiment classification task based on different deep learning models," *PeerJ Comput. Sci.*, vol. 9, p. e1578, Oct. 2023, doi: 10.7717/peerj-cs.1578.
- [18] J. Zhou, Z. Ye, S. Zhang, Z. Geng, N. Han, and T. Yang, "Investigating response behavior through TF-IDF and Word2vec text analysis: A case study of PISA 2012 problem-solving process data," *Heliyon*, vol. 10, no. 16, p. e35945, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35945.
- [19] M. Gao, M. Li, Z. Ling, J. Zhong, H. Ding, and Q. Wu, "Wavelet analysis text classification algorithm based on typical features of data samples," *PLoS One*, vol. 20, no. 6, p. e0319747, Jun. 2025, doi: 10.1371/journal.pone.0319747.
- [20] S. F. Taskiran, B. Turkoglu, E. Kaya, and T. Asuroglu, "A comprehensive evaluation of oversampling techniques for enhancing text classification performance," *Sci. Rep.*, vol. 15, Art. 21631, Jul. 2025, doi: 10.1038/s41598-025-05791-7.
- [21] J. S. Aguilar-Ruiz and M. Michalak, "Classification performance assessment for imbalanced multiclass data," *Sci. Rep.*, vol. 14, Art. 10759, May 2024, doi: 10.1038/s41598-024-61365-z.
- [22] S. Farhadpour, T. A. Warner, and A. E. Maxwell, "Selecting and Interpreting Multiclass Loss and Accuracy Assessment Metrics for Classifications with Class Imbalance: Guidance and Best Practices," *Remote Sens.*, vol. 16, no. 3, p. 533, Jan. 2024, doi: 10.3390/rs16030533.
- [23] A. Smailova, A. Kuzembayeva, R. Utkelbay, F. Kukeyeva, and K. Baizakova, "Public opinion monitoring technologies: how the state uses data to make political decisions," *Front. Polit. Sci.*, vol. 7, Art. 1680172, Dec. 2025, doi: 10.3389/fpos.2025.1680172.
- [24] T. Babad-Falk and S. Chun, "Public Opinion Classification on Government Policy Using Social Media: An Exploration of ChatGPT's Capabilities and Limitations," *The International FLAIRS Conference Proceedings*, vol. 38, no. 1, May 2025, doi: 10.32473/flairs.38.1.138905.
- [25] R. W. Harwenda, M. D. Angelo, I. Budi, A. B. Santoso, and P. K. Putra, "Sentiment Analysis on Government Public Policies: A Systematic Literature Review," *Dinasti Int. J. Educ. Manag. Soc. Sci.*, vol. 6, no. 5, pp. 4192–4211, Jul. 2025, doi: 10.38035/dijemss.v6i5.4699.

-
- [26] Y. Wang, M. L. Diabakanga Batatana, and M. H. Bikoumou Gambat, "Public perceptions of government policies to COVID-19: A comparative study in six African countries," *Heliyon*, vol. 10, no. 3, p. e24888, Jan. 2024, doi: 10.1016/j.heliyon.2024.e24888.
 - [27] A. A. Benson and R. Essah, "Sentiment Analysis of Ghanaian Social Media Discussion on Government Policy," *Int. J. Res. Innov. Appl. Sci.*, pp. 737–747, Aug. 2025, doi: 10.51584/IJRIAS.2025.100700067.
 - [28] S. Kumar et al., "Digital society social interactions and trust analysis model," *PeerJ Comput. Sci.*, vol. 8, p. e1129, 2022, doi: 10.7717/peerj-cs.1129.
 - [29] J. Amar and C. Lecourt, "Sovereign wealth fund governance: A trade-off between internal and external legitimacy," *Int. Bus. Rev.*, vol. 32, no. 6, p. 102193, Dec. 2023, doi: 10.1016/j.ibusrev.2023.102193.
 - [30] J. Diaz and A. L. P. Gajes, "Governance and Transparency in Sovereign Wealth Funds: A Systematic Review of Global Practices and Policy Insights for the Maharlika Investment Fund in the Philippines," *SSRN Electron. J.*, May 2025, doi: 10.2139/ssrn.5211840.
 - [31] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models," *JITEKI*, vol. 9, no. 3, pp. 746–757, Sep. 2023, doi: 10.26555/jiteki.v9i3.26490.
 - [32] B. Setiawan, "A Review of Sentiment Analysis Applications in Indonesia Between 2023-2024," *JIEET*, vol. 8, no. 2, pp. 71–83, May 2025, doi: 10.26740/jieet.v8n2.p71-83.