Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

June 2026. Vol. 5. No. 3; e-ISSN: 2808-4519

Sentiment Analysis of Social Media X Users Toward Finance Minister Purbaya Yudhi Sadewa Using the Support Vector Machine Algorithm

Adian Fahreza Surbakti[1*], Relita Buaton[2], Selfira[3]

[1,2,3]*Information System, STMIK Kaputama Binjai, Indonesia*
adianfhrza@gmail.com

Abstract

In this digital era, the rapid advancement of information and communication technology has transformed social media platforms particularly X (formerly Twitter) into a primary space for public discourse concerning government policies. The Minister of Finance, Purbaya Yudhi Sadewa, has become a focal point of public debate, garnering reactions ranging from appreciation to criticism regarding his management of national finances. However, manual sentiment analysis is impractical, time-consuming, and prone to subjectivity when handling the massive and continuously expanding volume of social media data. Therefore, an automated, machine learning-based approach is essential to process this big data into strategic insights for mapping public sentiment. This study aims to objectively analyze public sentiment toward the Minister of Finance by implementing the Support Vector Machine (SVM) algorithm within the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. The methodology includes data crawling, text preprocessing, and feature extraction using the TF-IDF (Term Frequency – Inverse Document Frequency) method. Analysis of 3,927 tweets reveals that public opinion is dominated by negative sentiment at 54.2%, followed by positive sentiment at 36.9% and neutral sentiment at 8.9%. The developed SVM model achieved a classification accuracy of 72.43%, demonstrating that this machine learning approach is both effective and reliable for mapping public perception. These findings indicate that the Minister of Finance, Purbaya Yudhi Sadewa, faces significant public scrutiny, and this data-driven analysis serves as a strategic tool for evaluating the policies under his administration.

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

1. Introduction

The rapid advancement of information and communication technology has fundamentally transformed the paradigm of social interaction and public discourse. Social media platforms, particularly X (formerly Twitter), have evolved into primary arenas for digital democracy, allowing the Indonesian public to express opinions, aspirations, and critiques regarding various public issues and government policies in real-time. The platform's capacity for rapid, brief, and dynamic information dissemination makes it a highly active hub for public discussion, where millions of posts are generated daily. Consequently, X serves as a transparent medium where individuals can directly voice their perspectives on the performance of public officials and the implementation of state policies[1]

As one of the most strategic institutions, the Ministry of Finance plays a vital role in determining the direction of fiscal policy, taxation, and state financial management. Purbaya Yudhi Sadewa, serving as the Minister of Finance, carries the significant responsibility of managing the national budget (APBN) and formulating economic strategies that directly impact all levels of society. Various policies, including tax reforms and subsidy adjustments, consistently emerge as focal points of intense debate on social media, eliciting a wide spectrum of public responses ranging from strong support to vocal opposition [2]

Understanding public sentiment toward the Minister of Finance is crucial, as these policies dictate national fiscal stability and public economic welfare. Media monitoring reveals a noticeable divergence in narratives: major outlets such as Tribunnews.com, Kumparan, and Kompas.com often report positive sentiments regarding Purbaya Yudhi Sadewa's performance. Conversely, platforms like Bisnis.com and Tempo.co frequently highlight negative sentiments, leading to a polarized perception among X users. Given that these sentiments serve as early indicators of policy acceptance or rejection, a systematic and accurate sentiment analysis is imperative[4].

While public feedback is essential for policy evaluation, manual sentiment analysis of such massive, high-velocity data is no longer practical, as it is time-consuming and susceptible to human subjectivity. Therefore, an automated machine learning approach is required to transform social media big data into strategic insights that can support more responsive governance [5].

To address these challenges, this study employs sentiment analysis utilizing the Term Frequency – Inverse Document Frequency (TF-IDF) method for text feature weighting. TF-IDF is widely recognized in text mining for its ability to assign higher weights to relevant terms while filtering out common, less informative words. This method is integrated with the Support Vector Machine (SVM) algorithm, which is renowned for its high accuracy in classifying unstructured text data into distinct sentiment categories: positive, negative, and neutral.[6].

This research aims to analyze public sentiment on platform X toward Minister of Finance Purbaya Yudhi Sadewa using the SVM algorithm. The urgency of this study is underscored by the need to understand public perception of a strategic figure whose decisions have a broad national economic impact, combined with the current scarcity of machine-learning-based studies specifically targeting high-ranking officials within the Ministry of Finance. By performing data collection, preprocessing, and model evaluation, this study seeks to provide an objective overview of public perception. Furthermore, the findings are expected to contribute to the development of automated sentiment analysis tools that assist stakeholders in making data-driven and public-responsive policy decisions[7].

2. Methodology

This study adopts a data mining approach based on the CRISP-DM framework, which encompasses six primary phases. CRISP-DM is the most widely utilized methodology in the development of data mining-based systems due to its systematic structure, flexibility, and applicability to a diverse range of research problems.

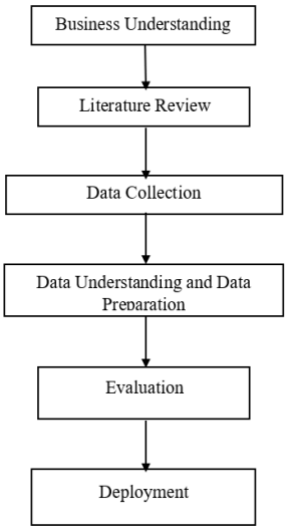

Fig. 1: Methodology

2.1. Data Collection

The data collection phase involves extracting textual data from the social media platform X. The dataset comprises posts and comments containing specific keywords related to the Minister of Finance, Purbaya Yudhi Sadewa. This process is conducted using the X API or authorized web scraping techniques, strictly adhering to the platform's terms of service and data usage policies. The acquired data is subsequently structured and stored in a dataset format, serving as the foundational input for the subsequent stages of the research.[1].

2.2. Data Preprocessing

In this phase, the raw data collected is processed to ensure its readiness for the classification task. The data preparation pipeline includes several critical steps: Data Cleaning, which involves the removal of URLs, symbols, emojis, and irrelevant characters; Text Preprocessing, encompassing case folding, tokenization, stopword removal, and stemming; and Sentiment Labeling (classifying data into positive, negative, and neutral categories), which is essential for the supervised learning approach. Subsequently, Feature Extraction is performed using the TF-IDF method to transform the textual data into a numerical format. This phase is crucial for producing clean, structured, and high-quality data, thereby establishing a robust foundation for the modeling process.[2].

2.3. Sentimen Anotation

Sentiment annotation is a critical process in supervised learning used to establish a 'ground truth' by systematically labeling textual data into positive, negative, and neutral categories. In this study, the process involves categorizing social media posts based on their emotional polarity toward the Minister of Finance, Purbaya Yudhi Sadewa, to differentiate between expressions of support, criticism, and objective statements. By ensuring consistent and precise labeling, this phase provides a high-quality dataset essential for training the Support Vector Machine (SVM) model to accurately classify the nuances of public discourse on the X platform[3].

2.4. Model Training Support Vector Machine

In the modeling and testing phase, the preprocessed data is utilized to develop a sentiment classification model employing the Support Vector Machine (SVM) algorithm. The dataset is partitioned into training and testing sets. The SVM model is subsequently trained using the training dataset and evaluated against the testing dataset to determine its classification performance. The resulting sentiment classification provides empirical insights into public opinion regarding the Minister of Finance, Purbaya Yudhi Sadewa, on the X platform[4].

2.5. Evaluation

The evaluation phase is conducted to assess the performance of the developed model. This evaluation utilizes key metrics, including accuracy, precision, recall, and F-score (or F1-score). These evaluation results serve as indicators of the SVM algorithm's success in classifying sentiment effectively. Furthermore, these findings provide the empirical basis for drawing conclusions and formulating recommendations for future research development [5].

3. Research Method

The research methodology adopted in this study is based on the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, which serves as the foundational structure for conducting this sentiment analysis. CRISP-DM is the most widely utilized standardized methodology in data mining and machine learning projects due to its highly structured and systematic approach. This study follows the six primary phases of the CRISP-DM lifecycle, which are as follows:

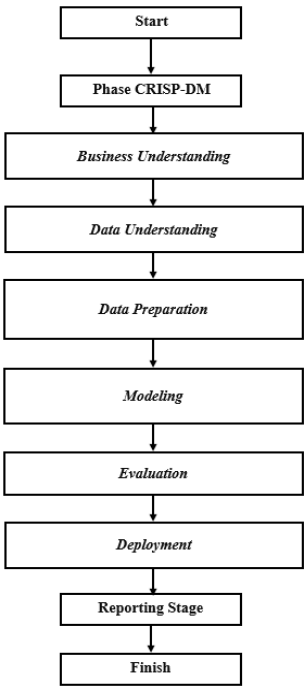

Fig. 2: Reaserch Framework

3.1. Data Collection

Data collection is conducted by extracting text-based data from the social media platform X, focusing on posts or comments containing keywords related to the Minister of Finance, Purbaya Yudhi Sadewa. This process utilizes the X API or authorized web scraping techniques in strict accordance with the platform's terms of service. The acquired data is subsequently structured into a dataset, which serves as the foundational input for the subsequent research phases.

3.2. Data Preprocessing

In this phase, the collected raw data is processed to ensure suitability for classification. The workflow includes data cleaning (removing URLs, symbols, emojis, and irrelevant characters), text preprocessing (case folding, tokenization, stopword removal, and stemming), sentiment labeling (positive, negative, and neutral), and feature extraction via TF-IDF. This stage aims to generate a clean, structured, and high-quality dataset, serving as the foundational input for the subsequent modeling process.

3.3. Data Spliting

Data splitting is the process of partitioning the dataset into training and testing sets, ensuring the SVM model learns from the former and is validated against unseen data in the latter. In this study, the labeled dataset regarding the Minister of Finance, Purbaya Yudhi Sadewa, is divided using a 70:30 ratio for training and testing, respectively. This split is essential for evaluating the model’s classification performance, utilizing key metrics such as accuracy, precision, recall, and F1-score.

3.4. Model Training

Following the cleaning phase, the data is analyzed using the Support Vector Machine (SVM) algorithm to model and classify sentiment from the processed tweets. A linear kernel is employed as the primary function to establish a linear decision boundary for data separation. This kernel functions by calculating the dot product between data vectors without mapping them to a higher-dimensional space. The linear kernel is particularly suitable when data is linearly separable and exhibits simple patterns, thereby enhancing computational efficiency.

3.5. Evaluation

The evaluation phase is conducted to assess the performance and robustness of the SVM classification model. This process involves calculating quantitative metrics based on the confusion matrix derived from the testing dataset, including accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model's predictions, while precision and recall provide insights into the model's ability to identify specific sentiment categories accurately. The F1-score is utilized as a balanced metric to evaluate the model's performance, particularly in cases where class distribution may be imbalanced. These metrics collectively determine the model's success in classifying public sentiment toward the Minister of Finance and serve as the empirical foundation for deriving conclusions and formulating recommendations for further research.

3.6. Integration

The integration phase involves the implementation of the developed SVM model into a functional framework capable of processing real-time or batch data from the X platform. This process ensures that the trained model is effectively deployed to transform raw textual inputs into actionable sentiment insights. By integrating the model into the system, the classification results—categorized as positive, negative, or neutral—become accessible to stakeholders, providing an automated and scalable mechanism for monitoring public perception. This final stage bridges the gap between theoretical model performance and practical utility, enabling decision-makers to utilize data-driven insights for responsive policy evaluation and public communication strategies.

4. Result and Discussion

This section outlines the performance of the SVM model in sentiment analysis, along with the distribution of sentiments reflected in Social Media X concerning the Minister of Finance Purbaya Yudhi Sadewa.

4.1. Accuracy

The model achieved an overall accuracy of 72.43%, demonstrating effective classification performance. Among the sentiment categories, the negative class yielded the best performance with an F1-score of 0.79, followed by the positive class at 0.72. Conversely, the neutral class showed a lower F1-score of 0.33, likely due to the inherent ambiguity and lack of clear emotional indicators in neutral text. Overall, these results confirm the model's effectiveness in analyzing sentiment toward Minister of Finance Purbaya Yudhi Sadewa, while also highlighting the need for further optimization in detecting neutral sentiments

Fig. 3: Confusion Matrix on Training Data

4.2. Sentiment Distribution

Visualization of the 3,927 collected tweets reveals that negative sentiment dominates the public discourse at 54.2%, followed by positive sentiment at 36.9% and neutral sentiment at 8.9%. This distribution indicates that over half of the analyzed content reflects criticism or dissatisfaction. While a significant portion of users still expresses support or appreciation, the prevalence of negative sentiment highlights a more critical public reception. The relatively low percentage of neutral sentiment suggests that users predominantly lean toward firm, expressive stances rather than objective commentary. Overall, these findings indicate that public opinion toward the Minister of Finance, Purbaya Yudhi Sadewa, was predominantly critical during the observation period.

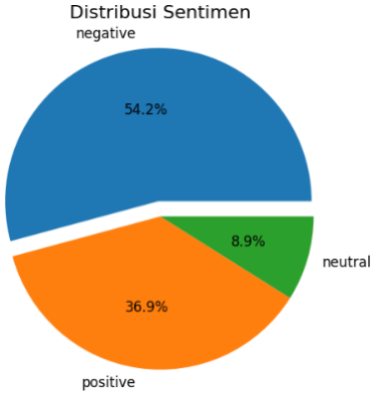

Fig. 4: Confusion Matrix on Training Data

4.3. Model Performance

The confusion matrix demonstrates the model's strong classification capability, with a high number of correct predictions: 507 for the negative class, 313 for the positive class, and 34 for the neutral class. This performance confirms that the model effectively identifies linguistic patterns associated with clear positive and negative sentiments. However, the lower performance in the neutral class is attributed to the presence of ambiguous, purely descriptive, or purely informational text, which blurs the semantic boundaries between classes. Despite this challenge, the model remains robust, as the majority of the test data was classified accurately, validating its overall effectiveness in sentiment analysis

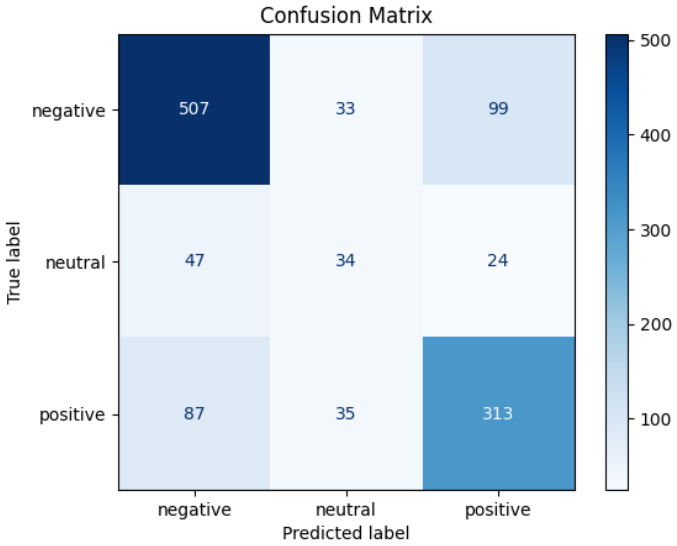

Fig. 5: Confusion Matrix on Training Data

5. Conclusion

This study successfully implemented the CRISP-DM framework and the Support Vector Machine (SVM) algorithm to analyze public sentiment toward the Minister of Finance, Purbaya Yudhi Sadewa, on the X platform. Through a systematic process—including web crawling, comprehensive text preprocessing, and TF-IDF feature extraction—the model achieved an overall accuracy of 72.43%, demonstrating its effectiveness in classifying unstructured social media data. The findings reveal that public discourse is predominantly critical, with 54.2% of sentiment categorized as negative, 36.9% as positive, and 8.9% as neutral. Ultimately, this research confirms that the integrated approach of data mining and machine learning provides a reliable and objective mechanism for mapping public opinion and policy-related dynamics in the digital space.

References

[1] Syarifuddinn, Muhammad. 2020. "Analisis Sentimen Opini Publik Terhadap Efek Psbb Pada Twitter Dengan Algoritma Decision Tree,Knn, Dan Naïve Bayes." INTI Nusa Mandiri 15(1): 87–94. doi:10.33480/inti.v15i1.1433..

[2] Chaerul, Muh, Gandung Triyono, Magister Ilmu Komputer, Fakultas Teknologi Informasi, and Universitas Budi Luhur. 2025. "Analisis Sentimen Kebijakan Pembatasan Subsidi Bahan Bakar Minyak Di Indonesia Tahun 2024 Menggunakan Algoritma Klasifikasi The Sentiment Analysis of the Fuel Subsidy Limitation Policy Using Support Vector Classifier and Random Forest Classifier Algorithm." Jurnal Pendidikan dan Teknologi Indonesia (JPTI) 5(5): 1471–84..

[3] Hariyanto, Muhadi, Muhammad Kholiq, Ahmad Yani, and Narti. 2021. "Inti Nusa Mandiri." Inti Nusa Mandiri 15(2): 133–38.

[4] Indrasetianingsih, A., E. M. P. Hermanto, M. Ilham, N. Rahmawati, and I. A. Hariyanti. 2025. "Inti Nusa Mandiri." LPPM Nusa Mandiri 19(2): 146–53.

[5] Afriza, Aulia, and Dede Latipah. 2025. "Analisis Sentimen Hate Speech Pada Pemerintahan Indonesia Di Twitter (X)." JINTIKOM: Jurnal Informasi … (X): 1–9. http://journal.umbogorraya.ac.id/index.php/jintikom/article/view/343%0Ahttp://journal.umbogorraya.ac.id/index.php/jintikom/article/download/343/268.

[6] Prahmana, I. G., Sitepu, K. A. Br., & Selfira. (2024). Analysis Sentiment On Social Media Instagram Towards Metaverse Games Saindbox Aplha 2 With Support Vector Machine Algorithm. Journal of Artificial Intelligence and Engineering Applications, 3(2).

[7] Sitepu, Juliani Br, Suci Ramadani, and Muammar Khadapi. 2025. "Global Research and Innovation Journal (GREAT) ANALISIS SENTIMEN PENGGUNA MEDIA TWITTER TERHADAP KURIKULUM MERDEKA MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)." 1: 172–77.