Public Sentiment Analysis on the Performance of the Ministry of Finance for the 2025-2029 Period Using the Support Vector Machine (SVM) Method on Social Media Data X

Annisa Tyas Wahyu Setyaningsih [1]*, Yoannes Romando Sipayung [2]*
* Program Studi S1 Teknik Informatika, Fakultas Komputer dan Pendidikan, Universitas Ngudi Waluyo
anisatyas254@gmail.com [1], yoannesromando@unw.ac.id [2]

Article Info

Article history:

Received 2026-01-14
Revised 2026-02-20
Accepted 2026-02-27

Keyword:

Sentiment Analysis,
Finance Minister,
SVM,
TF-IDF,
X (Twitter).

ABSTRACT

The change in leadership at the Ministry of Finance of the Republic of Indonesia for the 2025–2029 period, marked by the appointment of Purbaya Yudhi Sadewa as Minister of Finance, triggered various responses from the public, many of which were recorded on social media. This study aims to examine public sentiment trends regarding the performance of the Ministry of Finance and assess the ability of the Support Vector Machine (SVM) algorithm to classify public opinion. The research data was obtained from the X platform (formerly Twitter) between September and November 2025 and produced 1,164 documents that were declared valid after undergoing preprocessing. Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method, while the SVM model was optimized with a linear kernel through hyperparameter adjustment. The novelty of this study lies in its comprehensive analysis of sentiment polarity during the early transition phase of fiscal authority in the post–Sri Mulyani era, as well as in testing the robustness of the SVM model in handling ambiguity of economic terminology within X social media data. Through hyperparameter optimization and the integration of TF-IDF weighting, this research further explores how key fiscal policy terms such as taxation and national debt contribute to classification accuracy within the context of political opinion discourse. The analysis results show that negative sentiment dominated with a percentage of 81.1%, mainly related to issues of fiscal policy, taxation, and national debt. On the other hand, positive sentiment was recorded at 18.9% and reflected support and confidence in the new fiscal leadership. The best SVM model with regularization parameter C = 5.0 achieved an accuracy value of 86.27%, an F1-score of 0.86, and an Area Under the Curve (AUC) value of 0.91. These findings confirm that SVM is effective in sentiment classification under naturally imbalanced social media data and offer empirical observations about digital public perception relevant to the evaluation of government fiscal policy communication.

I. INTRODUCTION

The development of information technology has made social media, particularly the X platform (formerly Twitter), a strategic public space for expressing opinions, criticism, and support regarding various government issues. The real-time nature and short text format of this platform make it a relevant source of data for observing the dynamics of public response to crucial issues, such as economic policy, social restrictions, and strategic government programs [1], [2], [3]. In addition, the ability of the X platform to provide data directly enables more responsive analysis of changes in public sentiment, especially in the context of political events [4]. Opinions that appear on this platform are often used as early indicators in assessing the level of public trust and perception of the performance of state officials and government institutions [5], [6].

The Ministry of Finance of the Republic of Indonesia is one of the institutions that is always in the public spotlight, as its fiscal policies have a direct impact on national economic stability and public welfare [7]. During the 2025–2029 administration, public attention increased with the change in leadership, marked by the appointment of Purbaya Yudhi Sadewa as Minister of Finance, replacing Sri Mulyani Indrawati. Transitions involving strategic public officials generally elicit various responses in the digital realm, ranging from support to criticism of the direction of future policies [8], [9]. This type of analysis allows policymakers to detect public dissatisfaction early on and formulate appropriate improvement strategies, thereby supporting evidence-based improvements in public service quality [10]. Preliminary research on YouTube shows that the appointment of Purbaya Yudhi Sadewa sparked intense public discussion, with a wide spectrum of sentiments, ranging from cautious support and optimism to skepticism about future economic policies [11].

The large volume of unstructured public opinion on social media requires a systematic computational approach to extract meaningful information. Sentiment analysis, a component of text mining and natural language processing (NLP), classifies public opinion trends as positive, negative, or neutral [12]. The use of social media data for sentiment analysis presents distinct challenges, particularly due to the prevalence of informal language styles and substantial levels of textual noise. These conditions place the choice of classification algorithm as a determining factor in achieving reliable and accurate sentiment classification results.

Among the various approaches applied in sentiment analysis studies, Support Vector Machine (SVM) has gained considerable attention for its capability to manage high-dimensional feature spaces and construct optimal decision boundaries through the determination of an appropriate separating hyperplane. Empirical evidence from prior studies suggests that SVM demonstrates consistent and competitive performance when compared with alternative methods, including Naïve Bayes, across diverse application domains such as healthcare discussions, labor-related issues, regional election discourse, and evaluations of political figures [13], [14], [15], [16], [17]. Previous research on sentiment analysis in public regulations shows that although probabilistic methods such as Naïve Bayes are effective and computationally efficient for text classification, approaches such as SVM are still being explored to optimize class separation in complex datasets [18]. The effectiveness of SVM, however, remains influenced by data characteristics and the preprocessing stages applied.

Although sentiment analysis studies on public figures have been widely conducted, research that specifically analyzes public opinion on the performance of Finance Minister Purbaya Yudhi Sadewa using X social media data is still limited. Previous studies have mostly utilized platforms such as Instagram and YouTube with different methodological approaches [19]. In light of these considerations, this study focuses on identifying patterns of public sentiment regarding the performance of the Minister of Finance of the Republic of Indonesia during the 2025–2029 period, while simultaneously examining the classification capability of the SVM algorithm using data derived from social media X. The results are intended to provide empirical evidence on public perception and contribute to the development of academic research on sentiment analysis involving public figures in the Indonesian context.

II. METHOD

A. *Research Stages*

This research was conducted through a series of stages that were systematically designed to produce accurate and structured sentiment classification. The research process began with data collection, followed by data cleaning, text preprocessing, sentiment labeling, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), classification using the Support Vector Machine (SVM) algorithm, and model performance evaluation. All stages were arranged sequentially to maintain process consistency and ensure the validity of the research results [2], [17]. An overview of the research flow is shown in Figure 1.

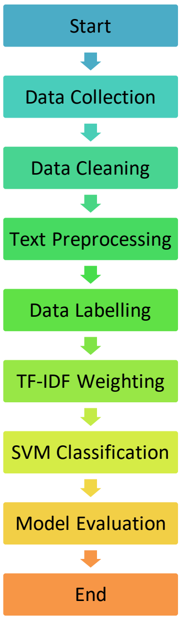

Figure 1. Research Flow Diagram

B. *Data Collection*

Research data were obtained from the social media platform X (formerly Twitter) using web scraping techniques implemented in a Google Colab environment with the Python programming language. The data collection process did not utilize the official X API; instead, publicly accessible posts were extracted based on predefined search parameters. Data extraction was conducted from September to November 2025, representing the initial phase of fiscal policy implementation following the leadership transition at the Ministry of Finance of the Republic of Indonesia.

The scraping process applied a structured keyword query to ensure contextual relevance. The search query used was:

"Purbaya" AND ("kinerja" OR "kerja" OR "menteri" OR "menkeu" OR "ekonomi" OR "investasi")

This query was designed to capture public discourse specifically related to performance and fiscal policy issues associated with the Minister of Finance. The dataset included tweet text and posting time information.

To maintain ethical research standards and protect user privacy, only publicly available posts were collected, and all personal identifiers such as usernames, user IDs, and location metadata were removed prior to analysis.

A total of 1,164 documents were declared valid after preprocessing. Although relatively modest in scale compared to large-scale social media studies, this dataset represents public opinion during a critical and narrowly defined transition period. The focused temporal scope ensures contextual relevance to early fiscal policy discourse, thereby capturing concentrated sentiment dynamics without contamination from unrelated political events.

To reduce potential bias from automated or bot accounts, additional filtering was applied during preprocessing. Accounts exhibiting abnormal posting frequency indicative of spamming behavior were excluded. This step was implemented to enhance data quality and maintain a human-centered representation of public opinion.

C. Data Cleaning

The data cleaning stage aims to select and retain attributes that are relevant to sentiment analysis. At this stage, data columns that do not directly contribute to determining sentiment orientation are eliminated so that the dataset focuses on comment text as the main corpus of the study. This step is taken to improve data processing efficiency and minimize interference from irrelevant attributes that could potentially affect the performance of the classification model [20].

D. Text Preprocessing

Text preprocessing is done to convert raw data into a more structured format that is ready for use in the sentiment classification process. The preprocessing stages follow the standard Indonesian text processing procedures commonly applied in similar studies [21], [22]. These stages include:

1. Noise Removal, which is the removal of non-text elements such as URLs, mentions, hashtags, numbers, symbols, emojis, duplicate data, and empty entries.
2. Case Folding, which converts all characters to lowercase to standardize the text format.
3. Normalization, which adjusts non-standard words, abbreviations, acronyms, and slang terms to their standard forms in accordance with the rules of Indonesian Spelling (EBI).
4. Tokenization, which breaks down text into words or tokens.
5. Stopword Removal, to eliminate common words that do not contribute significantly to the meaning of sentiment.
6. Stemming reduces morphological variation by returning inflected words to their base form.

E. Data Labeling

Text preprocessed text data were labeled into two sentiment categories: positive and negative. The labeling process was conducted manually (manual labeling) by the researcher to ensure contextual accuracy, particularly given the frequent use of sarcasm, informal language, and economic terminology in Indonesian social media discourse.

Sentiment categories were defined as follows:

- Positive: expressions of appreciation, support for fiscal policies, or optimistic expectations toward leadership performance.
- Negative: criticism related to taxation, national debt, dissatisfaction with performance, or pessimistic evaluations.

To ensure validity and reliability, labeling was guided by a structured annotation guideline. A subset of the labeled data was re-evaluated through cross-checking to minimize subjectivity and maintain labeling consistency.

F. TF-IDF Feature Extraction

Feature extraction was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method to convert text data into numerical representations. This method weights words based on their frequency of occurrence in documents and their level of importance across the corpus. The application of TF-IDF aims to highlight words that have high information value while suppressing the influence of general words that are less informative [3], [16]. The feature extraction dataset is then divided into training data and test data with a ratio of 80:20.

G. Classification Using SVM

The sentiment classification process was carried out using the Support Vector Machine (SVM) algorithm. This study applied a linear kernel to determine the optimal hyperplane that separates positive and negative sentiment classes. The selection of a linear kernel was based on the characteristics of high-dimensional text data generated from the TF-IDF feature extraction process. The SVM parameters are optimally configured to improve the model's generalization ability in sentiment classification tasks [13], [15].

H. Model Evaluation

The final stage of the research focuses on evaluating the performance of the sentiment classification model. The evaluation is carried out by utilizing a confusion matrix through a comparison between the model's predicted labels and the actual labels. Based on this matrix, model performance was measured using several evaluation metrics, including accuracy, precision, recall, and F1-score. These metrics were used to provide a comprehensive overview of the reliability and effectiveness of the model in classifying public sentiment [6], [23].

III. RESULT AND DISCUSSION

This section describes in detail the stages of data analysis, from data collection to the evaluation of classification model performance. The results are presented narratively so that readers can understand the relationship between the research context, the methods applied, and the implications of the findings. Quantitative evidence in the form of tables backs up each subsection, and images are used to describe data visualization.

A. Data Collection

Data collection was conducted through the social media platform X using the keywords “Minister of Finance” and “Purbaya Yudhi Sadewa” during the period from September to November 2025. This process yielded 1,474 raw tweets representing the public's spontaneous response to fiscal policy issues. The collected dataset reflects a broad spectrum of opinions, ranging from criticism of fiscal policy to support for the Minister of Finance's performance.

As shown in Table I, the sample tweets obtained show various public opinions, from casual comments to sharp criticism, including comparisons with previous policy periods. The raw data is the starting point for sentiment analysis and provides an initial overview of the patterns of responses that naturally emerge on social media.

The diversity of opinions in the dataset is an important factor in ensuring data representativeness. By involving many users and different time ranges, this study is able to map public sentiment trends more comprehensively. In addition, this process produces an initial database that is ready for cleaning and preprocessing so that sentiment analysis models can be trained using quality data.

TABLE I
EXAMPLE OR RAW DATA FROM CRAWLING

created_at	full_text
Fri Oct 24 00:40:39 +0000 2025	Jika Menkeu Purbaya dikatakan Presiden @prabowo Menteri koboy fenomenal berarti matanya selama ini tertutup kekuasaan rakyat dukung Purbaya krn sdh lama menduga org2 / kasus2 tertentu yg sebagian ada disekitarnya menghambat kemajuan neg bangunlah Tuama so siang ini !!!
Thu Oct 23 23:59:15 +0000 2025	setiap hari denger dongeng cerita menteri pak purbaya
Thu Oct 23 23:46:44 +0000 2025	Arek2 ini apa tidak tahu ya bahwa pak Menkeu Purbaya itu Termul? Militan pula. Pagi all// #JumatBerkah
Thu Oct 23 23:36:50 +0000 2025	Dedi Mulyadi dan Purbaya seru-seruan! Semoga ini bikin kita lebih peka sama tantangan ekonomi yuk dukung pemerintah untuk terus maju!

B. Data Cleaning

The data cleaning stage is carried out to simplify the dataset structure while ensuring that all documents are relevant to the sentiment analysis objective. Various administrative columns and interaction metrics that are not directly related to sentiment orientation, such as conversation_id_str, created_at, favorite_count, retweet_count, username, and other similar attributes, are removed. At this stage, only the full_text column is retained as the basis for analysis.

The cleaning process makes the dataset smaller and more focused on user opinion content. The cleaned dataset is shown in Figure 2, which shows the number of tweets ready for further processing. In addition to improving the efficiency of the analysis, data cleaning also plays a role in reducing potential interference from noise or outliers, such as spam and irrelevant comments, so that data quality is maintained and the analysis results are more reliable.

	full_text
0	Kenapa menteri menteri di kabinet merah putih ...
1	@AlbertKusen12 @susno2g Tidak legowo terhadap ...
2	DPR tidak perlu nyogok Masyarakat. Rakyat suda...
3	@CNNIndonesia Pak Menkeu doktor Purbaya Ku jam...
4	@regar_op0sisi Luhut dapat lawan sepadan... se...
...	...
1469	@dhaintly Tingginya beda woy Purbaya tinggi ge...
1470	Gak IG gak X sampe Yutub isinya muka pak Purba...
1471	Indikasi negara bobrok yaitu kita lihat sj per...
1472	Kalau purbaya kerja bener cukup diawasi konsis...
1473	Gapernah se aware ini soal ekonomi negara seme...

1474 rows × 1 columns

Figure 2. Data Cleaning Results

C. Text Preprocessing

The cleaned dataset then undergoes text preprocessing. This stage includes removing noise in the form of URLs, symbols, and special characters; filtering duplicate and empty data; case folding; normalizing non-standard words; removing stop stopwords, tokenization; and stemming. Through this series of processes, 1,164 valid documents were obtained, ready for use in analysis.

Examples of text transformation before and after preprocessing are presented in Table II. These changes show

how raw text is simplified into cleaner tokens without losing its contextual meaning. Thus, data complexity can be reduced while maintaining the accuracy of sentiment analysis.

TABLE II
TEXT TRANSFORMATION BEFORE AND AFTER PREPROCESSING

No	Raw Text	Preprocessed
1	Kenapa menteri menteri di kabinet merah putih tidak ada yang bisa tampil seperti Purbaya bernarasi melawan oligarki dan kebohongan	menteri kabinet merah putih
2	@AlbertKusen12 @susno2g Tidak legowo terhadap kritik Menkeu Purbaya berisiko membuat KDM terlihat defensif yang bisa meredupkan daya tarik konten viralnya di mata publik karena menimbulkan kesan ada yang ditutupi. Isu giro rugi karena dana Rp4 1T tak dapat bunga seperti deposito hilang potensi	tidak legowo kritik menkeu
3	DPR tidak perlu nyogok Masyarakat. Rakyat sudah pintar dan faham Itu Uang Rakyat sendiri. Rakyat tidak perlu receh-receh sogokan begituan. Rakyat perlu Lapangan Kerja dan Pembersihan para Pejabat Bandhit disegerakan. Maju terus pak Purbaya https://t.co/GbnasBD6sD	dpr tidak perlu sogok masyarakat rakyat
4	@CNNIndonesia Pak Menkeu doktor Purbaya Ku jamin Indonesia tidak akan pernah jadi bagus Jika tidak mau berubah ke negara industri/ Negara Produsen Ayo pak doktor Purbaya fokus bangkitkan industri Metal / peleburan besi baja nikel dsb Industri Non Metal termasuk kimia	pak menkeu doktor purbaya jamin indonesia

The distribution of documents after preprocessing is visualized in Figure 3. The visualization shows the proportion of clean documents compared to raw data and confirms the importance of this stage in ensuring that the resulting TF-IDF features are not distorted by irrelevant elements.

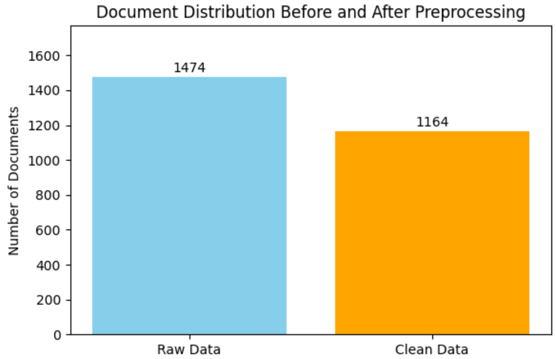

Figure 3. Text Preprocessing Results

D. Data Labeling

After preprocessing, documents were manually labeled into two sentiment classes: positive and negative. These two categories were chosen to emphasize clear polarity of opinion and reduce ambiguity, especially given the relatively limited size of the dataset. Manual labeling allowed researchers to understand the linguistic context, including sarcasm and implicit meaning, thereby maintaining the quality of the ground truth.

Several examples of labeled data are shown in Table III. This labeled data is then used as the basis for training the SVM model to be able to recognize public sentiment patterns more accurately.

TABLE III
EXAMPLES OF DATA AFTER SENTIMEN LABELING

No	text	label
1	menteri kabinet merah putih ada bisa tampil purbaya narasi lawan oligarki bohong	Negatif
2	legawa kritik menteri uang purbaya risiko kepala daerah defensif bisa redup daya tarik konten viralnya mata publik timbul kesan yang tutup isu giro rugi dana triliun rupiah dapat bunga deposito hilang potensi	Negatif
3	dewan wakil rakyat sogok masyarakat rakyat pintar paham uang rakyat rakyat perlu receh suap itu rakyat perlu lapang kerja bersih jabat bandit segera maju pak purbaya	Negatif
4	menteri uang doktor purbaya jamin indonesia akan menteri uang bagus tidak ubah negara industri negara produsen ayo doktor purbaya fokus bangkit industri logam lebur besi baja nikel bagai industri nonlogam kimia	Negatif

Analysis of the sentiment class distribution shows a dominance of negative comments, indicating a public tendency to criticize fiscal policy. This distribution is shown in Figure 4 and reinforces the interpretation of public opinion patterns.

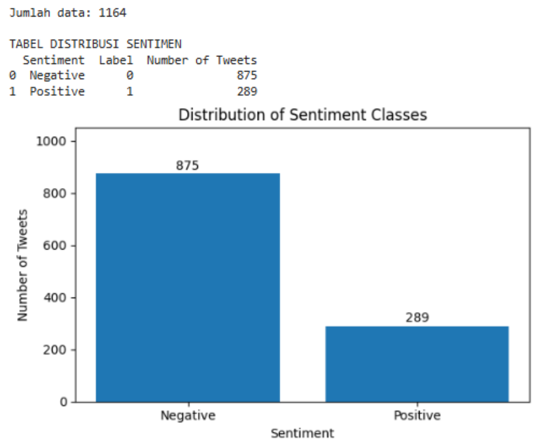

Figure 4. Data Labeling Results

E. TF-IDF Feature Extraction

The labeled documents are then converted into numerical form using the Term Frequency–Inverse Document Frequency (TF-IDF) method. This method emphasizes words that appear frequently in one document but rarely appear in other documents, thereby identifying terms with high information value.

The distribution of word weights in the corpus is visualized in Figure 5, which shows the dominant keywords in public comments. This stage produces features that are relevant to the SVM model while also assisting in the interpretation of public opinion patterns.

	Fitur (Kata/Bigram)	Bobot Rata-rata TF-IDF
7666	purbaya	0.045847
5746	menteri	0.032490
9859	uang	0.027843
5849	menteri uang	0.023182
6773	pak	0.020708
...	...	...
7452	prediksi ekonomi	0.000100
8589	saja saya	0.000100
1255	bodoh prediksi	0.000100
7222	pesimistis anda	0.000100
7221	pesimistis	0.000100

Figure 5. Visualization of TF-IDF Feature Weight Distribution

F. Classification Using SVM

TF-IDF feature vectors are used as input for the Support Vector Machine (SVM) algorithm with a linear kernel. This kernel was chosen based on its effectiveness in handling high-dimensional and sparse text data. To overcome class imbalance in the dataset, class weighting with a ratio of {0: 1.0, 1: 5.0} was applied so that the minority class received greater attention during the learning process.

This classification stage aims to build a basic model capable of separating positive and negative sentiment classes based on text patterns. The SVM model attempts to form an optimal hyperplane that maximizes the distance between classes, thereby improving generalization capabilities. The results of this stage form the basis for further optimization through hyperparameter adjustment.

G. Hyperparameter Adjustment

SVM model performance optimization was carried out by adjusting the regularization hyperparameter (C). Several C values, namely 0.1, 1.0, 5.0, and 10.0, were tested to see their effect on accuracy and F1-score values. The test results are presented in Table IV.

Based on Table IV, the configuration with a C value of 5.0 produced the best performance, as indicated by the highest accuracy and F1-score. A C value that is too small causes underfitting, while a larger value actually reduces performance due to over-regularization. Therefore, C = 5.0 was selected as the optimal parameter and used in the final evaluation stage.

TABLE IV
SVM HYPERPARAMETER TUNING RESULTS

C Value	Accuracy	F1-Score	Description
0.1	26.6%	0.1342	Underfitting
1.0	85.84%	0.8520	Good Fit
5.0	86.27%	0.8559	**Optimal (Best Fit)**
10.0	84.98%	0.8442	Performance Drop

H. Model Evaluation

Model evaluation was performed using 233 test data with optimal parameters (C = 5.0). Performance measurement was carried out using a confusion matrix to compare predicted labels with actual labels, followed by the calculation of precision, recall, F1 score, accuracy, and Area Under the Curve (AUC).

The confusion matrix provides a detailed representation of the classification performance for each sentiment class. As shown in Figure 6, the model successfully classified 166 negative tweets and 35 positive tweets. However, 9 negative tweets were misclassified as positive (false positives), and 23 positive tweets were misclassified as negative (false negatives). This indicates that the model performs well in identifying negative sentiment but has greater difficulty in detecting positive sentiment, especially when tweets contain implicit expressions or sarcasm.

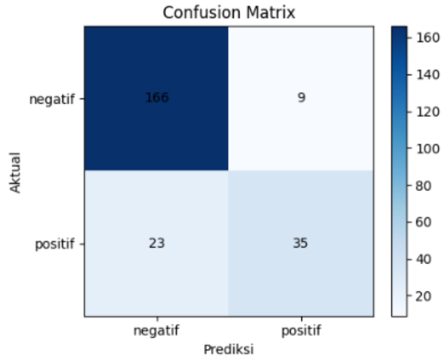

Figure 6. Confusion Matrix Visualization

Although the class distribution shows an imbalance (81.1% negative vs. 18.9% positive), this study still uses the original data to maintain the natural distribution of public opinion on social media X. To mitigate the risk of bias, the model evaluation did rely on accuracy and used the F1-score (0.86) and Area Under the Curve (AUC) metrics of 0.91. This high AUC value proves that the SVM model has excellent discrimination capabilities for distinguishing between the two sentiment classes, despite the imbalance in the amount of data.

In addition, class weighting with a ratio of {0: 1.0, 1: 5.0} was applied during training to provide greater learning emphasis on the minority class while maintaining the authenticity of the dataset distribution. The detailed classification metrics are presented in Table V.

TABLE V
SVM CLASSIFICATION REPORT

Sentiment Class	Precision	Recall	F1-Score	Support
Negative	0.88	0.95	0.91	175
Positive	0.80	0.60	0.69	58
Accuracy			0.86	233
Macro Avg	0.84	0.78	0.80	233
Weighted Avg	0.86	0.86	0.86	233

To further evaluate the relative performance of the proposed approach, a comparative experiment was conducted using three additional classification algorithms under identical preprocessing and TF-IDF feature extraction settings: logistic regression, naïve Bayes, and K-nearest neighbor (KNN). The results are summarized below:

TABLE VI
COMPARISON OF MODEL PERFORMANCE

Model	Accuracy	Precision	Recall	F1-Score	AUC
SVM (Linear)	0.8627	0.8577	0.8627	0.8559	0.9063
Logistic Regression	0.8412	0.8372	0.8412	0.8387	0.8791
Naïve Bayes	0.7725	0.8254	0.7725	0.6918	0.6920
KNN	0.7725	0.7455	0.7725	0.7403	0.6938

The comparison shows that SVM achieved the highest overall accuracy and F1-score among the evaluated models. Logistic regression demonstrated competitive performance but produced slightly lower discrimination capability, as reflected in its AUC value. Naïve Bayes and KNN showed comparatively weaker performance, particularly in F1-score and AUC metrics, indicating limitations in handling feature dependency and class imbalance within the dataset.

Consistent with prior studies comparing SVM and Naïve Bayes in ministerial sentiment analysis [13], the present findings confirm that SVM provides better classification accuracy and more stable performance. The margin maximization mechanism of SVM enables better separation of sentiment classes in high-dimensional feature space, making it a suitable and robust approach for analyzing fiscal policy discourse on social media.

These results show that SVM does a better job of classifying high-dimensional and sparse TF-IDF text features, which is in line with previous research on sentiment analysis in ministerial discourse. Therefore, the SVM model was selected as the primary classifier for the full dataset.

The ROC curve further confirms the model's strong discrimination capability with an AUC value of 0.91.

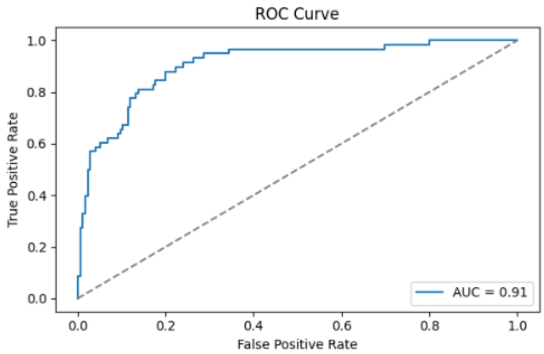

Figure 7. Model Evaluation Results

I. Prediction Results

The validated SVM model (C = 5.0) was then applied to classify the entire dataset. The prediction results show that negative sentiment dominated 81.1% of the total tweets, while positive sentiment accounted for 18.9%. This imbalance reflects the dominance of critical discourse during the fiscal leadership transition period.

Maintaining the natural distribution of data reinforces the ecological validity of the findings, as the results represent authentic patterns of public opinion rather than artificially balanced data.

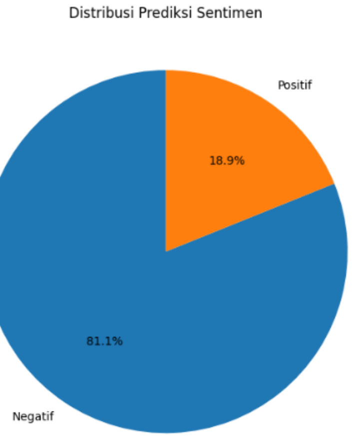

Figure 8. Prediction Results

J. Sentiment Interpretation

The dominance of negative sentiment suggests that concerns over the direction of economic policy and fiscal management primarily influence public discourse. This is reinforced by the word cloud visualization in Figure 8, which displays high-frequency words such as "people," "money," "taxes," and "debt" in negative-toned tweets.

Figure 9. Sentiment Interpretation Results

Conversely, tweets with positive sentiment highlight terms such as “support,” “trust,” and “work,” reflecting that there are groups within society who still have confidence in fiscal leadership. Comparative evaluations of policies, not mere knee-jerk reactions, influence public opinion, as evidenced by the emergence of names from previous administrations. These findings demonstrate that perceptions of policy outcomes shape public sentiment.

IV. Conclusion

This study successfully describes the pattern of public sentiment towards the performance of the Minister of Finance of the Republic of Indonesia, Purbaya Yudhi Sadewa, in the early phase of the 2025–2029 administration. The X platform's social media data analysis reveals that negative sentiment dominates public opinion, accounting for 81.1%. This dominance reflects the high level of public attention and concern regarding issues of economic stability, national debt management, and taxation policy. On the other hand, positive sentiment was recorded at 18.9% and was generally related to expressions of confidence and hope for the direction of the new fiscal leadership. From a methodological perspective, the Support Vector Machine (SVM) algorithm demonstrates solid performance in classifying unstructured text data. The application of a linear kernel combined with regularization parameter optimization at a value of C = 5.0 resulted in an accuracy rate of 86.27% and an Area Under the Curve (AUC) value of 0.91, indicating excellent sentiment class separation capabilities. Although the recall rate for negative sentiment is relatively high (0.95), the model still faces limitations in recognizing positive sentiment that is conveyed sarcastically or implicitly. Overall, these findings confirm that social media-based sentiment analysis can be used as a strategic tool to understand public perception and trust levels, while also demonstrating the relevance of strengthening fiscal policy communication in responding to the dynamics of public opinion.

References

[1] Mr. Pratama Srg Sistem Informasi and S. Kaputama, “Analisis Sentimen Pada Media Sosial Twitter Terhadap Penegakan Hukum Di Indonesia Tahun 2024 Menggunakan Metode Support Vector Machine (SVM).” [Online]. Available: https://journaledutech.com/index.php/great

[2] L. Rohmatun and A. Baita, “Machine Learning-Based Sentiment Analysis on Twitter (X): A Case Study of the ‘Kabur Aja Dulu’ Issue Using SVM,” 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[3] D. W. Syahputra, B. Rahayudi, and L. Muflikhah, “Analisis Sentimen Twitter terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat menggunakan Metode Support Vector Machine,” 2022. [Online]. Available: http://j-ptiik.ub.ac.id`

[4] R. N. Mauliza and Y. R. Sipayung, “Penerapan Text Mining Dalam Menganalisis Pendapat Masyarakat Terhadap Pemilu 2024 Pada Media Sosial X Menggunakan Metode Naive Bayes,” *Technomedia Journal*, vol. 9, no. 1, pp. 1–16, Feb. 2024, doi: 10.33050/tmj.v9i1.2212.

[5] M. Rizki Maulana and R. A. Putri, “Sistemasi: Jurnal Sistem Informasi Analisis Sentimen terhadap Komisi Pemilihan Umum 2024 di Indonesia melalui Twiter menggunakan Algoritma Support Vector Machine (SVM) Sentiment Analysis of the 2024 General Election Commission in Indonesia through Twiter using the Support Vector Machine (SVM) Algorithm.” [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[6] F. Fatkhurrohman, B. I. Nugroho, and N. Fadillah, “Analisis Sentimen Program Makan Bergizi Gratis Pemerintah RI Melalui Twitter Menggunakan Metode SVM,” *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 3, pp. 3906–3917, Aug. 2025, doi: 10.31004/riggs.v4i3.2533.

[7] P. T. Sejati, F. Al Zami, A. Marjuni, H. Indrayani, and I. D. Puspitarini, “Aspect-Based Sentiment Analysis for Enhanced Understanding of ‘Kemenkeu’ Tweets,” 2024. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[8] S. Berliani and S. Lestari, “Analisis Sentimen Masyarakat Terhadap Isu Pecat Sri Mulyani Pada Twitter Menggunakan Metode Naive Bayes Dan Support Vector Machine,” *Jurnal Sains dan Teknologi*, vol. 5, no. 3, pp. 951–960, Apr. 2024, doi: 10.55338/saintek.v5i3.2746.

[9] S. Ramadhani Putri *et al.*, “Sistemasi: Jurnal Sistem Informasi Analisis Sentimen Publik terhadap Nadiem Makarim sebagai Mendikbudrisktek menggunakan Support Vector Machine (SVM).” [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[10] A. A. Purnama and Y. R. Sipayung, “Sentiment Analysis of Public Service Using Naïve Bayes Classifier,” *Journal of Information Systems and Informatics*, vol. 7, no. 3, pp. 2439–2457, Sep. 2025, doi: 10.51519/journalisi.v7i3.1207.

[11] A. Dian Oktavianingsih, A. Fadlil Fauzi, J. Immanuel, O. DSimatupang, F. Amsury, and R. Fahlapi, “Analisis Sentimen Komentar YouTube terhadap Pengangkatan Menkeu Purbaya Menggunakan Algoritma SVM dan Naïve Bayes,” *Jurnal Cendekia Ilmiah*, vol. 5, no. 1, 2025.

[12] M. A. Gumilang, F. Abdillah, M. Y. Amin, and M. Hasan, “Sentiment Analysis of Indonesian Ministries Social Media: Citizen Responses Utilizing TextBlob Analyser,” *Jurnal Sosioteknologi*, vol. 23, no. 2, pp. 203–216, Aug. 2024, doi: 10.5614/sostek.itbj.2024.23.2.5.

[13] R. Pebrianto, S. N. Nugraha, A. Latif, and M. R. Firdaus, “Analisis Sentimen Twitter Terhadap Menteri Indonesia Dengan Algoritma Support Vector Machine Dan Naive Bayes,” vol. 17, pp. 1–12, 2022.

[14] G. Forda Nama, A. H. Shalihah, B. Wintoro, Y. Mulyani, and D. Despa, “Implementation of Naïve Bayes Classifier & Support Vector Machine Algorithm for Sentiment Classification using Twitter Data on Indonesian Presidential Candidates In 2024,” 2024. [Online]. Available: https://www.jisem-journal.com/

[15] A. Sentimen *et al.*, "Analisis Sentimen Masyarakat Di Media Sosial X Terhadap Kemenkes Dengan Naive Bayes dan SVM," *Jurnal Sains dan Teknologi*, vol. 7, no. 1, pp. 1–6, 2025, doi: 10.55338/saintek.v7i1.4615.

[16] A. M. AlHakim, P. D. Atika, and H. Herlawati, "Analisis Sentimen Masyarakat Terhadap PHK di Indonesia Pada Twitter Menggunakan Naïve Bayes dan Support Vector Machine (SVM)," *Journal of Students' Research in Computer Science*, vol. 6, no. 1, pp. 113–124, May 2025, doi: 10.31599/96sfw544.

[17] D. Shinta Angreni, S. Khaerawati Nur, and D. Luvi Jayanto, "Sentiment Analysis for the 2024 DKI Jakarta Gubernatorial Election Using a Support Vector Machine Approach," 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[18] D. Fajar Fiandaru and Y. Romando Sipayung, "ELSE (Elementary School Education Journal) Penerapan Metode Klasifikasi Naive Bayes untuk Analisis Sentimen terhadap Undang-Undang ITE di Media Sosial Twitter," 2025, [Online]. Available: https://journal.um-surabaya.ac.id/index.php/pgsd

[19] A. K. Author, "Analisis Sentimen Di Instagram Terhadap Menteri Keuangan Purbaya Yudhi Sadewa Menggunakan Metode Logistic Regression," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3S1, Oct. 2025, doi: 10.23960/jitet.v13i3S1.8002.

[20] 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, pp. 469–482, Apr. 2025, doi: 10.52436/1.jutif.2025.6.2.4068.

[21] I. Prayugah, U. Indahyanti, and N. Ariyanti, "Analisis Sentimen Publik Atas Respons Pemerintah Pada Serangan Ransomware Dengan Pendekatan Machine Learning Dan Smote," *JOISIE (Journal Of Information Systems And Informatics Engineering)*, vol. 8, no. 2, p. 333, Jan. 2025, doi: 10.35145/joisie.v8i2.4764.

[22] J. Br Sitepu, S. Ramadani, M. Khadapi, S. Informasi, and S. Kaputama, "Analisis Sentimen Pengguna Media Twitter Terhadap Kurikulum Merdeka Menggunakan Algoritma Support Vector Machine (SVM)." [Online]. Available: https://journaledutech.com/index.php/great

[23] T. Saputro and A. Hermawan, "Comparison of Naive Bayes and Support Vector Machine in Sentiment Analysis of Siwaslu Application," *bit-Tech*, vol. 8, no. 2, pp. 2004–2015, Dec. 2025, doi: 10.32877/bt.v8i2.3184.