RESEARCH ARTICLE OPEN ACCESS

ROS-Enhanced Sentiment Analysis of Social Media Discourse on Indonesia's 2025 Budget Efficiency Policy Using Support Vector Machine

Nadinta Kasih Amalia Suryono[1*], Toha Saifudin[2], Marisa Rifada[3], M. Fariz Fadhillah Mardianto[4]
[1,2,3,4]Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia
[1*]Email: nadinta.kasih.amalia-2022@fst.unair.ac.id

--************************----------------------------------

Abstract:

The 2025 State Budget (APBN) is a key instrument for the Indonesian government to sustain economic stability and support national development. Persistent challenges such as deficits, low absorption, and inefficiency prompted the issuance of Presidential Instruction No. 1 of 2025 on budget efficiency. While the policy seeks to reduce non-productive expenditures, it has also raised concerns about declining public service quality, sparking public debate. This study examines public sentiment toward the policy by analyzing 546 comments from the social media platform X, categorized into positive and negative classes. The methodological framework includes text preprocessing, TF-IDF weighting, and sentiment classification using Support Vector Machine (SVM). To address class imbalance, Random Oversampling (ROS) was applied to the training dataset under an 80:20 train-test split. Results show that SVM achieved strong performance on the original dataset (F1 Score 83.12%, Accuracy 87.04%), and its metrics remained relatively stable after ROS (F1 Score 81.64%, Accuracy 85.19%), indicating robustness to imbalance. These findings confirm that SVM provides consistent performance for sentiment analysis in policy-related discourse, while highlighting that oversampling yields minimal benefit. The study offers empirical insights into public responses to budget efficiency measures and methodological approaches for handling imbalanced data.

Keywords — **Sentiment Analysis, Random Oversampling, Machine Learning, X (Twitter).**

--************************----------------------------------

I. INTRODUCTION

In 2025 Indonesia faces a critical fiscal challenge in improving expenditure efficiency amid global economic uncertainty. Persistent budget deficits and low absorption rates prompted the government to issue Presidential Instruction No. 1 of 2025 which mandates the reduction of nonproductive spending and the strengthening of efficiency across ministries and agencies [1], [2]Policymakers present this directive as an essential reform to safeguard fiscal sustainability. Public responses remain divided. Supporters argue that efficiency measures are vital to protecting national finances. Critics caution that indiscriminate cuts may undermine public service delivery if transparency and prioritization are not ensured [3], [4].

These debates increasingly unfold on social media. With more than 25 million active users in Indonesia, X (formerly Twitter) has become a digital town square where citizens express frustrations, hopes, and critiques of government policy [5].

Researchers are turning to machine learning to capture this public pulse. Support Vector Machine algorithm has shown strong performance in classifying opinions into positive and negative categories, offering a data-driven lens into public discourse[6]Previous studies applying these methods to fiscal policy debates achieved promising accuracy rates, yet most relied on a single algorithm [7], [8].

This study addresses that gap by applying Support Vector Machine to sentiment data from X and enhanced with Random Oversampling to balance the dataset. By comparing the performance of these models, the research contributes to methodological advances in sentiment analysis while also providing insights into how Indonesians perceive the government's efficiency drive. In doing so, it bridges fiscal policy and artificial intelligence, showing how digital voices can inform smarter governance.

II. MATERIALS

The research employs a text mining methodology organized into sequential phases. Data were drawn from user comments on the social media platform X (previously known as Twitter) spanning January to October 2025. Sampling concentrated on discussions surrounding Indonesia's fiscal efficiency initiative as stipulated in Presidential Instruction No. 1 of 2025. To capture relevant content, keywords such as "efisiensi anggaran" and "efisiensi" were applied. The corpus was then partitioned into training and testing sets at an 80:20 proportion. The central analytical focus is sentiment polarity, classified into two categories: positive and negative.

III. METHOD

This study applies a structured text mining approach to analyze public sentiment. Figure I presents the flowchart broadly illustrating the procedure applied in this study to conduct sentiment analysis.

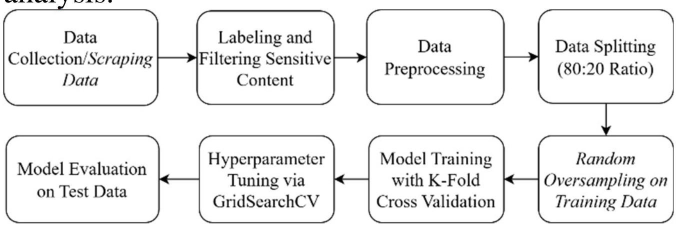

Fig. 1 Stages of Research

A. *Preprocessing Data*

Preprocessing steps were applied to prepare the textual data for classification. These included in the explanation below [9].

1) ***Cleaning:*** irrelevant elements such as URLs, emojis, numbers, and punctuation.
2) ***Normalization:*** to correct spelling variations.
3) ***Case folding:*** to convert all text into lowercase.
4) ***Stopword removal:*** to eliminate non-informative words.
5) ***Tokenization:*** to segment sentences into tokens.
6) ***Stemming:*** to reduce words to their root form.

B. *Random Oversampling*

Random Oversampling (ROS) is a commonly used oversampling technique in machine learning to address imbalanced datasets. The algorithm operates by randomly duplicating samples from the minority class within the training data, thereby increasing its size and balancing the overall class distribution [10].

C. *K-Fold Cross Validation*

Evaluating the performance of classification models is a critical step to ensure that predictive outcomes are both accurate and generalizable. Among the most widely adopted techniques in machine learning is K-Fold Cross Validation, which mitigates the bias associated with single data splits and provides a more stable estimate of model performance [11].

D. *Support Vector Machine*

Support Vector Machine (SVM) is a supervised learning algorithm widely applied in sentiment analysis due to its ability to construct optimal hyperplanes that separate classes in high dimensional spaces. In text mining, features are typically represented using Term Frequency-Inverse Document Frequency (TF-IDF), which assigns weights to words based on their frequency and rarity across documents [12]The TF-IDF formulation is expressed as [13]:

$$W_{(t,d)} = \mathrm{TF}_{(t,d)}.IDF_{(t)} \quad (1)$$

E. *Confusion Matrix*

A confusion matrix is a tabular representation that compares actual data with classification results to evaluate the performance of a model [14]. Table I presents the confusion matrix table.

TABLE I
CONFUSION MATRIX TABLE

Prediction	**Actual**	
	Positive	**Negative**
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

Model performance is assessed through several metrics derived from the confusion matrix. Accuracy measures the overall proportion of correct predictions. Recall evaluates the model's ability to identify positive cases, while Specificity reflects its capacity to recognize negatives. Precision indicates the correctness of positive predictions. To balance precision and recall, the F1-score is used as their harmonic mean, providing a more comprehensive measure of classification effectiveness.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$Specifity = \frac{TN}{FP+TN} \quad (4)$$

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

$$F1\ Score = 2 \times \frac{Presicion \times Recall}{Presicion+Recall} \quad (6)$$

IV. RESULT AND DISCUSSION

After the preprocessing stage, a total of 540 sentiment data points were obtained, representing public opinions in Indonesia regarding the 2025 budget efficiency issue. The distribution of sentiments is presented in Figure 2.

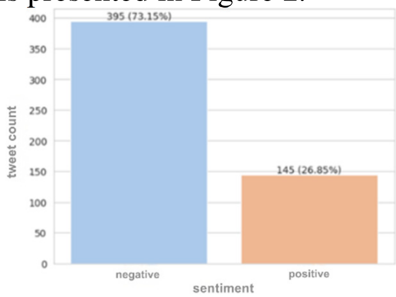

Fig. 2 Distribution of Sentiments

Based on Figure 2, public tweets (comments) predominantly express negative sentiment, accounting for 73.15%. This proportion indicates that public perception of the budget efficiency policy tends to be critical and generally unsupportive.

After the preprocessing stage, the resulting terms were obtained and prepared for further analysis based on Table 2 below.

TABLE 2
DATA AFTER PREPROCESSING

Actual Document	Terms after Preprocessing
Kebanyakan rapat...! Pak presiden memahami betul hal tersebut...... maka jawabannya adalah Efisiensi Anggaran	[banyak, rapat, presiden, paham, jawab]
@H******* Definisi memaksakan program gagal. Efisiensi anggaran dana pendidikan dialokasikan di MBG. Bayaran pegawai MBG gede sedangkan guru digerus. Ironi sih wkwkwk Indonesia emas tpi aspek pendidikan semua dikurangi wkwkwkk	[definisi, paksa, program, gagal, dana, pendidikan, mbg, pgawai, guru, habis, indonesia, emas, aspek, kurang]

At this stage, an imbalance was observed between positive and negative sentiment classes, indicating that the dataset was not evenly distributed shown in Figure 3.

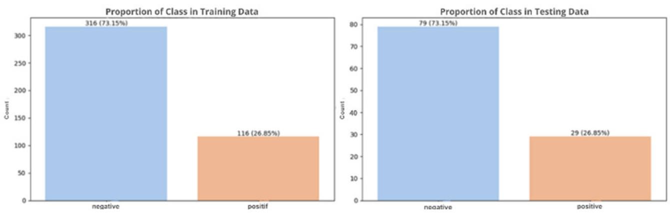

Fig. 3 Distribution Class from Training and Testing Data

This condition is critical to address, as it can affect the performance of classification models, particularly in accurately identifying minority classes.

The dataset was split into training and testing subsets using an 80:20 ratio. Class distribution remained imbalanced in both subsets, which risks biasing the model toward majority classes. To address this, ROS was applied to the training data, resulting in a more balanced sentiment distribution and improving the model's ability to learn from minority class instances.

Following data balancing, modeling was conducted using the Support Vector Machine (SVM) algorithm. Feature representation was based on Term Frequency–Inverse Document Frequency (TF-IDF) weighting for each word, with the results presented in a dedicated table.

TABLE 3
TF-IDF WEIGHT EACH TERMS

No	Term	TF-IDF
1	jawab	0.5106
2	banyak	0.4695
3	paham	0.4499
4	rapat	0.4205
⋮	⋮	⋮
1514	presiden	0.3734

The modeling process was further refined through hyperparameter optimization using GridSearchCV with a 10-fold cross validation scheme[15].

The best hyperparameter configurations identified for each algorithm are presented in the following table.

TABLE 4
DATA AFTER PREPROCESSING

Model	Hyperparameter	Value	Training Accuracy
SVM	C	4	95.27
	Coef	0	
	Kernel	Linier	
	γ	Auto	

After selecting the optimum hyperparameters, the next step was evaluating model performance on the testing dataset. This evaluation provides an objective measure of how well the tuned model generalizes to unseen data, with results presented in Table 5 below.

TABLE 5
DATA AFTER PREPROCESSING

Model	F1-Score	AUC	Accuracy
SVM	83.12	84.61	87.04
SVM with ROS	81.64	82.91	85.19

Table 5 summarizes the evaluation results of the classification models on the testing dataset. Without oversampling, the Support Vector Machine (SVM) achieved relatively high performance, with an F1-Score of 83.12% and accuracy of 87.04%, indicating balanced capability in recognizing both sentiment classes.

After applying Random Oversampling (ROS), the performance of SVM remained relatively stable, with only slight changes in metrics (F1-Score 81.64%, accuracy 85.19%), indicating that SVM not sensitive to class imbalance.

Overall, these results highlight that ROS insignificantly benefits to SVM because it is maintains consistent performance regardless of data balancing.

In addition to performance metrics, the evaluation results are further illustrated through the confusion matrix presented in Figure 4.

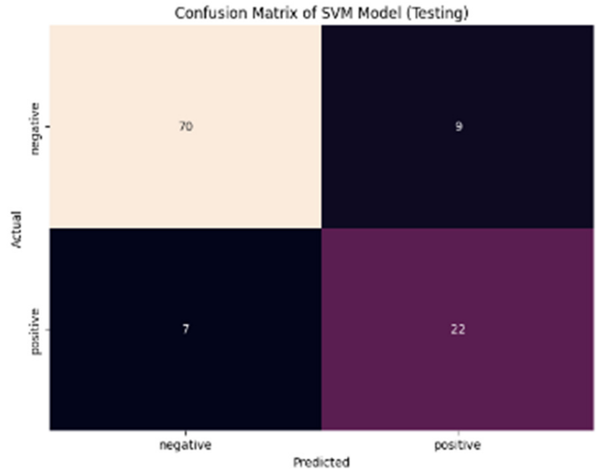

Fig. 4 Confusion Matrix

The visualization in Figure 5 and Figure 6 below presents a word cloud and a bar chart that illustrate the distribution of the most frequently appearing terms across all sentiments expressed by X (Twitter) users regarding the 2025 Budget Efficiency Policy.

Fig. 5 Word Cloud of Public Discourse on the 2025 Budget Efficiency Policy

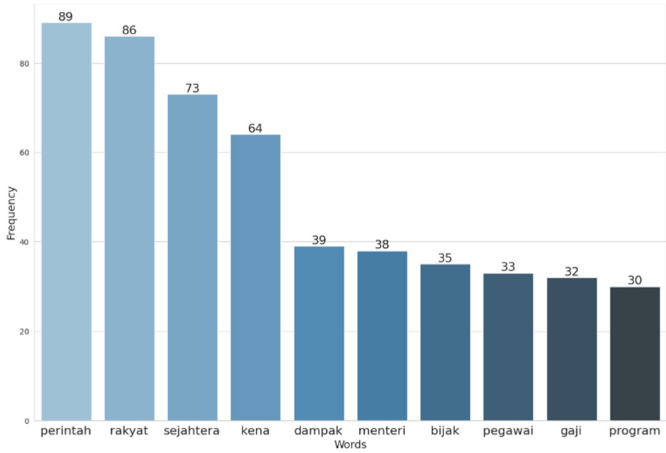

Fig. 6 Top 10 Terms Based on Frequency Related to the 2025 Budget Efficiency Policy

The analysis of word frequency shows that public conversations on X are dominated by several key terms that reflect users concerns regarding the 2025 Budget Efficiency Policy. The most frequent term is “perintah” (government/authority), indicating that discussions are strongly centered on governmental actions and decisions. This is followed by "rakyat" (citizens), which shows that users often frame the issue in relation to public welfare. The term "sejahtera" (prosperity) appears prominently, suggesting expectations for well-being that may be perceived as threatened by the policy. Meanwhile, "kena" (affected) reflects a sense of being impacted or burdened. The presence of "dampak" (impact) reinforces concerns about the consequences of the policy. The keyword "menteri" (minister) indicates that users associate responsibility with specific government officials. The term "bijak" (wise) suggests that users are evaluating whether policy decisions are considered prudent. Words such as "pegawai" (civil servants) and "gaji" (salary) highlight concerns about how the policy affects government employees and income stability. Finally, the term "program" (program) shows that users also discuss specific initiatives related to the policy.

V. CONCLUSIONS

The comparative evaluation of classification models for sentiment analysis on public discourse regarding the 2025 Budget Efficiency Policy demonstrates several key findings. SVM consistently achieved high accuracy and F1 Score, indicating robust and balanced performance in recognizing both sentiment classes, even under imbalanced data conditions. SVM's metrics remained stable before and after oversampling, suggesting its relative insensitivity to class imbalance, a finding echoed in other studies where SVM outperforms other probabilistic models and benefits less from oversampling techniques.

The analysis of word frequency further reveals that public sentiment is dominated by concerns about government actions, public welfare, and the direct impact of the policy, as evidenced by the prominence of terms such as "perintah," "rakyat," and "dampak." These insights underscore the importance of employing robust classification methods to ensure reliable and representative sentiment analysis outcomes. Overall, the use of SVM provides a consistent and effective approach for accurately capturing the nuances of public opinion in policy-related discussions on social media platforms.

ACKNOWLEDGMENT

The author extends sincere appreciation to the Faculty of Science and Technology, Universitas Airlangga, particularly the Department of Mathematics, for providing the facilities and academic support that made this sentiment analysis research possible.

REFERENCES

[1] (Kementerian Keuangan) Kemenkeu, "Akselerasi Pembangunan Ekonomi yang Inklusif dan Berkelanjutan," KERANGKA EKONOMI MAKRO DAN POKOK-POKOK KEBIJAKAN FISKAL TAHUN 2025, 2025. [Online]. Available: https://fiskal.kemenkeu.go.id/files/kemppkf/file/1716178459_final_kem_ppkf_2025_versi_publish.pdf

[2] Inpres, "Efisiensi Belanja dalam Pelaksanaan Anggaran Pendapatan dan Belanja Negara dan Anggaran Pendapatan dan Belanja Daerah Tahun Anggaran 2025," Instruksi Presiden (Inpres) Nomor 1 Tahun 2025.

[3] Suara.com, "Banyak Kritik Efisiensi Anggaran Ala Prabowo, Pengamat: Pemangkasan Bisa Berdampak Langsung ke Masyarakat."

[4] Sinergi Jatim, "Analisis Inpres Nomor 1 Tahun 2025 Efisiensi Anggaran atau Beban Baru bagi Pelayanan Publik?"

[5] Nouvan, "Negara dengan Pengguna X (Twitter) Terbanyak Juli 2025, Indonesia Urutan Keempat."

[6] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

[7] M. A. Alfaridzy, E. Haerani, and L. Oktavia, "KLASIFIKASI SENTIMEN MASYARAKAT TERHADAP EFISIENSI ANGGARAN PEMERINTAH MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER," *JUTECH: Journal Education and Technology*, vol. 6, no. 1, pp. 169–182, Jun. 2025.

[8] R. A. Hafika *et al.*, "PENERAPAN SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN TWITTER TERHADAP EFISIENSI ANGGARAN," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 4, Aug. 2025.

[9] J. C. Lapendy, A. A. C. Resky, A. Tenriola, D. F. Surianto, and U. S. Sidin, "Optimizing Sentiment Analysis of Electric Vehicles Through Oversampling Techniques on YouTube Comments," *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 14, no. 1, pp. 169–182, Mar. 2025, doi: 10.23887/janapati.v14i1.88205.

[10] R. D. Fitriani, H. Yasin, and Tarno, "PENANGANAN KLASIFIKASI KELAS DATA TIDAK SEIMBANG DENGAN RANDOM OVERSAMPLING PADA NAIVE BAYES (Studi Kasus: Status Peserta KB IUD di Kabupaten Kendal)," *Jurnal Gaussian*, vol. 10, no. 1, pp. 11–20, 2021, Accessed: Nov. 20, 2025. [Online]. Available: https://ejournal3.undip.ac.id/index.php/gaussian/

[11] V. Teodorescu and L. Obreja Brașoveanu, "Assessing the Validity of k-Fold Cross-Validation for Model Selection: Evidence from Bankruptcy Prediction Using Random Forest and XGBoost," *Computation*, vol. 13, no. 5, May 2025, doi: 10.3390/computation13050127.

[12] M. Q. H. Octava, D. G. Prasetyo Putri, F. M. Hilmy, U. Farooq, R. A. Nurhaliza, and A. Ganjar, "Web-based Sentiment Analysis System Using SVM and TF-IDF with Statistical Feature," in *2023 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 9–14. doi: 10.1109/3ICT60104.2023.10391734.

[13] F. R. Kareem and A. J. Obaid, "STFS_TF-IDF _HDSM: semantic and textual feature scaling using TF-IDF based sentiment analysis

with Hybrid Deep Sequential Model," *International Journal of Information Technology (Singapore)*, Jan. 2025, doi: 10.1007/s41870-024-02330-x.

[14] I. Markoulidakis and G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies (Basel)*, vol. 12, no. 7, Jul. 2024, doi: 10.3390/technologies12070113.

[15] I. K. Nti, O. Nyarko-Boateng, and J. Aning, "Performance of Machine Learning Algorithms with Different K Values in K-fold CrossValidation," *International Journal of Information Technology and Computer Science*, vol. 13, no. 6, pp. 61–71, Dec. 2021, doi: 10.5815/ijitcs.2021.06.05.