International Journal of Advances in Data and Information Systems
Vol. 6, No. 3, October 2025, pp. 763~772
ISSN: 2721-3056, DOI: 10.59395/ijadis.v6i3.1455

A Performance Enhancement Strategy for Sentiment Classification Models On Political Social Media Using Hyperparameter Tuning And Boosting

Andi Supriadi Chan[1], Meryatul Husna[2], Pandu Pratama Putra[3]
[1,2] Computer Engineering and Information, Politeknik Negeri Medan, Indonesia
[3]Department of Computer Engineering, Universitas Lancang Kuning, Indonesia

Article Info
Article history:
Received Jun 29, 2025
Revised Oct 23, 2025
Accepted Nov 17, 2025

Keywords:
Sentiment Analysis
Text Classification
Optuna
Machine Learning
XGBoost.

ABSTRACT
This study aims to develop an optimized machine learning-based sentiment classification model for election-related issues. A dataset comprising 10,001 entries was collected from the social media platform X and manually labeled into three sentiment classes: positive, negative, and neutral. The preprocessing stage involved text cleaning, stemming, and feature transformation using the Term Frequency-Inverse Document Frequency (TF-IDF) method. To address class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was employed. Three baseline classification algorithms—K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Gaussian Naive Bayes (GNB)—were initially evaluated to establish a performance benchmark. Model development proceeded by applying hyperparameter optimization using the Optuna framework and further enhancing the models via boosting with Extreme Gradient Boosting (XGBoost). Experimental results revealed that the combination of SVM with Optuna and XGBoost achieved the best performance, reaching 97% accuracy, precision, recall, and F1-score across all classes. In contrast, the KNN and GNB models experienced a notable decline in performance following hyperparameter tuning, although partial recovery was observed when combined with boosting. These findings suggest that hyperparameter tuning and boosting are not universally effective across all classifiers, yet their synergistic application significantly enhances performance in SVM-based models. This study highlights the importance of model-specific optimization strategies in building robust sentiment analysis systems, particularly for handling unbalanced public opinion data in social media contexts.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Andi Supriadi Chan,
Department of Computer Engineering and Information,
Politeknik Negeri Medan,
Jl. Almamater No 1, Medan, Indonesia.
Email: andisupriadi@polmed.ac.id

1. INTRODUCTION

The rapid growth of social media has led to a significant increase in the volume of digital data generated daily by users across various platforms [1]. Much of this data is unstructured, containing public opinions, sentiments, and social discussions that offer substantial potential for data-driven decision-making [2]. In this context, the classification of social media data has become essential for public issue monitoring, hoax detection, and mapping of public perceptions [3]. However, the main challenges in classifying social media data lie in its irregular text structure, high noise content, and severe class imbalance.

To address these challenges, this study employs three fundamental algorithms with different characteristics: K Nearest Neighbors (KNN), Support Vector Machine (SVM), and Gaussian Naive Bayes (GNB). KNN represents an instance-based method that relies on proximity between data points [4]. SVM is a margin-based approach that performs well with high dimensional data [5], while GNB is a probabilistic classifier known for its computational efficiency and is often used in basic text classification [6]. This diversity provides a strong foundation for a comprehensive evaluation of the effectiveness of balancing, tuning, and boosting techniques. To deal with class imbalance, the Synthetic Minority Over Sampling Technique (SMOTE) is applied. Unlike conventional oversampling methods that simply duplicate minority samples, SMOTE creates synthetic instances based on interpolation among the nearest neighbors in the minority class [7]. This allows the model to learn more representative patterns, reducing bias toward the majority class.

Subsequently, hyperparameter optimization is conducted using Optuna, a modern framework that utilizes the Tree Structured Parzen Estimator (TPE) to explore parameter space efficiently [8]. Compared to grid search, which is time consuming and resource intensive, Optuna adapts dynamically and identifies optimal parameter combinations with fewer computational costs [9]. Proper tuning not only improves accuracy but also enhances model robustness against data variability. To further enhance model performance, a boosting method is introduced using Extreme Gradient Boosting (XGBoost). XGBoost is known for its high efficiency, support for regularization to prevent overfitting, and ability to perform parallel processing. It is particularly well suited for complex data such as social media text, as it captures nonlinear relationships effectively [10].

Several previous studies have demonstrated the usefulness of classification methods for social media analysis, yet each has limitations. [11], used SVM and Naive Bayes for Twitter sentiment analysis and achieved an accuracy of 85% - 86%, but did not incorporate data balancing or parameter tuning. [12], applied KNN and Decision Tree for public opinion data and reported 76.5 % accuracy, but the model was highly sensitive to noise. [13], employed a Random Forest ensemble and reached 98.6% accuracy, though performance remained affected by imbalance. [14], used SMOTE with Naive Bayes and obtained a performance improvement of about 13%, though only one algorithm was evaluated. [15], implemented parameter tuning for SVM via grid search and achieved 89% accuracy, but the method required significant computation. [16], applied XGBoost to hoax detection and reported an F1 score of 0.69, but without balancing or baseline comparisons.

From these studies, it becomes clear that most prior research has not combined the three key strategies of data balancing, efficient hyperparameter tuning, and boosting in one integrated pipeline. Moreover, there is a lack of comprehensive comparisons across baseline, tuned, and boosted models on imbalanced social media datasets. In response to this gap, the present study aims to develop a classification pipeline that integrates SMOTE, Optuna, and XGBoost for three baseline models: KNN, SVM, and GNB. The evaluation is conducted in three phases: initial testing, post tuning, and post tuning with boosting. This approach is expected to yield classification models that are both accurate and fair when dealing with skewed data distributions commonly found in public opinion data on social media.

2. RESEARCH METHOD

This study adopts a systematic approach divided into four main stages: dataset preparation, model baseline evaluation, model development, and performance comparison. The goal of this pipeline is to construct and evaluate a text classification model for election-related issues by addressing data imbalance, optimizing model parameters, and implementing boosting techniques.

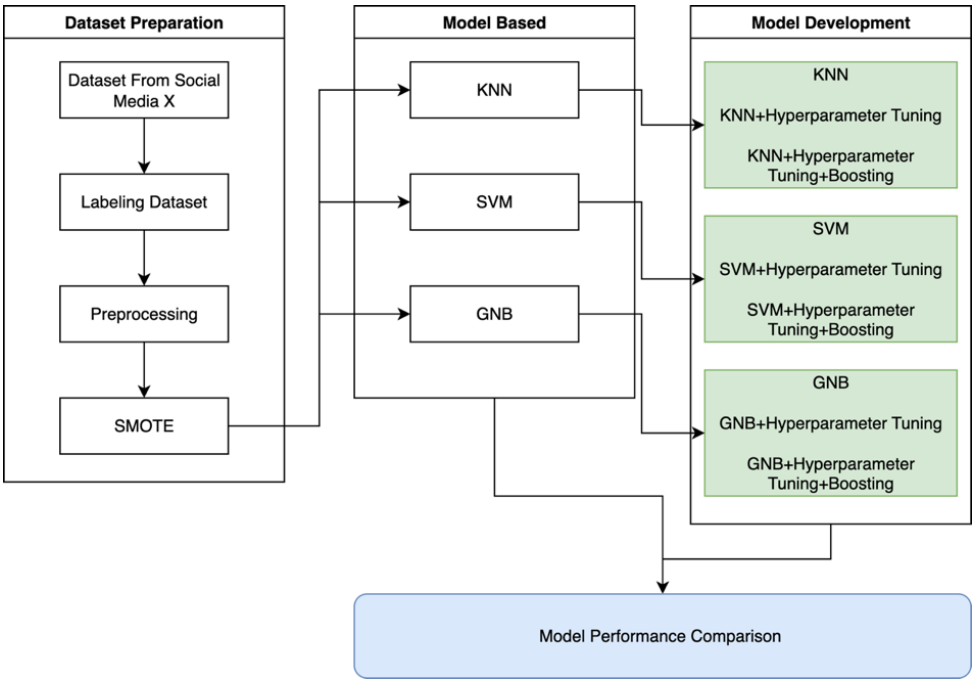


Figure 1. Methodology Flow

2.2. Dataset Preparation

The dataset used in this study was collected from the social media platform X (formerly Twitter) using the official API. Data collection focused on posts related to the 2024 Indonesian general election, covering the period from January to March 2024. Keywords and hashtags such as #Pemilu2024, #Capres, #KPU, #DebatPilpres, and #PartaiPolitik were employed to retrieve relevant tweets. After filtering out duplicates, retweets, and non-textual posts, a total of 10,001 unique entries were retained.

Manual labeling was conducted by three independent annotators with expertise in political communication. Each tweet was categorized into one of three sentiment classes: positive, negative, or neutral, based on its linguistic polarity and contextual meaning. To ensure reliability, Cohen's kappa coefficient ($\kappa = 0.86$) was calculated, indicating strong inter-annotator agreement.

The original class distribution before balancing consisted of 5,640 neutral (56.4%), 2,720 positive (27.2%), and 1,641 negative (16.4%) samples. After applying SMOTE (Synthetic Minority Over-sampling Technique), the dataset was balanced to 3,334 entries per class, yielding a total of 10,002 samples for model training and testing.

Text preprocessing was designed specifically for Indonesian-language content, involving lowercasing, tokenization, stopword removal using the Sastrawi library, normalization of colloquial words, and stemming. Non-alphabetic symbols, URLs, and user mentions were also removed. The resulting clean text was then transformed into numerical features using the TF-IDF (Term Frequency–Inverse Document Frequency) vectorization technique [17], [18].

2.3. Model Based

Following data transformation, three baseline classification algorithms were implemented. K-Nearest Neighbors (KNN) represents an instance-based method that classifies data based on proximity in vector space [19]. Support Vector Machine (SVM) applies a margin-based approach to construct optimal hyperplanes for maximum class separation [20]. Gaussian Naive Bayes (GNB) is a probabilistic model that assumes a Gaussian distribution and is computationally efficient for basic text classification [21]. These models were selected to evaluate different classification strategies—instance-based, margin-based, and probabilistic.

2.4. Model Development

This study implemented three baseline classifiers: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Gaussian Naive Bayes (GNB), and enhanced them using a two-phase development pipeline consisting of (1) hyperparameter optimization with Optuna and (2) ensemble enhancement using XGBoost.

It is important to clarify that XGBoost was not used to directly boost SVM or KNN, since XGBoost is inherently a tree-based boosting algorithm. Instead, the study adopted a stacked ensemble configuration, where the predictions from the tuned base models (KNN + Optuna, SVM + Optuna, GNB + Optuna) served as input features for the XGBoost meta-learner. In this setup, XGBoost acts as the final classifier that learns from the outputs of the optimized models, effectively combining their predictive strengths. The pipeline can be summarized as follows:
Preprocessing → TF-IDF → SMOTE → Model Training (KNN, SVM, GNB) → Hyperparameter Tuning (Optuna) → Stacking with XGBoost

This hybrid design allows XGBoost to integrate decision boundaries learned by margin-based (SVM), instance-based (KNN), and probabilistic (GNB) models, resulting in an adaptive ensemble capable of handling noisy and imbalanced textual data.

2.5. Model Performance Comparison

The final stage involves evaluating all models using standard classification metrics including accuracy, precision, recall, and F1-score. This evaluation aims to identify the most effective combination of approaches for classifying public opinion related to elections, particularly in imbalanced social media datasets.

3. RESULTS AND DISCUSSION

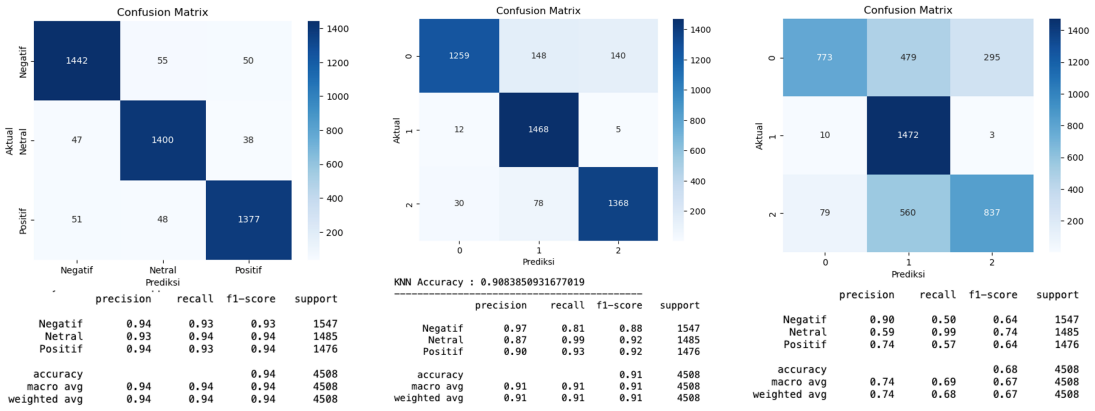
This section presents the experimental results obtained from the application of various classification models to the election-related public opinion dataset collected from the X social media platform. The evaluation includes three base models (KNN, SVM, and GNB), the models optimized through hyperparameter tuning using Optuna, and the final stacked ensemble using XGBoost as a meta-learner. Each model was tested using the preprocessed and balanced dataset of 10,002 text entries, obtained after applying SMOTE and transforming the text into numerical representations with TF-IDF.

The performance of each model was assessed using standard evaluation metrics: accuracy, precision, recall, and F1-score. These metrics provide a comprehensive perspective on how well each model distinguishes between the three sentiment classes: positive, neutral, and negative. The analysis was conducted in three stages: baseline evaluation, post-tuning evaluation, and stacked ensemble evaluation using XGBoost.

3.1. Base Model Evaluation

In the baseline stage, the three algorithms (KNN, SVM, and GNB) were trained on the TF IDF features of the balanced dataset. The Support Vector Machine (SVM) achieved the highest baseline performance with an accuracy of 94%, demonstrating stable precision and recall across all sentiment classes. Its margin-based optimization enabled better separation between sentiment boundaries, which is particularly effective for high-dimensional sparse data such as TF IDF.

The KNN model achieved an accuracy of 91%, slightly below SVM, indicating reasonable performance but with greater sensitivity to feature distance and data sparsity. The Gaussian Naive Bayes (GNB) model performed the weakest, achieving 68% accuracy. GNB's performance degradation stems from its assumption of independent Gaussian features, which does not align with the correlated term frequencies in TF IDF representation.



(a) Algorithm SVM (b) Algorithm KNN (c) Algorithm GNB

Figure 2. Base Algorithm Evaluation

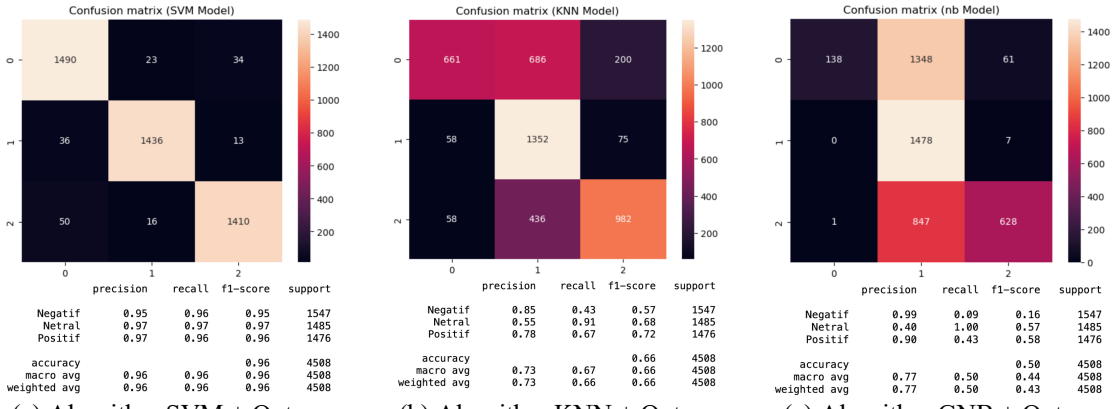
As shown in Figure 2, SVM consistently produced the highest and most balanced precision and recall values across all sentiment categories, particularly for the neutral and positive classes. KNN demonstrated moderate performance but was more sensitive to variations in distance metrics, which caused occasional misclassification of neutral posts as positive. In contrast, GNB showed a strong bias toward the majority class, indicating that its probabilistic assumptions were not well suited for TF IDF-based textual data. These findings confirm that SVM serves as the most reliable baseline model for sentiment classification in this study.

3.2. Evaluation After Hyperparameter Tuning with Optuna

In In the second stage, Optuna was applied to tune the key hyperparameters of each model. For SVM, Optuna optimized parameters such as the kernel type, regularization parameter (C), and gamma, leading to enhanced decision boundaries. For KNN, the optimal number of neighbors and weighting scheme were tuned, while for GNB, the smoothing parameter was adjusted.

After tuning, SVM with Optuna improved to 96% accuracy, with balanced precision, recall, and F1 scores across all classes. This shows that adaptive hyperparameter exploration using Tree Structured Parzen Estimator effectively refined the model's generalization on complex political sentiment data.

However, KNN with Optuna exhibited a decline in accuracy to 66%, likely due to overfitting or inappropriate weighting during parameter search, which can occur when dealing with sparse high dimensional data. Similarly, GNB with Optuna decreased to 50% accuracy, showing that tuning its smoothing parameter alone could not resolve the mismatch between Gaussian assumptions and TF IDF features. These results indicate that hyperparameter tuning does not universally improve all algorithms. Its effectiveness depends heavily on model characteristics and data representation.



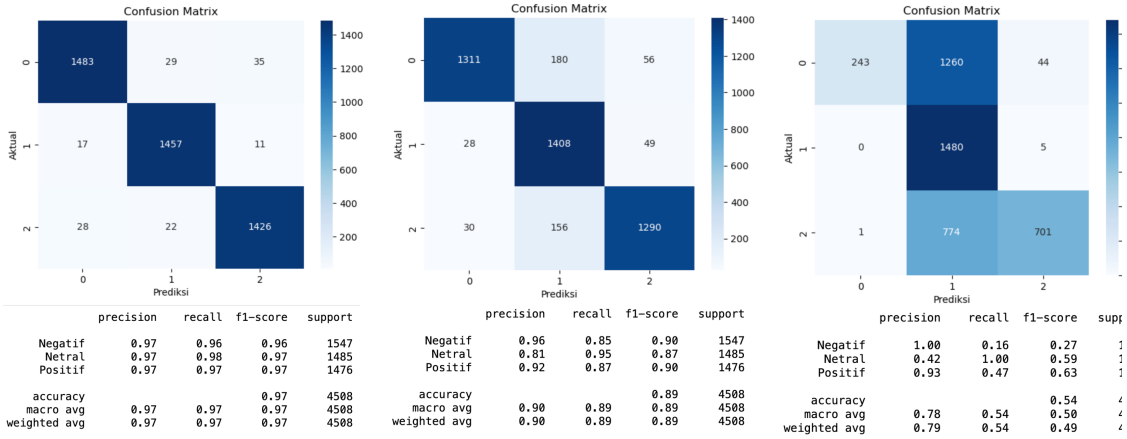
(a) Algorithm SVM + Optuna (b) Algorithm KNN + Optuna (c) Algorithm GNB + Optuna

Figure 3. Evaluation After Hyperparameter Tuning

As shown in Figure 3, SVM achieved more precise classifications for both positive and neutral sentiment categories, demonstrating improved boundary separation after tuning. The confusion matrix reveals that misclassifications between neutral and positive classes were significantly reduced. In contrast, KNN and GNB displayed increased confusion among minority classes, confirming that their tuned parameters did not generalize well to imbalanced or sparse feature spaces. These results highlight that SVM benefits the most from Optuna-based hyperparameter optimization, while KNN and GNB require additional ensemble or feature-selection mechanisms to achieve comparable stability.

3.3. Evaluation After Boosting with XGBoost

In the final phase, models that had undergone hyperparameter tuning were further enhanced using the XGBoost boosting technique. This phase aimed to evaluate the extent to which boosting could improve model performance.



(a) Algorithm SVM + Optuna + XGBoost

	precision	recall	f1-score	support
Negatif	0.97	0.96	0.96	1547
Netral	0.97	0.98	0.97	1485
Positif	0.97	0.97	0.97	1476
accuracy			0.97	4508
macro avg	0.97	0.97	0.97	4508
weighted avg	0.97	0.97	0.97	4508

(b) Algorithm KNN + Optuna + XGBoost

	precision	recall	f1-score	support
Negatif	0.96	0.85	0.90	1547
Netral	0.81	0.95	0.87	1485
Positif	0.92	0.87	0.90	1476
accuracy			0.89	4508
macro avg	0.90	0.89	0.89	4508
weighted avg	0.90	0.89	0.89	4508

(c) Algorithm GNB + Optuna + XGBoost

	precision	recall	f1-score	supp
Negatif	1.00	0.16	0.27	
Netral	0.42	1.00	0.59	
Positif	0.93	0.47	0.63	
accuracy			0.54	
macro avg	0.78	0.54	0.50	
weighted avg	0.79	0.54	0.49	

Figure 4. Evaluation After Hyperparameter Tuning + XGBoost

SVM + Optuna + XGBoost achieved the best performance, with all metrics—accuracy, precision, recall, and F1-score—reaching 97%. The confusion matrix shows balanced and accurate predictions across all classes. Boosting effectively enhanced SVM's generalization, making it the most stable and accurate model across all experiments. KNN + Optuna + XGBoost recorded an improved accuracy of 89%, up from the post-tuning result of 66%. Though still lower than the original base model (91%), the improvement illustrates that boosting can recover some of the lost performance.

GNB + Optuna + XGBoost performed poorly, with accuracy dropping to 54% and an average F1-score of just 50%. The confusion matrix indicates severe misclassifications, and the probabilistic nature of GNB proved incompatible with tree-based boosting, resulting in predictive degradation. These results highlight that SVM is the most suitable model for sentiment classification in imbalanced social media datasets, especially when combined with Optuna and XGBoost. KNN benefits from boosting but is less stable, while GNB is not suitable for this modeling pipeline.

3.4. Analysis and Comparative Review

Table 1 compares the proposed approach with prior sentiment analysis studies published between 2022 and 2025, covering various datasets such as COVID-19 opinions, product reviews, and general political discussions. The comparison highlights that although many studies used balancing techniques (e.g., SMOTE) and ensemble methods (e.g., stacking, voting), few integrated

data balancing, adaptive hyperparameter tuning, and stacking-based boosting into a single experimental pipeline. The current study, therefore, represents a more comprehensive optimization framework for sentiment classification on social media datasets related to political contexts.

Table 1. Comparison with Previous Studies

Researchers	Model	Balancing technique	Accuracy	Precision	Recall	F1-Score
[22]	Stacking Technique **Based**: CART, KNN, Naive Bayes, and Random Forest. **Meta learner**: Logistic Regression	SMOTE	88%	88%	88%	88%
[23]	Voting Technique (Naïve Bayes, SVM, C.45)	SMOTE	94%	94%	94%	94%
[24]	SVM	SMOTE	87%	87%	87%	87%
[25]	SVM + Grid Search	-	92%	92%	92%	92%
[26]	SVM + XGBoost	-	79%	74%	83%	77%
[27]	TF-IDF Vectorizer	-	95%	95%	95%	79%
This Study	SVM + Optuna + XGboost	SMOTE	97%	97%	97%	97%

Based on Table 1, which compares the results of this study with several previous studies, it can be seen that the approach proposed in this research yields the highest performance across all evaluation metrics. This study employed a combination of the SVM algorithm, hyperparameter tuning using Optuna, and boosting with XGBoost, applied to data balanced using SMOTE. The results show equal values of 97% for accuracy, precision, recall, and F1-score, making it the most superior approach compared to other methods applied in similar domains.

Compared to earlier studies that also used balancing techniques like SMOTE, such as study [22] implementing stacking with CART, KNN, Naive Bayes, and Random Forest, the accuracy and other metrics only reached 88%. Study [23], which used a voting ensemble based on Naive Bayes, SVM, and C4.5 along with SMOTE, achieved 94% in all metrics, still below this study's performance. Meanwhile, the single SVM approach with SMOTE in [24] only achieved 87% accuracy, indicating that model and tuning combinations are crucial in driving performance improvements.

Interestingly, study [25] using SVM with Grid Search tuning (without SMOTE) showed an accuracy of 92%, better than regular SVM, but still did not surpass the combination of Optuna and XGBoost used in this study. Study [26], which combined SVM with only XGBoost, reached only 79% accuracy, with unbalanced precision and F1-score values, showing the importance of data balancing and adaptive hyperparameter tuning. Meanwhile, study [27], which focused on using the TF-IDF vectorizer without specific model details, although recording high accuracy (95%), had a much lower F1-score (79%), indicating imbalanced predictions across classes.

Overall, this comparison shows that the layered and integrated approach applied in this study, which includes data balancing with SMOTE, parameter optimization using Optuna, and model enhancement using XGBoost, is the most effective strategy for classifying public opinion texts related to elections. This approach not only achieves the highest metric performance but also demonstrates stability across all aspects of classification, making it a superior solution compared to conventional methods and other ensemble approaches reported in previous studies.

4. CONCLUSION

This study demonstrates that developing sentiment classification models through a structured and sequential strategy, starting from preprocessing and data balancing to hyperparameter tuning and ensemble enhancement, produces varying results depending on the algorithm used. The Support Vector Machine (SVM) consistently emerged as the most robust and optimal model for analyzing public opinion data related to elections, particularly after being fine-tuned with Optuna and further enhanced with XGBoost. This configuration achieved 97 percent accuracy along with high precision, recall, and F1 scores, outperforming all other approaches examined in this study and those reported in previous research.

In contrast, tuning applied to KNN and GNB resulted in reduced baseline performance, although ensemble enhancement was able to partially recover the performance of the KNN model.

These findings indicate that not all algorithms benefit equally from optimization techniques, emphasizing the importance of aligning model development strategies with the intrinsic characteristics of both the data and the learning algorithm.

Future research can extend this work by integrating deep learning-based text representations such as Word2Vec, FastText, or BERT to achieve a more comprehensive semantic understanding of opinion contexts. Additionally, exploring neural network-based ensemble models such as LSTM, BiLSTM, or Transformer architectures could further improve classification performance on complex and dynamic social media data. The proposed framework may also be adapted for aspect-based sentiment analysis to identify public attitudes toward specific political issues, parties, or candidates. Evaluating the model on data from other platforms such as Facebook, Instagram, or TikTok will also be essential to assess its generalizability. Finally, applying semi-supervised or active learning methods could reduce dependence on manual labeling while maintaining model reliability.

REFERENCES

[1] W. Khan, M. A. Ghazanfar, M. A. Azam, A. Karami, K. H. Alyoubi, and A. S. Alfakeeh, "Stock market prediction using machine learning classifiers and social media, news," *J Ambient Intell Humaniz Comput*, vol. 13, no. 7, pp. 3433–3456, Jul. 2022, doi: 10.1007/s12652-020-01839-w.

[2] M. Orabi, D. Mouheb, Z. Al Aghbari, and I. Kamel, "Detection of Bots in Social Media: A Systematic Review," *Inf Process Manag*, vol. 57, no. 4, p. 102250, 2020, doi: 10.1016/j.ipm.2020.102250.

[3] E. Konovalova, G. Le Mens, and N. Schöll, "Social media feedback and extreme opinion expression," *PLoS One*, vol. 18, no. 11, pp. 1–20, Nov. 2023, doi: 10.1371/journal.pone.0293805.

[4] X. T. Dang and T. T. Le, "KNN-SMOTE: An Innovative Resampling Technique Enhancing the Efficacy of Imbalanced Biomedical Classification," in *Machine Learning and Other Soft Computing Techniques: Biomedical and Related Applications*, N. Hoang Phuong, N. T. Huyen Chau, and V. Kreinovich, Eds., Cham: Springer Nature Switzerland, 2024, pp. 111–121. doi: 10.1007/978-3-031-63929-6_11.

[5] V. Gupta and P. Rattan, "Improving Twitter Sentiment Analysis Efficiency with SVM-PSO Classification and EFWS Heuristic," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 698–715. doi: 10.1016/j.procs.2023.12.125.

[6] I. F. Ashari, M. C. Untoro, E. M. Sutowo, D. Salsabila, and D. A. Zhabiyan, "Hyperparameter Tuning Feature Selection with Genetic Algorithm and Gaussian Naïve Bayes for Diabetes Disease Prediction," *Jurnal Telematika*, vol. 17, no. 1, pp. 17–23, 2022, doi: 10.61769/jurtel.v17i1.488.

[7] F. Suandi *et al.*, "Enhancing Sentiment Analysis Performance Using SMOTE and Majority Voting in Machine Learning Algorithms," in *International Conference on Applied Engineering*, Atlantis Press, 2024, pp. 126–138. doi: 10.2991/978-94-6463-620-8_10.

[8] S. Hadianti and W. A. G. Kodri, "Optimization of The Machine Learning Approach using Optuna in Heart Disease Prediction," *Journal Medical Informatics Technology*, vol. 1, no. 3, pp. 59–64, Sep. 2023, doi: 10.37034/medinftech.v1i3.15.

[9] P. Srinivas and R. Katarya, "hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost," *Biomed Signal Process Control*, vol. 73, no. 1, pp. 1–10, Mar. 2022, doi: 10.1016/j.bspc.2021.103456.

[10] Y. Li, C. Stasinakis, and W. M. Yeo, "A Hybrid XGBoost-MLP Model for Credit Risk Assessment on Digital Supply Chain Finance," *Forecasting*, vol. 4, no. 1, pp. 184–207, Mar. 2022, doi: 10.3390/forecast4010011.

[11] R. Syahputra, G. J. Yanris, and D. Irmayani, "SVM and Naïve Bayes Algorithm Comparison for User Sentiment Analysis on Twitter," *Sinkron*, vol. 7, no. 2, pp. 671–678, May 2022, doi: 10.33395/sinkron.v7i2.11430.

[12] R. Pambudi and F. Madani, "Analysis of public opinion sentiment against COVID-19 in Indonesia on twitter using the k-nearest neighbor algorithm and decision tree," *Journal of Soft Computing Exploration*, vol. 3, no. 2, pp. 117–122, Sep. 2022, doi: 10.52465/joscex.v3i2.88.

[13] I. Setiawan *et al.*, "Utilizing Random Forest Algorithm for Sentiment Prediction Based on Twitter Data," in *Proceedings of the First Mandalika International Multi-Conference on Science and Engineering 2022, MIMSE 2022 (Informatics and Computer Science)*, Atlantis Press International BV, 2022, pp. 446–456. doi: 10.2991/978-94-6463-084-8_37.

[14] A. Saekhu and D. Intan Surya Saputra, "Enhancing Student Sentiment Classification on AI in Education using SMOTE and Naive Bayes," *Technology and Science (BITS)*, vol. 6, no. 4, pp. 2165–2174, 2025, doi: 10.47065/bits.v6i4.6469.

[15] M. Idris, A. Rifai, and K. D. Tania, "Sentiment Analysis of Tokopedia App Reviews using Machine Learning and Word Embeddings," *sinkron*, vol. 9, no. 1, pp. 210–219, Jan. 2025, doi: 10.33395/sinkron.v9i1.14278.

[16] A. Samih, A. Ghadi, and A. Fennan, "Enhanced sentiment analysis based on improved word embeddings and XGboost," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 2, pp. 1827–1836, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1827-1836.

[17] M. K. Anam *et al.*, "Sara Detection on Social Media Using Deep Learning Algorithm Development," *Journal of Applied Engineering and Technological Science*, vol. 6, no. 1, pp. 225–237, Dec. 2024, doi: 10.37385/jaets.v6i1.5390.

[18] V. Talasila, M. V Mohan, and N. M. R, "Enhancing Text-to-Image Synthesis with an Improved Semi-Supervised Image Generation Model Incorporating N-Gram, Enhanced TF-IDF, and BOW Techniques," *International Journal of Intelligent Systems and Applications in Engineering* , vol. 11, no. 7s, pp. 381–397, 2023, [Online]. Available: www.ijisae.org

[19] A. Masitha and M. Kunta Biddinika, "Preparing Dual Data Normalization for KNN Classfication in Prediction of Heart Failure," *KLIK*, vol. 4, no. 3, pp. 1227–1234, 2023, doi: 10.30865/klik.v4i3.1382.

[20] Y. A. Singgalen, "Performance Evaluation of Sentiment Classification Models: A Comparative Study of NBC, SVM, and DT with SMOTE," *Media Online*, vol. 4, no. 5, pp. 2539–2548, Apr. 2024, doi: 10.30865/klik.v4i5.1827.

[21] T. P. Handayani, W. Hasyim, and Nursetia, "Preliminary Evaluation of Gaussian Naive Bayes for Multi-Label Hate Speech and Abusive Language Detection on Indonesian Twitter," *Journal of International Multidisciplinary Research*, vol. 1, no. 1, pp. 159–165, Nov. 2023, doi: 10.62504/jimr532.

[22] P. P. Putra, M. K. Anam, A. S. Chan, A. Hadi, N. Hendri, and A. Masnur, "Optimizing Sentiment Analysis on Imbalanced Hotel Review Data Using SMOTE and Ensemble Machine Learning Techniques," *Journal of Applied Data Sciences*, vol. 6, no. 2, pp. 936–951, May 2025, doi: 10.47738/jads.v6i2.618.

[23] M. K. Anam, T. P. Lestari, H. Yenni, T. Nasution, and M. B. Firdaus, "Enhancement of Machine Learning Algorithm in Fine-grained Sentiment Analysis Using the Ensemble," *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 19, no. 2, pp. 159–167, Mar. 2025, doi: 10.37936/ecti-cit.2025192.257815.

[24] L. L. Van Fc, M. K. Anam, M. B. Firdaus, Y. Yunefri, and N. A. Rahmi, "Enhancing Machine Learning Model Performance in Addressing Class Imbalance," *COGITO Smart Journal*, vol. 10, no. 1, pp. 478–490, 2024.

[25] A. N. Ulfah, M. K. Anam, N. Y. S. Munti, S. Yaakub, and M. B. Firdaus, "Sentiment Analysis of the Convict Assimilation Program on Handling Covid-19," *JUITA: Jurnal Informatika*, vol. 10, no. 2, pp. 209–216, 2022, doi: 10.30595/juita.v10i2.12308.

[26] R. Rahmaddeni, M. K. Anam, Y. Irawan, S. Susanti, and M. Jamaris, "Comparison of Support Vector Machine and XGBSVM in Analyzing Public Opinion on Covid-19

Vaccination," *ILKOM Jurnal Ilmiah*, vol. 14, no. 1, pp. 32–38, Apr. 2022, doi: 10.33096/ilkom.v14i1.1090.32-38.

[27] C. G. Özmen and S. Gündüz, "Comparison of Machine Learning Models for Sentiment Analysis of Big Turkish Web-Based Data †," *Applied Sciences (Switzerland)*, vol. 15, no. 5, pp. 1–20, Mar. 2025, doi: 10.3390/app15052297.