Support Vector Machine Algorithms for Sentiment Analysis on the Inaugural Ceremony in the Capital City of Nusantara

F Putri[1], H Suparwito*[2]

[1,2]Informatics Department, Sanata Dharma University, Indonesia

E-mail: sasha@gmail.com[1], shirsj@usd.ac.id[2]

Abstract. The inauguration preview at IKN uploaded to the official Instagram account of President Joko Widodo (@jokowi) has garnered various reactions from the public, as seen in the comment columns that expressed all kinds of opinions, from positive to neutral and even negative. The objective of this paper is to provide an objective, systematic and real-time sentiment analysis of public using machine learning algorithms such as Naïve Bayes, and Support Vector Machine (SVM). The analysis process was done on two methods of labelling the data which are VADER and IndoBERT, as well as by implementing both on balancing a class imbalance using SMOTE. The performance result of SVM (RBF kernel) and IndoBERT labelling (95.05% accuracy, 0.951 F1-score) is the best one from evaluation process. The detect result can be analysed in terms of the frequency, which is positive posts and negative comments more than dominating. This work substantiates that such machine learning method, SVM for instance, can be employed to map and interpret the public opinion on government policy based on social media comment data

Keywords: sentiment analysis; ibu kota nusantara; vader; indobert; classification; support vector machine

1. Introduction

New social media sites have grown at a stunning rate and fundamentally changed the way in which public opinion is generated, shared, and observed. Social media's influence is not immune to its impact on international relations, as in the case of being migrants' symbolic spaces Instagram, one recent phenomenon among young urbanites and people in the community has been influential for conveying policies at almost every level of government ranging from national agenda to local events [1]. State-controlled accounts, and in particular the Instagram account of the 7th President of the Republic of Indonesia, Joko Widodo (@jokowi), strategically influence public discourses through dissemination of state action and policy-relevant content. One of the highest national events that received attention from the public was the inaugural ceremony in Capital City of Nusantara (IKN). The event's announcement through Instagram received thousands of public comments, which presented a range of opinions as well as emotions and attitudes about moving Mexico's federal capital city [2].

The vast amount of raw corpora stemming from Instagram comments is both an opportunity and a curse. Although these comments could be a rich source of public opinion, manual review is not feasible given the scale, subjectivity and time constraints. As a result, automated sentiment analysis is vital for systematically sifting, evaluating and interpreting public sentiment. Sentiment analysis is a type of common task in the field of Natural Language Processing (NLP) that determines sentiment polarity — e.g., positive, neutral or negative — as detected from text data. In public policy and government communication, sentiment analysis can provide researchers and policy makers an objective way to evaluate the response of citizens and the general attitude towards a particular national scheme.

Support Vector Machine (SVM) is widely used in sentiment analysis because it can handle high-dimensional and sparse textual data effectively [3][4]. In the last five years, many studies have shown that SVM is still a competitive classifier for performing SA in social media as on Twitter, Instagram, and online reviews. For Indonesian text sentiment analysis, SVM in combination with TF–IDF feature extraction and kernel-based optimization i.e., RBF and Polynomial kernels are often used in order to model the nonlinear nature of sentiment expression [5][6]. Based on experimental results presented in recent works, SVM accuracy lies in the range 80% to 90% and depends on dataset's features and pre-processing algorithms. We show that SVM is a reliable base and competitive classifier, as long as an adequately designed data pre-processing and feature engineering are made [7]. Some papers have also confirmed that SMOTE can enhance SVM classifiers performance. The class imbalance is a prevalent obstacle for sentiment in which one of the classes (positive or negative) is far more than the other. It has been evidenced that using SMOTE to be applied before training SVM improves recall and F1-score notably over minority classes sentiment. These findings further justify the use of SVM for sentiment analysis in noisy, imbalanced and informal social media text [8][9][10]. IndoBERT, in particular, represents a modern pioneer for Indonesian NLP research [11]. Our proposed model, IndoBERT, is a transformer-based language model that has been pre-trained over a large-scale Indonesian corpora; thus it should have better explicit understanding of semantics and syntax relations in contextual rather than simple bag-of-words. There are more recent works that claimed significant improvement in performance when adapting IndoBERT for detecting sentiment, reaching accuracy and F1-score above 90% on different datasets for Indonesian such as social media comments, product reviews and news articles [12]. The efficacy of IndoBERT is in its capability to represent context-dependent sentiment, process informal language and recognize implicit emotional cues typical of social media conversation. Accordingly, IndoBERT has been the standard approach for sentiment analysis in Indonesian research. However, IndoBERT is widely used in studies as an end-toend classifier; the possibility, of adopting it into a sentiment labeller for downstream classical machine learning models has not been deeply explored. This constraint opens up the possibility to further explore IndoBERT beyond simple classification, especially in a hybrid system that combines deep learning-based labelling with traditional classifiers like SVM.

Lexicon-based sentiment analysis methods remain popular due to their simplicity, transparency, and computational efficiency. VADER (Valence Aware Dictionary for Sentiment Reasoning) is one of the most widely used lexicon-based approaches, specifically designed to handle short and informal text commonly found on social media platforms [13][14]. VADER assigns sentiment polarity scores based on a predefined sentiment lexicon and heuristic rules that account for punctuation, capitalization, and emoticons. In recent Indonesian sentiment analysis studies, VADER is often applied through translation-based workflows or lexicon adaptation. Although VADER generally produces lower accuracy compared to supervised and deep learning-based methods, it remains valuable as a baseline approach and as an automatic labelling tool for large-scale datasets. Several studies have demonstrated that VADER-labelled datasets can be effectively used to train supervised classifiers, although the resulting performance is highly dependent on the quality of translation and lexicon coverage.

A growing number of recent studies have conducted comparative evaluations between traditional machine learning models and transformer-based approaches for sentiment analysis. These studies consistently report that IndoBERT outperforms SVM and lexicon-based methods in terms of accuracy and F1-score [15]. However, most comparisons focus on classifier-level performance rather than examining the interaction between labelling strategies and classifier learning behaviour. Only a limited number of studies explicitly investigate the effect of sentiment labelling methods on SVM performance [16]. Existing research often treats sentiment labels as fixed ground truth without considering how different labelling strategies may introduce varying degrees of noise and bias. Consequently, the influence of lexicon-based versus transformer-based labelling on classical machine learning classifiers remains insufficiently explored, particularly in the context of Indonesian social media data.

By positioning SVM as the fixed classifier and systematically comparing VADER-based and IndoBERT-based sentiment labelling within the same experimental setting, this study extends prior research and provides a more nuanced understanding of how labelling quality impacts sentiment classification performance. This approach not only bridges the methodological gap between traditional and modern NLP techniques but also offers practical insights for researchers seeking cost-effective yet accurate sentiment analysis solutions for large-scale Indonesian social media data.

This paper fills these gaps by conducting a controlled experiment on Support Vector Machine classifiers trained on the same Instagram dataset but labelled with two completely different sentiment labelling approaches: lexicon-based VADER and transformer-based IndoBERT. By keeping the classifier architecture and feature extraction approach constant while comparing different labelling strategies, we isolate the effect of label quality on SVM performance. Employing accuracy and F1-score as the major evaluation metrics guarantees a fair sketch of classification performances. Moreover, we further improve the soundness of experiments by using SMOTE to balance the data and k-fold cross-validation. In so doing, the study adds to empirical evidence of the importance of sentiment labelling strategy to improve classical machine learning performance for Indonesian social media sentiment analysis in public policy and governmental communication context.

The objective of this research is to experimentally analyse public sentiment towards the inaugural ceremony in Capital City of Nusantara through comment on Instagram in the official account @jokowi. The main aim is to assess and contrast the performance of SVM classifiers based on two sentiment labelling techniques, VADER-based and IndoBERT-based. We evaluate in terms of accuracy and F1-score as the main performance criteria. By conducting controlled experiments on real-world social media data, this study seeks to provide empirical insights into the impact of sentiment labelling strategies on SVM performance and contribute to the broader body of knowledge on Indonesian-language sentiment analysis, particularly in the context of public policy and government communication.

2. Methodology

The following is a research flowchart that explains the stages carried out during the research process. In general, this research flows can be seen in the following diagram.

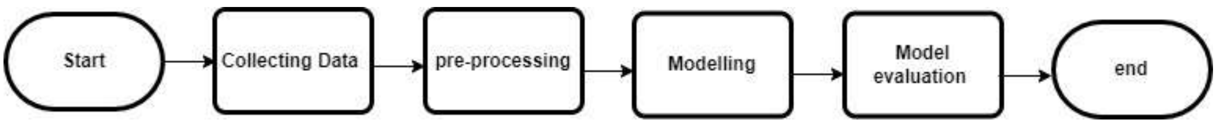

Figure 1. Research flowchart

2.1. Data

This study uses comment data from the official Instagram account @jokowi with a web scraping method using the Python programming language on Google Colab. The data was taken from posts published on August 10, 13, 16, and 17, 2024, which is within a week before the Inauguration Ceremony in the Capital City of Nusantara (IKN) and during the ceremony. A total of 8,003 comments were collected only from posts on August 10, 13, 16, and 17, 2024, as only these posts were related to the topic of the Inauguration Ceremony in the Capital City of Nusantara (IKN).

The comment data was collected on the same date, March 10, 2025, but at different times (timestamps) according to the scraping process that was carried out in stages for each post. This may affect the number and content of comments recorded, as the comments collected represent the conditions seven months after the original posts were published.

Table 1. An example of the data collected

Username	**Comment**
ra_yhan207	Kpn dpt rejeki bisa ke IKN
futiast_	wiiihhhhh keren . merdeka indonesiaku 🇮🇩💚💧
fg_zach	Upacara terkeren saat ini Pak 👍
im.rizkyyy	Cancel ajalah pak 17 Augustus kita tahun ini...kecewa kami pak
niaguslin	Tidak merdeka berhijab 🥺

2.2. Pre-processing

Once the raw data has been collected, the next step is pre-processing data, which aims to make previously unstructured data more structured and remove unused components. The pre-processing stages begin with emoticon conversion, cleaning, tokenizing, emoticon converter, case folding, tokenizing, normalization, stop-word removal, and stemming.

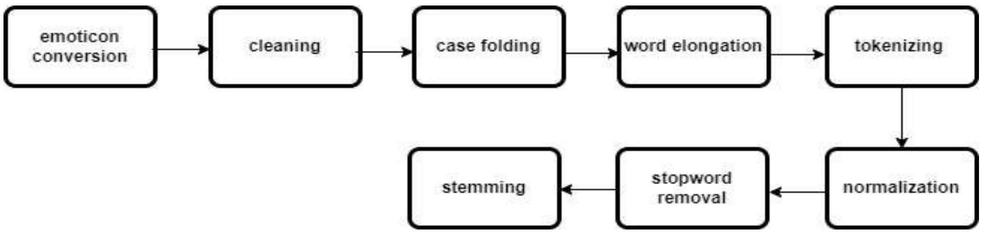

Figure 2. A detail of Pre-processing steps

Emoticon Conversion

In order to convert emoticons which, appear in comments into words that reflect their meanings, we do this after normalization. The process starts by reading the entire dataset, then we search for symbols or emoticons in the text. Then, we map each identified emoticon to a list of predefined emoticons. An emoticon 👍 is converted to a word "good" for instance.

Cleaning

The clean stage is centred on the detection & removal of noise or errors along with irrelevant data from the extracted text data. The process includes removal of duplicate data as well as unnecessary characters like the punctuation's, special characters like these (#$@"), HTML tags also numbers etc. Number: 1233456789 ##, Hashtags: #usernames and URLs www.example.com, Links: http://www.example.com, @UserName This stage aims at creating a more clean and structured text corpus for further operations.

Case Folding

Case folding is applied to normalize the text structure and turn all characters into lowercase. Following the cleaning process, each word is read for the text data and turned to be in lowercase from its original letter case so the problem of differences due to capital letters in matching is taken into account.

Word Elongation

This sub task focuses on reducing the excessive repetition of letters in a word without losing its meaning. After case-folding, the system detects words with repeating characters beyond those commonly found. The fact that we do not apply the deletion of repeated letters for a few words with correct repetition intrinsic to them like "tinggi" and "bangga". Letters that are repeated more than twice are also cut by only one letter e.g., Hebat (listing) becomes hebat (great).

Tokenization

Tokenization is a technique to segment sentences or other chunks of text into individual words, which are then represented as tokens. Each word is then tokenized to individual tokens after the word elongation process for subsequent text analysis.

Normalization

Normalization involves replacing the non-standard words to their standard form, as per dictionary used and in turn a GitHub dictionary. Each word is read and checked against the dictionary. If it is a word present in the dictionary, then that word can be replaced by its standard form if not again leave it unchanged.

Stop-word Removal

This step is to filter out words that contain little information with respect to the analysis. The word has to be formulated with respect to the stopword list used in the study after normalization. In the stop-list, a word is eliminated if found and kept otherwise.

Stemming

Stemming comes from linguistics, where it refers to the (albeit quite similar) process of Reduce words to their root or base form. Clearing words of their affixes. Following the stopword removal phase, each word is examined and compared with the stemming list. If a word is present in the dictionary, it will be stemmed.

2.3. Data Labelling

Another important aspect of this study is data labelling. Two sentiment labelling approaches were employed, namely VADER and IndoBERT, to provide comparative perspectives on sentiment classification.

Sentiment labelling of Instagram comments using VADER (Valence Aware Dictionary for Sentiment Reasoning) was performed through the Python NLTK library. Even though VADER has been developed for English texts, it could be used with Indonesian data with the aid of a suitable dictionary. VADER is a lexicon-based method so it requires translating all the texts to English before using it for labelling. The VADER lexicon is a dictionary with words and their sentiment scores which are utilized to obtain the aggregate sentiment polarity in text. According to these scores, we classify comments into positive (if the score > 0), neutral (if the score = 0), or negative (if the score < 0).

IndoBERT uses a state of the art based deep learning approach which does not depend on fixed dictionary for representing the sentence context and meaning. The model uses contextualized representations that were obtained by pre-training on large Indonesian language data. IndoBERT is a BERT derivative from the IndoNLU project pre-trained across a variety of Indonesian text sources such as news articles, Wikipedia and social media. This extensive pre-training helps the model obtain a strong understanding of Indonesian natural language structure and semantics. IndoBERT was selected because of its beperformance of understanding their Indonesian language context. Trained on local corpora, IndoBERT is well equipped to understand the nuances, structure and stylistic diversity of Indonesian, especially in online social media conversations that are less formal. This makes IndoBERT very suitable for sentiment classification tasks that involve nuanced and contextual Indonesian text.

2.4. Feature Extraction

This method is designed to allocate the weights for each word in pre-processed data. The word weighting is TF-IDF. In python, software library scikit-learn’s TfidfVectorizer in the feature extraction module is used for computing TF-IDF. These calculations are utilized later for training a model.

2.5. Data Balancing

Data that has undergone the labelling process often experiences an imbalance in the number of comments on each label. This condition can cause bias towards the majority label, so that the classification model tends to provide less accurate prediction results on the minority label. Therefore, a data balancing process is needed to ensure a more equitable distribution between labels.

The technique used for data balancing is SMOTE (Synthetic Minority Over-sampling Technique), which is a method that generates synthetic data on minority labels based on similarities between data.

2.6. Data Division

The data was split 80:20 for testing and training. For this, 1,601 data points (or patients) were used to test and the remaining of 6,402 were used in training and validation. Next, this training data is divided further by K-Fold Cross Validation technique on the basis of the K value under observation. This approach ensures that the model's behaviour is stable, not relying completely on a specific region of the data

2.7. SVM Modeling

The classification step was implemented with an SVM transcoding SVC module, from sklearn, to represent the SVM model. It is an SVM library in python programming language. The performance of the SVM model for RBF and polynomial kernels with $C \in [0.1, 1, 10,100]$ was evaluated. This testing was performed using the k-fold = 5 method, and on two labelling representations, which are VADER and IndoBERT

Table 2. experiment design parameters

Parameters	Value
K-fold	5
Kernels	Polynomial, RBF
C	0.1, 1, 10, 100
Labelling methods	VADER, IndoBERT

3. Results and Discussions

This section presents a comparative evaluation of Support Vector Machine (SVM) classifiers trained using two sentiment labelling approaches, namely lexicon-based VADER and transformer-based IndoBERT. The analysis focuses on the impact of labelling strategy, kernel selection, and regularization strength (*C*) on classification performance, evaluated using accuracy, precision, recall, and F1-score.

Table 3. SVM results using VADER

Kernel	c	Accuracy	Precision	Recall	F1 - score
RBF	0.1	84.04%	0.857	0.840	0.841
	1	93.97%	0.941	0.940	0.940
	10	94.50%	0.946	0.945	0.945
	100	94.45%	0.945	0.945	0.944
Polynomial	0.1	59.27%	0.786	0.593	0.573
	1	87.40%	0.897	0.874	0.874
	10	84.27%	0.885	0.843	0.844
	100	84.30%	0.885	0.843	0.844

Table 4. SVM results using IndoBERT

Kernel	c	Accuracy	Precision	Recall	F1 - score
RBF	0.1	85.83%	0.876	0.858	0.860
	1	94.36%	0.944	0.944	0.944
	10	95.05%	0.951	0.951	0.951
	100	94.95%	0.95	0.949	0.95
Polynomial	0.1	60.21%	0.789	0.602	0.583
	1	88.14%	0.900	0.881	0.882
	10	85.66%	0.890	0.857	0.859
	100	85.84%	0.892	0.858	0.86

After conducting several tests using two labelling methods and various parameters on the SVM model with a k-fold combination where k = 5, the best parameter combination was obtained. Table 4 presents the best parameter results from each combination for both labelling methods in each model.

Table 5. the best results of SVM algorithm

Algoritm	Labelling	*Kernel* SVM	C	Accuration	F1-score
SVM	VADER	RBF	10	94.50%	0.945
		Polynomial	1	87.40%	0.874
	IndoBERT	RBF	10	95.05%	0.951
		Polynomial	1	88.14%	0.882

Overall, models trained using sentiment labels obtained from VADER perform worse than those that are trained on IndoBERT-based labels for all experimental conditions. This performance gap is presented consistently in all evaluation part which suggests that the improvements are systematic and not ad-hoc. More specifically, IndoBERT-based models have superior F1-scores and find a better trade-off between precision and recall as well as reduced class bias. From a statistical learning perspective, the stable performance of IndoBERT labels might indicate enhanced inter-class separability in the feature space. Contextualized embeddings that are yielded by IndoBERT encompass the semantic dependencies and polarity inversions which lexicon-based methods do not capture, while mitigating label noise, enabling SVM to learn a more optimal maximum-margin decision boundary.

For the RBF kernel, both labelling methods appear sensitive to the regularization parameter C: using VADER labelling, accuracy ascends from 84.04% at C = 0.1 up to a maximum of 94.50% at C = 10 (accompanied by an increase in F1-score from 0.841 to 0.945). The result is the same on IndoBERT labelling and yet again, each model using IndoBERT always obtained better performance on all C values. The highest overall performance is achieved by using the IndoBERT labeling with an RBF kernel at C = 10 where it can reach Accuracy of 95.05% and F1-score of 0.951. Even though the absolute gain over the best VADER-based model is about 0.6%, but gains of this order are statistically significant in huge sentiment classification applications where marginal improvements result to substantial decreases in misclassification errors. Performance is degraded when C=100 for both labelling schemes. This behaviour is a sign of overfitting by miss-classification. It is worth mentioning that IndoBERT based-models face lesser degradation, indicating this model is more robust and tends to generalize better under high regularization.

Polynomial kernel models do perform much worse than RBF kernel across all C values. The labelling strategies achieved accuracies of less than 61% at C = 0.1; the model might not have enough capacity to learn non-linear sentiment patterns in light regularization conditions.

Compared to VADER-based models, the performance is much remarkably improved as increasing C to 1; IndoBERT-based models obtain an accuracy of 88.14% and F1-score of 0.882, rather than an accuracy of 87.40% and F1-score of 0.874 for VADER-based models. But increased factors of C does not achieve linear development either, and performance flattens out when the values of C are 10 or 100. This plateau indicates that the Polynomial feature is not expressive enough to capture sophisticated sentiment distributions in high dimensions.

Nevertheless, IndoBERT-based labelling rarely lose its edge throughout all Polynomial kernel setups with about 1–1.5% performance gap and proved that the quality of labeling holds more weight in classifying a case without compromise to a particular kernel type.

The experiments show that how to label the sentiment has a significant impact on SVM classification. In general, labelling induced by IndoBERT consistently results in higher accuracy and F1-score under kernels and regularization settings, thus leading to improved robustness and generalisation ability. These results indicate that contextual transformer-based models produce better sentiment annotations for downstream classical machine learning classifiers. In total, the best set identified by this study is a SVM with RBF kernel trained on IndoBERT labelled data at C = 10. This combination of parameters offers the best balance between bias and variance, resulting in the most consistent performance across performance measures. The results have shown that combining the contextual language models with classic classifiers can enhance the performance of sentiment analysis for Indonesian text on social media.

Moreover, the IndoBERT labelling method was also better than VADER in both models. The problem lies in the transformer architecture applied on IndoBERT, where all words in a sentence are processed at once. With the self-attention mechanism, IndoBERT could investigate how much all the words relate to each other in context rather than only according to word order. This ability allows IndoBERT to better understand sentence meaning in context, which contributes to enhancement for classification model engineering. Nonetheless, IndoBERT still has weaknesses on the prediction task; see Table 6.

Table 6. Examples of the weakness of IndoBERT

No	Comments	Manually	IndoBERT
1	Nyesel milih Jokowi	NEGATIF	NETRAL
2	Bangga? Amit2	NETRAL	POSITIF
3	Jangan lupa utang negara pres 🙌	NETRAL	POSITIF

The greatest strength of SVM is that it can create the optimal boundary for classification, especially non-linear separable data. The use of RBF kernel in our study means that SVM can identify more tense and contextual patterns between comment data. Besides, due to the working mechanism of SVM that maximizes the distance between classes, this model is more reliable and accurate in dealing with various types of input.

4. Conclusions

The SVM algorithms were effectively used in this research to assess the public attitude toward comments of inaugural ceremony held at IKN. Sentiments in the comments were labelled as positive, neutral, and negative sentiments using VADER concepts combined with IndoBERT label. Result shows that SVM with IndoBERT labelling performs the best, having 95.05% in accuracy score. SVM has the advantage of being able to process complex text data, particularly with RBF kernel. On the other hand, IndoBERT was more effective compared to VADER as it has a better grasp of Indonesian language context. To conclude, the use of SVM with IndoBERT offered state-of-the-art results for sentiment classification in this study.

5. References

[1] Cindy Mutia Annur, "Indonesia Jadi Negara dengan Pengguna Instagram Terbanyak ke-4 di Dunia," Katadata. Accessed: Oct. 09, 2024. [Online]. Available: https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/16d2a8a4aaa0bc2/indonesia-jadi-negara-dengan-pengguna-instagram-terbanyak-ke-4-di-dunia

[2] D. Pramana, M. Afdal, M. Mustakim, and I. Permana, "Analisis sentimen terhadap pemindahan ibu kota negara menggunakan algoritma Naive Bayes Classifier dan K-Nearest Neighbors," *Jurnal Media Informatika Budidarma*, vol. 7, no. 3, 2023, doi: 10.30865/mib.v7i3.6523.

[3] K. Han, W. Chien, C. Chiu, and Y. Cheng, "Application of support vector machine (SVM) in the sentiment analysis of Twitter dataset," *Applied Sciences*, vol. 10, no. 3, p. 1125, 2020, doi: 10.3390/app10031125.

[4] F. Rifaldy, Y. Sibaroni, and S. Prasetiyowati, "Effectiveness of Word2Vec and TF-IDF in sentiment classification on online investment platforms using support vector machine," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 2, pp. xxx–xxx, 2025, doi: 10.29100/jipi.v10i2.6055.

[5] S. Sofi and A. Selamat, "Aspect-based sentiment analysis: Feature extraction using latent Dirichlet allocation (LDA) and term frequency–inverse document frequency (TF-IDF) in machine learning," *Malaysian Journal of Information and Communication Technology (MyJICT)*, vol. 8, no. 2, 2023, doi: 10.53840/myjict8-2-102.

[6] L. Muflikhah and D. Haryanto, “High performance of polynomial kernel at SVM algorithm for sentiment analysis,” *Journal of Information Technology and Computer Science*, vol. 3, no. 2, 2018, doi: 10.25126/jitecs.20183260.

[7] A. Saputra, R. Saputro, and D. Saputra, “Enhancing sentiment analysis accuracy using SVM and slang word normalization on YouTube comments,” *Sinkron*, vol. 9, no. 2, 2025, doi: 10.33395/sinkron.v9i2.14613.

[8] M. Mujahid, E. Kına, F. Rustam, M. Villar, E. Alvarado, I. De La Torre Díez, and I. Ashraf, “Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering,” *Journal of Big Data*, vol. 11, no. 1, pp. 1–32, 2024, doi: 10.1186/s40537-024-00943-4.

[9] D. Salifu, L. Chepkemoi, E. Ibrahim, K. Nkoba, and H. Tonnang, “Data augmentation and machine learning algorithms for multi-class imbalanced morphometrics data of stingless bees,” *Heliyon*, vol. 11, p. e42214, 2025, doi: 10.1016/j.heliyon.2025.e42214.

[10] L. Martini, G. Pradipta, and R. Huizen, “Analysis of the impact of data oversampling on the support vector machine method for stroke disease classification,” *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, 2025, doi: 10.35882/jeeemi.v7i2.698.

[11] B. Wilie, K. Vincentio, G. Winata, S. Cahyawijaya, X. Li, Z. Lim, S. Soleman, R. Mahendra, P. Fung, S. Bahar, and A. Purwarianti, “IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding,” in *Proceedings of the Asia-Pacific Chapter of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing (AACL–IJCNLP)*, 2020, pp. 843–857, doi: 10.18653/v1/2020.aacl-main.85.

[12] H. Imaduddin, F. A’la, and Y. Nugroho, “Sentiment analysis in Indonesian healthcare applications using IndoBERT approach,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 8, pp. xxx–xxx, 2023, doi: 10.14569/ijacsa.2023.0140813.

[13] R. Jain, A. Kumar, A. Nayyar, K. Dewan, R. Garg, S. Raman, and S. Ganguly, “Explaining sentiment analysis results on social media texts through visualization,” *Multimedia Tools and Applications*, vol. 82, no. 15, pp. 22613–22629, 2023, doi: 10.1007/s11042-023-14432-y.

[14] V. Bonta, N. Kumaresh, and N. Janardhan, “A comprehensive study on lexicon-based approaches for sentiment analysis,” *Asian Journal of Computer Science and Technology*, vol. 8, no. S2, 2019, doi: 10.51983/ajcst-2019.8.s2.2037.

[15] G. Diri, P. Obiorah, and H. Du, “Comparative study of sentiment analysis techniques: Traditional machine learning vs. deep learning approaches,” in *Proceedings of the InSITE Conference*, 2025, doi: 10.28945/5490.

[16] A. Pimpalkar and R. Raj, “Influence of pre-processing strategies on the performance of ML classifiers exploiting TF-IDF and BoW features,” *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 9, no. 2, pp. 49–68, 2020, doi: 10.14201/adcaij2020924968.

6. Acknowledgement

The authors would also like to thank the Department of Informatics at Sanata Dharma University for providing facilities in smart computing laboratory that made this study become possible.