LEMBAGA PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT
UNIVERSITAS BINA SARANA INFORMATIKA
Alamat : Jl. Kramat Raya No. 98, Senen. Jakarta Pusat 10450
Telp. (021) 23231170 Fax (021) 21236158 Ext. 704, 705 e-mail: lppm@bsi.ac.id

SURAT TUGAS

0217/B.01/LPPM-UBSI/II/2026

Tentang

PENELITIAN YANG DIPUBLIKASIKAN DALAM PROSIDING INTERNATIONAL
Periode September 2025 - Februari 2026

Menulis Pada Prosiding IEEE International Conference on Advanced Information Scientific Development (ICAISD), 04-05 November 2025 | 19 February 2026 | Jakarta, Indonesia (ISBN : 979-8-3315-7499-4 | USB ISBN : 979-8-3315-7498-7 | ISBN : 979-8-3315-7500-7

Judul Makalah :

Sentiment Analysis on Indonesian Public Perception of Electric Vehicles Using IndoBERT and SMOTE

Menimbang :
1. Bahwa perlu di adakan pelaksanaan Kegiatan dalam rangka Seminar.
2. Untuk keperluan tersebut, pada butir 1 (satu) di atas, maka perlu dibentuk kelompok Seminar.

MEMUTUSKAN

Pertama : Menugaskan kepada saudara

1. 200908570 Sardiarinto
2. 200810954 Candra Agustina
3. 201510342 Angga Ardiansyah
4. 201103245 Rizal Amegia Saputra
5. 200209820 Sri Wasiyanti
6. 201610582 Raja Sabaruddin

Sebagai Penulis yang mempublikasikan Penelitianya.

Kedua : Mempunyai tugas sbb:
Melaksanakan Tugas yang diberikan dengan penuh rasa tanggung jawab.

Ketiga : Keputusan ini berlaku sejak tanggal ditetapkan, dengan ketentuan apabila dikemudian hari terdapat kekeliruan akan diubah dan diperbaiki sebagaimana mestinya.

Jakarta, 12 Februari 2026
LPPM Universitas Bina Sarana Informatika
Ketua

Agus Junaidi, M. Kom

Tembusan
- Ka. LPPM Universitas Bina Sarana Informatika
- Arsip
- Ybs

Sentiment Analysis on Indonesian Public Perception of Electric Vehicles Using IndoBERT and SMOTE

1st Sardiarinto
Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
Jakarta, Indonesia
sardiarinto.sdo@bsi.ac.id

2nd Candra Agustina
Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
Jakarta, Indonesia
candra.caa@bsi.ac.id

3rd Angga Ardiansyah
Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
Jakarta, Indonesia
angga.axr@bsi.ac.id

4th Rizal Amegia Saputra
Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
Jakarta, Indonesia
rizal.rga@bsi.ac.id

5th Sri Wasiyanti
Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
Jakarta, Indonesia
sri.siw@bsi.ac.id

6th Raja Sabaruddin
Sistem Informasi Akuntansi
Universitas Bina Sarana Informatika
Jakarta, Indonesia
raja.rjd@bsi.ac.id

2025 IEEE International Conference on Advanced Information Scientific Development (ICAISD) | 979-8-3315-7499-4/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICAISD68166.2025.11385356

***Abstract*—The rapid development of the electric vehicle (EV) industry in Indonesia has been driven by strong government support and global demand for sustainable transportation. However, public adoption remains limited, highlighting the importance of understanding societal perceptions toward EVs. This study applies sentiment analysis to examine Indonesian public opinion using a dataset of 764 articles collected from CNBC Indonesia, Otoshifter, and CNN Indonesia. The methodology combines traditional machine learning baselines with a Transformer-based IndoBERT model, enhanced by the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance. Experimental results demonstrate that although the model achieved a relatively high overall accuracy of 88.89%, its macro-level performance was significantly weaker, with a macro F1-score of 41.33% due to challenges in detecting minority sentiment classes. These findings underscore the need for balanced data representation and advanced modeling techniques to improve minority-class detection. The study contributes by providing a domain-specific dataset, validating IndoBERT's effectiveness in the EV domain, and highlighting the impact of class imbalance on sentiment classification in Indonesian texts.**

***Index Terms*—Electric Vehicles, Sentiment Analysis, IndoBERT, SMOTE, Indonesian Language Processing, Imbalanced Data, Natural Language Processing (NLP)**

I. Introduction

The electric vehicle (EV) industry in Indonesia has experienced significant growth in recent years, driven by strong government support and increasing global demand for sustainable transportation. The Indonesian government aims to establish a domestic EV ecosystem by 2025, including battery production, charging infrastructure, and vehicle assembly, supported by major investments such as the Hyundai–LG battery plant and VinFast's plan to install up to 100,000 charging stations nationwide [1] [2]. However, despite this momentum, EV adoption remains below 7% of total vehicles in the domestic market. Understanding public perception is therefore crucial to identifying both the barriers and enablers of EV adoption. [3] [4].

Sentiment analysis has emerged as a powerful tool to capture public attitudes toward EVs in real time, leveraging data from social media posts, news articles, and consumer reviews. Insights obtained from sentiment analysis can help manufacturers design more effective promotional strategies, assist policymakers in formulating incentive programs, and support investors in assessing market readiness [5] [6]. Nevertheless, sentiment analysis in the Indonesian language presents unique challenges, including the frequent use of slang and informal vocabulary, diverse sentence structures, and the issue of imbalanced datasets—where certain sentiment classes tend to dominate [7]. These characteristics require models that are sensitive to local linguistic contexts and capable of addressing class imbalance to minimize classification bias [8].

Recent advances in natural language processing (NLP) have introduced IndoBERT, a Transformer-based pre-trained language model specifically designed for the Indonesian language. IndoBERT has demonstrated strong performance in capturing semantic and syntactic nuances within Indonesian text, surpassing traditional machine learning approaches, particularly in domain-specific sentiment classification tasks [4] [9]. To further address the imbalance problem, oversampling techniques such as the Synthetic Minority Oversampling Technique (SMOTE) have been widely applied [10] [11]. In Indonesian sentiment classification, the integration of SMOTE has been shown to significantly improve model performance, with higher accuracy and more balanced F1-scores across sentiment classes.

This study contributes in three key areas: (1) the development of a domain-specific Indonesian dataset focusing on public perceptions of EVs, (2) the application of IndoBERT in combination with SMOTE to enhance sentiment classification performance, and (3) a comprehensive evaluation of the model's effectiveness in addressing the challenges of linguistic diversity and class imbalance in Indonesian sentiment analysis.

II. Related Work

Research on electric-vehicle (EV) adoption typically highlights two intertwined streams: technology diffusion in emerging markets and the role of public sentiment in shaping purchase intentions. Studies in the Indonesian context report that policy incentives, charging infrastructure, and total cost of ownership are persistent determinants of adoption, while concerns about driving range, battery longevity, and resale value remain barriers. In parallel, work that mines digital traces (news, social media, and app reviews) shows that public discourse around EVs oscillates with policy announcements and product launches, making sentiment analysis a practical lens for tracking readiness and pain points over time [2] [12].

A second stream examines sentiment analysis methods for mobility and energy topics [13] [14]. Global studies on EVs have leveraged Twitter, forum posts, and news headlines to quantify attitudes toward environmental benefits, brand perception, and infrastructure adequacy [14] [15]. These works consistently find that fine-grained, aspect-level sentiment (e.g., battery, charging, cost) is more informative than document-level polarity, and that temporal analysis captures shifts following policy changes or major events. However, most large-scale EV sentiment studies are conducted in English or Chinese, leaving a gap for high-quality analyses in Indonesian [5] [16].

Indonesian-language sentiment analysis introduces linguistic challenges that differ from high-resource languages. Informal morphology, slang and abbreviations, productive reduplication, and frequent code-mixing with English reduce the effectiveness of rule-based preprocessing and classical bag-of-words representations [17] [18]. Prior Indonesian NLP work shows that naively applying tokenization, stop-word removal, or stemming can harm downstream performance when slang or domain terms carry sentiment [19]. As a result, recent studies emphasize robust normalization (e.g., slang dictionaries), subword tokenization, and models that can learn contextual semantics directly from raw text [20].

Transformer-based models pre-trained for Indonesian, particularly IndoBERT and its social-media variant (often referred to as IndoBERTweet), have become de-facto baselines for Indonesian sentiment classification. Across domains such as product, service, and policy discourse, these models consistently outperform classical machine-learning methods (e.g., Naïve Bayes, SVM, logistic regression) and recurrent/CNN architectures, especially on noisy text. Fine-tuning IndoBERT with modest task-specific data yields strong gains due to its ability to encode subword morphology and long-range dependencies. Nevertheless, domain shift remains a concern: performance can drop when moving from product reviews to policy debate or from formal news to colloquial social media, motivating domain-aware fine-tuning and careful data curation.

A recurring practical issue in Indonesian sentiment datasets is class imbalance—for instance, neutral or positive classes dominating organic discussions, or negative spikes around specific events. The imbalanced-learning literature offers several remedies. Data-level methods such as SMOTE and its variants (Borderline-SMOTE, ADASYN) synthesize minority-class samples to balance training sets; algorithm-level methods such as class-weighted losses and focal loss adjust the optimization objective; and ensemble methods (e.g., EasyEnsemble) combine both. In text classification, SMOTE is most effective when applied on dense, semantically meaningful representations (e.g., TF-IDF with dimensionality reduction or sentence embeddings) rather than raw sparse vectors; when paired with modern encoders like BERT/IndoBERT, oversampling or class-weighted losses typically improves minority-class recall without unduly hurting precision. Recent Indonesian studies report that applying SMOTE (or related strategies) can raise macro-F1 and minority-class F1 substantially, though care is needed to avoid information leakage (e.g., oversampling before train or test split) and to monitor overfitting on synthetic samples.

Logistic Regression is a classical machine learning algorithm widely applied in classification tasks. Unlike linear regression, which predicts continuous outcomes, Logistic Regression models the probability of a data point belonging to a certain class by applying the logit or sigmoid function. This transformation maps the output into a range between 0 and 1, making it suitable for binary and multi-class classification problems [21].

In sentiment analysis, Logistic Regression is often used as a baseline model because it is simple, efficient, and interpretable. The algorithm assumes a linear relationship between the input features and the log-odds of the target variable. Despite its simplicity, it has proven effective in many text classification scenarios [22].

Comparative evaluations consistently show that pre-trained Transformers (BERT family) outperform classical baselines and LSTM/CNN models on Indonesian sentiment tasks, particularly under noisy, code-mixed conditions [23], [24]. That said, simple baselines remain useful for ablations and cost-sensitive deployments, and can serve as sanity checks for data quality and label balance. Error analyses across studies highlight common failure modes: sarcasm and irony, negation scope with intensifiers, entity-level sentiment shifts within a single post, and domain terms (e.g., EV-specific jargon such as *SoC*, *fast-charging*, *WLTP*) [25].

Positioning and gaps. Despite rapid progress, there remain notable gaps relevant to this work: (i) a lack of publicly documented, domain-specific Indonesian datasets focused on EV public perception; (ii) limited, controlled studies that combine IndoBERT with imbalanced-learning strategies (e.g., SMOTE vs. class-weighted/focal loss) and report per-class metrics (macro-F1, minority-class recall) under realistic noise; and (iii) insufficient aspect-level analysis that disentangles sentiment toward charging, cost, performance, and policy. Addressing these gaps, the present study (a) curates an EV-focused Indonesian dataset from public discourse, (b) fine-tunes IndoBERT with and without SMOTE to quantify the effect of oversampling on minority-class detection, and (c) reports comprehensive metrics and error analyses to clarify when and why oversampling helps in an Indonesian EV

context.

III. Methodology

A. Data Collection

The dataset consisted of **764 text documents** compiled from three major Indonesian digital news outlets: **CNBC Indonesia**, **Otoshifter**, and **CNN Indonesia**. These sources were selected to capture a diverse range of perspectives on electric vehicles (EVs) within the Indonesian context, including economic, technological, and consumer-oriented viewpoints.

Data acquisition was conducted using **web scraping techniques**, focusing on domain-specific keywords such as *"electric vehicle"*, *"EV Indonesia"*, and *"battery electric vehicle"*. The retrieved articles were then stored in a structured format (CSV/JSON) to facilitate subsequent preprocessing and analysis.

B. Preprocessing

Raw textual data underwent several preprocessing steps to ensure high-quality inputs for modeling:

1) **Text cleaning**: removal of URLs, hashtags, emojis, excessive punctuation, and non-alphabetic characters.
2) **Case folding**: converting all characters to lowercase.
3) **Tokenization**: segmenting text into individual tokens.
4) **Normalization**: standardizing slang and informal expressions into formal Indonesian equivalents (e.g., *gk* $\rightarrow$ *tidak*).
5) **Stopword removal**: eliminating high-frequency but semantically insignificant words (e.g., *yang*, *dan*, *atau*).
6) **Stemming**: reducing words to their root form (e.g., *berjalan* $\rightarrow$ *jalan*) [19].

C. Modeling

Two types of models were employed for sentiment classification:

1. Baseline Model
   Logistic Regression with TF-IDF was applied as the baseline model. The approach relies on traditional feature engineering, where text data are transformed into numerical representations using Term Frequency–Inverse Document Frequency (TF-IDF). This model establishes a benchmark for evaluating the performance of deep learning–based approaches.
2. Proposed Model
   IndoBERT, a Transformer-based pre-trained language model specifically designed for the Indonesian language, was fine-tuned on the EV sentiment dataset. Unlike the baseline model, IndoBERT captures contextual semantics and linguistic nuances, making it more suitable for handling informal morphology, slang, and code-mixing that frequently occur in Indonesian-language texts.

D. Handling Class Imbalance

As sentiment data tend to be imbalanced, with positive or neutral classes often dominating, the **Synthetic Minority Oversampling Technique (SMOTE)** was applied.

1) SMOTE generates synthetic samples of the minority classes by interpolating between existing instances using *k-nearest neighbors*.
2) This approach enhances class representation, reducing bias toward majority classes and improving overall classification performance.

E. Evaluation Metrics

Model performance was assessed using multiple evaluation metrics:

1) **Accuracy**: overall proportion of correctly predicted labels.
2) **Precision**: proportion of true positive predictions among all predicted positives.
3) **Recall**: proportion of correctly identified positive samples out of all actual positives.

IV. Results and Discussion

A. Confusion Matrix

As illustrated in Figure 1, the model configured with **IndoBERT embeddings + SMOTE + Logistic Regression** still demonstrates a bias toward the majority class (*neutral*). Out of 153 test samples, nearly all predictions fall into the neutral class, with only a small portion of *positive* samples correctly identified, while the *negative* class remains undetected. This outcome emphasizes the persisting challenge of **class imbalance** in the dataset.

B. Overall Metrics

Table 1 summarizes the overall performance of the model. While the accuracy is relatively high (0.8627), macro-averaged metrics are considerably lower. Precision (0.4726), Recall (0.4847), and F1-score (0.4785) reveal that the model struggles to consistently capture minority classes, despite its strength in identifying the majority class.

As presented in Table 1, the model achieved an overall accuracy of 0.8627, indicating that it performs reasonably well in predicting the dominant class. However, the macro-averaged metrics—precision (0.4726), recall (0.4847), and F1-score (0.4785)—highlight the model's limited ability to balance performance across all sentiment categories. This

TABLE I
Overall Metrics

Metric	Value
Accuracy	0.8627
Precision (Macro)	0.4726
Recall (Macro)	0.4847
F1-score (Macro)	0.4785

discrepancy suggests that while the model is reliable for majority-class predictions, it struggles to adequately capture minority sentiments, reinforcing the need for strategies to address class imbalance.

C. Per-Class Performance

The classification report in Table 2 highlights the variation in performance across classes. The *neutral* class dominates with high recall, while the *negative* class fails to be detected. The *positive* class achieves moderately good precision but low recall, resulting in a relatively weak F1-score. These findings reinforce the conclusion that the imbalance between classes heavily impacts model effectiveness.

TABLE II
CLASSIFICATION REPORT PER CLASS

Class	Precision	Recall	F1-score	Support
Negative	0.0000	0.0000	0.0000	3
Neutral	0.8890	0.9323	0.9101	133
Positive	0.5294	0.5294	0.5294	17
Macro Avg	0.4726	0.4847	0.4785	153
Weighted Avg	0.8350	0.8627	0.8479	153

As shown in Table 2, the model performs strongly in predicting the Neutral class, achieving high precision (0.8890), recall (0.9323), and F1-score (0.9101). However, performance significantly drops for the Positive class (F1-score 0.5294) and is completely ineffective for the Negative class, where all metrics are 0.0000 due to the very limited number of samples (n=3). This imbalance skews the overall results, as reflected in the macro-averaged F1-score of 0.4785, while the weighted average F1-score of 0.8479 indicates that the model is biased toward the majority Neutral class. These findings confirm that although the classifier achieves high overall accuracy, its ability to generalize across underrepresented sentiment categories remains inadequate, emphasizing the importance of applying resampling or class-balancing techniques.

D. Interpretation

Overall, while the model achieves a high accuracy (0.8627), per-class analysis indicates substantial limitations. The model is overly biased toward the *neutral* class, fails to identify the *negative* class, and performs only moderately well on the *positive* class. These results highlight the need for more robust strategies to address class imbalance, such as advanced oversampling, ensemble methods, cost-sensitive learning, or further fine-tuning of IndoBERT to improve the detection of minority classes.

V. CONCLUSION AND FUTURE WORK

The sentiment classification model was evaluated using accuracy, precision, recall, and F1-score. Results indicated an overall accuracy of 88.89% with strong performance in classifying neutral sentiment, while performance significantly declined for minority classes. The macro-averaged F1-score of 41.33% and the inability to correctly detect negative sentiment suggest that the model is less effective for underrepresented categories. These findings highlight the need for improvements in handling class imbalance to ensure more robust and equitable sentiment analysis across all categories.

Future work should address this limitation by: (i) employing resampling and data augmentation strategies to increase representation of minority classes, (ii) integrating cost-sensitive and ensemble learning approaches to reduce misclassification bias, and (iii) leveraging advanced contextual embeddings such as BERT or RoBERTa to capture richer semantic nuances. Additionally, evaluations should incorporate macro- and micro-averaged metrics alongside stratified cross-validation to ensure reliability and generalizability. By adopting these approaches, subsequent studies can contribute to the development of more balanced, resilient, and domain-adaptive sentiment analysis systems capable of supporting real-world applications.

REFERENCES

[1] P. Indonesia - Automotive Team, "The Road Ahead: Indonesia's Electric Vehicle Readiness and Consumer Insights 2024," 2024. [Online]. Available: www.pwc.com/id .

[2] D. F. Hakam and S. Jumayla, "Electric vehicle adoption in Indonesia: Lesson learned from developed and developing countries," *Sustainable Futures*, vol. 8, Dec. 2024, doi: 10.1016/j.sftr.2024.100348.

[3] Y. A. Singgalen, "Analisis Sentimen Konsumen terhadap Food, Services, and Value di Restoran dan Rumah Makan Populer Kota Makassar Berdasarkan Rekomendasi Tripadvisor Menggunakan Metode CRISP-DM dan SERVQUAL," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 4, Mar. 2023, doi: 10.47065/bits.v4i4.3231.

[4] Y. A. Singgalen, "Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data," *Article in Journal of Information System Research*, vol. 6, no. 2, 2025, doi: 10.47065/josh.v6i2.6505.

[5] S. Wang, M. Li, R. Assitant, B. Yu, S. Bao, and Y. Chen, "Investigating The Impacting Factors on The Public's Attitudes Towards Autonomous Vehicles Using Sentiment Analysis from Social Media Data," 2021. Accessed: Aug. 20, 2025. [Online]. Available: https://arxiv.org/pdf/2108.03206.

[6] S. Wang, M. Li, R. Assitant, B. Yu, S. Bao, and Y. Chen, "Investigating The Impacting Factors on The Public's Attitudes Towards Autonomous Vehicles Using Sentiment Analysis from Social Media Data," 2021. Accessed: Aug. 20, 2025. [Online]. Available: https://arxiv.org/pdf/2108.03206.

[7] R. N. S. M. Jen, S. N. Kapita, and M. Fhadli, "Comparison of Normalization of Indonesian Slang Words Using the FastText & Word2vec Model with the Natural Language Processing Approach," *Nusantara Science and Technology Proceedings*, pp. 40–49, Apr. 2025, doi: 10.11594/nstp.2025.4805.

[8] M. A. Fauzi, R. F. N. Firmansyah, and T. Afirianto, "Improving sentiment analysis of short informal Indonesian product reviews using synonym based feature expansion," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 16, no. 3, pp. 1345–1350, Jun. 2018, doi: 10.12928/TELKOMNIKA.v16i3.7751.

[9] F. Koto Jey Han Lau Timothy Baldwin, "INDOBERTWEET: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization," 2020. [Online]. Available: https://huggingface.co/huseinzol05/.

[10] D. A. Kristiyanti, S. A. Sanjaya, V. C. Tjokro, and J. Suhali, "Dealing imbalance dataset problem in sentiment analysis of recession in Indonesia," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 2, pp. 2058–2070, Jun. 2024, doi: 10.11591/ijai.v13.i2.pp2060-2072.

[11] 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.

[12] T. W. Sasongko, U. Ciptomulyono, B. Wirjodirdjo, and A. Prastawa, "Identification of electric vehicle adoption and production factors based on an ecosystem perspective in Indonesia," *Cogent Business and Management*, vol. 11, no. 1, 2024, doi: 10.1080/23311975.2024.2332497.
[13] E. C. M. Torres and L. G. de Picado-Santos, "Sentiment Analysis and Topic Modeling in Transportation: A Literature Review," Jun. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/app15126576.
[14] J. T. Ren, "Harnessing Public Sentiment: A Literature Review of Sentiment Analysis in Energy Research," 2025.
[15] Y. Özkara *et al.*, "Analysing Social Media Discourse on Electric Vehicles with Machine Learning," *Applied Sciences (Switzerland)*, vol. 15, no. 8, Apr. 2025, doi: 10.3390/app15084395.
[16] M. K. Walke, "Evaluating Global News Sentiment on Electric Vehicle Adoption Using Gemini 2.0," 2025.
[17] A. Fikri Aji *et al.*, "One Country, 700+ Languages: NLP Challenges for Underrepresented Languages and Dialects in Indonesia," Long Papers, 2022.
[18] A. F. Hidayatullah, R. A. Apong, D. T. C. Lai, and A. Qazi, "Corpus creation and language identification for code-mixed Indonesian-Javanese-English Tweets," *PeerJ Comput Sci*, vol. 9, 2023, doi: 10.7717/PEERJ-CS.1312.
[19] R. Budiarto, P. Siber, and S. Negara, "Text Preprocessing for Optimal Accuracy in Indonesian Sentiment Analysis Using a Deep Learning Model with Word Embedding," 2021. [Online]. Available: https://www.researchgate.net/publication/365799401.
[20] M. F. Adilazuarda, S. Cahyawijaya, G. I. Winata, P. Fung, and A. Purwarianti, "IndoRobusta: Towards Robustness Against Diverse Code-Mixed Indonesian Local Languages," Nov. 2023, [Online]. Available: http://arxiv.org/abs/2311.12405.
[21] J. Daniel and J. H. Martin, "Speech and Language Processing," 2024.
[22] S. A. Salloum, R. Alfaisal, A. Basiouni, K. Shaalan, and A. Salloum, "Effectiveness of Logistic Regression for Sentiment Analysis of Tweets About the Metaverse," in *Communications in Computer and Information Science*, Springer Science and Business Media Deutschland GmbH, 2024, pp. 32–41. doi: 10.1007/978-3-031-65996-6_3.
[23] Dhendra and V. Gayuh Utomo, "Benchmarking IndoBERT and Transformer Models for Sentiment Classification on Indonesian E-Government Service Reviews," *Jurnal Transformatika*, vol. 23, no. 1, pp. 86–95, Jul. 2025, doi: 10.26623/transformatika.v23i1.12095.
[24] A. F. Hidayatullah, R. A. Apong, D. T. C. Lai, and A. Qazi, "Word Level Language Identification in Indonesian-Javanese-English Code-Mixed Text," in*Procedia Computer Science*, Elsevier B.V., 2024, pp. 105–112. doi: 10.1016/j.procs.2024.10.183.
[25] D. Suhartono, W. Wongso, and A. Tri Handoyo, "IdSarcasm: Benchmarking and Evaluating Language Models for Indonesian Sarcasm Detection," *IEEE Access*, vol. 12, pp. 87323–87332, 2024, doi: 10.1109/ACCESS.2024.3416955.