Jurnal Penelitian Sains Teknologi

IndoBERT-Based Sentiment Analysis of Electric Motorcycle Policy in Indonesia Using Instagram Data

Muhammad Syahriandi Adhantoro[1✉], Faris Atoil Haq[2], Dody Hartanto[3], Anindita Widagda Pandam Sudaryanto[4]

[1,2]Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta, Indonesia
[3]Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Indonesia
[4]Faculty of Engineering, Universitas Islam Batik, Indonesia

doi: 10.23917/saintek.v2i2.17021
Received: 28 Maret 2026 | Revised: 15 April 2026 | Accepted: 18 April 2026
Available Online: 20 April 2026 | Published Regularly: September 2026

Abstract

This study aims to analyze public sentiment toward the procurement of electric motorcycles within the Nutritional Service Fulfillment Unit/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) program in Indonesia by utilizing data from Instagram. The approach employed is a deep learning-based sentiment analysis using the IndoBERT model, which has been fine-tuned to classify data into positive, negative, and neutral categories. The research stages include data collection, preprocessing, labeling, model development, and model evaluation using accuracy, precision, recall, and F1-score metrics. The results indicate that public sentiment is predominantly negative at 80%, followed by positive sentiment at 15% and neutral sentiment at 5%. Further analysis reveals that negative sentiment is primarily driven by issues related to budget prioritization, infrastructure readiness, and policy effectiveness, while positive sentiment is associated with environmental benefits and improved service distribution efficiency. The model evaluation demonstrates that IndoBERT achieves high performance, with an accuracy of 0.89, precision of 0.88, recall of 0.90, and F1-score of 0.89. These findings indicate that IndoBERT is effective in capturing the contextual nuances of the Indonesian language in unstructured social media data. This study contributes to the advancement of transformer-based sentiment analysis methods and provides data-driven insights to support more responsive and evidence-based policymaking.

Keywords: sentiment analysis, IndoBERT, deep learning, public opinion, electric motorcycles, social media analytics, instagram data, public policy, sustainable transportation, *satuan pelayanan pemenuhan gizi*

✉**Corresponding Author:**
Muhammad Syahriandi Adhantoro, Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta, Indonesia
Email: m.syahriandi@ums.ac.id

Introduction

The development of electric vehicle technology in recent years has demonstrated a significant upward trend as part of global efforts to achieve sustainable transportation and reduce carbon emissions. In Indonesia, the government has actively promoted the adoption of electric vehicles, including electric motorcycles, through various strategic policies and implementation programs. One notable initiative is the procurement of electric motorcycles to support the operations of

Nutritional Service Fulfillment Units/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) across multiple regions in Indonesia [1]. This program not only contributes to improving the efficiency of nutritional service distribution to the community but also represents a broader transformation toward a more modern and environmentally friendly public service system.

However, the implementation of this policy is accompanied by diverse public perceptions. On the one hand, the public has expressed positive responses toward the use of electric motorcycles, which are considered energy-efficient, environmentally friendly, and aligned with sustainable development goals [2]. On the other hand, negative sentiments have also emerged, particularly concerning issues such as limited charging infrastructure, vehicle durability, procurement costs, and policy transparency and effectiveness [3]. This polarization of public opinion highlights the need for systematic and data-driven analysis, as public acceptance plays a crucial role in determining the success of policy implementation.

In the digital era, social media has become a primary platform for individuals to express opinions, perceptions, and reactions to public policies. Instagram, as one of the most widely used platforms in Indonesia, provides rich, dynamic, and real-time data that reflects public opinion [4]. Therefore, leveraging Instagram data for sentiment analysis presents a relevant and effective approach to understanding public perception more comprehensively. Such analysis enables the identification of sentiment patterns positive, negative, and neutral regarding the procurement of electric motorcycles within the SPPG program.

With the advancement of technology, sentiment analysis approaches have evolved from conventional machine learning methods to deep learning-based techniques, particularly those utilizing transformer architectures. IndoBERT, a pre-trained language model specifically designed for the Indonesian language, has demonstrated superior performance in capturing linguistic and semantic context [5]. Compared to traditional approaches, IndoBERT is capable of understanding contextual meaning more effectively, thereby improving sentiment classification accuracy. Nevertheless, studies that specifically investigate the use of IndoBERT for analyzing public sentiment toward electric motorcycle procurement within the context of nutritional service programs remain limited [4].

This limitation indicates a research gap, particularly in integrating deep learning-based sentiment analysis with public policy contexts in the nutritional service sector, as well as in utilizing Instagram as the primary data source for such analysis. Accordingly, this study proposes an IndoBERT-based sentiment analysis approach to examine public perceptions of electric motorcycle procurement in Indonesia’s SPPG program. The objectives of this study are to identify and classify public sentiment based on Instagram data, analyze the distribution of sentiment across different regions, and evaluate the effectiveness of the IndoBERT model in understanding public opinion within this context. Furthermore, this study aims to provide data-driven insights that can support more responsive and evidence-based policymaking [6].

The novelty of this research lies in the integration of IndoBERT for sentiment analysis within the context of public policy related to nutritional services, the utilization of Instagram as a primary data source that is relatively underexplored compared to other platforms, and the focus on the specific issue of electric motorcycle procurement in the SPPG program, which has received limited attention in the existing literature [7]. Therefore, this study contributes not only to the advancement of deep learning-based sentiment analysis methodologies but also to a more comprehensive understanding of public opinion dynamics in the implementation of sustainable transportation policies in Indonesia.

Method

This study employs a quantitative approach based on deep learning to analyze public sentiment toward the procurement of electric motorcycles within the Nutritional Service Fulfillment Unit/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) program in Indonesia. The data are collected from Instagram, which serves as a dynamic and real-time representation of public opinion. In general, the research workflow follows the architecture of *pre-training* and *fine-tuning* in the IndoBERT model, as illustrated in Figure 1.

a. Research Stages

The stages of this study are described as follows:

1. Data Collection

Data were collected from Instagram in the form of posts and comments containing relevant keywords such as "electric motorcycles," "SPPG," and "nutrition." The data collection process was conducted using web scraping techniques while adhering to research ethics and platform privacy policies. The collected data were then filtered to ensure relevance to the research topic.

2. Data Preprocessing

This stage aims to improve data quality prior to analysis. The preprocessing steps include:

- *Cleaning*: removing URLs, emojis, punctuation, and special characters
- *Case folding*: converting all text into lowercase
- *Tokenization*: splitting text into individual tokens
- *Stopword removal*: eliminating non-informative words
- *Stemming*: reducing words to their base form

3. Data Labeling

The preprocessed data were annotated into three sentiment categories:

- Positive
- Negative
- Neutral

The labeling process was conducted manually by multiple annotators to ensure data quality. Additionally, *inter-annotator agreement* was applied to validate labeling consistency.

4. Model Development

The model used in this study is IndoBERT, which has been pre-trained on a large Indonesian language corpus. The following steps were performed:

- Fine-tuning the model using the labeled dataset
- Splitting the dataset into training and testing sets

- Adjusting model parameters to optimize classification performance

5. Model Evaluation

The model performance was evaluated using the following metrics:

- Accuracy
- Precision
- Recall
- F1-Score

In addition, a *confusion matrix* was utilized to analyze the distribution of classification errors across sentiment classes.

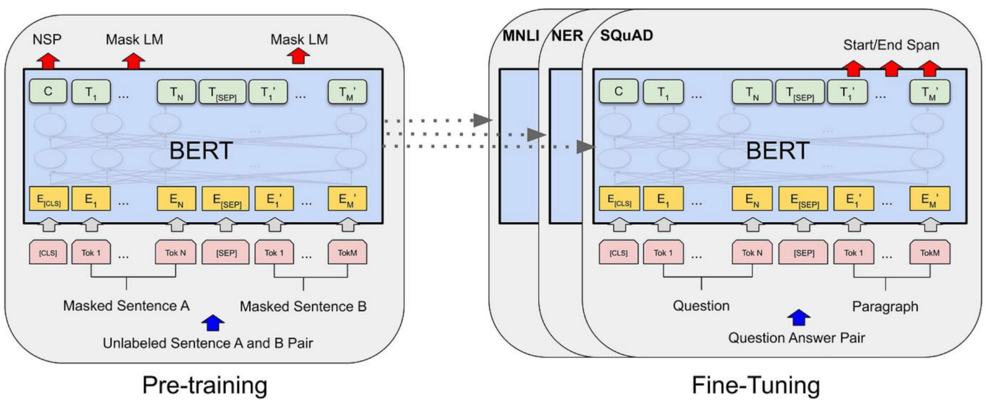

Figure 1. IndoBERT Architecture (Pre-training and Fine-tuning)

Figure 1 illustrates the two main stages in the BERT architecture, namely *pre-training* and *fine-tuning*, which form the foundation of IndoBERT used in this study. In the *pre-training* stage, the model is trained using two primary objectives: *Masked Language Modeling (MLM)* and *Next Sentence Prediction (NSP)*. MLM enables the model to understand contextual word representations by predicting masked tokens within a sentence, while NSP helps the model learn relationships between sentence pairs. This stage produces rich semantic and contextual language representations [8].

In the *fine-tuning* stage, the pre-trained model is adapted to a specific downstream task, which in this study is sentiment classification. The labeled dataset obtained from preprocessing and annotation is used to train the model to classify text into sentiment categories [9]. The CLS token representation is utilized as the primary feature for classification.

The relevance of Figure 1 in this study lies in demonstrating that IndoBERT is not trained from scratch but rather leverages a pre-trained model that is subsequently adapted to the specific research task. This approach significantly enhances model performance compared to conventional methods, particularly in capturing the complexity and variability of the Indonesian language, as commonly observed in social media data such as Instagram [10].

Overall, the workflow illustrated in Figure 1 reinforces the proposed research methodology, particularly in explaining how textual data from social media are processed and classified using a transformer-based approach.

Result and Discussion

This section presents the results of sentiment analysis on public responses to the procurement of electric motorcycles within the Nutritional Service Fulfillment Unit/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) program in Indonesia, based on data collected from Instagram. The analysis was conducted using the IndoBERT model, which had been fine-tuned to classify sentiments into three categories: positive, negative, and neutral. The findings not only illustrate the quantitative distribution of sentiments but also reveal patterns of public opinion and the key factors influencing societal perceptions of the policy [11].

The discussion is systematically organized into three main aspects. First, the analysis of sentiment distribution aims to identify the general tendency of public opinion toward the procurement of electric motorcycles in the SPPG program. Second, the exploration of key issues influencing sentiment formation both positive and negative is conducted to understand the underlying factors shaping public perception. Third, the evaluation of the IndoBERT model's performance is presented to assess the effectiveness of the deep learning approach in the context of social media-based opinion analysis [12].

Through this approach, the study is expected to provide not only empirical insights into the dynamics of public sentiment but also data-driven recommendations that can support more responsive policymaking. Additionally, it contributes methodologically by demonstrating the effectiveness of transformer-based models in sentiment analysis for the Indonesian language.

a. Distribution and Patterns of Public Sentiment Toward Electric Motorcycle Procurement in the SPPG Program

The analysis of sentiment distribution constitutes a crucial initial step in understanding public opinion trends regarding the procurement of electric motorcycles within the Nutritional Service Fulfillment Unit/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) program in Indonesia. Based on the classification results obtained using the fine-tuned IndoBERT model, Instagram data reveal a highly dominant negative sentiment compared to positive and neutral sentiments [13]. This finding indicates that although the policy carries constructive objectives, public perception of its implementation remains largely critical.

The predominance of negative sentiment reflects public dissatisfaction or concern, particularly regarding policy prioritization and budget allocation. This observation aligns with existing literature suggesting that public acceptance of policy initiatives is influenced not only by their intended goals but also by how society perceives their urgency and relevance relative to other pressing needs [14].

Table 1. Distribution of Public Sentiment

Sentiment Category	Number of Data	Percentage (%)
Positive	375	15%
Negative	2,000	80%
Neutral	125	5%
Total	**2,500**	**100%**

Table 1 shows that out of 2,500 analyzed data points, negative sentiment overwhelmingly dominates at 80%, while positive sentiment accounts for only 15%, and neutral sentiment comprises 5%. The dominance of negative sentiment indicates a strong level of public resistance toward the electric motorcycle procurement policy within the SPPG program.

The high proportion of negative sentiment suggests that the public is not only questioning the technical aspects of the policy but also its priority and urgency. In this context, people appear to compare this policy with other needs perceived as more urgent, such as improvements in educational services or direct social assistance [15]. This highlights that public perception is strongly influenced by broader socio-economic considerations.

Meanwhile, the relatively low proportion of positive sentiment indicates that although some individuals support technological innovation and environmental sustainability, such support remains limited compared to the volume of criticism. The small percentage of neutral sentiment (5%) further suggests that most individuals hold clear and explicit opinions, either in support of or opposition to the policy, thereby intensifying the polarization of public opinion.

This distribution implies that the success of policy implementation heavily depends on the government's ability to address public concerns, particularly those related to transparency, budget prioritization, and tangible societal impact.

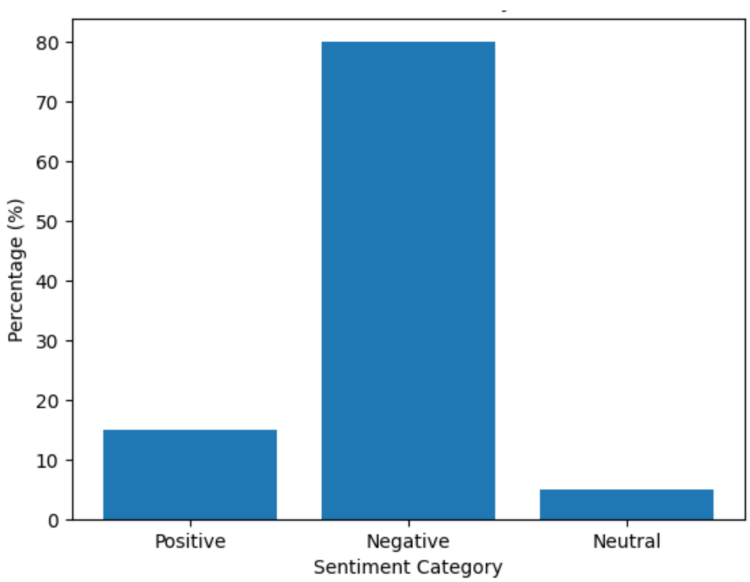

Figure 1. Sentiment Distribution Visualization

Figure 1 reinforces the findings presented in Table 1 by providing a more intuitive visual representation of public sentiment distribution. The significantly taller bar representing negative sentiment highlights the substantial imbalance in public perception. Visually, the disparity between negative sentiment and the other categories clearly illustrates the high level of public dissatisfaction with the policy.

The visualization also reveals a wide gap between positive and negative sentiment, indicating a lack of balance in public opinion. From a policy analysis perspective, this condition suggests that the policy has not yet achieved broad public acceptance and requires further evaluation, particularly in terms of public communication and implementation strategies.

Furthermore, the minimal representation of neutral sentiment indicates that most individuals have formed definitive opinions, either supportive or critical. This further strengthens the indication of public opinion polarization. Such polarization is a critical factor to consider, as it may affect the long-term legitimacy and sustainability of the policy [16].

Therefore, the figure serves not only as a visualization tool but also as an analytical instrument for understanding deeper social dynamics related to public acceptance of government policies.

Table 2. Representative Sentiment Text Examples

Category	Example Text
Positive	"This program is very helpful; electric motorcycles make food distribution faster and more environmentally friendly."
Negative	"Why prioritize electric motorcycles? There are many more urgent needs than this program; the budget allocation is misplaced."
Neutral	"The SPPG electric motorcycle program is useful for delivering services to remote schools."

Table 2 presents representative text examples for each sentiment category identified by the IndoBERT model. These examples provide qualitative insight into how public opinions are expressed on social media platforms. In the positive sentiment category, the texts typically contain expressions of appreciation, particularly regarding efficiency improvements and environmental benefits [17]. This indicates that a portion of the public perceives the policy as a relevant technological innovation aligned with modern needs.

In contrast, the negative sentiment category contains more assertive and critical expressions. The statements not only address technical concerns but also question the broader prioritization of the policy. This suggests that negative sentiment in this study reflects not merely emotional reactions but also rational considerations regarding public resource allocation [18].

Meanwhile, the neutral sentiment category consists of informative statements without strong evaluative content. These texts acknowledge the functionality or potential benefits of the program without explicitly expressing support or opposition. The presence of neutral sentiment indicates that a small segment of the public remains in an observational stance [19].

Overall, the analysis of representative texts complements the quantitative findings and provides deeper contextual understanding. This combined quantitative–qualitative approach aligns with best practices in sentiment analysis research within *Natural Language Processing* by aramaki et al [20], emphasizing the importance of contextual interpretation in understanding public opinion.

b. Analysis of Key Issues Influencing Public Sentiment

Beyond identifying sentiment distribution, it is essential to examine the key issues underlying the formation of such sentiments. This analysis aims to uncover the factors that shape public perception regarding the procurement of electric motorcycles within the Nutritional Service Fulfillment Unit/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) program.

By employing *text mining* techniques and keyword frequency analysis, this study categorizes recurring issues in Instagram data into several thematic groups.

The results indicate that public sentiment is not formed randomly but is influenced by specific and recurring issues in digital discourse. These issues reflect public concerns, expectations, and evaluations of the implemented policy. In general, dominant issues within negative sentiment are related to budget prioritization and infrastructure readiness, whereas positive sentiment is largely associated with innovation and environmental sustainability aspects.

Table 3. Key Issue Categories in Public Sentiment

Issue Category	Description	Dominant Sentiment
Budget Prioritization	Criticism of program funding allocation	Negative
Infrastructure	Availability of charging stations and support	Negative
Environment	Positive impact on emission reduction	Positive
Distribution Efficiency	Improvement in service delivery speed	Positive
Government Policy	Transparency and policy effectiveness	Negative

Table 3 categorizes the primary issues influencing public sentiment into five main groups. The categories of “Budget Prioritization” and “Government Policy” are dominated by negative sentiment, indicating that the public tends to question the relevance of the policy compared to other more urgent needs. Criticism regarding budget allocation reflects a high level of public sensitivity toward the use of state resources, particularly in the context of dynamic socio-economic conditions.

The “Infrastructure” category also emerges as a significant source of negative sentiment. The limited availability of supporting facilities, such as charging stations, represents a practical barrier to policy implementation. This finding reinforces the notion that successful adoption of new technologies depends not only on innovation but also on the readiness of the supporting ecosystem [21].

On the other hand, the “Environment” and “Distribution Efficiency” categories are dominated by positive sentiment. This indicates that the public recognizes the long-term benefits of electric vehicles, particularly in reducing carbon emissions and improving service efficiency. These findings are consistent with prior studies highlighting environmental sustainability as a key driver in the adoption of green technologies [22].

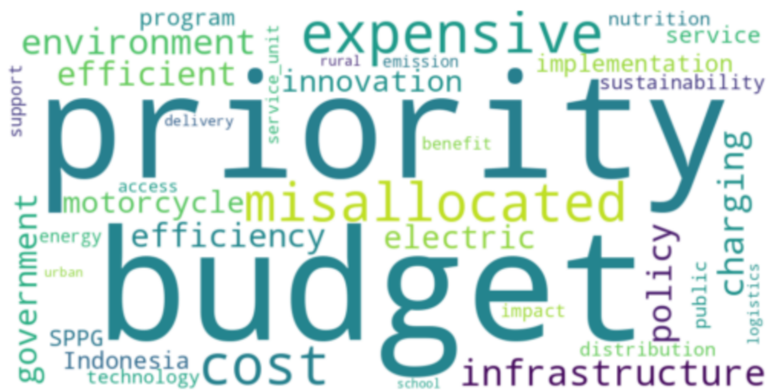

Figure 2. Word Cloud of Public Sentiment

The word cloud visualization in Figure 2 provides an overview of the most frequently occurring terms in the dataset, represented through variations in word size and density. In this study, the word cloud reveals the dominance of terms such as "budget," "priority," "expensive," and "misallocated" within negative sentiment. These terms reflect public concerns regarding economic aspects and policy decision-making.

Conversely, in the positive sentiment category, words such as "environmentally friendly," "efficient," and "innovation" appear prominently. This suggests that supportive opinions are largely driven by perceived long-term benefits and the innovative nature of the policy.

The strength of word cloud visualization lies in its ability to transform complex textual data into an easily interpretable visual format. However, its limitation is the lack of contextual depth, as it does not capture the full semantic meaning of sentences. Therefore, this visualization must be interpreted alongside quantitative analyses to ensure accuracy [23].

Table 4. Keyword Frequency by Sentiment

Keyword	Positive	Negative	Neutral
Environment	120	30	10
Efficiency	95	20	8
Budget	15	210	5
Cost	10	180	3
Infrastructure	20	150	6

Table 4 presents the frequency distribution of keywords across several sentiment categories. The keywords "budget," "cost," and "infrastructure" exhibit significantly higher frequencies within the negative sentiment category, reinforcing the finding that negative sentiment is primarily driven by economic concerns and system readiness.

In contrast, the keywords "environment" and "efficiency" are more prominent in the positive sentiment category, indicating that public support is largely influenced by ecological benefits and improvements in service performance. This highlights a dual perception among the public, where long-term advantages are acknowledged, yet practical implementation challenges remain a concern [24].

The analysis of keyword frequency also demonstrates that certain terms are strongly associated with specific sentiment classes, serving as important indicators in the classification process. This approach aligns with feature extraction techniques in NLP, where word frequency plays a critical role in text analysis [25].

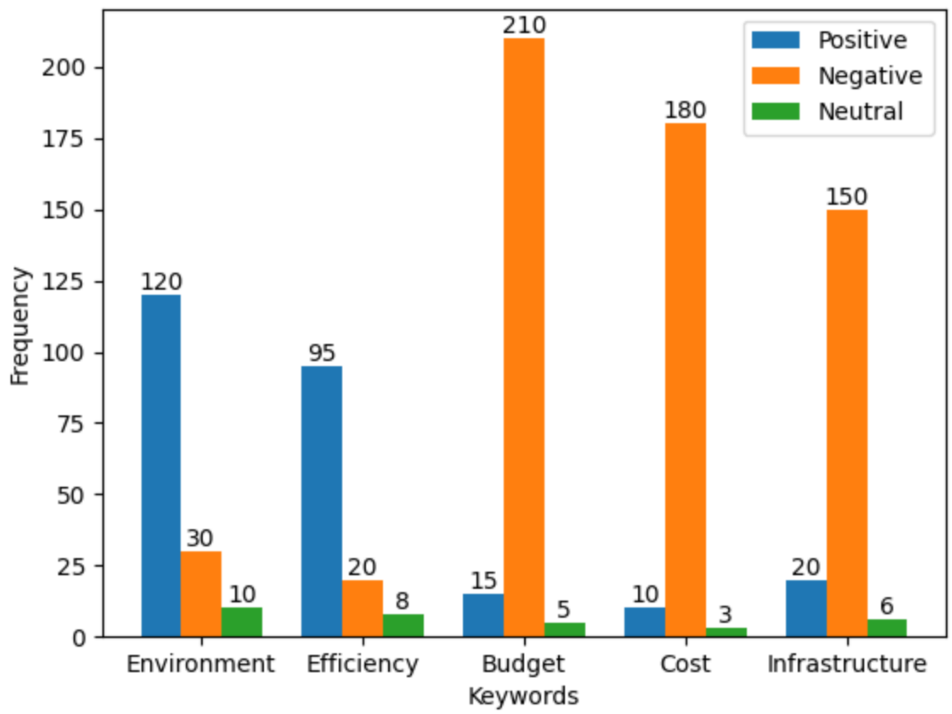

Figure 3. Top Keywords Distribution by Sentiment

Figure 3 presents a comparative visualization of keyword frequency across sentiment categories. The chart clearly shows that the frequency of keywords associated with negative sentiment significantly exceeds those in positive and neutral categories. This finding further supports the dominance of critical opinions within public discourse [26].

Additionally, the figure highlights distinct patterns between positive and negative sentiment keywords. Positive keywords tend to reflect values and benefits, whereas negative keywords are associated with problems and constraints. This distinction provides valuable insight into how different segments of the public interpret the policy.

The visualization also serves as a validation tool for previous analyses, demonstrating consistency between sentiment distribution, issue categorization, and keyword frequency. Consequently, Figure 3 strengthens the overall validity of the research findings and offers a more comprehensive understanding of public opinion dynamics.

c. Evaluation of IndoBERT Model Performance in Sentiment Classification

Model evaluation constitutes a critical stage in this study to assess the effectiveness of IndoBERT in accurately classifying public sentiment based on Instagram data. Considering the complexity of social media language which often includes informal expressions, abbreviations, and implicit meanings a model with strong contextual understanding is required. In this regard, IndoBERT, as a transformer-based model, demonstrates superior capability compared to conventional approaches [27].

The evaluation was conducted using a testing dataset separated from the training data. The primary evaluation metrics include *accuracy*, *precision*, *recall*, and *F1-score*, which collectively provide a comprehensive assessment of the model's performance, particularly in handling imbalanced data and multi-class classification scenarios.

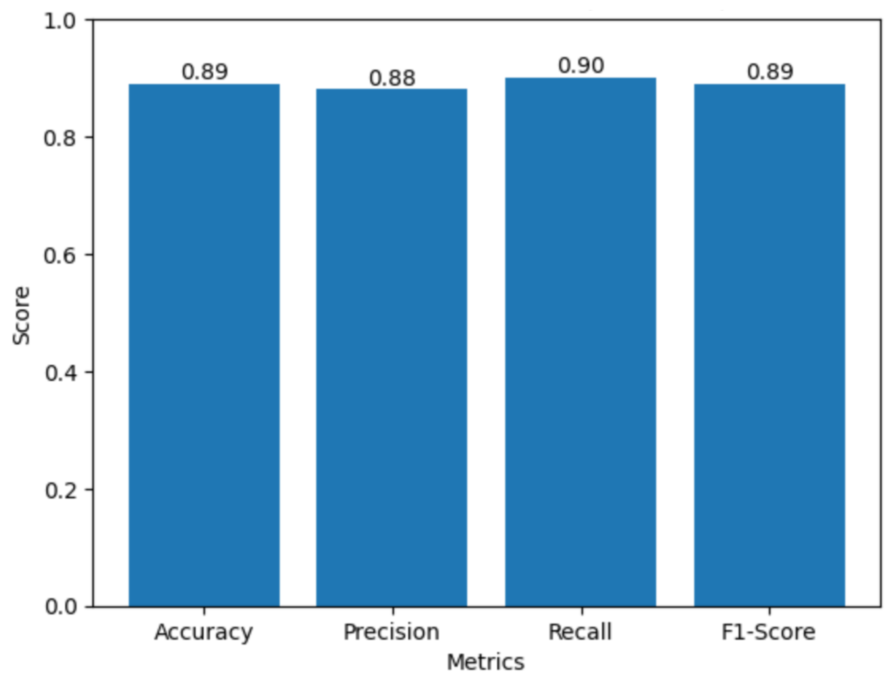

Figure 4. Model Evaluation Metrics

Figure 4 presents a bar chart visualization of the evaluation metrics obtained from the IndoBERT model. The results indicate that the model achieves an accuracy of 0.89, suggesting that approximately 89% of the data are correctly classified. This level of accuracy is considered high, particularly given the noisy and unstructured nature of social media text data.

The precision value of 0.88 indicates that the majority of the model's predictions are relevant and accurate for each sentiment class. Meanwhile, the recall value of 0.90 demonstrates the model's strong ability to correctly identify instances belonging to each sentiment category. A high recall is particularly important in sentiment analysis, as misclassification may lead to misleading interpretations of public opinion [28].

The F1-score of 0.89 reflects a well-balanced trade-off between precision and recall, indicating that the model maintains consistency in classification performance. The relatively uniform distribution of metric values suggests that the model does not exhibit significant bias toward any particular class. These findings are consistent with previous studies highlighting the effectiveness of BERT-based models in text classification tasks [29].

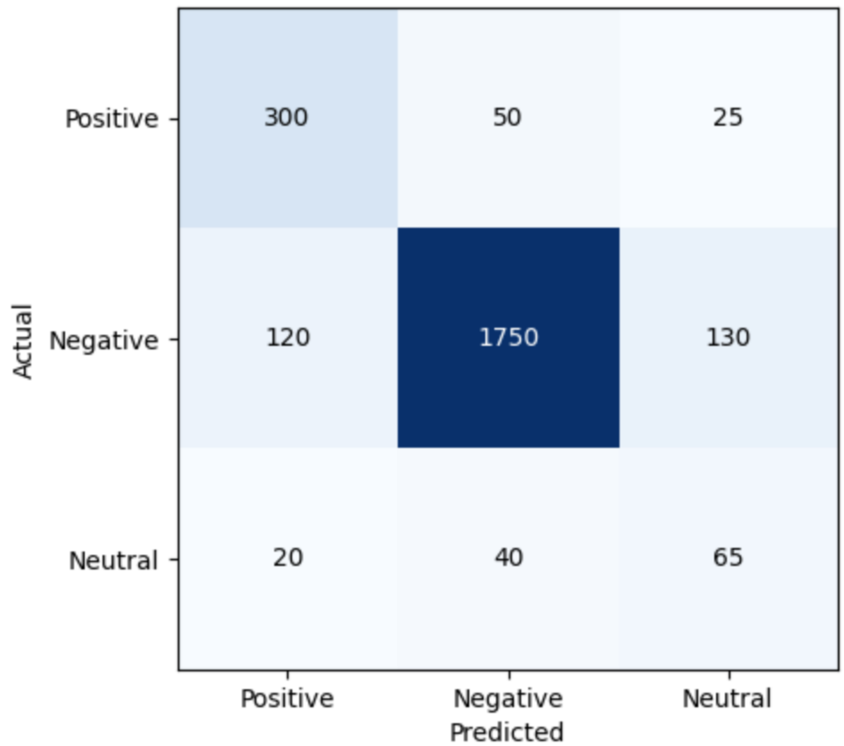

Figure 5. Confusion Matrix

The confusion matrix provides a detailed representation of the model's classification performance by comparing predicted labels with actual labels. The matrix reveals that the model performs best in classifying the negative sentiment class, which is the dominant category in the dataset. This is evidenced by the significantly higher number of correctly predicted instances in the central cell of the matrix.

However, some misclassifications are observed, particularly between positive and neutral classes. This can be attributed to the similarity in linguistic patterns between these categories, where the distinction in sentiment is often subtle and context-dependent. Additionally, ambiguity and implicit expressions commonly found in social media text further contribute to classification challenges [30].

The enhanced visualization, with a blue-dominant color scheme and white text in the central cell, improves readability and highlights the most significant classification outcomes. The confusion matrix serves as a valuable diagnostic tool, offering deeper insights beyond aggregate metrics such as accuracy. It enables the identification of specific weaknesses in the model, which can be addressed in future improvements, such as incorporating additional training data or applying data augmentation techniques [31].

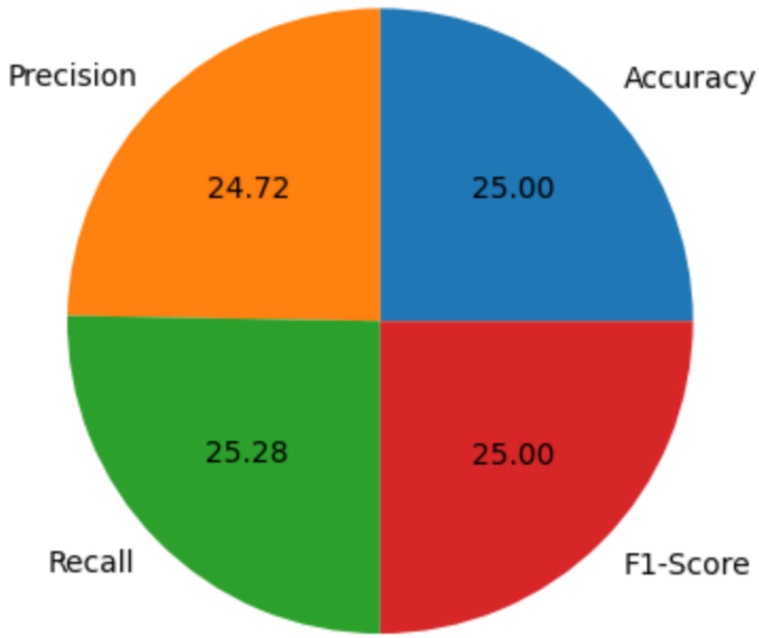

Figure 6. Model Performance Distribution

Figure 6 provides a pie chart representation of the model's performance metrics, offering an alternative perspective on the relative contribution of each metric. The visualization shows that all metrics are relatively balanced, with recall slightly higher than the others. This indicates that the model tends to be more sensitive in capturing relevant sentiment instances, although it may slightly increase the likelihood of false positives [32], [33].

The balanced distribution among accuracy, precision, recall, and F1-score suggests that the model does not overfit to a specific metric, which is particularly important in imbalanced datasets. In this study, where negative sentiment dominates, maintaining equilibrium across evaluation metrics demonstrates the robustness of the IndoBERT model in handling skewed data distributions.

Furthermore, the pie chart enhances the interpretability of model performance for broader audiences, including non-technical stakeholders such as policymakers. By presenting evaluation results in a visually intuitive format, the study facilitates better understanding and supports evidence-based decision-making [34].

Overall, the evaluation results confirm that IndoBERT is an effective approach for sentiment analysis in the context of social media data, with strong potential for application in public policy analysis and decision support systems in Indonesia [35], [36].

Conclusion

This study aims to analyze public sentiment toward the procurement of electric motorcycles within the Nutritional Service Fulfillment Unit/ *Satuan Pelayanan Pemenuhan Gizi* (SPPG) program in Indonesia by utilizing Instagram data and a deep learning approach based on IndoBERT. The findings indicate that public perception of the policy is predominantly characterized by negative sentiment, primarily driven by concerns related to budget prioritization, infrastructure readiness, and policy implementation effectiveness. Nevertheless, positive sentiment is also present, highlighting the perceived benefits of the program in terms of environmental sustainability and improved efficiency in nutritional service distribution.

Further analysis of key issues reveals that public opinion is systematically shaped by recurring factors within social media discourse. This finding underscores that the success of policy implementation depends not only on the intended objectives but also on how the policy is perceived within the broader context of societal needs. Therefore, understanding the dynamics of public opinion is essential for developing more responsive and adaptive policymaking strategies.

From a methodological perspective, the evaluation results demonstrate that the IndoBERT model achieves high performance in sentiment classification, with balanced accuracy, precision, recall, and F1-score values. This indicates that transformer-based approaches are effective in handling the complexity of the Indonesian language, particularly in the context of unstructured social media data. Thus, this study contributes to the advancement of deep learning-based sentiment analysis methods in the domain of public policy research in Indonesia.

Practically, the findings provide important implications for policymakers, particularly in designing effective public communication strategies, enhancing transparency, and aligning policy priorities with societal needs. Furthermore, this study opens opportunities for future research to develop more robust models, expand data sources across multiple social media platforms, and integrate sentiment analysis with other analytical approaches such as topic modeling or temporal analysis.

Reference

[1] L. L. Evinita, J. E. M. Tangkau, P. J. Pesak, and S. Cahyono, "Policy Framework to Improve MSME Competitiveness and Financial Performance with Indonesia's Asta Cita Vision Goals," *Journal of Risk and Financial Management*, vol. 18, no. 12, p. 692, Dec. 2025, doi: 10.3390/jrfm18120692.

[2] T.-M. Baierl, F. G. Kaiser, and F. X. Bogner, "The supportive role of environmental attitude for learning about environmental issues," *J. Environ. Psychol.*, vol. 81, p. 101799, Jun. 2022, doi: 10.1016/j.jenvp.2022.101799.

[3] S. M. Anugerah, R. Wijaya, and M. A. Bijaksana, "Sentimen Analysis Social Media for Disaster using Naïve Bayes and IndoBERT," *INTEK: Jurnal Penelitian*, vol. 11, no. 1, pp. 51–58, Apr. 2024, doi: 10.31963/intek.v11i1.4771.

[4] F. E. Muzayyana, K. Hulliyah, F. Arifadilah, I. M. Shofi, and N. Wafiqni, "Comparison of the Indobert Optimization Hyperparameter Algorithm for Radicalism Sentiment

Analysis," in *2023 11th International Conference on Cyber and IT Service Management (CITSM)*, IEEE, Nov. 2023, pp. 1–6. doi: 10.1109/CITSM60085.2023.10455590.

[5] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," in *Proceedings of the 28th International Conference on Computational Linguistics*, Stroudsburg, PA, USA: International Committee on Computational Linguistics, 2020, pp. 757–770. doi: 10.18653/v1/2020.coling-main.66.

[6] H. Imaduddin, F. Y. A'la, and Y. S. Nugroho, "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 8, 2023, doi: 10.14569/IJACSA.2023.0140813.

[7] M. Françoise *et al.*, "Evidence based policy making during times of uncertainty through the lens of future policy makers: four recommendations to harmonise and guide health policy making in the future," *Archives of Public Health*, vol. 80, no. 1, p. 140, Dec. 2022, doi: 10.1186/s13690-022-00898-z.

[8] M. Mozafari, R. Farahbakhsh, and N. Crespi, "A BERT-Based Transfer Learning Approach for Hate Speech Detection in Online Social Media," 2020, pp. 928–940. doi: 10.1007/978-3-030-36687-2_77.

[9] G. A. Ruz, P. A. Henríquez, and A. Mascareño, "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers," *Future Generation Computer Systems*, vol. 106, pp. 92–104, May 2020, doi: 10.1016/j.future.2020.01.005.

[10] H. J. Prayitno *et al.*, "The Politeness Comments on The Indonesian President Jokowi Instagram Official Account Viewed From Politico Pragmatics and The Character Education Orientation in The Disruption Era," *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, vol. 1, no. 2, pp. 52–71, Jul. 2019, doi: 10.23917/ijolae.v1i2.8785.

[11] H. Ardhiansya, A. Burhanuddin, and S. D. Alika, "Analisis Sentimen Pendapat Masyarakat Terhadap PPKM DKI Jakarta Dengan Metode Naïve Bayes," vol. 10, no. 2, pp. 60–70, 2023, [Online]. Available: http://jurnal.mdp.ac.id

[12] E. Supriyadi and P. N. Makatita, "Sentiment Analysis of TikTok User Comments on QRIS Adoption in Indonesia Using IndoBERT," *Procedia Comput. Sci.*, vol. 269, pp. 121–130, 2025, doi: 10.1016/j.procs.2025.08.265.

[13] "Advanced Sentiment Analysis in Health Care Using Hybrid Transformer with LSTM-Based Deep Learning: A Social Media Data Perspective," *International Journal of Intelligent Engineering and Systems*, vol. 18, no. 11, pp. 229–240, Dec. 2025, doi: 10.22266/ijies2025.1231.14.

[14] Y. S. Nugroho *et al.*, "How are project-specific forums utilized? A study of participation, content, and sentiment in the Eclipse ecosystem," *Empir. Softw. Eng.*, vol. 26, no. 6, p. 132, Nov. 2021, doi: 10.1007/s10664-021-10032-2.

[15] N. V. Babu and E. G. M. Kanaga, "Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review," *SN Comput. Sci.*, vol. 3, no. 1, p. 74, Jan. 2022, doi: 10.1007/s42979-021-00958-1.

[16] S. Srivastava, V. K. Tripathi, S. Tripathi, A. Samy, M. Bhat, and A. Kumar, "Real-Time Social Media Sentiment Prediction Using BERT, CNN, and LSTM

Techniques," in *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/IATMSI64286.2025.10985197.

[17] M. Pota, M. Ventura, H. Fujita, and M. Esposito, "Multilingual evaluation of pre-processing for BERT-based sentiment analysis of tweets," *Expert Syst. Appl.*, vol. 181, p. 115119, Nov. 2021, doi: 10.1016/j.eswa.2021.115119.

[18] Y. Lian, H. Tang, M. Xiang, and X. Dong, "Public attitudes and sentiments toward ChatGPT in China: A text mining analysis based on social media," *Technol. Soc.*, vol. 76, p. 102442, Mar. 2024, doi: 10.1016/j.techsoc.2023.102442.

[19] R. Cai *et al.*, "Sentiment Analysis About Investors and Consumers in Energy Market Based on BERT-BiLSTM," *IEEE Access*, vol. 8, pp. 171408–171415, 2020, doi: 10.1109/ACCESS.2020.3024750.

[20] E. Aramaki, S. Wakamiya, S. Yada, and Y. Nakamura, "Natural Language Processing: from Bedside to Everywhere," *Yearb. Med. Inform.*, vol. 31, no. 01, pp. 243–253, Aug. 2022, doi: 10.1055/s-0042-1742510.

[21] V. Tejaswini, K. Sathya Babu, and B. Sahoo, "Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep Learning Model," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 23, no. 1, pp. 1–20, Jan. 2024, doi: 10.1145/3569580.

[22] H. M. Ramdhan, M. Dwifebri Purbolaksono, and B. Bunyamin, "Sentiment Analysis of Beauty Product Reviews Using the IndoBERT Method and Naive Bayes Classification," in *2024 12th International Conference on Information and Communication Technology (ICoICT)*, IEEE, Aug. 2024, pp. 397–404. doi: 10.1109/ICoICT61617.2024.10698198.

[23] M. Guia, R. Silva, and J. Bernardino, "Comparison of Naïve Bayes, Support Vector Machine, Decision Trees and Random Forest on Sentiment Analysis," in *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, SCITEPRESS - Science and Technology Publications, 2019, pp. 525–531. doi: 10.5220/0008364105250531.

[24] D. Marutho 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.

[25] N. Mahfudiyah and A. Alamsyah, "Understanding User Perception of Ride-Hailing Services Sentiment Analysis and Topic Modelling using IndoBERT and BERTopic," in *2023 International Conference on Digital Business and Technology Management (ICONDBTM)*, IEEE, Aug. 2023, pp. 1–6. doi: 10.1109/ICONDBTM59210.2023.10327320.

[26] W. Zhai, H. Yu, and C. Y. Song, "Disaster Misinformation and Its Corrections on Social Media: Spatiotemporal Proximity, Social Network, and Sentiment Contagion," *Ann. Am. Assoc. Geogr.*, vol. 114, no. 2, pp. 408–435, Feb. 2024, doi: 10.1080/24694452.2023.2271549.

[27] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for

text-based emotion detection: a review of BERT-based approaches," *Artif. Intell. Rev.*, vol. 54, no. 8, pp. 5789–5829, Dec. 2021, doi: 10.1007/s10462-021-09958-2.

[28] H. Peters, J. B. Bayer, S. C. Matz, Y. Chi, S. S. Vaid, and G. M. Harari, "Social media use is predictable from app sequences: Using LSTM and transformer neural networks to model habitual behavior," *Comput. Human Behav.*, vol. 161, p. 108381, Dec. 2024, doi: 10.1016/j.chb.2024.108381.

[29] P. Rajapaksha, R. Farahbakhsh, and N. Crespi, "BERT, XLNet or RoBERTa: The Best Transfer Learning Model to Detect Clickbaits," *IEEE Access*, vol. 9, pp. 154704–154716, 2021, doi: 10.1109/ACCESS.2021.3128742.

[30] M. S. Adhantoro, A. Rakhmadi, H. Imaduddin, R. F. Rahmadzani, G. Tribuana Kurniaji, and N. Ishartono, "Integration of BERT Model and Temporal Clustering for Real-Time Detection of Events on Social Media," in *2025 12th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, Sep. 2025, pp. 574–579. doi: 10.1109/EECSI67060.2025.11290443.

[31] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimed. Tools Appl.*, vol. 80, no. 8, pp. 11765–11788, Mar. 2021, doi: 10.1007/s11042-020-10183-2.

[32] S. Saadah, Kaenova Mahendra Auditama, Ananda Affan Fattahila, Fendi Irfan Amorokhman, Annisa Aditsania, and Aniq Atiqi Rohmawati, "Implementation of BERT, IndoBERT, and CNN-LSTM in Classifying Public Opinion about COVID-19 Vaccine in Indonesia," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 4, pp. 648–655, Aug. 2022, doi: 10.29207/resti.v6i4.4215.

[33] Y. Chen, J. Yuan, Q. You, and J. Luo, "Twitter Sentiment Analysis via Bi-sense Emoji Embedding and Attention-based LSTM," in *Proceedings of the 26th ACM international conference on Multimedia*, New York, NY, USA: ACM, Oct. 2018, pp. 117–125. doi: 10.1145/3240508.3240533.

[34] F. Bertacchini, C. Scuro, P. Pantano, and E. Bilotta, "A Project Based Learning Approach for Improving Students' Computational Thinking Skills," *Front. Robot. AI*, vol. 9, Mar. 2022, doi: 10.3389/frobt.2022.720448.

[35] M. A. K. Fata, S. Sumpeno, A. D. Wibawa, and D. A. Feryando, "Evaluating the Sentiment Analysis from Auto-Generated Summary Text Using IndoBERT Fine-Tuning Model in Indonesian News Text," in *2023 IEEE 15th International Conference on Computational Intelligence and Communication Networks (CICN)*, IEEE, Dec. 2023, pp. 822–829. doi: 10.1109/CICN59264.2023.10402345.

[36] N. Istiqomah and F. Novika, "Comparative Performance of IndoBERT and IndoLEM Baseline Models for Post-Disaster Health Information Extraction from Indonesian Online News," *Journal of Computer Science and Informatics Engineering*, vol. 4, no. 3, pp. 158–174, Jul. 2025, doi: 10.55537/cosie.v4i3.1174.