Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.3, June 2026, pp. 2613~2619
e-ISSN: 2548-6861

Comparative Sentiment Analysis of Indonesian Social Media Opinions on Fuel Subsidy Policy Using IndoBERT and NusaBERT

Dini Ambarwati [1], Pungkas Subarkah [2]*, Septi Nurhayati [3]
* Informatika, Universitas Amikom Purwokerto
diniambarwati70@gmail.com [1], subarkah@amikompurwokerto.ac.id [2],septi.nh@amikompurwokerto.ac.id [3]

Article Info

Article history:

Received 2026-05-09
Revised 2026-05-23
Accepted 2026-06-11

Keyword:

BBM Policy,
IndoBERT,
NusaBERT,
Public Sentiment,
Transformer.

ABSTRACT

Rising geopolitical tensions in the Middle East have triggered public concerns in Indonesia regarding fuel subsidy policies and fuel availability. This study aims to compare the performance of IndoBERT and NusaBERT in classifying Indonesian public sentiment on social media related to fuel subsidy policies. Data were collected from X (Twitter) and Instagram comments between October and November 2025 using keywords such as "BBM", "Pertalite", "fuel subsidy", and "Middle East conflict". After filtering duplicate, spam, and irrelevant content, a total of 1,500 opinion texts were manually annotated into positive, neutral, and negative sentiment classes and divided using an 80:20 train-test split configuration. The preprocessing stage included case folding, text cleansing, slang word normalization, emoji removal, duplicate filtering, and tokenization. Both Transformer models were fine-tuned using the AdamW optimizer with a learning rate of 2e-5, batch size of 16, and 3 training epochs. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results show that NusaBERT achieved better performance than IndoBERT, obtaining an accuracy of 96.3% and weighted F1-score of 96.3%, while IndoBERT achieved an accuracy of 92.5%. Additional evaluation through confusion matrix and error analysis indicates that both models still face challenges in handling sarcasm, ambiguous expressions, and mixed-context sentences commonly found in informal Indonesian social media text. The findings suggest that NusaBERT is more effective for Indonesian social media sentiment classification due to its stronger adaptation to informal language patterns.

I. INTRODUCTION

National energy security is one of the fundamental pillars in maintaining macroeconomic stability and the smooth running of community activities in Indonesia. Fuel Oil (BBM) plays a very crucial role, not only as the main driver of the transportation and logistics sectors but also has a direct multiplier effect on inflation rates and people's purchasing power. Considering that the fulfillment of Indonesia's domestic oil needs is still heavily tied to import schemes and the international market, the stability of domestic energy security becomes highly vulnerable to sudden external shocks. The escalation of geopolitical conflicts in the Middle East directly impacts the instability of global crude oil prices and the world's energy supply chain [1]. This can be illustrated by the spike in global benchmark oil price fluctuations during the crisis period, as visually represented in Figure 1.

For Indonesia, the fluctuation in world oil prices puts significant pressure on the State Revenue and Expenditure Budget (APBN), particularly on the allocation for fuel oil (BBM) energy subsidies [2]. This situation triggers various speculations among the public regarding potential price increases, quota restrictions, and even fuel scarcity at gas stations. Information uncertainty often drives *panic buying* and social unrest expressed by the public through various social media platforms [3]. Therefore, the main problem identification in this study is the need for a mechanism to automatically map and analyze public sentiment and opinions regarding fuel availability policies

during this crisis, to prevent disinformation and the escalation of panic.

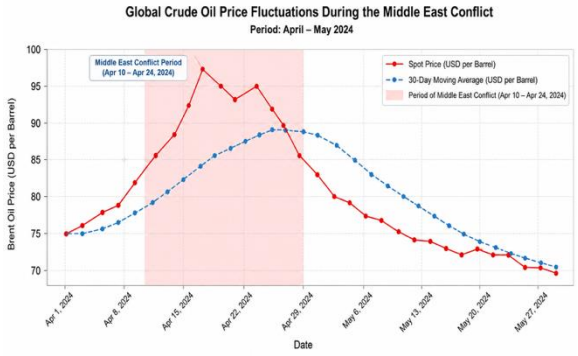

Figure 1. Global crude oil price fluctuations during the Middle East conflict escalation.

A literature review shows that sentiment analysis has been widely used to extract public opinion on government policies and social issues across various domains . A summary of the comparison of several relevant previous studies to identify the research gap is presented in Table I.

TABLE I
COMPARISON OF PREVIOUS STUDIES (STATE-OF-THE-ART)

Study	Method	Limitations / Gap
Samantri et al. [4]	SVM	Only uses traditional NLP; limited accuracy on slang text.
Wulandari et al. [5]	Naïve Bayes	Unable to capture bidirectional sentence context.
Wilie et al. [6]	IndoBERT	Model trained on formal/standard text corpus.
Cahyawijaya et al. [7]	NusaBERT	Not specifically tested on energy panic (fuel/BBM) cases.

Traditional methods such as Support Vector Machine (SVM) and Naive Bayes are often used [4], [5], but they have limitations in understanding complex sentence semantic contexts and slang. Previous studies also reported that conventional machine learning methods such as SVM, Logistic Regression, and Naïve Bayes generally achieved lower classification performance compared to Transformer-based architectures, particularly when handling contextual meaning, slang expressions, and informal sentence structures in Indonesian social media text. Along with the development of Natural Language Processing (NLP), the Transformer architecture has proven to provide better classification performance [8]. This architecture forms the basis for the development of the BERT (*Bidirectional Encoder Representations from Transformers*) model, which can read text sequences simultaneously from both directions to deeply understand linguistic context [9]. The results of this BERT-based model has been tested in Indonesia; for instance, Previous studies indicate that IndoBERT outperforms conventional machine learning approaches, including SVM and Random Forest, in sentiment classification tasks. Its ability to understand bidirectional contextual information enables it to achieve relatively high accuracy, reaching approximately 80% under appropriate pre-processing conditions [10], [11]. However, research by Wilie et al. indicates that IndoBERT was trained on a formal text corpus [6]. On the other hand, Cahyawijaya et al. developed a model that is more adaptive to everyday language and local dialects through initiatives like the one underlying NusaBERT [7]. Although both models are promising, there is no comprehensive study specifically comparing them in the case of energy policy sentiment, which is heavily laden with informal language [12], even though an accurate understanding of public opinion is crucial given that oil price fluctuations have a long history of disrupting national economic stability and demand rapid public communication responses from energy companies [13]. The novelty of this study lies in the comparative evaluation of IndoBERT and NusaBERT specifically within the context of Indonesian energy policy discourse during geopolitical crises, which remains underexplored in previous sentiment analysis research.

Based on this research gap, the objective of this study is to implement and compare the performance of the IndoBERT and NusaBERT Transformer algorithms in classifying public sentiment (positive, negative, and neutral) regarding fuel availability policies following the Middle East conflict. Performance evaluation is measured using *Accuracy*, *Precision*, *Recall*, and *F1-Score* metrics. The theoretical benefit derived from the results of this study is providing empirical evidence regarding the adaptability of the NusaBERT model to informal Indonesian texts compared to IndoBERT. Practically, the results of this sentiment visualization and classification can be used as an *early warning system* and a foundation for the government and relevant stakeholders (such as Pertamina and the Ministry of Energy and Mineral Resources) in formulating more precise and responsive public communication strategies to maintain national energy security [14]. In addition, the study may still be influenced by potential social media biases such as coordinated buzzer activity, echo chambers, and undetected automated accounts.

II. METHOD

This study was conducted through several systematic stages to ensure valid and comprehensive sentiment analysis results. In general, the research framework or workflow is illustrated in Figure 2.

Figure 2. Workflow Diagram.

Based on Figure 2, the operational stages of the study include the process of collecting raw data containing public opinions, followed by a text preprocessing stage to remove noise and structure the data. The cleaned data are then used in the implementation and fine-tuning stage of Transformer-based language models. This sequence is concluded with a performance evaluation stage to measure and compare the classification accuracy of each model.

A. Data Collection

The data used in this study consist of primary public opinion texts extracted specifically from social media platforms X (formerly Twitter) and Instagram comment sections. The data collection (crawling) process was focused on the period following the escalation of geopolitical conflicts, which coincided with issues related to adjustments in national energy policies. Relevant keywords such as "BBM," "subsidy," "Middle East conflict," and "Pertalite" were used during data retrieval. The data crawling process was conducted between October and November 2025 using keyword-based searches on X (Twitter) and Instagram public comment sections. Initial data collection produced approximately 2,300 raw opinion texts. Several filtering stages were then applied, including duplicate removal, spam filtering, advertisement removal, and elimination of irrelevant content. Accounts suspected to be automated bots were excluded through manual inspection based on repetitive posting patterns and abnormal interaction frequency. Through the extraction and selection process, this study successfully compiled a dataset of 1,500 text samples. All data samples were manually annotated by two independent annotators with backgrounds in linguistics and social media discourse analysis. The annotation process categorized each opinion into positive, neutral, or negative sentiment classes. In cases of disagreement, a discussion-based reconciliation process was conducted to determine the final label. The final dataset distribution consisted of 502 positive, 498 neutral, and 500 negative opinions. The relatively balanced class distribution was obtained after the filtering and selection process rather than through synthetic oversampling or undersampling techniques. This approach was intended to minimize potential evaluation bias during model training and testing.

B. Data Preprocessing

Text data from social media are generally unstructured and contain a significant amount of noise. Therefore, the preprocessing stage is crucial before feeding the data into Transformer-based models. The preprocessing steps applied in this study include:

1) Case Folding, which converts all characters into lowercase;
2) Cleansing, which removes elements with no semantic value such as URLs, user mentions (@username), hashtags, special characters, and punctuation;
3) Normalization, which corrects informal words or abbreviations (e.g., "yg" to "yang," "gmn" to "bagaimana") using an Indonesian slang dictionary;
4) Tokenization, which splits sentences into smaller units (tokens) that can be processed by the model;
5) Emoji and emoticon handling, where emojis and excessive emoticons were removed because they could introduce noise into the text representation process; and
6) Duplicate text checking, which ensured that repeated opinions and reposted content were excluded from the final dataset.

C. Transformer Modeling: IndoBERT and NusaBERT

This study employs a fine-tuning approach on two Transformer-based pre-trained language models specifically designed for the Indonesian language. The first is IndoBERT, a model trained on the Indo4B dataset, which consists of billions of words sourced from news articles, Wikipedia, and other formal text corpora. IndoBERT utilizes the Bidirectional Encoder Representations from Transformers (BERT) architecture, enabling it to understand contextual information in both directions (left-to-right and right-to-left).

The second model is NusaBERT, a Transformer variant developed through the NusaCrowd initiative, with a stronger focus on understanding conversational language, regional dialects, and informal styles commonly found in Indonesian social media. In this study, both models are implemented using the HuggingFace Transformers library with the following baseline hyperparameters: a learning rate of 2e-5, a batch size of 16, and 3 training epochs [15]. The experiments were conducted using Google Colab Pro with NVIDIA Tesla T4 GPU support and 16 GB RAM. Fine-tuning was implemented using the HuggingFace Transformers library and PyTorch framework. Both models utilized the AdamW optimizer with a maximum sequence length of 128 tokens.

The average training time for each model ranged from 20 to 30 minutes depending on the model architecture and tokenization complexity.

D. Model Performance Evaluation

To compare the effectiveness of IndoBERT and NusaBERT, model performance is evaluated using a Confusion Matrix. The evaluation is based on four primary classification metrics: Accuracy, Precision, Recall, and F1-Score. The mathematical formulas used to compute each metric are presented as follows:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$
$$Precision = TP / (TP + FP)$$
$$Recall = TP / (TP + FN)$$
$$F1\text{-}Score = 2 \times (Precision \times Recall) / (Precision + Recall)$$

Where True Positive (TP) represents correctly predicted positive instances, True Negative (TN) represents correctly predicted negative instances, False Positive (FP) represents incorrectly predicted positive instances (Type I error), and False Negative (FN) represents incorrectly predicted negative instances (Type II error). The F1-Score is used as the primary indicator, as it represents the harmonic mean of Precision and Recall, making it particularly useful in cases where the dataset may be imbalanced [16]. In addition to weighted evaluation metrics, macro-level evaluation was also considered to ensure balanced classification performance across all sentiment categories. The results indicate that both IndoBERT and NusaBERT maintain relatively consistent precision, recall, and F1-score values among positive, neutral, and negative classes, suggesting that the models do not exhibit significant bias toward any specific sentiment category.

E. Ethical Considerations

This study only utilized publicly accessible social media content from X (Twitter) and Instagram. Personal identifiers such as usernames, profile photos, and account information were anonymized during preprocessing to protect user privacy. The data collection process was conducted solely for academic research purposes and followed general platform usage policies regarding publicly available data.

III. RESULTS AND DISCUSSION

A. Distribution and Trends of Public Sentiment

Based on the data extraction process conducted during the period following the escalation of conflict in the Middle East, a dataset of 1,500 public opinion text samples was obtained, all of which had undergone preprocessing. The classification results indicate that public sentiment toward fuel (BBM) availability and policy is distributed in a highly proportional and balanced manner.

As illustrated in Figure 3, there are 500 negative sentiment opinions, predominantly reflecting public concern and anxiety regarding potential price increases and fuel shortages. On the other hand, 502 opinions are classified as positive, mostly expressing appreciation for the anticipatory measures taken by the government and Pertamina in maintaining national buffer stock reserves. The remaining 498 opinions are categorized as neutral, generally consisting of objective, analytical views and factual observations from netizens.

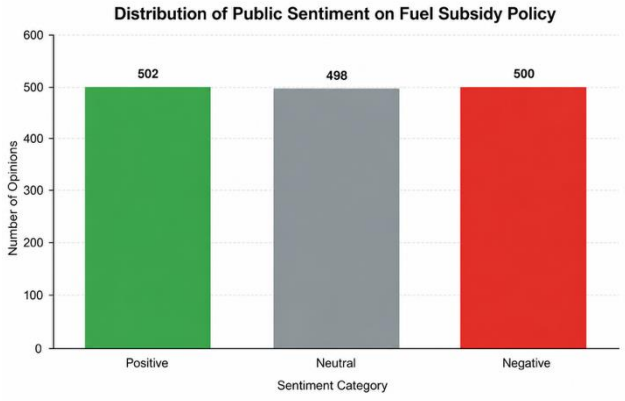

Figure 3. Distribution of Public Sentiment on Fuel (BBM) Policy

Furthermore, the dynamics of public opinion over time can be observed through time series analysis, which shows a fluctuating pattern. At the early stage of the global energy crisis, the intensity of negative sentiment often surged due to the spread of information regarding global crude oil prices surpassing new psychological thresholds, thereby triggering the potential for panic buying. However, the sentiment trend gradually shifted toward positive and neutral as a result of extensive official clarifications and communication interventions by the government, particularly regarding assurances of Pertalite availability quotas. The fluctuation of daily public opinion throughout the data collection period is visually represented in Figure 4.

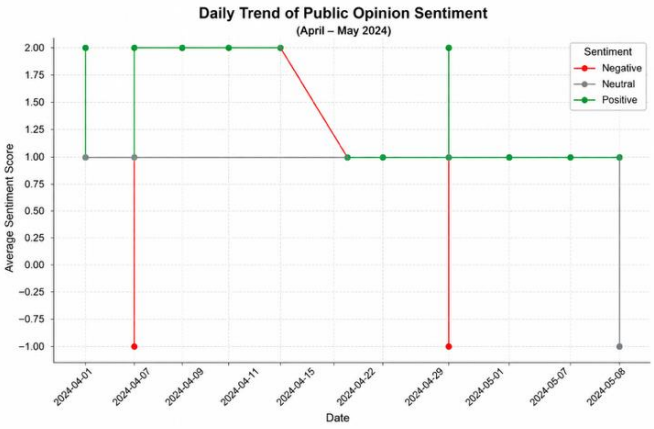

Figure 4. Daily public sentiment trends following the Middle East conflict, showing a shift from negative toward positive sentiment as a result of government information interventions.

B. Keyword Analysis (Word Cloud)

To validate the semantic context of public opinions, feature extraction was performed using Word Cloud visualization techniques. This method displays the most frequently occurring words in the corpus, where the size of each word represents its frequency. The visualization in Figure 5 illustrates the mapping of these keywords.

Figure 5. Word Cloud visualization of public sentiment

Referring to Figure 5, words such as "government," "BBM," "war," "subsidy," and "secure" appear as dominant terms. This confirms that public discourse is primarily centered on expectations for the government to take strategic policies to secure fuel subsidy quotas from the impact of war, while also providing a sense of security for the working class and MSME actors who are vulnerable to energy inflation shocks.

C. Model Training Process (Fine-Tuning)

Before evaluation, the Transformer models undergo a fine-tuning process using 80% of the dataset (1,200 samples). This computational learning process is configured using standard hyperparameters, including a learning rate of 2e-5, a batch size of 16, and 3 training epochs.

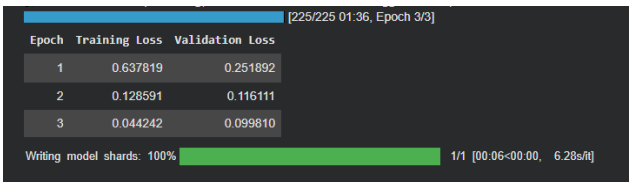

Figure 6. Training log of the NusaBERT fine-tuning process

Based on the training log shown in Figure 6, it can be observed that as the number of epochs increases, both Training Loss and Validation Loss decrease in a structured manner (from 0.25 in the first epoch to 0.09 in the third epoch). The reduction in loss values indicates that the NusaBERT architecture successfully learns the informal language patterns in the dataset without experiencing overfitting. The calibrated model is then used to evaluate the remaining 20% of the test data.

D. Model Evaluation and Performance Comparison

The core of this study is to compare the reliability of the Transformer architectures IndoBERT and NusaBERT in performing sentiment classification tasks. Based on the evaluation results on the testing dataset, both models demonstrate strong performance. A comprehensive summary of the evaluation results is presented in Table II.

TABLE II
COMPARISON OF TRANSFORMER MODEL PERFORMANCE

Evaluation Metrics	IndoBERT (%)	NusaBERT (%)
Accuracy	92.5	96.3
Precision	92.1	96.4
Recall	92.8	96.3
F1-Score	92.4	96.3

The comparative data in Table II empirically demonstrate that the NusaBERT architecture consistently outperforms the baseline IndoBERT model across all evaluation metrics. NusaBERT achieves the highest Accuracy score of 96.3% and an F1-Score of 96.3%. In comparison, IndoBERT records an Accuracy of 92.5% and an F1-Score of 92.4%. These results indicate that NusaBERT's domain-specific pre-training, which incorporates social media corpora and informal language, is more effective than IndoBERT, which is primarily trained on formal text. The percentage comparison of the evaluation metrics for both algorithms is further illustrated through a bar chart in Figure 7.

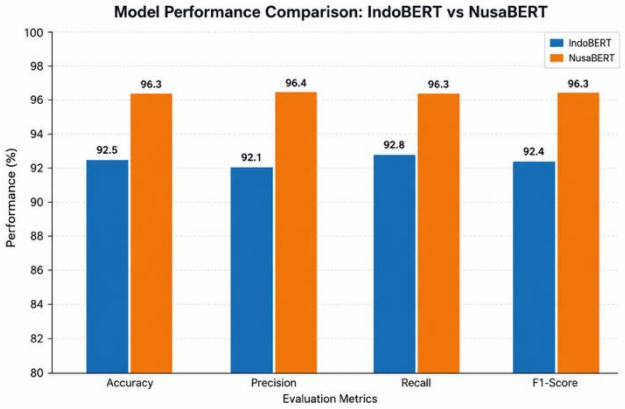

Figure 7. Comparative graph of evaluation metrics between IndoBERT and NusaBERT in classifying Indonesian text sentiment.

E. Discussion and Policy Implications

The superior performance of NusaBERT in this study can be justified by its pre-training architecture. Public opinions regarding fuel (BBM) policies on social media often do not follow formal Indonesian grammar standards (EYD). Many sentences contain slang, spontaneous abbreviations (e.g., "gmn nih bbm", "hrg mnyak naek"), and sarcasm. NusaBERT, which is specifically designed and trained using informal language corpora and local Indonesian dialects

(through the NusaCrowd initiative), demonstrates better capability in capturing lexical context within noisy text compared to IndoBERT, which is primarily trained on formal documents such as Wikipedia and news articles. From a practical perspective, these findings have important implications for policymakers. The NusaBERT algorithm is potentially suitable for implementation as the core engine of a government social listening dashboard. With high classification accuracy, relevant institutions can detect early signs of mass panic (panic buying) in real time following global geopolitical shocks. This enables mitigation strategies such as market operations and public communication management to be executed before negative sentiment escalates into a domestic economic crisis.

However, further integration with real-time monitoring infrastructure and broader multi-platform validation is still required before large-scale governmental deployment can be fully implemented.

F. Data Preprocessing Results (Text Preprocessing))

Before being processed by Transformer models, the quality of social media text data significantly affects the final classification results. Raw text contains substantial noise, such as user mentions, links, and informal language that may not be recognized by the model's standard vocabulary. Table IV presents representative samples of text transformation results during the preprocessing stage.

TABLE III
EXAMPLES OF TEXT PREPROCESSING RESULTS

Raw Text	Clean Text
@rakyat_kecil Perang di Timur Tengah bikin ngeri, harga minyak dunia naik!! smg pemerintah gak naikin harga BBM subsidi :((	perang di timur tengah buat ngeri harga minyak dunia naik semoga pemerintah tidak naikkan harga bbm subsidi
Waduh, denger kabar BBM mau naik gara2 perang Iran-Israel. Makin susah cari uang klo beneran naik.	waduh dengar kabar bbm mau naik gara gara perang iran israel makin susah cari uang kalau sungguhan naik
Tadi ke pom bensin sepi2 aja tuh, stok pertamax jg full. Stop nyebarin hoax bbm langka deh #InfoBBM	tadi ke pom bensin sepi saja itu stok pertamax juga penuh berhenti sebarkan kebohongan bbm langka deh

Based on Table III, the cleaning process is proven to effectively simplify lexical structures without removing the semantic context of the sentences. The normalization of abbreviations such as "smg" to "semoga" and "klo" to "kalau" is particularly crucial in helping models like IndoBERT trained on formal corpora map these words into appropriate mathematical vector representations.

G. Confusion Matrix and Error Analysis

To comprehensively understand how NusaBERT, as the best-performing model, makes predictions, the Confusion Matrix is further analyzed. The Confusion Matrix provides a detailed mapping between actual labels (ground truth) and predicted labels generated by the model. The distribution of NusaBERT's predictions is illustrated in detail in Figure 8.

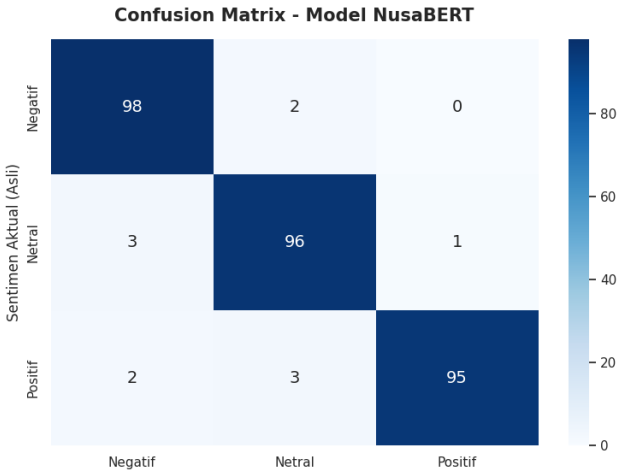

Figure 8. Heatmap visualization of the Confusion Matrix for the NusaBERT model, showing prediction accuracy along the main diagonal.

Referring to Figure 8, the cells along the main diagonal with the darkest blue intensity represent the True Positive (TP) values for each sentiment class. NusaBERT demonstrates very high classification sensitivity, correctly predicting 98 negative sentiment opinions out of 100 actual samples. Similarly, for the positive class, the model accurately predicts 95 opinions. These impressive results confirm that the NusaBERT architecture exhibits exceptional stability in detecting and classifying clear sentiment polarities (strong positive and strong negative).

However, the error analysis stage reveals that the greatest computational challenge lies in the classification of the neutral class. Based on the matrix, out of 100 actual neutral test samples, 3 opinions are misclassified as negative, and 1 opinion is incorrectly classified as positive. These prediction errors (Type I and Type II) are generally caused by the semantic complexity of noisy social media text, such as sarcasm, irony, or conditional sentence structures.

As a representative case, the test sentence: "Anggaran subsidi BBM pasti bengkak lagi tahun ini kalau konflik berkepanjangan" is contextually annotated as Neutral by human evaluators, as it reflects an analytical and factual assumption related to macroeconomic conditions. However, the model automatically classifies it as Negative due to the high attention weights assigned to pessimistic lexicons such as "bengkak" (increase/swelling) and "konflik" (conflict), without recognizing it as a conditional statement.

Despite the presence of such minor contextual interpretation anomalies, the overall error margin of the model remains within an acceptable threshold. These findings suggest NusaBERT as a highly robust algorithm in the context of Indonesian Natural Language Processing (NLP) research.

IV. CONCLUSION

This study compared the performance of IndoBERT and NusaBERT for classifying Indonesian public sentiment on social media regarding fuel subsidy policies during geopolitical tensions in the Middle East. The experimental results indicate that NusaBERT achieved better classification performance than IndoBERT across all evaluation metrics, with an accuracy of 96.3% and weighted F1-score of 96.3%. The findings suggest that NusaBERT is more effective in handling informal Indonesian social media text, including slang expressions and non-standard linguistic patterns commonly found in online public discussions.

In addition to quantitative evaluation, confusion matrix and error analysis revealed that both models still face challenges in handling sarcasm, ambiguous expressions, and mixed-context opinions. Nevertheless, the overall performance demonstrates the potential of Transformer-based approaches for supporting social media sentiment monitoring related to public policy issues.

From a practical perspective, the findings of this study may contribute to the development of social listening systems for monitoring public reactions toward fuel subsidy policies and other strategic governmental issues. However, further integration with real-time monitoring infrastructure and broader multi-platform validation is still required before large-scale implementation can be fully applied.

Despite achieving strong classification performance, this study has several limitations. First, the dataset size remains relatively limited for large-scale Transformer fine-tuning and was collected only from X (Twitter) and Instagram. Second, the study did not perform external validation using different public policy datasets, limiting the generalizability of the models across domains. Third, sarcasm, irony, and mixed-context opinions remain difficult for both models to classify accurately. In addition, this study did not include statistical significance testing such as McNemar or paired t-test analysis. Furthermore, this research only utilized descriptive keyword extraction through Word Cloud visualization and did not implement advanced topic modeling or co-occurrence analysis techniques.

Future studies are encouraged to utilize larger multi-platform datasets, integrate topic modeling or aspect-based sentiment analysis, and evaluate model robustness across different public policy domains. Further research may also explore real-time monitoring systems integrated with interactive dashboards for governmental policy evaluation.

REFERENCES

[1] C. D. Aprida, "Menyulam Ketegangan Timur Tengah Dampak Ekonomi Global dari Konflik Iran-Israel," vol. 1, no. 1, pp. 25–36, 2025.

[2] D. Soesanto Edy, Utami Dewi Puspita, "Dampak Fluktuasi Harga Minyak Dunia Terhadap Ekonomi di Indonesia," *Ris. Ilm.*, vol. 2, no. 1, pp. 231–242, 2025, doi: https://doi.org/10.62335 Jurnal.

[3] E. M. Dilasari, "Pengaruh Harga Minyak Goreng dan Panic Buying Terhadap Keputusan Pembelian dalam Perspektif Bisnis Syariah," 2023.

[4] J. Jtik, J. Teknologi, and M. Samantri, "Perbandingan Algoritma Support Vector Machine dan Random Forest untuk Analisis Sentimen Terhadap Kebijakan Pemerintah Indonesia Terkait Kenaikan Harga BBM Tahun 2022," vol. 8, no. 1, pp. 1–9, 2024.

[5] N. Wulandari, Y. Cahyana, and H. H. Handayani, "Sentiment Analysis on the Relocation of the National Capital (IKN) on Social Media X Using Naive Bayes and K-Nearest Neighbor (KNN) Methods," vol. 9, no. 3, pp. 724–731, 2025.

[6] B. Wilie *et al.*, "IndoNLU : Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," pp. 843–857, 2020.

[7] S. Cahyawijaya *et al.*, "NusaCrowd : Open Source Initiative for Indonesian NLP Resources," 2022.

[8] Liu, M. Ott, N. Goyal, et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint arXiv:1907.11692, 2019.

[9] M. C. Kenton, L. Kristina, and J. Devlin, "BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding," pp. 4171–4186, 2019.

[10] N. Nur, A. Aryanti, and O. Suria, "Analisis Sentimen Terhadap Pemutusan Hubungan Kerja di Indonesia : Komparasi IndoBERT dengan SVM,Random Forest, dan Decision Treedengan Optimasi TF-IDF," *Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 1158–1176, 2025.

[11] P. Subarkah, A. N. Ikhsan, E. Anggraeni, and A. P. Sabaniyah, "Sentiment Perspective of Government ' s Free Nutritious Meal Policy on Social Media X using Indo-BERT and Bi-LTSM," vol. 7, no. 2, pp. 201–210, 2025.

[12] M. Y. Dhinora and E. Mailoa, "Analisa Tweet Mahasiswa untuk Deteksi Gejala Depresi dengan Penerapan Natural Language Processing," vol. 6, no. 2, pp. 1193–1211, 2025.

[13] W. Syahrina, "Strategi Komunikasi Public Relations dalam Mengelola Isu Lingkungan pada Perusahaan Energi Syahrina Wirdani," *Adm. Publik Univ. Medan Area*, pp. 1–8.

[14] A. W. V Hutabarat, N. Luh, S. Saraswati, and K. Suryadi, "Analisis Sentimen Data Ulasan Pengguna MyPertamina di Twitter dengan Metode Machine Learning dan Deep Learning," vol. 13, no. 1, pp. 145–154, 2024.

[15] L. Afuan, N. Hidayat, H. Hamdani, and H. Ismanto, "Optimizing BERT Models with Fine-Tuning for Indonesian Twitter Sentiment Analysis," vol. 2, pp. 248–267, 1959, doi: 10.58346/JOWUA.2025.I2.016.

[16] E. Bagli and G. Visani, "Metrics for multi-class classification: An overview," pp. 1–17, 2020.