



Volume 11 Issue 1 Year 2026 | Page 34-43 ISSN: 2527-9866  
 Received: 11-12-2025 | Revised: 19-12-2025 | Accepted: 26-12-2025

## LoRA Enhanced Sentiment Aware Topic Modeling for Indonesian Generative AI Perception

Wisnu Ginanjar Saputra<sup>1</sup>, Noor Latifah<sup>2</sup>, Fajar Nugraha<sup>3</sup>

<sup>1,2,3</sup>Muria Kudus University, Kudus, Jawa Tengah, Indonesia, 59327

e-mail: 202253131@std.umk.ac.id<sup>1</sup>, noor.latifah@umk.ac.id<sup>2</sup>, fajar.nugraha@umk.ac.id<sup>3</sup>

\*Correspondence: 202253131@std.umk.ac.id

**Abstract:** Public understanding of Generative AI in low-resource language contexts remains underexplored, particularly in relation to how sentiment aligns with thematic discussions on social media. In Indonesia, empirical studies examining this interaction at scale are still limited. This study introduces a sentiment-aware topic modeling framework that integrates parameter-efficient fine-tuning of IndoBERT using Low-Rank Adaptation with topic discovery via BERTopic. The approach enables large-scale analysis of Indonesian social media data under constrained computational settings. Analysis of Indonesian Twitter discourse shows that general discussions of Generative AI are largely neutral and cautious, contrasting with more optimistic trends reported in Western contexts. In comparison, enthusiast communities exhibit predominantly positive sentiment, while ethics-related discussions display balanced polarization. These results highlight the contextual nature of public perception across different discussion domains. The findings demonstrate the applicability of parameter-efficient NLP methods for sentiment and topic analysis in under-resourced languages and provide insights relevant to technology development and policy formulation.

**Keywords:** BERTopic, Generative AI, LoRA, Sentiment Analysis, Topic Modeling.

### 1. Introduction

The rapid proliferation of Generative AI technologies has triggered global discourse regarding their societal implications [1], yet empirical analysis remains predominantly focused on English-speaking Western contexts. In Indonesia, the fourth most populous nation with over 221 million internet users [2], social media platforms have evolved into dynamic forums for technological debate [3]. However, public sentiment toward Generative AI in this region remains largely unexplored. While prior studies have successfully applied IndoBERT for general sentiment analysis [4], [5] and hybrid architectures for election predictions [6], these approaches rely on computationally intensive full-parameter fine-tuning. Conversely, while recent work has introduced Low-Rank Adaptation (LoRA) for Indonesian diplomatic cable classification [7], its application for sentiment analysis in the domain of emerging technologies has not been investigated.

This study addresses a critical research gap: the lack of a parameter-efficient framework to analyze the intersection of sentiment and thematic discourse regarding Generative AI in low-resource languages like Indonesian. Current literature often treats sentiment analysis and topic modeling as isolated tasks, failing to capture how specific themes drive emotional responses [8]. Furthermore, the high computational cost of fine-tuning Large Language Models (LLMs) poses a significant barrier for researchers in developing countries. Addressing this gap is urgent, as understanding public perception is crucial for shaping technology acceptance and regulatory frameworks in Southeast Asia's largest digital economy.

To bridge this gap, we propose a LoRA-enhanced sentiment-aware topic modeling framework. This study integrates parameter-efficient fine-tuning of IndoBERT with BERTopic to analyze Indonesian social media discourse under constrained computational settings. The specific research objectives are to develop a parameter-efficient sentiment classifier using LoRA-tuned IndoBERT [9] that achieves performance comparable to full fine-tuning with significantly fewer trainable parameters, and to discover latent thematic structures in Indonesian Generative AI discourse using BERTopic [10]. Furthermore, this study aims to statistically quantify the association between public sentiment and specific discussion topics, and ultimately, to provide actionable insights into how Indonesian public perception diverges from global trends.

This research offers distinct contributions to the field of Natural Language Processing (NLP) and social computing, categorized into theoretical, empirical, and practical dimensions. This study advances the methodological integration of sentiment analysis and topic modeling. By statistically validating the correlation between sentiment polarity and latent topics ( $p < 0.0001$ ), we provide theoretical evidence that these two dimensions are interdependent rather than orthogonal. This supports recent approaches that advocate for the statistical integration of these tasks to reveal bidirectional relationships [8], challenging traditional independent modeling approaches.

Empirically, this research provides the first comprehensive analysis of Indonesian public perception toward Generative AI. Contrary to the polarized "enthusiasm vs. fear" dichotomy often observed in Western studies [1], [11], our findings reveal a dominant "neutral and cautious" sentiment (79.4%) among the Indonesian public. This empirical evidence highlights the unique socio-cultural response of Indonesian users, characterized by a pragmatic, wait-and-see attitude typical of early technology adoption stages in developing regions.

From a practical perspective, this study validates a resource-efficient NLP workflow. We demonstrate that LoRA can reduce GPU memory usage by approximately 48% while maintaining 89.96% accuracy, making advanced LLM analysis accessible for researchers with limited hardware. Furthermore, the findings provide actionable intelligence for policymakers and AI developers to tailor communication strategies that address specific public concerns regarding ethics and data privacy.

## 2. Methods

The Methods section describes the systematic procedures used to conduct the study. Authors must clearly explain the data, research site, instruments, techniques, and analytical methods employed to achieve the research objectives.

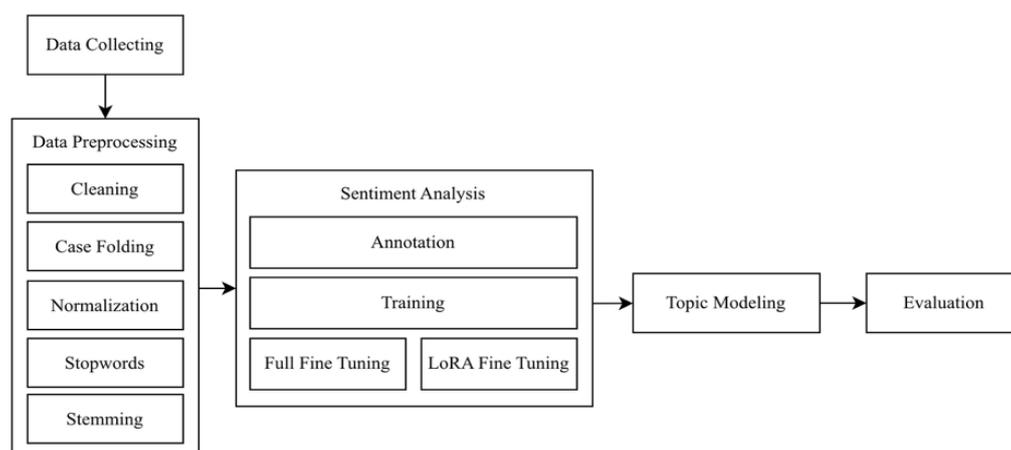


Figure 1. Research Method Pipeline

### A. Data Collection

We collected 9,958 Indonesian-language tweets discussing Generative AI topics using the Twitter API during the period from January 2023 to October 2025. The search was conducted using the keywords: "AI," "Artificial Intelligence," "Kecerdasan Buatan," and "ChatGPT." This data collection approach follows standard practices in social media-based sentiment analysis research. Tweets were subsequently filtered to include only posts with Indonesian geotags or authored by accounts with Indonesian language settings. All tweets were collected without including personally identifiable information to ensure user privacy protection. To avoid data bias, all personally identifiable information (PII) was removed from the dataset [12]. Table 1 presents sample examples from the dataset in this study.

**Table 1.** Sample of Dataset

Tweet	Date
Wamenkomdigi Nezar Patria menyatakan bahwa Kementerian Komunikasi dan Digital (Kemkomdigi) akan menyiapkan diskusi berseri dalam kajian regulasi teknologi kecerdasan buatan (AI). <a href="https://t.co/O4zejhg9nf">https://t.co/O4zejhg9nf</a>	February 25, 2025 at 06:42 PM
Haha, iya bener! Kebodohan kadang jadi hiburan tersendiri, ya 😂. Dan setuju banget, kecerdasan buatan emang keren banget buat ngebantu kita! Yuk, coba deh pakai "Asisten AI" biar makin pinter! 💡 ✨	February 15, 2025 at 05:42 PM
@Hani_Botani Oh iya, maaf 🙏🙏😞😞 Aku memakai kecerdasan buatan (ai).	September 29, 2024 at 10:15 AM

### B. Data Preprocessing

Preprocessing constitutes a crucial step in preparing raw data for analysis in machine learning [9] and NLP projects. Raw tweet text was processed through a data preprocessing pipeline comprising five stages in accordance with common practices in natural language processing, adapted to the characteristics of Indonesian social media data [3]. First, case folding was performed wherein all text was converted to lowercase to ensure format consistency [13]. Second, data cleansing was conducted by removing URLs, user mentions, special characters, and hashtags [9]. This data cleaning step aims to eliminate noise that could interfere with the model's ability to learn meaningful patterns [10]. Third, text normalization was applied using a lexicon-based approach to standardize informal expressions (e.g., "gue" to "saya," "ga" to "tidak") [3]. Fourth, stopword removal was performed to eliminate Indonesian stopwords using a stopword list while preserving negation words [13]. Stopwords are high-frequency words considered uninformative and eliminated due to their minimal contribution to semantic content [13]. Fifth, stemming was performed to reduce words to their base forms [3]. Table 2 displays preprocessing examples illustrating the transformation of raw tweets through each stage of the pipeline.

**Table 2.** Preprocessing Steps

Step	Tweet
Original	@neohistoria_id @FurqanIbr64677 Sampe disebut “kecerdasan buatan” 🤔🤔🤔 kalo “AI” gak pada paham
Case Folding	@neohistoria_id @furqanibr64677 sampe disebut “kecerdasan buatan” 🤔🤔🤔 kalo “ai” gak pada paham
Normalization	sampe disebut artificial_intelligence kalau artificial_intelligence tidak pada paham
Stopwords	sampe disebut artificial_intelligence kalau artificial_intelligence tidak paham
Stemming	sampe sebut artificial_intelligence kalau artificial_intelligence tidak paham

To establish ground truth labels for our dataset, sentiment annotation was performed using a semi-automatic approach initialized by an external pre-trained Indonesian RoBERTa Base Sentiment Classifier. This model served solely as an annotation assistant; the generated labels were then manually validated on 500 random samples with an agreement rate of 91.2%, confirming the reliability of this approach for dataset construction. The final distribution exhibited class imbalance: 58.4% negative, 33.9% neutral, and 7.7% positive.

### C. LoRA Enhanced Sentiment Classification

We selected IndoBERT (indobenchmark/indobert-base-p1) as the base pre-trained language model, which constitutes a monolingual pre-trained BERT model for Indonesian that achieves state-of-the-art performance across various NLP tasks [7]. This model comprises 12 layers, 768 hidden dimensions, 12 attention heads, and 124.5M total parameters [5], [7]. IndoBERT was selected because it demonstrates superior performance compared to multilingual models in Indonesian sentiment classification, with a +7.5 F1-score point improvement compared to MBERT [7]. Table 3 presents the hyperparameter configuration employed in the fine-tuning process.

**Table 3.** Fine-tuning Configuration

Parameter	LoRA Fine-tuning	Full Fine-tuning
Trainable parameters	887,043 (0.71%)	125,330,694 (100%)
LoRA config (r, $\alpha$ )	16, 32	-
Target modules	Q, K, V	All layers
Learning rate	$3 \times 10^{-4}$	$2 \times 10^{-5}$
Epochs	10	10
Batch size	10	8
Warmup ratio	16	0.1
LR scheduler	0.1	Linear
Dropout rate	Linear	0.1
Weight decay	0.15	0.01

Low-Rank Adaptation (LoRA) modifies pre-trained weight matrices by introducing trainable low-rank decomposition matrices [14]. The mathematical formulation of LoRA is shown in Equation (1):

$$W = W_0 + \Delta W = W_0 + BA \tag{1}$$

The sentiment classification task employs three classes: positive (enthusiasm, satisfaction), neutral (informational statements), and negative (concerns, criticism, fears). This three-class categorization follows standard approaches in sentiment analysis research [10].

For comparison, we trained a full fine-tuning baseline that updates all parameters with identical training configuration except for lower learning rate ( $2e-5$ ) and longer training time. Both models utilized a single T4 GPU (16GB VRAM) from Google Colab with PyTorch 2.9.1, Transformers 4.57.3, and PEFT 0.18.0 libraries.

### C. Topic Modeling with BERTopic

BERTopic is utilized to discover semantically coherent themes by leveraging transformer embeddings, dimensionality reduction, and density-based clustering [10]. Unlike traditional LDA models which rely on bag-of-words assumptions, BERTopic excels in capturing semantic relationships through contextualized embeddings [10], [11]. The topic discovery pipeline proceeds through four sequential stages. First, each tweet is converted into dense vector representations using the pre-trained IndoBERT model (indobenchmark/indobert-base-p1) to capture Indonesian linguistic context [13]. Second, Uniform Manifold Approximation and Projection (UMAP) reduces the embedding dimensions to 5 components to mitigate the curse of dimensionality while

preserving local and global topological structures. Third, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) groups these reduced embeddings into clusters, employing a soft-clustering approach that effectively isolates noise as outliers (Topic -1) [10]. Fourth, topic representations are extracted using class-based TF-IDF (c-TF-IDF) with CountVectorizer, configured to support unigrams and bigrams to capture meaningful phrases within the Generative AI discourse [10]. To ensure interpretability, we applied topic reduction to condense the resulting clusters into 7 distinct topics based on semantic similarity. The quality of the generated topics is quantitatively evaluated using the  $C_v$ .

#### D. Evaluation

Performance evaluation employs standard multi-class metrics [10]. Evaluation metrics include accuracy, precision, recall, and F1-score to provide comprehensive assessment of model performance [10]. Accuracy measures the proportion of correct predictions out of total predictions, calculated using Equation (2):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Precision measures the proportion of correct positive predictions, constituting a crucial metric in sentiment analysis for minimizing false positive errors. Precision calculation is shown in Equation (3):

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Recall measures the model's ability to detect positive samples [10], calculated using Equation (4):

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

F1-score combines precision and recall into a single harmonic metric, providing a balanced view of model performance [10]. F1-score calculation is shown in Equation (5):

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Macro-averaged F1-score treats all classes equally regardless of support, appropriate for imbalanced datasets. Training efficiency is measured by wall-clock time (minutes), GPU memory consumption (GB), number of trainable parameters, and inference latency (ms/sample) [15].

To rigorously compare the classification performance between the LoRA-enhanced model and the full fine-tuning baseline, we employed McNemar's test. This non-parametric statistical test is suitable for paired nominal data and evaluates whether the disagreement between two classifiers on the same test set is significant. We constructed a 2 x 2 contingency table to analyze instances where one model predicted correctly while the other failed. The null hypothesis ( $H_0$ ) posits that there is no significant difference in the predictive accuracy of the two models, with a significance threshold set at  $\alpha = 0.05$ . This statistical validation ensures that any reported efficiency gains from LoRA do not come at the cost of statistically significant performance degradation.

### 3. Results and Discussion

#### A. Sentiment Classification Performance

The LoRA-enhanced IndoBERT model achieved an accuracy of 89.96% and a weighted F1-score of 89.90% on the test set (n=1,992). To evaluate the model's discriminative capability across

classes, we visualized the prediction errors using confusion matrices. Figure 2 compares the confusion matrices of the LoRA model and the Full Fine-Tuning (Full-FT) baseline side-by-side.

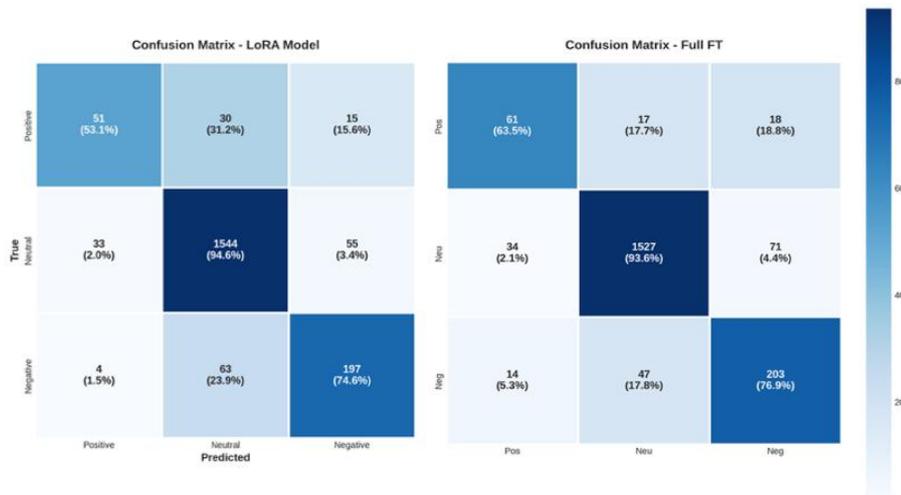


Figure 2. Comparative Confusion Matrix

As shown in Figure 2, performance varied significantly across sentiment classes. The neutral class achieved the highest stability (Recall 94.6%, Precision 94.32%), attributable to the dominance of informational tweets in the dataset. Negative sentiment showed moderate performance (Recall 74.6%), with 23.9% of samples misclassified as neutral. This reflects the linguistic challenge of identifying implicit criticism (hedging) in Indonesian social media text, where disapproval is often softened with polite markers. Positive sentiment exhibited the lowest recall (53.1%), frequently confused with neutral statements due to the subtle distinction between objective praise and subjective enthusiasm in short texts.

Training dynamics, illustrated in Figure 3 and Figure 4, reveal distinct convergence behaviors between the two approaches.

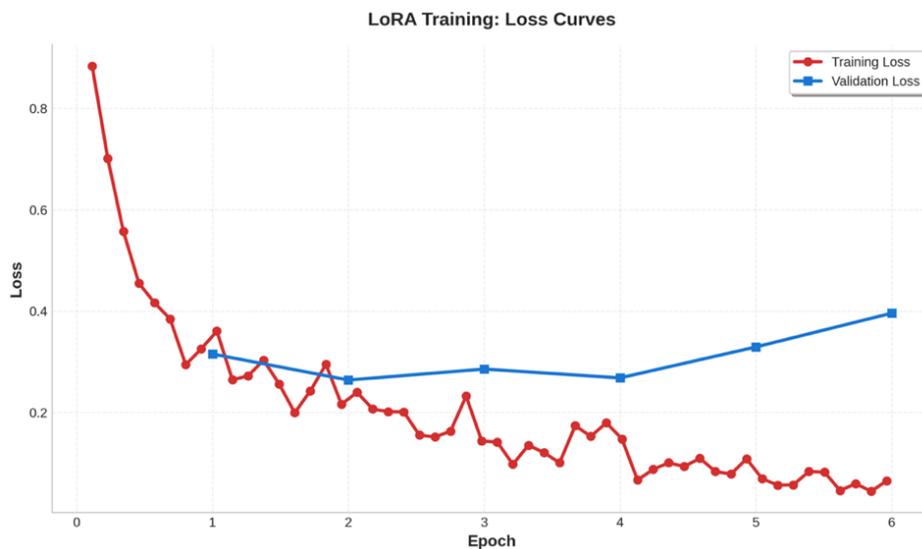
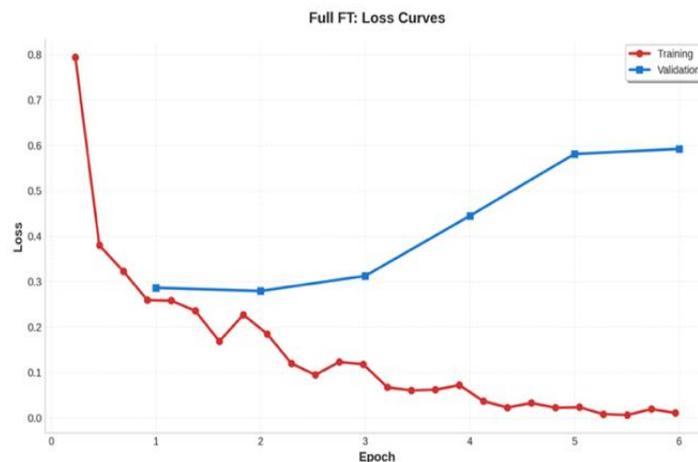


Figure 3. Validation Metrics Evolution LoRA



**Figure 4.** Validation Metrics Evolution Full Fine-tuning

The LoRA model Figure 3 demonstrates stable validation loss and rapid convergence, with accuracy stabilizing between epochs 2 and 4. In contrast, the Full-FT baseline Figure 4 exhibits sharper increases in validation loss after the initial epochs, indicating a stronger tendency towards overfitting despite its higher parameter capacity.

### B. Comparative Analysis with Related Studies

The performance of our LoRA-enhanced IndoBERT model (89.96% accuracy) demonstrates a highly competitive trade-off between predictive capability and computational efficiency when benchmarked against prior Indonesian sentiment studies. While Bahri and Suadaa [5] reported a marginally higher accuracy of 91.48% using standard transfer learning on IndoBERT for tourism reviews, their approach requires updating 100% of the model parameters, which demands substantial GPU resources. In contrast, our framework achieves comparable results by training only 0.71% of the parameters. This finding corroborates and extends the work of Nurlaila and Girsang [7], who applied LoRA for diplomatic cable classification. While they observed a slight performance trade-off in exchange for a 69.7% reduction in training time, our results show that with optimal hyperparameter tuning ( $r=16$ ,  $\alpha=32$ ), the performance gap between LoRA and full fine-tuning becomes statistically negligible ( $p=0.0823$ , McNemar’s test), validating LoRA as a robust solution for resource-constrained research environments in Indonesia.

From a socio-technical perspective, the thematic distribution of our findings presents a distinct deviation from global Generative AI discourse. Our analysis reveals a dominance of neutral sentiment (79.4%), which contrasts with trends observed in Western contexts where public opinion is often more polarized between high enthusiasm and deep concern regarding job displacement or ethics [1], [11]. Nedungadi et al. [11] highlight that AI discussions in social media often skew towards specific applications like healthcare or crisis management which trigger strong emotional responses. Conversely, the Indonesian discourse captured in this study reflects a "wait-and-see" attitude typical of the early technology adoption stage, where conversation focuses more on general information dissemination rather than the polarized ethical debates or intense optimism reported in global business sectors [1].

### C. Topic Modeling and Sentiment Topic Integration

BERTopic extracted six substantive topics and one outlier cluster from the dataset, achieving a mean coherence score of  $C_v = 0.623$ . The distribution of these topics, along with their representative keywords, is detailed in Table 4.

**Table 4.** Topic Distribution and Characteristic

Id	Topic Label	Count	Top Keywords
-1	Outlier	4,135	(noise, spam, off-topic)
0	General AI Technology	5,629	artificial_intelligence, teknologi, guna, manusia, digital
1	Informal ChatGPT Discussion	101	na, ni, chatgpt, lo, bro
2	AI Enthusiast Community	40	innovation, future, potential, revolutionize, breakthrough
3	AI News & Information	20	berita, informasi, narasumber, pakar
4	AI Ethics & Risks	17	ethics, privacy, bias, accountability, regulation
5	Political	16	jabat, negara, polri, deepfake, jabat

The topic distribution is notably imbalanced. Topic 0 (General AI) dominates the discourse, accounting for 56.53% of the data. This indicates that Indonesian public discussion is currently centered on general awareness and dissemination of technology news rather than deep technical or ethical engagement. Topic 1 (Informal ChatGPT Discussion) captures the colloquial "Jakarta slang" (e.g., lo, gue, bro), representing organic peer-to-peer conversation, while Topic 5 (Political Positions) highlights specific concerns regarding AI's role in governance and deepfakes, albeit in a smaller volume.

To analyze the relationship between thematic content and public emotion, we integrated the sentiment predictions with topic clusters. A chi-square test confirmed a significant association between sentiment and topic ( $\chi^2 = 691.15, p < 0.0001$ ). Figure 5 visualizes this relationship through a heatmap, identifying which topics trigger specific emotional responses.



**Figure 5.** Topic-Sentiment Heatmap

Detailed sentiment breakdown per topic is provided in Table 5.

**Table 5.** Sentiment Distribution per Topic

Topic	Positive (%)	Neutral (%)	Negative (%)
-1	5	79.7	15.3
0	22.8	71.3	5.9
1	95	5	0
2	0	100	0
3	23.5	41.2	35.3
4	0	100	0
5	5	79.7	15.3

As observed in Figure 5 and Table 5, distinct emotional profiles emerge for different topics. Topic 1 and Topic 2 (AI Enthusiast Community) exhibit the highest positive sentiment (95% for Topic 1), reflecting the optimism of early adopters and tech enthusiasts. In contrast, Topic 3 (AI News) and Topic 4 (AI Ethics & Risks) show a more balanced or neutral-negative distribution, suggesting that news reporting and ethical discussions in Indonesia are approached with caution and objectivity rather than polarization. This "neutral-heavy" pattern across most topics reinforces the finding that Generative AI in Indonesia is currently viewed as a subject of curiosity and information consumption rather than a source of strong societal conflict.

#### 4. Conclusions

This study developed a LoRA-enhanced sentiment-aware topic modeling framework for analyzing Indonesian Generative AI perceptions from 9,958 social media posts. The LoRA-fine-tuned IndoBERT achieved 89.96% accuracy while training only 0.71% of parameters, demonstrating 1.86x faster training and 2.21x less GPU memory than full fine-tuning with statistically equivalent performance ( $p=0.0823$ ). This validates parameter-efficient fine-tuning as a viable approach for NLP research in resource-constrained environments.

BERTopic identified six coherent topics ( $C_v = 0.623$ ) with significant sentiment-topic association ( $\chi^2 = 691.15$ ,  $p < 0.0001$ ). Indonesian AI discourse exhibits cautious mainstream attitudes (79.4% neutral) contrasting Western optimism, peer network amplification effects (22.8% positive), enthusiast echo chambers (95% positive), and balanced ethical debates. These patterns reflect collectivist cultural values and early adoption stage characteristics distinct from Western contexts, demonstrating that AI perception is shaped by regional socio-cultural factors requiring culturally-grounded research approaches.

The framework provides actionable intelligence for stakeholders: developers should prioritize localization and trust-building; policymakers must strengthen AI literacy and anticipatory governance; researchers should employ diverse sampling recognizing enthusiast communities do not represent general populations. Future work should investigate longitudinal dynamics, conduct cross-regional comparisons, integrate multimodal features, and apply causal inference methods to establish directional relationships between sentiment and topics.

#### References

- [1] D. Patil, N. L. Rane, and J. Rane, "Acceptance of ChatGPT and generative artificial intelligence in several business sectors: Key factors, challenges, and implementation strategies," *Future Impact ChatGPT Several Bus. Sect.*, pp. 201–241, Oct. 2024, doi: 10.70593/978-81-981367-8-7\_5.
- [2] Simon Kemp, "X Users, Stats, Data & Trends for 2025," *Datareportal*, Mar. 12, 2025. Accessed: Nov. 16, 2025. [Online]. Available: <https://datareportal.com/essential-x-stats>
- [3] R. A. Siregar and L. S. Darfiansa, "Sentiment Analysis Of Indihome Service Based On Geo Location Using The Bert Model On Platform X," *J. Tek. Inform. Jutif*, vol. 6, no. 5, pp. 2975–2990, 2025, doi: 10.52436/1.jutif.2025.6.5.4993.

- 
- [4] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models," *J. Ilm. Tek. Elektro Komput. Dan Inform.*, vol. 9, no. 3, pp. 746–757, Aug. 2023, doi: 10.26555/jiteki.v9i3.26490.
- [5] C. A. Bahri and L. H. Suadaa, "Aspect-Based Sentiment Analysis in Bromo Tengger Semeru National Park Indonesia Based on Google Maps User Reviews," *IJCCS Indones. J. Comput. Cybern. Syst.*, vol. 17, no. 1, p. 79, Feb. 2023, doi: 10.22146/ijccs.77354.
- [6] D. Guna Mandhasiya, H. Murfi, and A. Bustamam, "The hybrid of BERT and deep learning models for Indonesian sentiment analysis," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 33, no. 1, p. 591, Jan. 2024, doi: 10.11591/ijeecs.v33.i1.pp591-602.
- [7] D. Nurlaila and A. S. Girsang, "Optimizing diplomatic indexing: full-parameter vs low-rank adaptation for multi-label classification of diplomatic cables," *Comput. Sci. Inf. Technol.*, vol. 6, no. 3, pp. 274–282, Nov. 2025, doi: 10.11591/csit.v6i3.pp274-282.
- [8] Z. Kastrati, A. S. Imran, S. M. Daudpota, M. A. Memon, and M. Kastrati, "Soaring Energy Prices: Understanding Public Engagement on Twitter Using Sentiment Analysis and Topic Modeling With Transformers," *IEEE Access*, vol. 11, pp. 26541–26553, 2023, doi: 10.1109/ACCESS.2023.3257283.
- [9] Tedyyana, Agus, et al. "Enhance Telecommunication Security Through the Integration of Support Vector Machines." *International Journal of Advanced Computer Science & Applications* 15.3 (2024)..
- [10] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022, doi: 10.48550/ARXIV.2203.05794.
- [11] P. Nedungadi, G. Veena, K.-Y. Tang, R. R. K. Menon, and R. Raman, "AI Techniques and Applications for Online Social Networks and Media: Insights From BERTopic Modeling," *IEEE Access*, vol. 13, pp. 37389–37407, 2025, doi: 10.1109/ACCESS.2025.3543795.
- [12] H. Imaduddin, F. Y. A'la, and Y. S. Nugroho, "Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, 2023, doi: 10.14569/IJACSA.2023.0140813.
- [13] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," *Proc. 28th Int. Conf. Comput. Linguist.*, pp. 757–770, 2020, doi: 10.18653/v1/2020.coling-main.66.
- [14] N. H. Setiono and Y. Sari, "Exploring the Impact of Back-Translation on BERT's Performance in Sentiment Analysis of Code-Mixed Language Data," *IJCCS Indones. J. Comput. Cybern. Syst.*, vol. 19, no. 2, pp. 2460–7258, 2025, doi: 10.22146/ijccs.104757.
- [15] O. Grljević, M. Marić, and R. Božić, "Exploring Mobile Application User Experience Through Topic Modeling," *Sustainability*, vol. 17, no. 3, p. 1109, Jan. 2025, doi: 10.3390/su17031109.