International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue. 4, July-Aug 2026, page no. 71-84

To cite this article: Nazla Dzaalika Ainaya, Netti Herawati and Misgiyati, Subian Saidi (2026). FINE-TUNING INDOBERT USING FOCAL LOSS FOR IMBALANCED NEWS HEADLINE CLASSIFICATION, International Journal of Applied Science and Engineering Review (IJASER) 7 (4): 71-84 Article No. 291 Sub Id 427

FINE-TUNING INDOBERT USING FOCAL LOSS FOR IMBALANCED NEWS HEADLINE CLASSIFICATION

Nazla Dzaalika Ainaya, Netti Herawati and Misgiyati, Subian Saidi

Department of Mathematics, University of Lampung, Lampung, Indonesia

*Corresponding Author: Netti Herawati
Department of Mathematics, University of Lampung, Lampung, Indonesia
netti.herawati@fmipa.unila.ac.id.

DOI: https://doi.org/10.52267/IJASER.2026.7407

ABSTRACT

News headline classification plays an important role in identifying event-related information from large volumes of online news. However, performance often declines when datasets are imbalanced, as non-event news dominates and causes models to favor the majority class, reducing minority detection accuracy. This study applied the IndoBERT model with two fine-tuning approaches, cross entropy loss and focal loss to classify Indonesian news headlines into event and non-event categories, aiming to evaluate the effectiveness of focal loss on imbalanced data. Using a supervised learning framework, the dataset was split into training, validation, and testing sets, and both models were trained with identical hyperparameters: learning rate 2×10^{-5}, batch size 16, and 3 epochs. Evaluation employed accuracy, precision, recall, F1-score, confusion matrix, and ROC analysis. Results showed that cross entropy achieved 84% accuracy but lower minority-class recall, while focal loss improved performance to 95% accuracy with more balanced precision and recall. The findings indicate that focal loss helps the model focus on difficult samples, enhancing robustness and minority-class detection in imbalanced Indonesian news headline classification.

KEYWORDS: IndoBERT, Text Classification, Imbalanced Dataset, Focal Loss, Event Detection.

1. INTRODUCTION

The rapid development of information technology and internet usage has significantly increased digital news consumption. According to Telecommunication Statistics in Indonesia, internet penetration reached

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

72.78% of the population in 2024, indicating a growing dependence on online media as a primary information source [1]. The increasing volume of online news production, particularly in the form of short headlines, creates challenges in automatically identifying whether a headline represents an actual event or merely general information. Therefore, automatic text classification methods are required to efficiently distinguish event-related and non-event news content. In the field of Natural Language Processing (NLP), transformer-based models have achieved substantial progress in text understanding tasks [2]. Bidirectional Encoder Representations from Transformers (BERT) is a transformer-based model designed for pre-training on unlabeled text data using a bidirectional learning approach that captures semantic context from both preceding and succeeding words within a sentence [3]. Its Indonesian adaptation, IndoBERT, has demonstrated strong performance across various Indonesian NLP benchmarks due to its ability to capture contextual relationships between words effectively [4]-[7].

Despite these advances, text classification often faces the challenge of class imbalance, where one class contains significantly more samples than another [8]. Imbalanced data can bias classification models toward the majority class, reducing performance in detecting minority instances [9]. In this study, the news headline dataset consists of 12,290 non-event samples and 4,547 event samples, creating a skewed distribution that may negatively affect recall and F1-score for the minority class. Several approaches have been proposed to address imbalance problems, including modification of loss functions during model training. Focal loss, introduced by Lin et al. (2017), extends cross entropy loss by reducing the contribution of easily classified samples while emphasizing difficult samples [10]. This mechanism allows the model to focus more on minority-class learning and has been proven effective in imbalance classification scenarios [11]. Based on these considerations, this research applies IndoBERT with two fine-tuning strategies, namely cross entropy loss as a baseline and focal loss as an imbalance-aware optimization approach, to classify Indonesian news headlines into event and non-event categories. The study aims to analyze model performance under imbalanced data conditions using evaluation metrics including accuracy, precision, recall, F1-score, confusion matrix, and Receiver Operating Characteristic (ROC) curve analysis. Through this approach, the research seeks to provide an effective transformer-based classification framework capable of improving minority-class detection in Indonesian news headline classification.

2. METHODS

This study employed an experimental research design to compare two loss functions, namely cross entropy loss and focal loss, in fine-tuning the IndoBERT model for text classification with imbalanced dataset class. The objective was to analyze the effect of each loss function on classification performance in identifying event and non-event news headlines. The overall research stages consisted of data preparation, preprocessing, data splitting, model fine-tuning, performance evaluation, and comparative analysis.

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

2.1 Dataset Size and Sampling Technique

The dataset consisted of Indonesian news headlines labeled into two categories: event (minority class) and non-event (majority class). The dataset had an imbalanced class distribution (Figure 1). The total dataset contained 16,837 headlines, specifically contained 12,290 non-event headlines (class 0) and 4,547 event headlines (class 1).

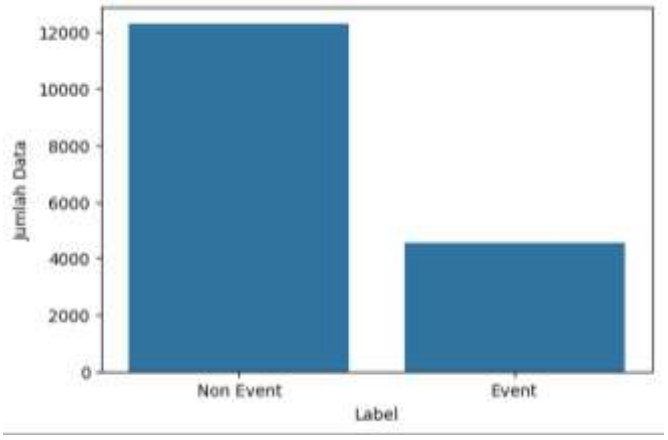


Figure 1. Distribution of event vs non-event datasets

This distribution indicates that approximately 73% of the data belonged to the majority class, while only 27% represented the minority class. The imbalance scenario was intentionally maintained to evaluate the effectiveness of focal loss in handling skewed class distributions and improving model sensitivity toward minority-class samples [12]. To analyze the effect of data partitioning on model performance, the dataset splitting process was conducted using three different scenarios, namely 70:30, 80:20, and 90:10. Each scenario represents the proportion between training data and testing data used during the experiment. The training set was used to fine-tune IndoBERT parameters, while the validation set was utilized for monitoring model performance during training, tuning hyperparameters, and preventing overfitting through early stopping mechanisms. The test set was strictly reserved for final performance evaluation and was not involved in any stage of model training or parameter adjustment [13].

Additionally, the splitting procedure was conducted after preprocessing but prior to model training to ensure consistent text representation across all subsets. Random seed initialization was applied during dataset partitioning to maintain experiment reproducibility [14]. Furthermore, each splitting scenario was executed three times to reduce randomness bias during data partitioning and to obtain more stable evaluation results. The final classification metrics reported in this study were calculated using the average

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

results from the three experimental runs. This structured data partitioning ensured reliable evaluation results and minimized bias in performance measurement.

2.2 Data Preprocessing

Before model training, several preprocessing steps were applied:

- Text normalization (lowercasing).
- Removal of unnecessary symbols and special characters.
- Tokenization using the IndoBERT tokenizer.
- Padding and truncation to ensure uniform input length.

Each headline was converted into token IDs, attention masks, and segment embeddings compatible with the IndoBERT architecture.

2.3 MODEL ARCHITECTURE

This study employed the IndoBERT-base-p1 pretrained language model for Indonesian text classification. IndoBERT is based on the transformer architecture consisting of multi-head self-attention layers that capture bidirectional contextual relationships between words [15]. Input headlines were tokenized into token IDs, attention masks, and segment embeddings before being processed by the encoder. The contextual representation of the special classification token [CLS] was used as the sentence-level feature vector. A fully connected classification layer was added on top of the encoder to perform binary classification (event and non-event). The classifier projected a 768-dimensional hidden representation into two output neurons followed by a softmax activation to generate prediction probabilities. All pretrained parameters and the classification layer were fine-tuned during training.

2.4 Fine Tuning Model

Two fine-tuning scenarios were implemented to evaluate model performance under imbalanced data conditions:

a. Cross Entropy Loss

Cross entropy loss was used as the standard optimization objective. This function minimizes the difference between predicted probability distributions and true labels, treating all samples equally during learning. The cross-entropy loss formula is defined as:

$$CE(p_t) = -\log(p_t) \tag{1}$$

where:

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

p_t is the predicted probability for the correct class,

a. Focal Loss

Focal loss is a modification of cross entropy loss that introduces a modulating factor to reduce the loss contribution of easily classified samples and focus more on hard examples. The focal loss formula is defined as:

$$FL - \quad \alpha_t(1 \quad p_t)^{\gamma}\log(p_t) \tag{2}$$

where:

p_t is the predicted probability for the correct class,

γ is the focusing parameter,

α_t is the balancing factor for class weighting.

2.5 Training Configuration

Model training was conducted using the PyTorch and HuggingFace Transformers frameworks. The model was trained with a learning rate of 2×10^{-5}, a batch size of 16, and 3 training epochs. A warmup ratio of 0.1 was applied to gradually increase the learning rate at the beginning of training. The maximum input sequence length was set to 32 tokens. The optimization process used the AdamW optimizer with a dropout rate of 0.1 to reduce overfitting. In addition, a linear learning rate scheduler was employed throughout the training process to stabilize parameter updates and improve convergence.

2.6 Model Evaluation

Model performance was evaluated using multiple classification metrics, accuracy, precision, recall, and F1-score.

Additionally, performance visualization was conducted using confusion matrix to analyze classification errors per class and Receiver Operating Characteristic (ROC) Curve and Area Under Curve (AUC), to measure the model's discrimination capability. Evaluation was performed on the unseen test dataset to ensure unbiased performance measurement.

3. RESULT AND DISCUSSION

Below are presented the results of each Indobert model's performance:

3.1 IndoBERT Performance Using Cross Entropy Loss

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

a) Data Splitting 70:30

The first research is evaluated IndoBERT fine-tuned using cross entropy loss as a baseline model on the test dataset containing 5,052 Indonesian news headlines. The evaluation results are summarized in Table 1.

Table 1. Performance of IndoBERT with Cross Entropy Loss (Data Splitting 70:30)

	Precision	Recall	F1-score	Support
0 (non-event)	0.85	0.95	0.90	3368
1 (event)	0.79	0.55	0.65	1364
Accuracy			0.84	5052

b) Data Splitting 80:20

Next is evaluate IndoBERT fine-tuned using cross entropy loss on the test dataset containing 3,368 Indonesian news headlines. The evaluation results are summarized in Table 2.

Table 2. Performance of IndoBERT with Cross Entropy Loss (Data Splitting 80:20)

	Precision	Recall	F1-score	Support
0 (non-event)	0.85	0.95	0.90	2458
1 (event)	0.80	0.55	0.65	910
Accuracy			0.84	3368

c) Data Splitting 90:10

Last evaluate IndoBERT fine-tuned using cross entropy loss on the test dataset containing 1,684 Indonesian news headlines. The evaluation results are summarized in Table 3.

Table 3. Performance of IndoBERT with Cross Entropy Loss (Data Splitting 90:10)

	Precision	Recall	F1-score	Support
0 (non-event)	0.85	0.95	0.90	1229
1 (event)	0.79	0.55	0.65	455
Accuracy			0.84	1684

International Journal of Applied Science and Engineering Review

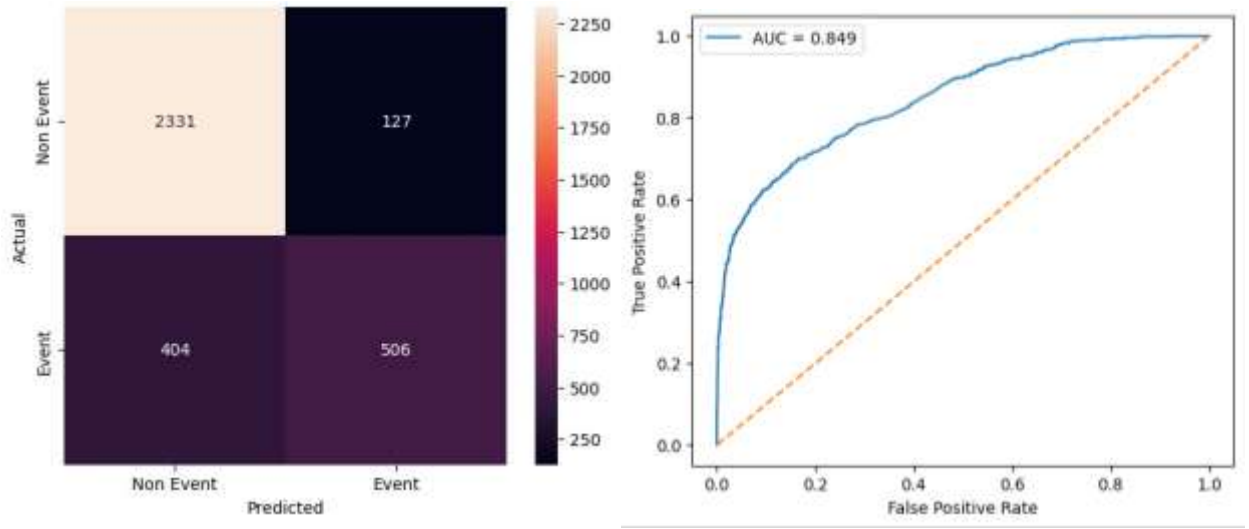
IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

The model summary achieved an accuracy of 84% across the three data splitting scenarios, showed relatively consistent performance on all evaluation metrics. No significant differences were observed between the splitting configurations, indicating that variations in the proportion of training and testing data did not substantially affect the overall classification performance.

Despite the stable accuracy results, the evaluation metrics revealed that the model was still more effective in recognizing the non-event class than the event class. The non-event class consistently obtained high recall values, while the event class produced relatively lower recall and F1-score values. This indicates that the model still struggled to detect minority-class samples when trained using cross entropy loss, causing many event headlines to be misclassified as non-event. Overall, the results suggest that the choice of loss function had a greater impact on classification performance than the data splitting configuration itself.



(a) (b)

Figure 2. (a) Confusion Matrix and (b) ROC Curve of IndoBERT model using Cross Entropy Loss

Figure 2 presents the confusion matrix and ROC curve of IndoBERT using cross entropy loss on the 80:20 data splitting scenario. The visualization shows a similar classification pattern in all scenarios, where the model was more dominant in correctly predicting the non-event class compared to the event class. Although most non-event samples were successfully classified, a relatively large number of event samples were still predicted as non-event, indicating that the model sensitivity toward the minority class remained limited.

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

This result is consistent with the evaluation metrics obtained from the classification reports, where the recall and F1-score values for the event class were significantly lower than those of the non-event class. The ROC curve also produced relatively similar AUC values across all splitting scenarios, indicating that changes in data partition proportions did not significantly influence the model's discriminative capability. Overall, these findings suggest that the limitations of the model were primarily caused by the characteristics of cross entropy loss, which tends to prioritize overall prediction accuracy and majority-class learning during the training process.

3.2 IndoBERT Performance Using Focal Loss

The second experiment evaluated IndoBERT fine-tuned using focal loss to address the class imbalance problem in Indonesian news headline classification.

a) Data Splitting 70:30

The evaluation results of IndoBERT using focal loss on the 70:30 data splitting scenario are presented in Table 4.

Table 4. Performance of IndoBERT with Focal Loss (Data Splitting 70:30)

	Precision	Recall	F1-score	Support
0 (non-event)	0.97	0.97	0.97	3688
1 (event)	0.93	0.90	0.91	1364
Accuracy			0.95	5052

b) Data Splitting 80:20

Next is evaluate IndoBERT fine-tuned using focal loss on the test dataset containing 3,368 Indonesian news headlines. The evaluation results are summarized in Table 5.

Table 5. Performance of IndoBERT with Focal Loss (Data Splitting 80:20)

	Precision	Recall	F1-score	Support
0 (non-event)	0.97	0.97	0.97	2458
1 (event)	0.92	0.91	0.92	910
Accuracy			0.95	3368

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

c) Data Splitting 90:10

Last evaluate IndoBERT fine-tuned using focal loss on the test dataset containing 1,684 Indonesian news headlines. The evaluation results are summarized in Table 6.

Table 6. Performance of IndoBERT with Cross Entropy Loss (Data Splitting 90:10)

	Precision	Recall	F1-score	Support
0 (non-event)	0.97	0.96	0.97	1229
1 (event)	0.90	0.91	0.91	455
Accuracy			0.95	1684

The evaluation results across the three data splitting scenarios showed that IndoBERT fine-tuned using focal loss consistently achieved high classification performance. All scenarios produced the same overall accuracy of 95%, indicating that changes in training and testing proportions did not significantly affect model performance. Overall, the evaluation metrics demonstrate that focal loss successfully improved model sensitivity toward the event class while maintaining high performance on the non-event class.

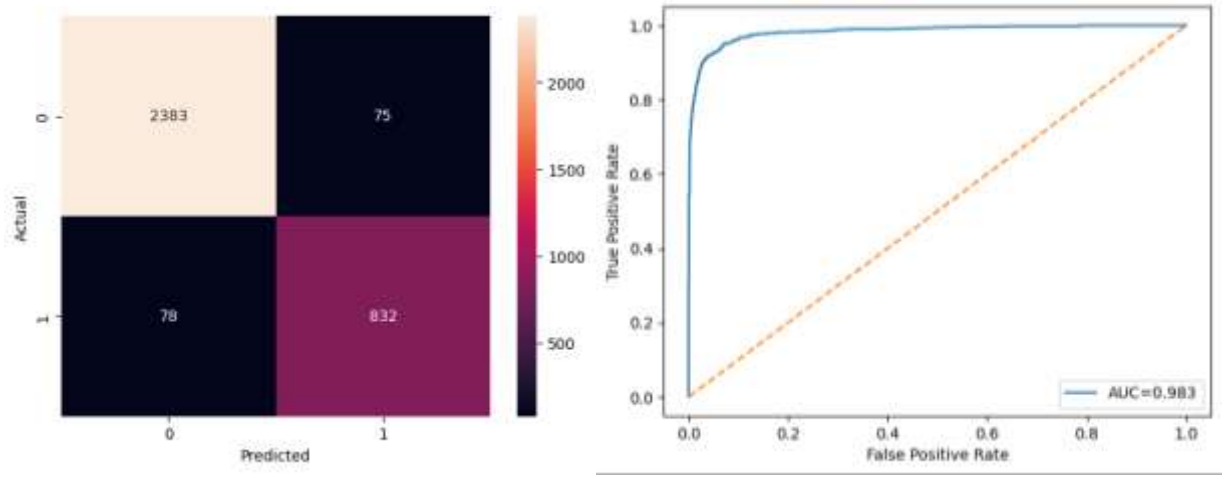
Compared to cross entropy loss, the resulting precision, recall, and F1-score values became more balanced across both classes, indicating that the model was no longer heavily biased toward the majority class. The confusion matrix and ROC curve in Figure 3 also confirm fewer false negatives and stronger discrimination capability across all splitting scenarios and presents IndoBERT model using focal loss on the 80:20 data splitting scenario, which produced the most optimal and stable performance among all splitting configurations.

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84



(a) (b)

Figure 3. (a) Confusion Matrix and (b) ROC Curve of IndoBERT model using Focal Loss

The confusion matrix shows that the model successfully classified 2383 non-event (0) headlines and 832 event (1) headlines correctly, while only a small number of samples were misclassified. Several non-event headlines were predicted as event, and a few event headlines were predicted as non-event. However, the number of misclassifications between both classes remained relatively balanced, indicating that the model no longer showed a strong bias toward the majority class.

Compared to the baseline model, the number of false negatives decreased significantly, demonstrating that focal loss improved the model's sensitivity toward the minority (event) class while still maintaining strong performance on the majority class. Furthermore, the ROC curve demonstrates excellent discriminative capability with an AUC value of 0.983. The curve positioned close to the upper-left corner indicates high true positive rate and low false positive rate, confirming that focal loss enabled the model to distinguish event and non-event classes more accurately and in a more balanced manner.

3.3 Comparison with Data Balancing Technique (SMOTE)

To further validate the effectiveness of focal loss in handling imbalanced data, an additional experiment was conducted using the Synthetic Minority Over-sampling Technique (SMOTE). Unlike focal loss, which improves learning through loss optimization, SMOTE balances the dataset by generating synthetic samples for the minority class before the training process.

International Journal of Applied Science and Engineering Review

IJASER

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

Table 7. Performance of IndoBERT with Balancing Dataset Technique

	Precision	Recall	F1-score	Support
0 (non-event)	0.95	0.94	0.94	2459
1 (event)	0.84	0.87	0.85	910
Accuracy			0.92	3369

The evaluation results on Table 7 showed that IndoBERT trained using SMOTE achieved an accuracy of 0.92, indicating improved performance compared to the baseline cross entropy loss. However, the performance still remained below the IndoBERT model using focal loss, which achieved an accuracy of 0.95 without applying any data balancing technique. In addition, the F1-score and recall values of the event class obtained using focal loss were also consistently higher. These findings demonstrate that optimizing the learning process through focal loss was more effective than balancing the dataset using synthetic data generation. Therefore, focal loss proved capable of handling class imbalance more efficiently while preserving the original data distribution.

3.4 Overall Model Evaluation and Performance Comparison

Performance comparison was conducted between IndoBERT using cross entropy loss, IndoBERT using focal loss, and IndoBERT combined with the SMOTE balancing technique. The evaluation focused on the model's ability to classify the minority (event) class using accuracy, precision, recall, F1-score, and AUC metrics. The overall comparison results are presented in Table 8.

Table 8. Comparative Performance Evaluation of IndoBERT Models

Evaluation Metrics	Model without Data Balancing						Model with Data Balancing		
	Cross Entropy Loss			Focal Loss			SMOTE		
	70:30	80:20	90:10	70:30	80:20	90:10	70:30	80:20	90:10
Accuracy	0.84	0.84	0.84	0.95	0.95	0.95	0.92	0.92	0.92
Precision	0.79	0.80	0.79	0.91	0.92	0.90	0.82	0.84	0.82
Recall	0.55	0.55	0.55	0.91	0.91	0.91	0.89	0.87	0.89
F1-score	0.65	0.65	0.65	0.91	0.92	0.91	0.85	0.85	0.85
AUC	0.851	0.851	0.846	0.986	0.987	0.986	0.967	0.975	0.965

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

value

Based on Table 8, IndoBERT using focal loss consistently achieved the best performance across all splitting scenarios without applying any data balancing technique. The most significant improvement occurred in the minority (event) class, where the recall value increased from 0.55 using cross entropy loss to 0.91 using focal loss. In addition, the model using focal loss achieved a stable accuracy of 95% and an AUC value reaching 0.987 on the 80:20 data splitting scenario, indicating excellent discriminative capability in distinguishing event and non-event classes.

Meanwhile, the SMOTE-based model also improved classification performance compared to cross entropy loss, particularly in the recall of the event class, which increased to 0.89. However, the overall performance still remained below the IndoBERT model using focal loss. These findings indicate that optimizing the learning process through loss function modification was more effective than balancing the dataset using synthetic data generation. Overall, the results demonstrate that focal loss successfully improved the sensitivity of IndoBERT toward minority-class detection while maintaining balanced classification performance, without requiring additional data balancing techniques.

4. CONCLUSION

This study evaluated the performance of IndoBERT for Indonesian news headline classification under imbalanced data conditions by comparing cross entropy loss, focal loss, and data balancing using SMOTE. The experimental results showed that the model using cross entropy loss achieved an accuracy of 84%, but still demonstrated limitations in detecting the minority (event) class, particularly in recall and F1-score values. The application of SMOTE improved the classification performance with an accuracy of 92% and better minority-class detection compared to cross entropy loss. However, the model still relied on synthetic data generated during the balancing process.

Among all approaches, focal loss produced the best and most balanced classification performance without requiring any data balancing technique. The model consistently achieved 95% accuracy with high precision, recall, F1-score, and AUC values across all experiments. In addition, focal loss significantly reduced false negatives and improved the model's sensitivity toward the minority (event) class while maintaining strong performance on the majority (non-event) class. These findings demonstrate that focal loss is more effective than both cross entropy loss and SMOTE for handling class imbalance in IndoBERT-based text classification.

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

REFERENCES

[1] Central Statistics Agency, "Telecommunication Statistics in Indonesia" 2024. [Online]. Available: https://www.bps.go.id/id/publication/2025/08/29/beaa2be400eda6ce6c636ef8/telecommunication-statistics-in-indonesia-2024.html. [Accessed: March. 7, 2026]

[2] R.C. Tarumingkeng, "Natural Language Processing (NLP)". RUDYCT e-PRESS, Bogor, Indonesia, 2024.

[3] S. Shreyashree, P. Sunagar, S. Rajarajeswari, & A. Kanavalli, "A literature review on bidirectional encoder representations from transformers," In Proceedings of ICICIT, pp. 305-320, 2021.

[4] F.M. Habibi, R.G. Guntara, & A. Nuryadin, "Komparasi Model IndoBERT dan IndoGPT untuk Analisis Sentimen pada Produk E-commerce," Jurnal Algoritma, XXII (2), pp. 2249-2259, 2025.

[5] M. Amien, & G.F. Gunawan, "BERT dan Bahasa Indonesia: Studi tentang Efektivitas Model NLP Berbasis Transformer," Journal of Interdisciplinary Research, I (2), pp. 132-140, 2024.

[6] C.J. L. Tobing, I.L. Wijayakusuma, & L.P.I. Harini, "Perbandingan kinerja IndoBERT dan mBERT untuk deteksi berita hoaks politik dalam bahasa Indonesia," JST (Jurnal Sains dan Teknologi), XIV (1), pp. 114–123, 2025.

[7] M.R. Manoppo, I.C. Kolang, D.N.N. Fiat, R.M.C. Mawara, A.D.P. Sumarno, A. Yusupa, & V. Tarigan, "Analisis sentimen publik di media sosial terhadap kenaikan PPN 12% di Indonesia menggunakan IndoBERT," Jurnal Kecerdasan Buatan dan Teknologi Informasi, IV (2), pp. 152–163, 2025.

[8] S.F. Taskiran, B. Turkoglu, E. Kaya, & T. Asuroglu, "A comprehensive evaluation of oversampling techniques for enhancing text classification performance," Scientific Reports, XV (1), p. 21631, 2025.

[9] A.A. Khan, O. Chaudhari, & R. Chandra, "A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation," Expert Systems with Applications, (244), p. 122778, 2024.

[10] T.Y. Lin, P. Goyal, R. Girshick, K. He, & P. Dollár, "Focal loss for dense object detection," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 2980–2988, 2017.

[11] G.A. Syafarina, I.I. Purnomo, & M. Hasbi, "Klasifikasi buah dan sayuran multi-label menggunakan CNN: Mengatasi class imbalance dengan focal loss," Jurnal CoSciTech (Computer Science and Information Technology), VI (3), pp. 561–567, 2025.

[12] J. Singh, C. Beeche, Z. Shi, O. Beale, B. Rosin, J. Leader, & J. Pu, "Batch-balanced focal loss: a hybrid solution to class imbalance in deep learning," Journal of Medical Imaging, X (5), p. 051809, 2023.

[13] H. Babaei, M. Zamani, & S. Mohammadi, "The impact of data splitting methods on machine learning models: A case study for predicting concrete workability," Machine Learning for Computational Science and Engineering, I (1), p. 21, 2025.

ISSN: 2582-6271

Vol. 7, Issue.4, July-Aug 2026, page no. 71-84

[14] L. Wegmeth, T. Vente, L. Purucker, & J. Beel, "The effect of random seeds for data splitting on recommendation accuracy," in Proceedings of the Perspectives Workshop at RecSys, 2023.

[15] A. Apandi, A.B. Mutiara, & D. Dharmayanti, "Generating image captions in Indonesian using a deep learning approach based on vision transformer and IndoBERT architectures," Journal of Applied Data Sciences, VI (2), pp. 1190–1201, 2025.