



Deep Sentiment Analysis of Halal Tourism: An Enhanced IndoBERT with Attention-Based and Coupling Latent Sampling on User-Generated Feedback

Rio Andika Malik¹, Marta Riri Frimadani²

¹Digital Business, Faculty of Economy, Business and Social Sciences, Universitas Perintis Indonesia, Padang, Indonesia

²Integrated Natural Resources Management, Postgraduate Program, Andalas University, Padang, Indonesia

*rioandikamalik@upertis.ac.id

Abstract

Public sentiment analysis regarding government policies, particularly in niche sectors like Halal tourism, often faces significant computational challenges due to data scarcity and severe class imbalance. Traditional augmentation methods, such as Synthetic Minority Over-sampling Technique (SMOTE) applied to raw text or TF-IDF vectors, often degrade semantic integrity, while standard fine-tuning of pre-trained models like IndoBERT tends to bias predictions toward majority classes. To address these limitations, this study proposes IndoBERT-LSS (Latent Space Sampling), a novel decoupled two-stage deep learning framework. The first stage employs a Representation Learning approach integrating an Attention-Pooling mechanism with Supervised Contrastive Loss (SupCon) to enforce compact intra-class clustering in the embedding space. The second stage introduces a Latent Space Sampling strategy, where SMOTE is applied to the extracted high-dimensional embeddings rather than the raw text, followed by a final classification using a Multi-Layer Perceptron. Validated on a dataset of 1,051 textual responses regarding Halal tourism in Pariaman, Indonesia, the proposed model achieved an accuracy of 93.84% and a macro F1-score of 0.94. Notably, the model demonstrated exceptional robustness in identifying minority classes, achieving a 0.99 F1-score for neutral sentiments. These results conclude that decoupling representation learning from feature balancing in the latent space significantly enhances model performance on imbalanced short-text datasets compared to standard fine-tuning baselines.

Keywords: contrastive learning; halal tourism; imbalanced data; IndoBERT; SMOTE

How to Cite: R. A. Malik and M. R. Frimadani, "Deep Sentiment Analysis of Halal Tourism: An Enhanced IndoBERT with Attention-Based and Coupling Latent Sampling on User-Generated Feedback," *J. RESTI (Rekayasa Sist. Dan Teknol. Informasi)*, vol. 10, no. 4, pp. 1107–1116, 2026.

Permalink/DOI: <https://doi.org/10.29207/resti.v10i4.7378>

Received: December 16, 2025

Accepted: May 23, 2026

Available Online: August 19, 2026

This is an open-access article under the [CC BY 4.0 License](#)

Published by [Ikatan Ahli Informatika Indonesia](#)

1. Introduction

The global tourism industry has witnessed a paradigm shift [1] with the rapid emergence of Halal tourism [2] as a significant economic driver [3],[4]. This industry is a strategic governmental objective for countries with a majority of Muslims, such as Indonesia [5], in order to promote inclusive economic growth [6], rather than just a market trend [7]. To capitalize on the region's rich Islamic cultural legacy [8], the local government of Pariaman City, West Sumatra, has actively pushed Halal tourism legislation [9],[10]. However, complicated socio-political processes are frequently triggered [11] by the implementation of such identity-based policies [12]. The public's reactions are rarely consistent [13], ranging from fervent support to doubts about economic inclusion [14] and worries about possible cultural exclusivity [15],[16]. To create

responsive and successful policies [17], stakeholders must accurately capture and analyze these varied public views [18].

Public perception analysis has historically relied on manual qualitative techniques [19], [20] like focus group discussions (FGD) [21] and small-scale surveys [22]. Although informative, these approaches are not suitable for real-time policy evaluation because they are time-consuming, resource-intensive [23], and sensitive to observer bias [24]. As a result using Deep Learning and Natural Language Processing (NLP) [25], [26] has become the norm for automated sentiment analysis [27]. By utilizing bidirectional context awareness, pre-trained Transformer-based models especially IndoBERT [28], [29], have attained state-of-the-art performance in the Indonesian language [30], [31]. Despite their success, there are two basic computational

issues [32] with using these models on real-world stakeholder data that have not yet been thoroughly addressed in the literature [33].

The issue of class disparity is the first and most important obstacle [34]. Survey and social media statistics on public opinion are inherently biased, with majority feelings (like "Agree") frequently predominating and minority sentiments (like "Neutral" or "Undecided") being under-represented [35]. When IndoBERT is fine-tuned using Cross-Entropy loss, it frequently experiences "minority collapse," in which the model ignores the subtle but important minority voices and biases its predictions towards the majority class in order to minimise global error [36]. Lack of data and noise in short-text inputs present the second difficulty [37], [38]. Survey responses are frequently succinct and devoid of context [39], which makes it challenging for models to differentiate between semantically similar categories like "Undecided" and "Disagree"[40].

Data augmentation strategies are commonly used by researchers to overcome class imbalance. In conventional machine learning, the Synthetic Minority Over-sampling Technique (SMOTE) is a widely used approach [41]. Nevertheless, synthetic samples that frequently lack semantic coherence are produced when SMOTE is applied directly [42] to high-dimensional sparse text representations (such TF-IDF or Bag-of-Words) [43]. On the other hand when the dataset size is limited, text-level augmentation techniques like synonym substitution or back-translation run the risk of changing the original sentiment or adding linguistic noise [44]. To summarize the core problem addressed in this research is the computational difficulty of accurately classifying public sentiment regarding Halal tourism policies due to three intertwined issues. First, the severe class imbalance that leads to minority collapse, second the high semantic ambiguity and lack of context in short-text Indonesian survey responses, and the semantic degradation and linguistic noise caused by conventional text-level augmentation approaches.

There is a significant research gap in developing a framework that handles class imbalance without compromising the semantic integrity of the text. The majority of recent research that has tried to integrate Deep Learning with sampling techniques focuses on end-to-end training [45], where feature extraction and classification are combined [46]. This conventional approach results in less-than-ideal feature separation for minority classes. Consequently, there remains a critical void in exploring decoupled frameworks that mathematically stabilize and separate textual features in the latent space *before* any synthetic augmentation is applied.

This paper suggests a novel technique to overcome these gaps called IndoBERT-LSS (Latent Space Sampling). We provide a decoupled two-stage system that regards class balance and representation learning as

distinct goals [40], in contrast to conventional methods [47]. The primary contributions of this research are threefold:

The primary contributions of this research are threefold. First, we propose the IndoBERT-LSS framework, which introduces a decoupled two-stage architecture that strictly separates representation learning from class balancing, directly addressing the minority collapse flaw in standard end-to-end fine-tuning [48]. Second, we integrate an Attention-Pooling mechanism with Supervised Contrastive Loss (SupCon) to mathematically enforce compact and distinct intra-class clustering in the latent space specifically tailored for short, ambiguous Indonesian texts [49]. Third, we empirically prove that applying Latent Space Sampling (SMOTE on dense embeddings rather than raw text) preserves semantic integrity [50]. This provides a technically viable and highly accurate analytical framework for stakeholders to capture marginalized minority voices in public policy debates, such as Halal tourism [51].

2. Methods

This research employs a quantitative experimental approach using Deep Learning. This study proposes IndoBERT-LSS (Latent Space Sampling), a deep learning framework designed to handle imbalanced short-text classification. Unlike traditional end-to-end architectures, our methodology decouples the feature extraction process from the class balancing mechanism.

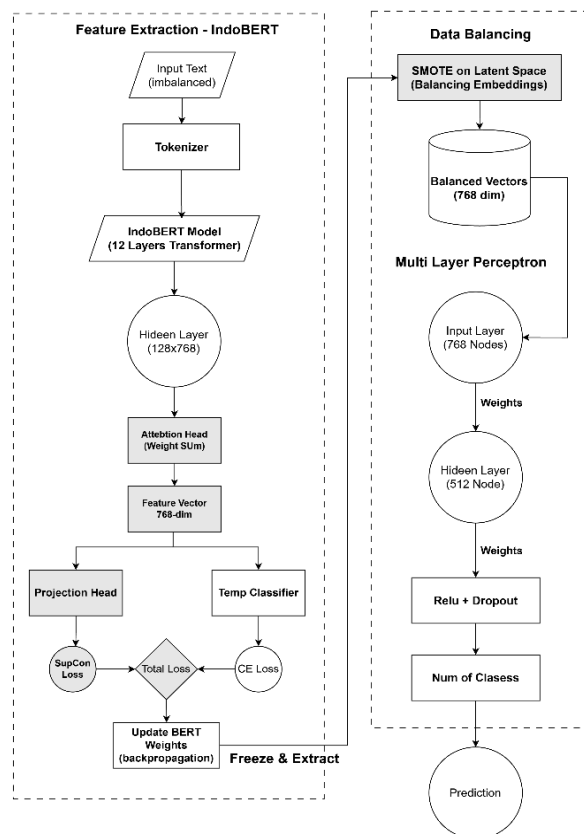


Figure 1. The proposed IndoBERT-LSS Architecture consisting of two decoupled phases.

The framework is executed in two distinct phases representation learning using Attention-based Supervised Contrastive Loss, and Latent Space Balancing using SMOTE followed by classification. The overall architecture is illustrated in Figure 1.

The proposed IndoBERT-LSS Architecture consisting of two decoupled phases. Phase 1 (left panel) focuses on Representation Learning, where the IndoBERT backbone is fine-tuned using an Attention-Pooling mechanism and Supervised Contrastive Loss to optimize feature clustering. Once trained, the backbone weights are frozen. Phase 2 (right panel) illustrates the Latent Space Sampling process, where extracted embeddings are balanced using SMOTE before being fed into the final Multi-Layer Perceptron (MLP) classifier for prediction.

2.1 Dataset Acquisition and Preprocessing

This study utilizes a primary dataset consisting of 1,051 textual responses gathered through a mixed-method approach, involving online surveys and structured interviews conducted in Pariaman City, West Sumatra. The respondents represent three key stakeholder groups: tourists, Micro, Small, and Medium Enterprises (MSMEs) owners, and government officials. Each response is labeled with one of four sentiment classes: Agree, Disagree, Undecided, and Neutral.

To ensure the reliability and validity of the ground truth, the sentiment labeling process was performed independently by two domain experts in tourism and public policy. Any discrepancies in labeling were resolved through a consensus discussion, resulting in a high inter-annotator agreement. Initial exploratory data analysis of the training set revealed a severe class imbalance prior to any data augmentation.

Prior to model training, the total dataset of 1,051 responses was split into a training set (80%, comprising 840 samples) and a testing set (20%, comprising 211 samples) using stratified sampling to maintain the original class proportions. It is a fundamental machine learning principle that data augmentation must strictly be applied only to the training set to prevent data leakage.

Table 1. Training Class Distribution Before Latent SMOTE

Sentiment Class	Number of Samples
Neutral	134
Undecided	223
Agree	244
Disagree	239
Total	840

As detailed in Table 1, initial exploratory data analysis revealed a severe class imbalance, with the Neutral and Undecided classes being significantly underrepresented compared to the polarized Agree and Disagree sentiments. To ensure the quality of input features for the model, a rigorous preprocessing pipeline was applied. The process began with case folding where all text was converted to lowercase to maintain uniformity. Next, noise removal was performed to eliminate non-

alphanumeric characters, URLs, and excessive punctuation using regular expressions (regex).

Following this, tokenization was executed using the indobert-base-uncased tokenizer, which is optimized for the Indonesian language, with a maximum sequence length of 128 tokens to capture sufficient context from the short-text responses. Finally, in the data splitting stage, the dataset was partitioned into a training set (80%) and a testing set (20%) using stratified sampling to preserve the original class distribution in both subsets.

2.2 Phase 1: Representation Learning (Feature Extraction)

The primary objective of this phase is to train a robust feature extractor capable of capturing subtle semantic nuances. We employ IndoBERT-base-uncased as the backbone encoder [52]. IndoBERT was explicitly chosen as the sole backbone encoder for this framework because it currently represents the state-of-the-art for Indonesian Natural Language Processing [53]. While multilingual models (such as mBERT or XLM-RoBERTa) are available, they often dilute language-specific nuances [54]. In contrast, IndoBERT was pre-trained extensively on a massive, dedicated corpus of Indonesian text (over 220 million words). This specialized pre-training makes it exceptionally superior at capturing the local linguistic nuances, informalities, and implicit cultural contexts inherently present in Indonesian user-generated feedback and survey responses, thereby providing the most robust foundational embeddings for the subsequent contrastive learning phase [55]. Unlike standard BERT implementations that rely solely on the [CLS] token for classification, we introduce a custom Attention Head Pooling mechanism to capture context-aware representations.

To achieve the optimal conditions for feature extraction, a rigorous hyperparameter tuning process was conducted empirically. We tested various configurations of learning rates ($1e^{-5}$, $2e^{-5}$, and $3e^{-5}$) and batch size (16 and 32). The best performance was achieved using the AdamW optimizer with a learning rate of $2e^{-5}$ with a weight decay of 0.01 and a batch size of 16, which provided stable gradient updates without exhausting GPU memory.

To stabilize early training, a linear learning rate scheduler with a 10% warmup proportion was applied, alongside gradient clipping (max norm of 1.0). Empirical trials indicated that training over 8 epochs was the optimal sweet spot; it was sufficient for the Supervised Contrastive Loss (with the temperature parameter set to $\tau = 0.07$) to converge and tightly cluster the embeddings without causing the model to overfit on the training data. Simultaneously, a temporary classifier computed the Cross-Entropy (CE) Loss. The total loss was computed as a weighted sum of the SupCon and CE loss. In our implementation, both

objectives were found to be equally critical during this phase. Therefore, the loss weight ratio (λ) between SupCon and CE was set to 1:1 (equal weighting).

Furthermore, for the final classification phase (Phase 3) using the SMOTE-balanced latent features, the Multi-Layer Perceptron (MLP) was trained using the Adam optimizer with a learning rate of $1e^{-3}$, a dropout rate of 0.2, and a larger batch size of 32 over 20 epochs, ensuring the classifier efficiently learned the unbiased decision boundaries."

Pretend that the sequence of hidden states from the final layer of IndoBERT is $H = \{h_1, h_2, \dots, h_L\}$. Each vector h_i is a dense representation of a word with a dimension of $d=768$. To determine which words are most important for sentiment analysis, we calculate an attention score (α_i) for each word vector (h_i) using Equations 1 and 2 [56].

$$\mu_i = \tanh(W_\omega h_i + b_\omega) \quad (1)$$

$$\alpha_i = \frac{\exp(\mu_i^T \cdot \mu_c)}{\sum_{j=1}^L \exp(\mu_j^T \cdot \mu_c)} \quad (2)$$

W_ω and b_ω are the learnable weights and bias of the projection layer. μ_c is a context vector that learns to recognize informative words (e.g., "agree", "bad"). $\tanh$ is the non-linear activation function and α_i represents the importance weight of the i -th word, where all weights sum to 1.

Finally, the sentence-level feature vector v [57] is obtained by computing the weighted sum of all word vectors as show in Equation 3 [58]:

$$v = \sum_{i=1}^L \alpha_i h_i \quad (3)$$

This results in a single vector v (size 768) that summarizes the entire sentence by focusing on the most relevant words. To optimize this representation, the feature vector is passed through a Projection Head (a non-linear MLP) and normalized. In order to enhance the structure of the feature space, we utilize Supervised Contrastive Loss (SupCon). The SupCon loss enforces compact clustering of embeddings belonging to the same class. Simultaneously, a temporary classifier computes the Cross-Entropy (CE) Loss. The total loss is a weighted sum of SupCon and CE loss, updated via backpropagation.

Unlike standard Cross-Entropy which only focuses on decision boundaries, SupCon explicitly pulls embeddings of the same class closer together while pushing apart embeddings from different classes [59], [60]. The loss function is defined in Equation 4 [61]

$$\mathcal{L}_{sup} = \sum_{i \in I} \frac{-1}{|P(i)|} \sum_{p \in P} \log \frac{\exp(z_i \cdot z_p / \tau)}{\sum_{a \in A(i) \setminus \{i\}} \exp(z_i \cdot z_a / \tau)} \quad (4)$$

z_i and z_p are the normalized feature vectors of the anchor and positive sample respectively, τ is the temperature parameter (set to 0.07) and $P(i)$ represents the set of positive samples in the batch. This optimization ensures that the learned embeddings form

compact, separable clusters, providing a high-quality latent space for the subsequent balancing phase.

2.3 Phase 2: Latent Space Sampling and Classification

In the second phase, the training objective shifts from representation learning to robust classification. To preserve the semantic clusters formed in Phase 1, the weights of the IndoBERT backbone and the Attention mechanism are frozen. We perform a "Freeze & Extract" operation, transforming the textual input space into a fixed latent feature space.

This leads to the Data Balancing stage (right panel of Figure 1). The core innovation of this framework lies in how it handles data imbalance. Standard approaches often apply augmentation on the raw text (e.g., synonym replacement), which risks altering the sentence's original meaning or damaging its grammatical structure.

To circumvent this, we employ Latent Space Sampling using the Synthetic Minority Over-sampling Technique (SMOTE) on these latent embeddings (feature vectors). To ensure algorithmic reproducibility, the SMOTE algorithm was configured using $k=5$ nearest neighbors and a fixed random state Unlike text augmentation, Latent SMOTE mathematically synthesizes new vectors for minority classes (Neutral, Undecided) by interpolating between existing vector neighbors. This results in a Balanced Feature Set where all classes have equal representation. Finally, these balanced vectors are fed into a Multi-Layer Perceptron (MLP) classifier as illustrate in Figure 2.

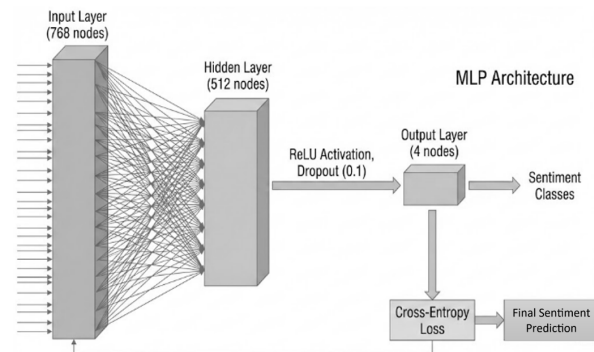


Figure 2. Visualization of the Latent Space Sampling process.

Figure 2 shows the MLP architecture consists of Input layer with 768 nodes (matching the embedding dimension), Hidden Layer with 512 nodes with ReLU activation and Dropout (0.1) to prevent overfitting. Output Layer is 4 nodes representing the sentiment classes. This classifier is trained using standard Cross-Entropy Loss to produce the final sentiment prediction are based on a fair and balanced representation of all sentiment categories.

Furthermore, to rigorously evaluate the effectiveness of the proposed IndoBERT-LSS framework, we established explicit baseline models for comparative analysis. The primary baseline is a standard, end-to-end fine-tuned IndoBERT model trained solely with

standard Cross-Entropy loss on the raw, imbalanced dataset (without SupCon and Latent SMOTE). Additionally, traditional deep learning architectures, specifically Stacked Bi-LSTM with Attention and Bi-GRU paired with FastText embeddings, were included as secondary baselines. Defining these models establishes the comparative foundation for the empirical evaluation discussed in the Results section.

2.4 Evaluation Metrics

To rigorously assess the predictive capability of the IndoBERT-LSS model, we employed a suite of standard classification metrics: Accuracy, Precision, Recall, and F1-Score [62], [63]. The evaluation is conducted exclusively on the held-out test set, which contains original, non-synthetic data to ensure an unbiased assessment of real-world performance [64], [65].

Given the imbalanced nature of the test distribution (where *Neutral* and *Undecided* classes are minorities), relying solely on Accuracy can be misleading, as it may hide poor performance on minority classes. Therefore, we prioritized the Macro-Average F1-Score as the primary metric for robustness.

The metrics are defined as follows, based on the Confusion Matrix elements such as True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) [66]. The confusion matrix's data can be used to compute a number of evaluation metrics. Accuracy measures the overall correctness of the model across all classes. Precision quantifies the model's exactness or reliability by answering the question of all instances predicted as a certain class, how many were actually that class with a high precision indicates a low False Positive rate. Recall (Sensitivity) measures the model's completeness by identifying how many of all actual instances of a class the model correctly identified with a high recall indicates a low False Negative rate, which is crucial for capturing minority opinions.

Furthermore, the F1-Score represents the harmonic mean of Precision and Recall. This metric is particularly vital for imbalanced datasets as it penalizes extreme values, meaning a high F1-Score is only achievable if both Precision and Recall are high. Lastly, the Macro-Average F1-Score is the arithmetic mean of the per-class F1-Scores. Unlike the Weighted Average, which is dominated by the majority class, the Macro-Average treats all classes equally, regardless of their support size.

3. Results and Discussions

This section presents the empirical evaluation of the IndoBERT-LSS framework. The experiments were designed to verify two primary hypotheses: first, that the integration of Attention-Pooling and Supervised Contrastive Loss (SupCon) creates a more discriminative feature space than standard methods; and second, that applying SMOTE on these latent features significantly improves classification performance on

minority classes. The results are analyzed sequentially, starting from the representation learning phase to the final classification metrics, followed by a critical discussion on the model's limitations.

3.1 Effectiveness of Representation Learning (Phase 1)

This phase specifically addresses the second gap identified in the introduction regarding the noise and lack of context in short-text responses. The Attention-Pooling mechanism allowed the model to focus on sentiment-bearing tokens, effectively reducing semantic ambiguity between closely related classes like 'Neutral' and 'Disagree'.

The training dynamics captured in Figure 3 provide strong empirical evidence for the stability of the proposed representation learning phase. As observed in the training logs, the IndoBERT backbone initialized with a high loss of 4.1276 in the first epoch. Throughout the 8-epoch cycle, the loss trajectory exhibited a sharp and consistent decline without divergence, reaching a final value of 1.4138. This substantial loss reduction of approximately 66% quantitatively confirms that the encoder successfully adapted to the contrastive objective, efficiently mapping the input text into a lower-dimensional manifold where semantic similarities are maximized.

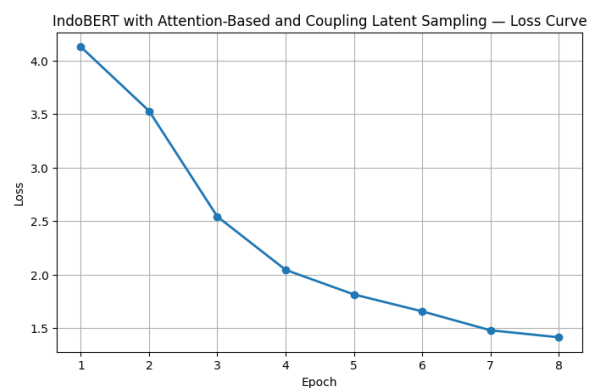


Figure 3. Training log of Phase 1

Figure 3 show training log of Phase 1 showing the rapid convergence of SupCon Loss from 4.12 to 1.41 over 8 epochs. The first key finding relates to the quality of the feature extraction process in Phase 1. During the training of the IndoBERT backbone using the hybrid SupCon loss, we observed a rapid and stable convergence. Within just 8 epochs, the training loss decreased significantly, indicating that the model successfully learned to minimize the intra-class variance. This stability is attributed to the Attention Head Pooling mechanism, which allowed the model to dynamically focus on sentiment-bearing tokens (e.g., "agree", "doubt") while suppressing noise from stop words.

Qualitatively, this phase succeeded in transforming the disjointed raw text into a structured vector space. Unlike standard Cross-Entropy training which allows decision boundaries to be ambiguous, the SupCon loss forced the embeddings of the same sentiment class to

form compact, distinct clusters. This geometric separation in the latent space is the critical factor that enabled the subsequent SMOTE algorithm to work effectively. By ensuring that the "Neutral" and "Undecided" samples were tightly clustered, the synthetic vectors generated by SMOTE in Phase 2 were guaranteed to be semantically consistent, avoiding the generation of noisy or "hallucinated" features often seen in text-level augmentation.

3.2 Classification Robustness and Minority Class Performance

The efficacy of the proposed two-stage framework is corroborated by the training logs captured during the experiment. The training process exhibits three distinct phases that collectively contribute to the model's final robustness. First, in the Representation Learning phase (Phase 1), the IndoBERT backbone demonstrated stable convergence. The training loss decreased consistently from an initial value of 4.1276 in Epoch 1 to 1.4138 in Epoch 8. This steady decline indicates that the Attention-Pooling mechanism and Supervised Contrastive Loss successfully optimized the latent space, pulling semantically similar samples into compact clusters even before any oversampling was applied in Figure 4.

```
[PHASE 2] Embedding Extraction & SMOTE Application...
Extracting Embeddings: 100% | 53/53 [04:46:00:00, 5.40s/it]
Extracting Embeddings: 100% | 14/14 [01:12:00:00, 5.15s/it]
Dimensions before SMOTE: (840, 768)
Class distribution before SMOTE: [134 223 244 239]
Dimensions after SMOTE: (976, 768)
Class distribution after SMOTE: [244 244 244 244]
(All classes are now balanced!)

[PHASE 3] Training the Final Classifier on SMOTE Data...
Classifier Epoch 5/20, Loss: 0.0004
Classifier Epoch 10/20, Loss: 0.0002
Classifier Epoch 15/20, Loss: 0.0001
Classifier Epoch 20/20, Loss: 0.0000
```

Figure 4. Log SMOTE Embedding Classifier

By applying SMOTE in the latent space rather than on raw text, we successfully addressed the third research gap concerning the degradation of semantic integrity often caused by traditional text-level augmentation. The logs in Phase 2 confirm that the synthetic features generated were mathematically consistent with the learned semantic clusters, ensuring high-quality data balancing without introducing linguistic noise.

Second, the logs in Phase 2 provide a quantitative confirmation of the data imbalance problem and its resolution. As shown in the extraction logs, the original distribution of the training set was highly skewed: the minority class (Class 0 / 'Neutral') contained only 134 samples, whereas the majority classes (Class 2 and 3) contained over 240 samples. This severe imbalance (ratio ~1:1.8) explains why traditional models often fail to recognize neutral sentiments. The application of Latent Space SMOTE effectively rectified this disparity. The logs show the total training samples increasing from 840 to 976, resulting in a perfectly balanced distribution of [244, 244, 244, 244] across all four classes. This confirms that our sampling strategy

successfully synthesized 110 high-quality minority features for Class 0 without altering the textual input.

Third, the Classification Phase demonstrates the benefits of learning from this balanced feature set. The MLP classifier achieved rapid convergence, with the loss reaching a near-zero value (0.0000) by Epoch 20. This indicates that the balanced latent features produced in the previous steps were highly discriminative and linearly separable, allowing the classifier to learn unbiased decision boundaries with high precision.

Following the successful stabilization of the feature space and the rectification of class imbalance during the training phases, the final model was evaluated on the held-out test set to determine its generalization capability. This evaluation provides the ultimate validation of the proposed two-stage framework, ensuring that the synthetic balancing in the latent space translates into actual predictive robustness on real-world data.

The most significant finding of this study is the model's superior performance on the imbalanced test set. As shown in Table 2, the proposed IndoBERT-LSS achieved a global accuracy of 93.84%. However, the true strength of the model is revealed in the class-wise performance metrics. Traditional models typically exhibit a "minority collapse," where performance on underrepresented classes (like 'Neutral') drops significantly.

Table 2. Performance Metrics of IndoBERT-LSS on Test Data

Sentiment Class	Precision	Recall	F1-score	Support
Neutral	1.00	0.97	0.99	34
Undecided	0.88	0.89	0.88	56
Agree	1.00	1.00	1.00	61
Disagree	0.90	0.90	0.90	60
Overall Accuracy			93.84	211

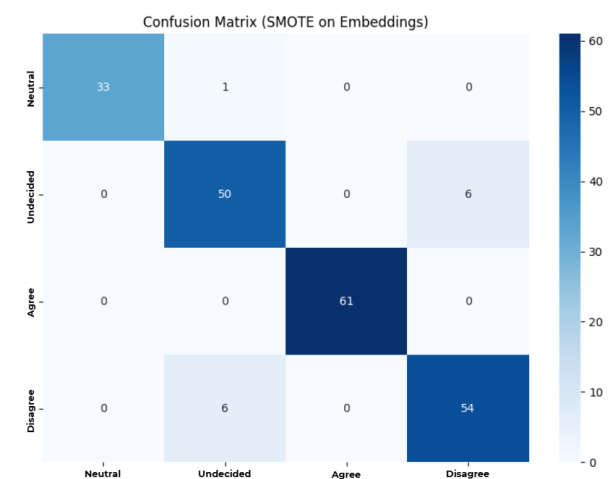


Figure 5. Confusion Matrix

In contrast, our model achieved a near-perfect F1-Score of 0.99 for the 'Neutral' class and 0.88 for the 'Undecided' class. This result empirically validates the Latent Space Sampling strategy. By balancing the data in the vector space, the final MLP classifier was able to

learn unbiased decision boundaries. The Confusion Matrix further corroborates this, as shown in Figure 5.

Confusion Matrix on the test set showing robust performance across all classes. The misclassifications were minimal and largely confined to the boundary between 'Undecided' and 'Disagree'. Specifically, the model misclassified 6 'Undecided' samples as 'Disagree'. This error is linguistically justifiable, as expressions of doubt in Indonesian culture often carry implicit disagreement (polite refusal), creating a natural semantic overlap. Crucially, there was zero confusion between the contradictory classes of 'Agree' and 'Disagree', demonstrating the model's high discriminative power.

3.3 Comparison with Baseline Models and Previous Studies

To rigorously demonstrate the contribution and novelty of the proposed IndoBERT-LSS framework, we conducted two levels of comparative analysis, an internal baseline comparison using the identical Halal tourism dataset, and an external comparison with recent state-of-the-art studies. First, to validate the architectural superiority of our approach, we benchmarked IndoBERT-LSS against several traditional deep learning architectures and a standard fine-tuned IndoBERT model using the exact same dataset and testing conditions.

Table 3. Baseline Comparison on the Same Dataset

Model	Accuracy	Classification	Precision	Recall	F1-Score
Stacked Bi-LSTM + Attention	0.893	Neutral	0.91	0.97	0.94
		Undecided	0.92	0.87	0.90
		Agree	0.92	0.90	0.91
		Disagree	0.84	0.86	0.85
BI-GRU + FastText + Attention	0.868	Neutral	0.79	1.00	0.89
		Undecided	0.91	0.78	0.84
		Agree	0.89	0.95	0.92
		Disagree	0.85	0.82	0.82
BI-GRU + Fine-Tuning Embedding	0.8732	Neutral	0.84	1.00	0.91
		Undecided	0.90	0.81	0.85
		Agree	0.89	0.95	0.92
		Disagree	0.85	0.78	0.81
Fine-Tuning + IndoBERT	0.927	Neutral	0.94	1.00	0.97
		Undecided	0.96	0.81	0.88
		Agree	0.92	1.00	0.96
		Disagree	0.90	0.92	0.91
IndoBERT- LSS (Our Proposed Model)	0.938	Neutral	1.00	0.97	0.99
		Undecided	0.88	0.89	0.88
		Agree	1.00	1.00	1.00
		Disagree	0.90	0.90	0.90

As shown in Table 3 traditional recurrent networks (Stacked Bi-LSTM and Bi-GRU combined with FastText or Fine-Tuned Embeddings) achieved accuracies ranging between 86.8% and 89.3% [67]. Upgrading to a standard Fine-Tuning IndoBERT model significantly improved the accuracy to 92.7%. However, the proposed IndoBERT-LSS model outperformed all baselines, achieving the highest global accuracy of 93.8%. This explicitly represents a 1.14% overall accuracy improvement over the standard IndoBERT baseline. The key contribution of our model lies in its precision with minority classes while standard IndoBERT struggles with semantic overlap, our decoupled approach using Latent Space Sampling pushed the F1-Score for the critical 'Neutral' class to a near-perfect 0.99.

Second, to contextualize our contribution within the broader literature, we compared our results with recent sentiment analysis models proposed by other researchers, as summarized in Table 4.

While direct dataset-to-dataset comparisons are challenging due to different domains, comparing the methodological achievements reveals a clear progression. Previous studies utilizing SMOTE with traditional deep learning, such as Anam et al. (SMOTE

+ GRU + BiLSTM) and Putra et al. (Meta-learners), plateaued at accuracies of 89.00% and 88.00%, respectively, primarily because text-level SMOTE often introduces linguistic noise.

Table 4. Comparison with Previous State-of-the-Art Research

Researcher	Model	Dataset	Accuracy
Bharti et al. [37]	CNN+ BiGRU+ SVM	ISEAR, WASSA, and Emotion-stimulus	80.11%
Anam et al. [68]	SMOTE+ GRU+ BiLSTM	Twitter	89.00%
Putra et al. [69]	XGBoost + RF + LSTM meta learner	Kaggle	88.00%
Owen et al. [70]	SBERT+ TextConvonet	Customized Data e-wallet service providers	86.90%
This Research	IndoBERT + Latent Space Sampling	Customized Survey on Tourism	93.84%

Furthermore, advanced models like SBERT applied to custom data by Owen et al. achieved 86.90%. In contrast, our research demonstrates that moving the sampling process into the latent space (Latent SMOTE) combined with Attention-based IndoBERT breaks this plateau, achieving an exceptional accuracy of 93.84%.

This confirms the novelty of our framework as a highly effective solution for customized, imbalanced textual data.

3.4 Limitations and Future Work

The near-perfect F1-score of 0.99 for the 'Neutral' class empirically resolves the first gap regarding 'minority collapse' in imbalanced datasets. This result demonstrates that decoupling representation learning from balancing allows the model to capture the nuanced voices of minority stakeholders which are typically ignored by standard fine-tuning methods.

Despite the promising results, this study has several limitations that warrant further investigation. First, the two-stage nature of the framework introduces higher computational complexity compared to end-to-end fine-tuning. The process requires freezing weights and extracting features for the entire dataset before training the classifier, which may be less efficient for real-time applications with streaming data. Second, the dataset used is domain-specific to Halal tourism in a specific region (Pariaman). While the linguistic features are representative of Indonesian sentiment, the model's generalization capability to other domains (e.g., political elections or product reviews) has not been explicitly tested.

Future research should focus on optimizing the computational pipeline, potentially by integrating the sampling mechanism into an end-to-end differentiable process. Additionally, expanding the dataset to cover multi-domain sentiments would provide a more comprehensive validation of the Latent Space Sampling technique. We also recommend exploring the application of this framework to other low-resource languages where labeled data is scarce and class imbalance is prevalent.

4. Conclusions

This study addresses the critical challenge of analyzing public sentiment from imbalanced, short-text data regarding Halal tourism policies in Indonesia. By proposing the IndoBERT-LSS (Latent Space Sampling) framework, we successfully demonstrated that decoupling representation learning from class balancing yields superior predictive robustness compared to traditional end-to-end fine-tuning.

The primary contribution of this work lies in the novel integration of an Attention-Pooling mechanism with Supervised Contrastive Loss (SupCon) to rigorously structure the feature space, followed by the application of SMOTE on latent embeddings rather than raw text. This two-stage approach effectively neutralized the risk of semantic drift and linguistic noise typically associated with text-level augmentation. Empirically, the model achieved a state-of-the-art accuracy of 93.84% on the held-out test set.

More importantly, it successfully overcame the "minority collapse" phenomenon observed in standard

baselines, attaining a near-perfect F1-score of 0.99 for the underrepresented 'Neutral' class. These findings directly answer the research objectives by confirming that optimizing the geometric structure of embeddings is a prerequisite for effective synthetic sampling in NLP.

Theoretically, this study extends the application of contrastive learning to low-resource languages, offering a reproducible blueprint for handling severe class imbalance without compromising feature quality. Practically, the model provides the Pariaman local government with a high-fidelity analytical tool to capture nuanced stakeholder feedback including hesitant or neutral voices that are often marginalized thereby facilitating more inclusive and data-driven policymaking. While the current decoupled pipeline introduces a slight computational overhead due to the feature extraction step, it offers a necessary trade-off for accuracy. Future work will focus on optimizing this pipeline into an end-to-end differentiable architecture and testing its generalization capabilities across broader multi-domain datasets to further enhance its scalability for real-time social sensing systems.

Acknowledgements

The authors thank the Government of Pariaman City for their help in implementing this research, as well as all those who provided useful insights and assistance during the writing process. The authors state that there is no conflict of interest.

References

- [1] J. Chen, J. Wu, D. Wang, and B. Stantic, "Beyond static rankings: A tourist experience-driven approach to measure destination competitiveness," *Tour. Manag.*, vol. 106, no. August 2024, p. 105022, 2025, doi: [10.1016/j.tourman.2024.105022](https://doi.org/10.1016/j.tourman.2024.105022).
- [2] A. Ariyanto, M. Adam, S. Musnadi, and M. A. Djalil, "Empowering and Influencing Factors of Intellectual and Spiritual Capital in Improving Organizational Performance at Tourism Institutions," *J. Ecohumanism*, vol. 3, no. 7, pp. 5074–5085, 2024, doi: [10.62754/joe.v3i7.4616](https://doi.org/10.62754/joe.v3i7.4616).
- [3] L. Alfarajat, "Halal Food and Medical Tourism: Muslim Patients' Experiences and Satisfaction in South Korea," *J. Relig. Health*, vol. 63, no. 5, pp. 3274–3290, 2022, doi: [10.1007/s10943-022-01727-x](https://doi.org/10.1007/s10943-022-01727-x).
- [4] Repelita, N. Iskandar, and Mursal, "Halal Without Label: Implementation of Maqāsid Al-Sharī'ah in Community-Based Tourism in Kerinci, Indonesia," *Malaysian J. Syariah Law*, vol. 12, no. 1, pp. 192–205, 2024, doi: [10.33102/mjssl.vol12no1.477](https://doi.org/10.33102/mjssl.vol12no1.477).
- [5] A. Holder, L. Ruhanen, G. Walters, and M. Mkono, "'I think ... I feel ...': using projective techniques to explore socio-cultural aversions towards Indigenous tourism," *Tour. Manag.*, vol. 98, no. January 2022, p. 104778, 2023, doi: [10.1016/j.tourman.2023.104778](https://doi.org/10.1016/j.tourman.2023.104778).
- [6] S. J. Page and M. Duignan, "Progress in Tourism Management: Is urban tourism a paradoxical research domain? Progress since 2011 and prospects for the future," *Tour. Manag.*, vol. 98, no. March, p. 104737, 2023, doi: [10.1016/j.tourman.2023.104737](https://doi.org/10.1016/j.tourman.2023.104737).
- [7] I. M. Permana and F. Adam, "Impact of Islamic attributes and destination image on tourist satisfaction and revisit intentions in halal tourism: A study in Banda Aceh," *Int. J. Adv. Appl. Sci.*, vol. 11, no. 4, pp. 118–127, 2024, doi: [10.21833/ijaas.2024.04.014](https://doi.org/10.21833/ijaas.2024.04.014).

- [8] S. Pongsakornrungrungsilp, P. Pongsakornrungrungsilp, T. Chouyakaew, H. Niyomdech, S. Chansamran, and Y. He, "Integrating Sustainability and Cultural Sensitivity: Clustering Muslim Tourist Lifestyles in the Andaman Coastal Cluster, Thailand," *Sustain.*, vol. 16, no. 21, pp. 1–15, 2024, doi: [10.3390/su16219542](https://doi.org/10.3390/su16219542).
- [9] Syufa'at, E. Z. Mutaqin, A. Dharin, and Mutholaah, "Implementation of Chse Certificate for Tourism Object Management in Banyumas (Halal Tourism Perspective)," *Rev. Gest. Soc. e Ambient.*, vol. 18, no. 9, pp. 1–18, 2024, doi: [10.24857/rgsa.v18n9-069](https://doi.org/10.24857/rgsa.v18n9-069).
- [10] D. Rusyaida, "Local wisdom as Halal tourism model in West Sumatera: an 'Urf perspective," *Indones. J. Islam Muslim Soc.*, vol. 14, no. 2, pp. 259–287, 2024, doi: [10.18326/ijims.v14i2.259-287](https://doi.org/10.18326/ijims.v14i2.259-287).
- [11] R. Polus, R. Nautiyal, A. Nautiyal, and P. Kari, "Caught in the maze : Exploration of the ' tourist trap ' phenomenon," *Tour. Manag.*, vol. 106, no. June 2024, p. 105009, 2025, doi: [10.1016/j.tourman.2024.105009](https://doi.org/10.1016/j.tourman.2024.105009).
- [12] L. Li, "Insights into People ' s Perceptions Towards Urban Public Spaces Through Analysis of Social Media Reviews : A Case Study of Shanghai †," pp. 1–19, 2025, doi: [10.3390/buildings15173033](https://doi.org/10.3390/buildings15173033).
- [13] A. Boukis, D. Caruelle, and L. Harris, "Dealing with travel influencers : Opportunity or threat ? ☆," *Tour. Manag.*, vol. 108, no. March 2024, p. 105112, 2025, doi: [10.1016/j.tourman.2024.105112](https://doi.org/10.1016/j.tourman.2024.105112).
- [14] L. Xie-carson, T. Magor, P. Benckendorff, and K. Hughes, "All hype or the real deal? Investigating user engagement with virtual influencers in tourism," *Tour. Manag.*, vol. 99, no. April, p. 104779, 2023, doi: [10.1016/j.tourman.2023.104779](https://doi.org/10.1016/j.tourman.2023.104779).
- [15] Sunarmo, M. Thurisaina, A. Amri, and A. Sagita, "Effects of Halal Awareness, Service Quality and Promotion on Visitor Decisions Through Destination Imagery as a Mediator," *Indones. J. Halal Res.*, vol. 6, no. 1, pp. 13–25, 2024, doi: [10.15575/ijhar.v6i1.19489](https://doi.org/10.15575/ijhar.v6i1.19489).
- [16] A. Sutono, Sumaryadi, W. Rahtomo, F. F. Puksi, and O. Permana, "Re-modelling Muslim-friendly tourism value chain ecosystem: the case study of West Sumatera, Indonesia," *Cogent Soc. Sci.*, vol. 10, no. 1, p., 2024, doi: [10.1080/23311886.2024.2361536](https://doi.org/10.1080/23311886.2024.2361536).
- [17] V. Arefieva, R. Egger, and J. Yu, "A machine learning approach to cluster destination image on Instagram," *Tour. Manag.*, vol. 85, no. March, p. 104318, 2021, doi: [10.1016/j.tourman.2021.104318](https://doi.org/10.1016/j.tourman.2021.104318).
- [18] G. Hughes, S. E. Shaw, J. Bowley, J. L. Darbyshire, and T. J. Stephens, "Choosing wisely? A frame analysis of the evolution and adoption of shared decision making in the UK," *SSM - Qual. Res. Heal.*, vol. 8, no. December 2024, p. 100579, 2025, doi: [10.1016/j.ssmqr.2025.100579](https://doi.org/10.1016/j.ssmqr.2025.100579).
- [19] K. Miazek and K. Bocian, "When AI is fairer than humans: The role of egocentrism in moral and fairness judgments of AI and human decisions," *Comput. Hum. Behav. Reports*, vol. 19, no. May, p. 100719, 2025, doi: [10.1016/j.chbr.2025.100719](https://doi.org/10.1016/j.chbr.2025.100719).
- [20] J. Cuadrado, E. Martinez, J. C. Martinez-Santos, and E. Puertas, "The media framing dataset: Analyzing news narratives in Mexico and Colombia," *Data Br.*, vol. 58, p. 111284, 2025, doi: [10.1016/j.dib.2025.111284](https://doi.org/10.1016/j.dib.2025.111284).
- [21] V. A. Cordero-Vinueza, T. van Dijk, C. Lamker, and F. Niekerk, "Making child-friendly cities: A quest for local agenda-setting," *Cities*, vol. 161, no. March, p. 105882, 2025, doi: [10.1016/j.cities.2025.105882](https://doi.org/10.1016/j.cities.2025.105882).
- [22] F. M. Suud, K. S. Rouzi, M. Al-Fairusy, T. Na'imah, H. Harinawati, and R. B. Yasa, "Tourism Education Development on the Coastline: Challenges Empirical Study on Islamic Law Province," *J. Ilm. Peuradeun*, vol. 12, no. 3, pp. 1289–1310, 2024, doi: [10.26811/peuradeun.v12i3.1172](https://doi.org/10.26811/peuradeun.v12i3.1172).
- [23] T. Pang, J. Liu, L. Han, H. Liu, and D. Yan, "A Deep Learning-Based Analysis of Customer Concerns and Satisfaction: Enhancing Sustainable Practices in Luxury Hotels," *Sustain.*, vol. 17, no. 8, 2025, doi: [10.3390/su17083603](https://doi.org/10.3390/su17083603).
- [24] D. Erdoğan et al., "Developing a Deep Learning-Based Sentiment Analysis System of Hotel Customer Reviews for Sustainable Tourism," *Sustain.*, vol. 17, no. 13, 2025, doi: [10.3390/su17135756](https://doi.org/10.3390/su17135756).
- [25] Y.-W. Chen and S.-C. Pei, "Always Clear Days: Degradation Type and Severity Aware All-In-One Adverse Weather Removal," *IEEE Access*, vol. 13, no. January, pp. 7650–7662, 2023, doi: [10.1109/ACCESS.2025.3526168](https://doi.org/10.1109/ACCESS.2025.3526168).
- [26] F. Marchesani and G. Testa, "Understanding municipalities and visitors tourism perspectives : An AI-driven analysis of Instagram posts in Italian and French villages," *Tour. Manag.*, vol. 112, no. August 2025, p. 105279, 2026, doi: [10.1016/j.tourman.2025.105279](https://doi.org/10.1016/j.tourman.2025.105279).
- [27] A. S. Ekakristi, A. F. Wicaksono, and R. Mahendra, "Intermediate-task transfer learning for Indonesian NLP tasks," *Nat. Lang. Process. J.*, vol. 12, no. May, p. 100161, 2025, doi: [10.1016/j.nlp.2025.100161](https://doi.org/10.1016/j.nlp.2025.100161).
- [28] G. Kałek, R. Kozik, A. Pawlicka, M. Pawlicki, and M. Choraś, "In depth analysis for securing the truth: Addressing the fake news challenge with graph neural networks," *Neurocomputing*, vol. 654, no. June, p. 131327, 2025, doi: [10.1016/j.neucom.2025.131327](https://doi.org/10.1016/j.neucom.2025.131327).
- [29] K. Mnassri, R. Farahbakhsh, and N. Crespi, "Multilingual Hate Speech Detection: A Semi-Supervised Generative Adversarial Approach," *Entropy*, vol. 26, no. 4, 2024, doi: [10.3390/e26040344](https://doi.org/10.3390/e26040344).
- [30] G. Z. Nabilah, S. Y. Prasetyo, Z. N. Izdihar, and A. S. Girsang, "BERT base model for toxic comment analysis on Indonesian social media," *Procedia Comput. Sci.*, vol. 216, no. 2022, pp. 714–721, 2022, doi: [10.1016/j.procs.2022.12.188](https://doi.org/10.1016/j.procs.2022.12.188).
- [31] A. Jazuli, Widowati, and R. Kusumaningrum, "Optimizing Aspect-Based Sentiment Analysis Using BERT for Comprehensive Analysis of Indonesian Student Feedback," *Appl. Sci.*, vol. 15, no. 1, pp. 1–28, 2025, doi: [10.3390/app15010172](https://doi.org/10.3390/app15010172).
- [32] S. Kang, "LLM-Guided Weighted Contrastive Learning with Topic-Aware Masking for Efficient Domain Adaptation: A Case Study on Pulp-Era Science Fiction," *Electron. MDPI*, vol. 14, no. 17, pp. 1–24, 2025, doi: [10.3390/electronics14173351](https://doi.org/10.3390/electronics14173351).
- [33] R. Pratiwi et al., "The Collaboration of Penta Helix to Develop Halal Tourism Villages in Batang, Cental Java," *Int. J. Sustain. Dev. Plan.*, vol. 19, no. 7, pp. 2753–2761, 2024, doi: [10.18280/ijstdp.190732](https://doi.org/10.18280/ijstdp.190732).
- [34] J. Amarta, D. Tyogo, D. Utomo, and B. Ignasio, "ScienceDirect ScienceDirect Deep Learning And Machine Learning For Indonesian Online Gambling Website Promotion Detection," *Procedia Comput. Sci.*, vol. 269, pp. 979–992, 2025, doi: [10.1016/j.procs.2025.09.040](https://doi.org/10.1016/j.procs.2025.09.040).
- [35] G. Z. Nabilah, S. Y. Prasetyo, Z. N. Izdihar, and A. S. Girsang, "ScienceDirect ScienceDirect 7th International Conference on Computer Science and Computational Intelligence 2022 BERT base model for toxic comment analysis on BERT base model for toxic comment analysis on Indonesian social media Indonesian social media," *Procedia Comput. Sci.*, vol. 216, no. 2022, pp. 714–721, 2023, doi: [10.1016/j.procs.2022.12.188](https://doi.org/10.1016/j.procs.2022.12.188).
- [36] F. Dornaika, Z. Ibrahim, J. Charafeddine, and A. Bosaghzadeh, "Towards dynamic self-training for scalable semi-supervised learning on graphs," *Neurocomputing*, vol. 654, no. August, p. 131223, 2025, doi: [10.1016/j.neucom.2025.131223](https://doi.org/10.1016/j.neucom.2025.131223).
- [37] S. K. Bharti et al., "Text-Based Emotion Recognition Using Deep Learning Approach," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: [10.1155/2022/2645381](https://doi.org/10.1155/2022/2645381).
- [38] N. R. P. Syallya, A. A. Pravitarsari, and A. Helen, "NLP-Based Intent Classification Model for Academic Curriculum Chatbots in Universities Study Programs," *J. RESTI*, vol. 9, no. 1, pp. 111–117, 2025, doi: [10.29207/resti.v9i1.6276](https://doi.org/10.29207/resti.v9i1.6276).
- [39] S. Dehghan and M. F. Amasyali, "SupMPN: Supervised Multiple Positives and Negatives Contrastive Learning Model for Semantic Textual Similarity," *Appl. Sci.*, vol. 12, no. 19, 2022, doi: [10.3390/app12199659](https://doi.org/10.3390/app12199659).
- [40] X. Huang et al., "Optimizing sleep staging on multimodal time series: Leveraging borderline synthetic minority oversampling technique and supervised convolutional contrastive learning," *Comput. Biol. Med.*, vol. 166, no. August, p. 107501, 2023, doi: [10.1016/j.compbiomed.2023.107501](https://doi.org/10.1016/j.compbiomed.2023.107501).
- [41] D. Gonzalez-Cuautle et al., "Synthetic minority oversampling technique for optimizing classification tasks in botnet and intrusion-detection-system datasets," *Appl. Sci.*, vol. 10, no. 3, 2020, doi: [10.3390/app10030794](https://doi.org/10.3390/app10030794).
- [42] A. Kummer, T. Ruppert, T. Medvegy, and J. Abonyi, "Machine learning-based software sensors for machine state monitoring

- The role of SMOTE-based data augmentation,” *Results Eng.*, vol. 16, no. November, 2022, doi: [10.1016/j.rineng.2022.100778](https://doi.org/10.1016/j.rineng.2022.100778).
- [43] M. Yusoff, Y. Mahmud, P. A. R. Azmi, and M. T. M. Sallehud-din, “The improvement of SMOTE-ENN-XGBoost through Yeo Johnson strategy on Dissolved Gas Analysis dataset,” *Energy Reports*, vol. 13, no. April, pp. 6281–6290, 2025, doi: [10.1016/j.egyr.2025.05.013](https://doi.org/10.1016/j.egyr.2025.05.013).
- [44] M. Ali, “Classification of imbalanced travel mode choice dataset with SMOTE and prediction using interpretable machine learning,” *Sustain. Futur.*, vol. 10, no. August, p. 101119, 2025, doi: [10.1016/j.sfr.2025.101119](https://doi.org/10.1016/j.sfr.2025.101119).
- [45] I. Aruleba and Y. Sun, “Enhanced credit risk prediction using deep learning and SMOTE-ENN resampling,” *Mach. Learn. with Appl.*, vol. 21, no. April, p. 100692, 2025, doi: [10.1016/j.mlwa.2025.100692](https://doi.org/10.1016/j.mlwa.2025.100692).
- [46] H. Jakha, S. El Houssaini, M.-A. El Houssaini, S. Ajjaj, and A. Hadir, “Optimizing Sentiment Analysis in Multilingual Balanced Datasets: A New Comparative Approach to Enhancing Feature Extraction Performance with ML and DL Classifiers,” *Appl. Syst. Innov.*, vol. 8, no. 4, p. 104, 2025, doi: [10.3390/asi8040104](https://doi.org/10.3390/asi8040104).
- [47] Z. He and H. Wang, “Fake News Detection Based on Contrastive Learning and Cross-Modal Interaction,” *Electron. MDPI*, vol. 14, no. 7, pp. 1–24, 2025, doi: [10.3390/electronics14173351](https://doi.org/10.3390/electronics14173351).
- [48] H. Honda and M. Hagiwara, “Context-dependent neuro-symbolic AI through self-supervised learning with large language models,” *Neurocomputing*, vol. 654, no. August, p. 131269, 2025, doi: [10.1016/j.neucom.2025.131269](https://doi.org/10.1016/j.neucom.2025.131269).
- [49] E. Hashmi, S. Y. Yayilgan, M. M. Yamin, M. Abomhara, and M. Ullah, “Self-supervised hate speech detection in Norwegian texts with lexical and semantic augmentations,” *Expert Syst. Appl.*, vol. 264, no. April 2024, p. 125843, 2025, doi: [10.1016/j.eswa.2024.125843](https://doi.org/10.1016/j.eswa.2024.125843).
- [50] J. W. Bi, Y. Liu, Z. P. Fan, and J. Zhang, “Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews,” *Tour. Manag.*, vol. 70, no. March 2018, pp. 460–478, 2019, doi: [10.1016/j.tourman.2018.09.010](https://doi.org/10.1016/j.tourman.2018.09.010).
- [51] L. C. Cheng, K. Chen, M. C. Lee, and K. M. Li, “User-Defined SWOT analysis – A change mining perspective on user-generated content,” *Inf. Process. Manag.*, vol. 58, no. 5, p. 102613, 2021, doi: [10.1016/j.ipm.2021.102613](https://doi.org/10.1016/j.ipm.2021.102613).
- [52] L. Anthony, “Concordancing with AI: Applications of word and sentence embeddings,” *Appl. Corpus Linguist.*, vol. 5, no. 3, p. 100164, 2025, doi: [10.1016/j.acorp.2025.100164](https://doi.org/10.1016/j.acorp.2025.100164).
- [53] L. Kong, W. Yang, and F. Wei, “Implicit Emotion Analysis Based on Improved Supervised Contrastive Learning,” *Electron.*, vol. 12, no. 10, pp. 1–12, 2023, doi: [10.3390/electronics12102172](https://doi.org/10.3390/electronics12102172).
- [54] Q. Wang, W. Zhang, T. Lei, and D. Peng, “Grouped Contrastive Learning of Self-Supervised Sentence Representation,” *Appl. Sci.*, vol. 13, no. 17, 2023, doi: [10.3390/app13179873](https://doi.org/10.3390/app13179873).
- [55] H. Xiao, F. Dornaika, J. Charafeddine, and J. Bi, “Metric learning-enhanced semi-supervised Graph Convolutional Network for multi-view learning,” *Inf. Fusion*, vol. 124, no. October 2024, p. 103420, 2025, doi: [10.1016/j.inffus.2025.103420](https://doi.org/10.1016/j.inffus.2025.103420).
- [56] C. Oham, S. Nair, S. Bhattasali, and P. Resnik, “A predictive coding model for online sentence processing S,” *J. Mem. Lang.*, vol. 146, no. August 2025, p. 104705, 2026, doi: [10.1016/j.jml.2025.104705](https://doi.org/10.1016/j.jml.2025.104705).
- [57] B. Zhu and W. Pan, “Heliyon Chinese text classification method based on sentence information enhancement and feature fusion,” *Heliyon*, vol. 10, no. 17, p. e36861, 2024, doi: [10.1016/j.heliyon.2024.e36861](https://doi.org/10.1016/j.heliyon.2024.e36861).
- [58] O. Wallaart, F. Frascar, and F. Van Der Knaap, “Knowledge-Based Systems Weakly-supervised sentence-based aspect category and sentiment classification,” vol. 317, no. March, 2025, doi: [10.1016/j.knosys.2025.113403](https://doi.org/10.1016/j.knosys.2025.113403).
- [59] S. Begum, M. U. Ahmed, and S. Barua, “Intelligent Systems with Applications Attention-based fuzzy neural networks for self-supervised data annotation Md Rakibul Islam,” *Intell. Syst. with Appl.*, vol. 28, no. November, p. 200610, 2025, doi: [10.1016/j.iswa.2025.200610](https://doi.org/10.1016/j.iswa.2025.200610).
- [60] T. Wu, M. Wang, Y. Xi, and Z. Zhao, “Neurocomputing Intent recognition model based on sequential information and sentence features,” *Neurocomputing*, vol. 566, no. June 2023, p. 127054, 2024, doi: [10.1016/j.neucom.2023.127054](https://doi.org/10.1016/j.neucom.2023.127054).
- [61] K. Bannigan, “Results in Engineering EnCo SupCon : Entropy-driven supervised contrastive learning for discriminative feature representations from remote sensing imagery,” *Results Eng.*, vol. 29, no. November 2025, p. 108311, 2026, doi: [10.1016/j.rineng.2025.108311](https://doi.org/10.1016/j.rineng.2025.108311).
- [62] L. Alexander, R. Gian, A. Ananta, J. Harefa, L. Alexander, and K. Jingga, “ScienceDirect ScienceDirect Evaluating IndoBERT for Indonesian Hoax News Detection : A Evaluating IndoBERT for Ensemble Indonesian Hoax News Detection : Comparative Study with and Models Comparative Study with Ensemble and CNN-LSTM Models,” *Procedia Comput. Sci.*, vol. 269, pp. 1625–1633, 2025, doi: [10.1016/j.procs.2025.09.105](https://doi.org/10.1016/j.procs.2025.09.105).
- [63] W. Christian et al., “Leveraging Leveraging IndoBERT IndoBERT and and DistilBERT DistilBERT for for Indonesian Indonesian emotion emotion Leveraging IndoBERT and DistilBERT for Indonesian classification in reviews classification in e-commerce reviews emotion classification in e-commerce reviews ScienceDirect,” *Procedia Comput. Sci.*, vol. 269, pp. 321–330, 2025, doi: [10.1016/j.procs.2025.08.284](https://doi.org/10.1016/j.procs.2025.08.284).
- [64] M. Alfhi, “ScienceDirect Online Gambling Promotion Detection in Indonesian YouTube Comments Using Semi-Supervised IndoBERT Classification,” *Procedia Comput. Sci.*, vol. 269, pp. 1269–1278, 2025, doi: [10.1016/j.procs.2025.09.068](https://doi.org/10.1016/j.procs.2025.09.068).
- [65] E. Supriyadi and P. Nurhuda, “ScienceDirect ScienceDirect Sentiment Analysis of TikTok User Comments on QRIS Sentiment Analysis TikTok User Comments on QRIS Adoption in of Indonesia Using IndoBERT Adoption in Indonesia Using IndoBERT,” *Procedia Comput. Sci.*, vol. 269, pp. 121–130, 2025, doi: [10.1016/j.procs.2025.08.265](https://doi.org/10.1016/j.procs.2025.08.265).
- [66] R. A. Malik, F. Riza, and S. Defit, “Quickly Assess the Acceptability Sentiment of White Paracetamol Intake Using KNN-SMOTE Based On Receptive Deciding,” *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 15, no. 1, p. 51, 2024, doi: [10.24843/lkjiti.2024.v15.i01.p05](https://doi.org/10.24843/lkjiti.2024.v15.i01.p05).
- [67] S. M. Octafia, R. A. Malik, A. Weriframayeni, and D. Delpa, “IndoBERT-SupCon : A Supervised Contrastive Learning Model for Analyzing Public Perception on Halal Tourism,” vol. 7, no. 1, pp. 218–231, 2026, doi: [10.47738/jads.v7i1.1045](https://doi.org/10.47738/jads.v7i1.1045).
- [68] M. K. Anam et al., “Improved Performance of Hybrid GRU-BiLSTM for Detection Emotion on Twitter Dataset,” *J. Appl. Data Sci.*, vol. 6, no. 1, pp. 354–365, 2025, doi: [10.47738/jads.v6i1.459](https://doi.org/10.47738/jads.v6i1.459).
- [69] P. P. Putra, M. K. Anam, A. S. Chan, A. Hadi, and N. Hendri, “Optimizing Sentiment Analysis on Imbalanced Hotel Review Data Using SMOTE and Ensemble Machine Learning Techniques,” *J. Appl. Data Scie.*, vol. 6, no. 2, pp. 936–951, 2025, doi: [10.47738/jads.v6i2.618](https://doi.org/10.47738/jads.v6i2.618).
- [70] C. Owen and D. Suhartono, “Intent Classification in Artificial Intelligence-Based Customer Service Chatbot for E-Wallet Service Providers,” *Int. J. Comput. Digit. Syst.*, vol. 17, no. 1, pp. 1–9, 2025, doi: [10.12785/ijcds/1571046212](https://doi.org/10.12785/ijcds/1571046212).