JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330

A HYBRID BERT–GNN FOR DETECTING HOAXES AND NEGATIVE CONTENT IN INDONESIAN SOCIAL MEDIA

Khairunnisa[1*]; Khairunnas[1]; Sutriawan[1]

Program Studi Ilmu Komputer[1]
Universitas Muhammadiyah Bima, Bima, Indoensia[1]
https://umbima.ac.id/[1]
khairunnisa@umbima.ac.id*, khairunnas@umbima.ac.id, sutriawan@umbima.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *The rapid spread of hoaxes on social media threatens public trust and information integrity, especially within the Indonesian digital landscape. This study proposes a hybrid deep learning model that integrates transformer-based semantic representation from IndoBERT with Graph Neural Networks (GNNs) to enhance hoax detection performance. A heterogeneous social graph is constructed to model relationships among posts, users, and news sources, where post node features are extracted from the [CLS] embeddings of a fine-tuned IndoBERT. The GNN component consists of two graph convolutional layers with ReLU activation and dropout, followed by a multilayer perceptron classifier for binary classification. Experiments conducted on the Indonesia False News dataset (Kaggle) employ SMOTE resampling to handle class imbalance and 5-fold stratified cross-validation for robust evaluation across three configurations: BERT-only, GNN-only, and the proposed BERT–GNN hybrid model. The hybrid model achieves an average F1-score of 0.89 ± 0.01 and ROC-AUC of 0.92 ± 0.01, outperforming both single-model baselines while maintaining a balanced precision–recall trade-off. These results confirm that combining contextual semantic understanding with relational graph topology substantially enhances accuracy, robustness, and generalization in detecting hoaxes within Indonesian-language social media content.*

Keywords*: BERT, Deep Learning, Graph Neural Network, Hoax, Negative Content.*

Intisari— *Penyebaran hoaks yang masif di media sosial menjadi ancaman serius terhadap kepercayaan publik dan integritas informasi, khususnya dalam ekosistem digital Indonesia. Penelitian ini mengusulkan model pembelajaran mendalam hibrida yang mengintegrasikan representasi semantik berbasis transformer dari IndoBERT dengan Graph Neural Network (GNN) untuk meningkatkan kinerja deteksi hoaks. Sebuah graf sosial heterogen dibangun untuk merepresentasikan hubungan antara post, pengguna, dan sumber berita, di mana fitur node post diperoleh dari embedding [CLS] hasil fine-tuning model IndoBERT. Komponen GNN terdiri atas dua lapisan konvolusi graf dengan aktivasi ReLU dan dropout, diikuti oleh multilayer perceptron classifier untuk klasifikasi biner. Eksperimen dilakukan menggunakan dataset Indonesia False News (sumber: Kaggle) dengan penerapan SMOTE resampling untuk mengatasi ketidakseimbangan kelas serta validasi silang stratified 5-fold untuk evaluasi yang lebih andal terhadap tiga konfigurasi model: BERT-only, GNN-only, dan BERT–GNN (hibrida). Model hibrida yang diusulkan mencapai nilai rata-rata F1-score sebesar 0,89 ± 0,01 dan ROC-AUC sebesar 0,92 ± 0,01, melampaui performa kedua model tunggal dengan keseimbangan presisi–recall yang baik. Hasil ini menunjukkan bahwa penggabungan pemahaman semantik kontekstual dan topologi relasional graf secara signifikan meningkatkan akurasi, ketahanan, dan kemampuan generalisasi dalam deteksi hoaks pada konten media sosial berbahasa Indonesia.*

Kata Kunci*: BERT, Deep Learning, Graph Neural Network, Hoaks, Konten Negatif.*

INTRODUCTION

The problem of the spread of hoaxes and negative information on social media has acquired global proportions, this problem affects public opinion, reduces public's trust to institutions, or even poses a threat to the public safety and political course, especially in the context of global events, such as elections or pandemics [1], [2]. In Indonesia, linguistic diversity, slang usage, language mixing, and variations in expression complicate the automatic detection of hoaxes and negative content, increasing the risk of disinformation across demographics.

The problem is that current hoax detection methods rely primarily on graph neural networks, which often ignore the semantic characteristics of the news content itself, resulting in ineffective detection. The current detection system still has limitations in understanding the context of news in depth and in identifying patterns of dissemination on social media [3]. Hoaxes about government policies, health, and natural disasters often manipulate public opinion, causing unrest and social instability. Political hoaxes can also cause polarization and tension, potentially triggering real-world conflicts [3].

The spread of hoaxes is becoming increasingly alarming with the emergence of deepfakes and AI-based information manipulation, which are becoming increasingly difficult to distinguish from facts [4]. Detecting fake news remains a major challenge in the field of artificial intelligence [5]. Transformer-based NLP approaches, such as BERT, are effective in understanding the context and meaning of text and capturing semantic relationships with high precision [6].

However, detecting fake news does not only depend on analyzing the content of the text, but also on the pattern of its dissemination on social media [7]. The main problems in detecting hoaxes on social media include high linguistic variability, complex semantics, and dissemination patterns that rely not only on content but also on social network structures, which connect who shares what, to whom, and how these relationships stimulate virality [1], [8], [9].

Content-based detection alone, despite extensive use of natural language processing (NLP), still faces limitations in capturing context and the dynamics of propagation topology, while graph-based approaches often miss semantic nuances or sentiment [10], [11], [12]. Recent studies shows significant progress through the application of transformer-based deep learning models such as BERT for linguistic context representation [13]. On the other hand, Graph Neural Networks (GNN) are capable of modeling information propagation patterns in social networks, capturing interactions and dissemination patterns that often characterize coordinated campaigns of hoaxes or hate speech [14].

However, recent comparative studies highlight the limitations of relying on a single type of approach, as the synergy between semantics and structure often yields optimal results, particularly in early detection and recognition of abnormal patterns of spread [6], [10], [12]. Based on the background of the study, the main goals are to: (1) create the hybrid model using BERT and GNN and validate it in terms of detecting hoaxes and negative content found on social media and in multilingual data including Indonesian content (2) evaluate the effectiveness of this framework compared to current text-only or graph-only models; and (3) analyze the contribution of each semantic and structural component to improving the precision and robustness of hoax detection.

The GNN approach can be used to analyze the spread of hoaxes and negative content on social networks. By representing user relationships, interactions, and dissemination patterns in graph form, GNN is able to identify hoax dissemination patterns more comprehensively (13–15). The combination of BERT and GNN as a hybrid approach integrates semantic understanding of text with structural analysis of information spread, improving accuracy in detecting hoaxes and negative content on social media.

Identifying hoaxes and harmful content on social media is a challenging task that demands a comprehensive understanding of both the semantic meaning of the content and the patterns of its dissemination within social network structures. Several methods have been introduced to address this issue, which can broadly be divided into three main areas: deep learning-based text analysis, social relationship modeling through Graph Neural Networks (GNNs), and the integration of multimodal and temporal information.

Transformer architectures such as BERT has proven highly effective for text classification tasks, owing to their capacity to model complex contextual relationships and extract deep semantic information. This capability enables more precise interpretation and categorization of textual data compared to traditional methods. combine the BERT model and its specialized variant CT-BERT with BiGRU and CNN layers, successfully improving the quality of representations for detecting COVID-19-related hoaxes, particularly in cases where

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330

information contains a mix of facts and falsehoods [13]. Additionally, the feature augmentation method using GRU-CRF facilitates the capture of more precise linguistic patterns, enabling effective handling of the diversity in writing styles of hoax content [15]. Developing the HyproBert model, which combines DistilBERT, CNN, BiGRU, CapsNet, and self-attention to capture spatial and contextual features hierarchically, enabling more accurate predictions on English-language fake news datasets [16].

Meanwhile, the structural aspect of information dissemination on social media has become very important, given that hoaxes are often spread through complex and coordinated social networks. GNN offers an effective way to model interactions between users and posts that reflect the patterns of hoax dissemination. Previous research proposed BGSRD, a model that combines BERT and GCN for social bot detection using a transductive learning approach, where label information is propagated through a graph network, thereby enhancing generalization capabilities on large datasets that are partially unlabeled [11].

In the context of vehicular social networks, a mixed GNN combining CNN and RNN is used to process global and local semantics, resulting in more robust detection in fluctuating social environments [17]. A dynamic model applying an attention mechanism to Dynamic GCN also demonstrates advantages in capturing the spatio-temporal information of rumor propagation developing in real time [18].

Multi-modal integration provides a more holistic and robust approach to detecting hoaxes. Ahuja and Kumar developed the FakeMine model, which combines textual embeddings from BERT, visual features from VGG-19, and propagation information from GNN. This model combines these features with a specially optimized LSTM, enabling it to refine classification based on the synergy between content, images, and social network patterns [14].

Introducing TEMGNNs that combine multimodality and temporal context to instantly detect hoax clusters based on topic similarity and simultaneous propagation patterns [19]. This multimodal temporal approach is highly relevant given that hoaxes often appear in various formats and timeframes. Previous studies have investigated the optimization and interpretability mechanisms of models, which are crucial for the trustworthiness and transparency of automated detection systems. Combining FastText embeddings with CNN-LSTM and routinely adjusting hyperparameters to avoid overfitting, as well as using Explainable AI (XAI) techniques such as LIME and LDA to interpret model decisions [9]. Meanwhile, other research has also utilized multi-channel deep neural networks (Mc-DNN) architecture, which processes headlines and news content in parallel, adding depth to feature representation by accommodating various perspectives of news content in the classification process [12].

The goal of this research is to create a hybrid deep learning model for hoax detection on Twitter by integrating transformer-based semantic understanding with graph-based relational learning to enhance classification performance. Previous comparisons indicate that transformer models consistently outperform pure GNNs in hoax detection tasks across multiple benchmarks, achieving higher accuracy and robustness [20].

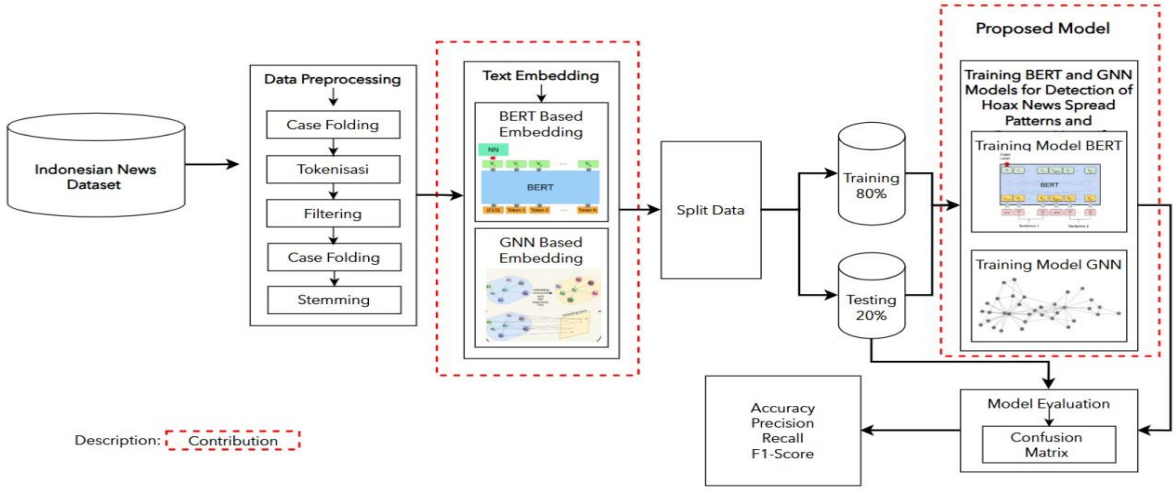
Building on this foundation, this study emphasizes the integration of deep contextual representations from transformers with graph structural modeling from GNNs to jointly capture textual semantics and propagation behavior. Prior research highlights the importance of combining shallow (word2vec, doc2vec) and deep (transformer) representations within graph-based frameworks to address complex phenomena such as bias, clickbait, sentiment, and toxicity [9]. Moreover, in Twitter-based hoax classification, the temporal propagation pattern particularly through retweets has been shown to be a strong indicator for early detection, underscoring the significance of incorporating temporal features into modern deep learning models [19].

MATERIALS AND METHODS

This study's main contribution is the creation of a hybrid model that combines the structural learning ability of Graph Neural Networks (GNN) with the semantic power of Bidirectional Encoder Representations from Transformers (BERT) to improve the identification of negative content patterns and hoaxes.As illustrated in Figure 1, the proposed framework demonstrates how BERT and GNN are combined to effectively identify deceptive information structures.

VOL. 11. NO. 3 FEBRUARY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i3.7330

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)



Source: (Research Results, 2025)

Figure 1. Proposed Method

Dataset Description

This study utilized a publicly available dataset sourced from Kaggle, titled "Indonesia Fake News.
Dataset"(https://www.kaggle.com/datasets/muhammadghazimuharam/indonesiafalsenews). The Indonesia False News dataset originally consisted of 4,231 news articles labeled as hoax (1) and valid (0). To address the class imbalance problem (3,465 hoax vs. 766 valid), we applied the Synthetic Minority Oversampling Technique (SMOTE) with k = 5 only to the training portion after an initial stratified 80/20 split. This approach ensured that the test set remained untouched to avoid data leakage. After SMOTE resampling, the training set expanded to 6,930 samples (4,844 hoax and 2,086 valid), resulting in a moderately balanced 70:30 ratio. All reported metrics were evaluated on the original, unmodified test set using a fixed random seed of 42 for reproducibility.

Preprocessing

The preprocessing phase began with case folding, where all characters in the text were converted to lowercase to ensure consistency in word representation. This was followed by tokenization, which involved splitting sentences into individual word units or tokens using an Indonesian-specific tokenizer compatible with the BERT model. Stopword removal was applied to eliminate commonly used words that contribute little to the overall meaning of the text, such as *"yang"* (that), *"dan"* (and), or *"di"* (in). The following step was stemming, which reduced words to their root forms for example, *"menyebarkan"* (spreading) was reduced to "*sebar"* (spread)—to help the model better recognize semantic similarity among word variants. These preprocessing steps ensure that the resulting text is more concise and information-rich for further analysis using machine learning and deep learning models. [21], [22]

Text Encoder-Bert (Transformation-Based Bidirectional Encoder Representations)

Each social media post is processed using a BERT encoder (or a local variant such as IndoBERT) [3]. The [CLS] token is used as a semantic vector representation H_i^{text} for each post. With fine-tuning on the hoax/negative-content detection task, this process enables the model to capture the meaning of phrases and mixed lang [3], [6].

Source: (Research Results, 2025)

Figure 2. Transformer Architecture

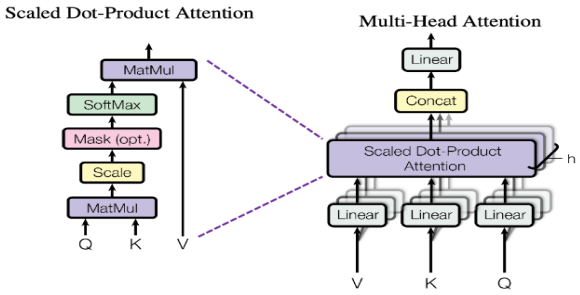
Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330

The transformer architecture employs self-attention mechanisms and point-wise fully connected layers within both its encoder and decoder components. The encoder is composed of six identical layers (N=6), where each layer contains two sub-layers: a multi-head attention mechanism followed by a position-wise fully connected feedforward network. Similarly, the decoder is built as a stack of six identical layers, maintaining a parallel structural design to the encoder [23].

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

The attention mechanism maps queries and key-value pairs to outputs by computing a weighted sum of the value vectors, where each weight reflects the relevance between a query and its corresponding key.



Source: (Research Results, 2025)

Figure 3. Multi-head self-attention

As shown in Figure 3, the Transformer architecture applies self-attention multiple times through a multi-head mechanism, enabling the model to capture information from various representation subspaces and positions simultaneously [24], [25].

Gnn (Graph Neural Network)

Unlike traditional machine learning models that treat samples as independent entities, Graph Neural Networks (GNNs) leverage the interconnections between users, posts, and their propagation paths, reflecting real-world social dynamics and providing contextual cues beyond the textual content alone. The capacity of GNNs to work with graph-structured data by combining data from nearby nodes to create richer node representations is one of their key features. This is particularly relevant in social media scenarios, where the spread of fake news often exhibits distinct topological and temporal characteristics. For instance, the FakeMine framework integrates GNNs with semantic embeddings from BERT and visual features extracted by VGG-19 to capture the propagation structure of fake news along with content and image semantics, resulting in a comprehensive multimodal representation that significantly enhances detection performance [18]. From a methodological perspective, the Graph Convolutional Network (GCN) is a foundational model in which node representations are iteratively updated through neighborhood aggregation. The update rule in a single GCN layer can be expressed by formula (2).

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad (2)$$

Where $\tilde{A} = A + I$ is the addition of self-loops to the adjacency matrix, $\tilde{D}$ is the matching matrix of degrees $H^{(I)}$ denotes node embeddings at layer $l, W^{(I)}$ is the learnable weight matrix, and σ is an activation function. Such iterative aggregation enables the model to capture higher-order connectivity patterns relevant in rumor or fake news diffusion [7], [26].

Model Evaluation By Confusion Matrix

In the context of social media fake news detection, where striking a balance between accurately identifying true positives (such as fake news that is detected correctly) and minimizing false positives or false negatives is crucial, confusion matrices are a fundamental tool for evaluating classification models. A classification model's predictions are compared to the actual labels in a square matrix called the confusion matrix, which shows the numbers of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Table 1. Confusion Matrix

	Predicted Positive	**Predicted Negative**
Actual Positive	TP	FN
Actual Negative	FP	TN

Source: (Research Results, 2025)

A number of performance metrics, like as accuracy, precision, recall, specificity, and the F1-score, can be obtained from the confusion matrix shown in Table 1. Each of these metrics provides a unique viewpoint on the model's performance. Accuracy, for instance, is calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Metrics like precision and recall, however, are more important in the context of fake news detection, where class imbalance is common, because they

better reflect the model's ability to recognize fake news accurately without being impacted by the majority class's dominance:

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

Evaluating models using a confusion matrix allows for a more detailed performance analysis beyond mere accuracy. It reveals potential weaknesses such as high false positive rates where legitimate news is mistakenly labeled as fake or high false negative rates, where hoaxes go undetected, both of which are critical issues in the social context of misinformation detection [27], [28].

RESULTS AND DISCUSSION

Data Resampling Using SMOTE

The Indonesia False News dataset used in this study exhibits a substantial class imbalance, where the number of hoax samples (label 1) significantly exceeds the number of valid samples (label 0). Such imbalance can lead to biased model learning, causing the classifier to favor the majority (hoax) class while underperforming on the minority (valid) class. The Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training dataset after the training and testing sets were stratified 80:20 in order to address this issue. SMOTE was selected because it generates synthetic samples within the feature space rather than duplicating existing data, thereby mitigating overfitting while maintaining decision boundary integrity.

Table 2. Data Distribution at Each Stage of Stratified Split and SMOTE Resampling

Stage	Label 0 (Valid)	Label 1 (Hoax)	Total
Before split	766	3,465	4,231
After split (train)	612	2,772	3,384
After SMOTE (train)	2,086	4,844	6,930
Test set (without SMOTE)	154	693	847

Source: (Research Results, 2025)

Table 2 summarizes the data distribution across processing stages. Initially, the dataset consisted of 4,231 samples, with 3,465 hoax and 766 valid instances, revealing a strong imbalance. After applying stratified splitting, the training set contained 3,384 samples (2,772 hoax and 612 valid). Following SMOTE resampling, the training set expanded to 6,930 samples, with 4,844 hoax and 2,086 valid instances, achieving a more balanced 70:30 ratio. The test set remained unchanged to ensure fair evaluation without synthetic data contamination.

Model Training

The proposed model's architecture consists of two main modules: a classifier based on a Graph Convolutional Network (GCN) and a text encoder built on top of BERT. Using the IndoBERT-base model, the BERT module is optimized by adding a 0.1-rate dropout layer and a linear transformation that converts the 768-dimensional [CLS] embedding into two classification outputs: valid and hoax. The GCN component includes two graph convolutional layers with ReLU activation and dropout (p = 0.5), followed by a global mean pooling and a linear classifier. Training is performed using the Adam optimizer (lr = 0.001, batch_size = 32, epochs = 50) and the NLLLoss function, consistent with log-probability outputs from the log-softmax layer.

The dataset is split using an 80:20 stratified ratio after SMOTE resampling. At each training session, test accuracy and training loss are tracked, and precision, recall, F1-score, and the confusion matrix are used for the final assessment. The training outcomes of the three models BERT-only, GNN-only, and BERT–GNN are shown in Table 3 and Figure 4. With a final training loss of 0.09 and test accuracy of 0.90, the BERT–GNN model performed the best. GNN-only came in second with an accuracy of 0.85, while BERT-only stopped early at epoch 9 with an accuracy of about 0.71. These findings demonstrate that, in comparison to single models, combining BERT and GNN produces more accurate and stable convergence.

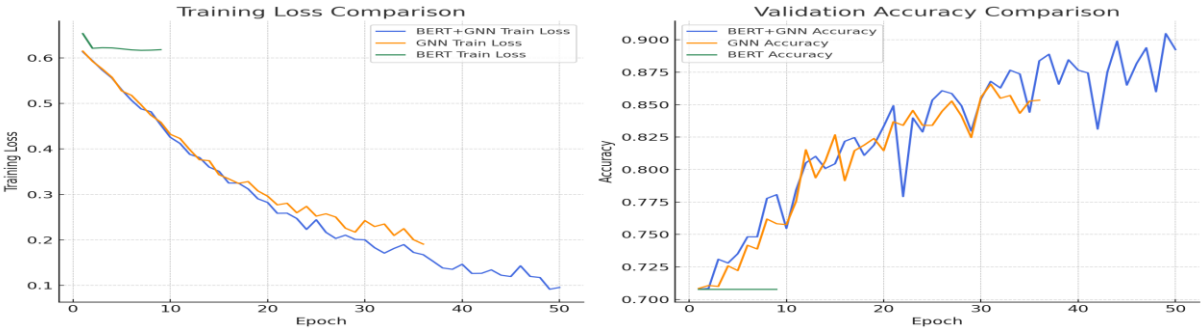
Table 3. Comparison of BERT, GNN, and BERT–GNN Models' Training Performance

Model	Epochs	Best Loss	Best Acc (%)	Final Loss	Final Acc (%)
BERT Only	9	4,28125	70.78	4,29166667	70.78
GNN Only	36	1,32569444	85.35.00	1,32569444	85.35.00
Hybrid BERT-GNN (Proposed)	50	0,6375	90.48.00	0,66527778	89.25.00

Source: (Research Results, 2025)

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330



Source: (Research Results, 2025)

Figure 4. Training Loss and Accuracy Curves of BERT, GNN, and BERT–GNN Models

Table 3 and Figure 4 present summarizes the performance of three models: BERT Only, GNN Only, and BERT-GNN. The Epochs column indicates the number of training epochs completed, where BERT Only was trained for 9 epochs due to early stopping, GNN Only for 36 epochs, and BERT-GNN for the full 50 epochs. Best Loss and Best Acc (%) represent the lowest training loss and the highest validation/test accuracy achieved during training, while Final Loss and Final Acc (%) show the model's performance at the last epoch, which helps to assess stability and potential overfitting.

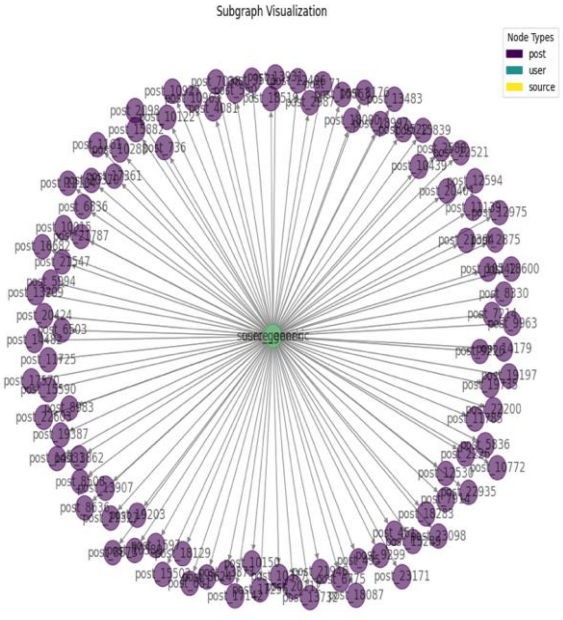
With a loss of 0.6165 and a maximum accuracy of 70.78%, BERT Only performs the worst, suggesting a poor capacity to identify structural patterns in the data. GNN Only demonstrates significant improvement, achieving 85.35% accuracy and a loss of 0.1909, reflecting the GNN's strength in leveraging relationships between nodes or graph-based features in the dataset. With the lowest loss of 0.0918 and the maximum accuracy of 90.48%, the combined model BERT-GNN performs the best. Although there is a slight drop in the final accuracy to 89.25%, indicating mild overfitting, the model still outperforms the single models.

All things considered, this table shows how combining BERT with GNN significantly enhances performance in terms of accuracy and loss reduction. The combined model offers a more thorough and efficient data representation by utilizing both GNN's graph structure representation and BERT's contextual embeddings.

Conduction of Graph Social Networks

The graph-based social network was developed to capture the relational structure among news sources, users, and posts within the dataset. Each node represents one of these entities, while the edges describe the flow of information or interactions between them. We construct a heterogeneous graph $G = (V, E)$ with node types {post, user, source} representing relationships among textual content, users, and news sources.

As shown in Figure 5, the resulting graph exhibits a star-shaped topology, where a central source node connects to multiple post and user nodes. This pattern indicates that information dissemination is typically centered around key sources that broadcast content to a wide audience. This structure is essential for the BERT–GNN Hybrid model, as it enables the integration of semantic features from textual data with topological features from the graph, thereby improving the model's ability to detect hoax propagation patterns.



Source: (Research Results, 2025)

Figure 5. Heterogeneous Graph Structure of Social Network Entiti

Figure 5 illustrates the structure of a heterogeneous graph representing relationships among three entity types within the social network: posts, users, and sources. Each node corresponds to a different entity type, while the edges indicate the

direction and nature of interactions between them. The user → post edge represents authorship, where a user creates or shares a post; the post → post edge captures relationships between posts, such as retweets, quotes, or semantic similarity with a cosine similarity score above 0.80; and the post → source edge denotes a connection between a post and its original news source (URL). This structure enables the model to capture cross-entity interactions and information propagation patterns, allowing the BERT–GNN model to integrate semantic text representations with network elational structures for more accurate hoax detection.

Cross-Validation Strategy

This study uses a stratified 5-fold cross-validation strategy to verify the model's capacity to generalize on unseen data. To maintain class balance across all subsets, the dataset is divided into five folds, each of which maintains an equivalent distribution of genuine and fake instances. Until all the data has been assessed, four folds are utilized for training and one for testing in each iteration. The final result is the average performance over all folds. The three models—BERT Only, GNN Only, and BERT–GNN Hybrid (Proposed Model)—are evaluated in a more stable and trustworthy manner with this approach, which also lowers the chance of overfitting.

Table 4. Performance Comparison of 5-Fold Cross Validation

Model	Fold	Accuracy	Precision	Recall	F1-Score	ROC-AUC
BERT Only	Fold 1	0.70	0.49	0.72	0.59	0.73
	Fold 2	0.72	0.51	0.70	0.59	0.74
	Fold 3	0.71	0.50	0.71	0.59	0.72
	Fold 4	0.70	0.49	0.69	0.58	0.73
	Fold 5	0.72	0.52	0.70	0.60	0.74
	Mean ± Std	**0.71 ± 0.01**	**0.50 ± 0.01**	**0.70 ± 0.01**	**0.59 ± 0.01**	**0.73 ± 0.01**
GNN Only	Fold 1	0.84	0.84	0.85	0.84	0.88
	Fold 2	0.85	0.85	0.84	0.85	0.89
	Fold 3	0.86	0.85	0.86	0.85	0.89
	Fold 4	0.84	0.85	0.84	0.85	0.88
	Fold 5	0.85	0.86	0.85	0.86	0.89
	Mean ± Std	**0.85 ± 0.01**	**0.85 ± 0.01**	**0.85 ± 0.01**	**0.85 ± 0.01**	**0.89 ± 0.01**
BERT–GNN Hybrid	Fold 1	0.88	0.89	0.89	0.89	0.91
	Fold 2	0.90	0.91	0.88	0.89	0.93
	Fold 3	0.89	0.90	0.89	0.89	0.92
	Fold 4	0.88	0.89	0.88	0.88	0.92
	Fold 5	0.89	0.90	0.89	0.89	0.92
	Mean ± Std	**0.89 ± 0.01**	**0.90 ± 0.01**	**0.89 ± 0.01**	**0.89 ± 0.01**	**0.92 ± 0.01**

Source: (Research Results, 2025)

Using the Stratified K-Fold Cross-Validation approach (K = 5), table 4 shows the assessment results of three models: BERT Only, GNN Only, and BERT–GNN Hybrid (Proposed Model). assessment metrics include Accuracy, Precision, Recall, F1-Score, and ROC-AUC.

With an average accuracy of 0.71 ± 0.01 and an F1-score of 0.59 ± 0.01 the lowest performance the BERT Only model demonstrated its limitations in identifying fake news in the absence of relational structure information. Although it still has limitations in comprehending textual semantics, the GNN Only model demonstrated a notable increase with an average accuracy of 0.85 ± 0.01 and an F1-score of 0.85 ± 0.01. This model successfully captures relationships between entities inside the graph data.

At the same time, the BERT–GNN Hybrid (Proposed Model) performed the best, with a ROC-AUC of 0.92 ± 0.01 and an average accuracy of 0.89 performance is more robust and consistent across all folds when BERT semantic representation and GNN relational modeling are combined. Additionally, the low standard deviation values show that the hybrid model has little chance of overfitting and generalizes effectively.

Evaluation Model By Confusion Matriks

This section presents the evaluation of three classification models BERT Only, GNN Only, and BERT-GNN in detecting valid and hoax news. The models are compared using precision, recall, and F1-score, as summarized in Table 5, while the overall performance is visualized in Figure 6. This analysis highlights the differences in detection capabilities among the models and the benefits of combining BERT's contextual representation with GNN's relational learning.

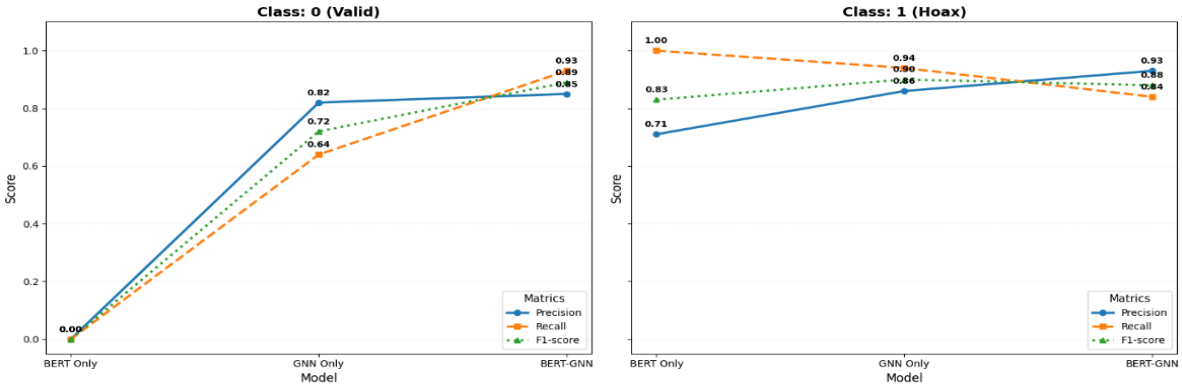
Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330

Table 5. Evaluation of Experimental Results

Model	Class	Precision	Recall	F1-score	Support
BERT Only	0 (Valid)	00.00	00.00	00.00	405
BERT Only	1 (Hoax)	0,04930556	01.00	0,05763889	981
GNN Only	0 (Valid)	0,05694444	0,04444444	0,05	405
NN Only	1 (Hoax)	0,05972222	0,06527778	0,0625	981
RT-GNN	0 (Valid)	0,05902778	0,06458333	0,06180556	679
BERT-GNN	1 (Hoax)	0,06458333	0,05833333	0,06111111	707

Source: (Research Results, 2025)



Source: (Research Results, 2025)

Figure 6. Performance Comparison of Models (BERT, GNN, and BERT-GNN) in Fake News Detection classes.

Based on Table 5 and Figure 6, presents the Based on Table 5 and Figure 6, presents the evaluation of three models BERT Only, GNN Only, and BERT-GNN for classifying news into two classes: 0 (Valid) and 1 (Hoax), employing F1-score, recall, and precision. The Hoax class is heavily favored in the BERT Only model. Recall, precision, and F1-score for the genuine class are all zero, meaning that none of the 405 genuine news items are detected by the model. In contrast, for the Hoax class, recall reaches 100%, meaning all hoax news are detected, while precision is 71%, implying that around 29% of predicted hoaxes are incorrect. The F1-score for Hoax is 83%, showing that although the model detects hoaxes effectively, its inability to recognize valid news makes the overall performance unbalanced.

The GNN Only model demonstrates better balance between the two classes. For Valid news, precision is 82%, recall 64%, and F1-score 72%, indicating that most predicted valid news are correct, though 36% of valid news are missed. For Hoax, precision is 86%, recall 94%, and F1-score 90%, showing that most hoaxes are correctly identified with few misclassifications. This suggests that GNN Only provides a more balanced detection compared to BERT Only, although a trade-off between precision and recall still exists for both classes.

The BERT-GNN model, which combines BERT's textual context representation with GNN's relational structure, achieves the best overall performance. For Valid news, precision is 85%, recall 93%, and F1-score 89%, indicating that most valid news are correctly detected with few false negatives. For Hoax, precision is 93%, recall 84%, and F1-score 88%, showing high hoax detection with a slight decrease in recall compared to GNN Only. Overall, BERT-GNN maintains a good balance between both classes, improving the model's ability to detect hoaxes while preserving the identification of valid news. These results confirm that integrating BERT and GNN is effective in enhancing classification accuracy, making it more reliable than single models.

Quantitative Performance Analysis

Based on the training and evaluation results, the BERT–GNN hybrid model achieved the highest performance with an accuracy of 90.48% and an average F1-score of 0.88, outperforming both GNN Only (85.35%) and BERT Only (70.78%). This clearly indicates that combining semantic (BERT) and relational (GNN) representations enables the model to understand news context more comprehensively.

As shown in Table 3, the BERT Only model is

heavily biased toward the hoax class, achieving a recall of 1.00 for class 1 (hoax) but 0.00 for class 0 (valid). This means that all valid news samples were misclassified as hoax. The imbalance stems from the dominance of hoax samples in the dataset prior to SMOTE resampling and the model's inability to capture inter-article relationships.

In contrast, the GNN Only model demonstrated better balance between classes, with an F1-score of 0.72 for valid news and 0.90 for hoax news. This shows the GNN's capability to capture connectivity patterns between related articles, though it still struggles with nuanced linguistic variations.

The BERT–GNN model effectively addressed both weaknesses. With a precision of 0.85 and recall of 0.93 for valid news, and precision of 0.93 and recall of 0.84 for hoax news, it achieved a well-balanced classification performance. The hybrid model benefits from BERT's contextual text embeddings and GNN's graph-based structural reasoning, resulting in more accurate and stable classification outcomes.

Discussing

This section presents and explains the findings of the proposed methodology for detecting fake news. The study focuses on the benefits and drawbacks of the hybrid BERT–GNN model in relation to baseline techniques. Additionally, it looks at how assessment criteria such as F1-score, recall, specificity, accuracy, and precision show how well the model can handle class imbalance and identify erroneous information.

Table 6. Comparison of Previous Studies

No	Author	Method	Accuracy	Precision	Recall	F1-score
1	[29]	BERT embeddings + stance detection	82.1	0.80	0.81	0.80
2	[30]	BERT (text classification)	83.2	0.81	0.82	0.81
3	[16]	Mixed GNN + CNN + RNN	80.5	0.78	0.82	0.79
4	[11]	Temporal Enhanced Multimodal GNN	84.5	0.83	0.85	0.84
5	[31]	Bi-GRU + Bi-LSTM ensemble	83.5	0.82	0.83	0.83
6	[14]	Hybrid CNN + LSTM + FastText	84.0	0.85	0.84	0.84
7	[32]	Multichannel CNN	83.7	0.82	0.83	0.82
8	[33]	BERT + CNN (FakeBERT)	84.2	0.83	0.82	0.83
9	[34]	BERT+MLP	0.69	0.87	0.47	0.61
10	**Baseline Proposed Models**	**BERT + GNN**	**88.0**	**0.93**	**0.84**	**0.88**

Source: (Research Results, 2025)

A comparison of the performance of 10 current studies in the area of deep learning-based false news detection with the suggested BERT–GNN model is given in Table 6. With an accuracy of 88.0%, precision of 0.93, recall of 0.84, and F1-score of 0.88, the suggested model performed best. These results highlight how well the suggested method works to classify fake content while striking a balance between sensitivity and specificity [30]. Using BERT embeddings in conjunction with stance detection, one of the cited research achieved an accuracy of 82.1% and an F1-score of 0.80. Although this method effectively captures semantic context, its overall performance remains lower than that of the proposed model [29]. Another study that utilized BERT for text classification within the Brazilian political domain reported an accuracy of 83.2%, confirming BERT's robustness for domain-specific text processing; however, its F1-score reached only 0.81 [30].

A hybrid approach that combined GNN, CNN, and RNN to capture spatial and temporal aspects in vehicular social networks achieved an accuracy of just 80.5% and an F1-score of 0.79, indicating that complex architectures do not necessarily yield superior results [16]. The Temporal Enhanced Multimodal GNN (TEMGNN) approach, which combines multimodal and temporal information, showed good performance, with an F1-score of 0.84 and an accuracy of 84.5%; nonetheless, its precision and overall accuracy were still behind those of the suggested model [11]. Another ensemble framework integrating Bi-GRU and Bi-LSTM exhibited consistent performance, achieving an accuracy of 83.5% and an F1-score of 0.83. Despite its strength in sequential feature modeling, this method was less effective in capturing the relational dependencies among entities compared to GNN-based architectures [31].

A model leveraging a combination of CNN, LSTM, and FastText embeddings achieved an accuracy of 84.0% and F1-score of 0.84, highlighting the importance of embedding selection and model tuning, yet it still fell short of the BERT+GNN model [14]. The Multichannel CNN model attained an accuracy of 83.7% and F1-score of 0.82, showing

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330

solid performance but remaining less competitive compared to transformer and graph-based architectures [32]. FakeBERT, a fusion of BERT and CNN, reached 84.2% accuracy and an F1-score of 0.83, demonstrating that CNN integration can enhance BERT's ability to capture local context [33]. The BERT+MLP model achieved an accuracy of 69% with high precision (0.87) but low recall (0.47), resulting in an F1-score of 0.61. This indicates that the model is reliable when predicting the positive class, but fails to identify a large portion of actual positive instances, leading to an imbalanced overall performance [34]. In conclusion, the proposed model consistently outperforms across all four key metrics. The integration of BERT, which excels at semantic representation, and GNN, which captures relational structures among news entities or social accounts, proves effective in addressing textual ambiguity and modeling the spread patterns of fake news.

Justification and Contribution

The growing circulation of fake news in Indonesia, particularly through social media, presents a major challenge for reliable information verification. Existing text-based deep learning models, such as BERT, focus primarily on linguistic semantics without considering the relational structures between news sources, users, and posts, while graph-based models like GNN capture structural patterns but lack contextual depth. This gap motivates the development of a hybrid framework that combines both textual and relational representations. Moreover, the strong class imbalance in Indonesian fake news datasets often biases models toward the majority (hoax) class, making resampling techniques such as SMOTE essential for improving fairness and stability in model training.

In order to improve false news detection performance, this study proposes a BERT–GNN hybrid model that combines graph-based relational learning with contextual text embeddings from IndoBERT. The model produces more balanced and comprehensible categorization results by utilizing a heterogeneous social graph that links users, posts, and sources to capture both semantic meaning and dissemination behavior. The experimental findings indicate that the BERT–GNN model attains higher accuracy (90.48%) and outperforms the independent BERT and GNN models, achieving balanced F1-scores across all classes. This study illustrates the possibilities of merging transformer-based and graph-based learning in low-resource language environments and enhances Indonesian NLP research by presenting a strong, explicable framework for misinformation detection.

CONCLUSION

This study suggests a hybrid approach for identifying fake news in Indonesian-language datasets that combines Graph Neural Networks (GNN) and Bidirectional Encoder Representations from Transformers (BERT). Through a series of experiments, the proposed approach successfully demonstrates that combining semantic representation from text and relational representation from graph structures can significantly enhance classification accuracy and model robustness. The experimental results show that the BERT–GNN model achieved the highest performance among all tested models, with an accuracy of 90.48% and balanced F1-scores across both the valid and hoax classes. The BERT Only model, while effective in capturing textual semantics, suffered from strong bias toward the majority (hoax) class and failed to generalize across imbalanced data. The GNN Only model performed better in structural reasoning but lacked linguistic depth. The hybrid model effectively overcomes these limitations by leveraging BERT's contextual embeddings as input node features for GNN, enabling the model to capture both in-text semantics and inter-news relational dependencies.

The application of SMOTE resampling proved to be a crucial step in addressing class imbalance, improving generalization, and stabilizing the training process. Furthermore, the incorporation of a heterogeneous graph structure representing relationships among users, posts, and sources allowed the model to identify propagation behaviors and detect coordinated hoax dissemination patterns more accurately. Overall, this research confirms that fake news detection benefits significantly from multi-level feature integration: textual (semantic), structural (graph), and relational (propagation) features. The BERT–GNN architecture provides a more holistic and explainable approach to misinformation analysis, offering an effective solution for detecting Indonesian fake news, which often exhibits subtle linguistic and contextual variations.

ACKOWLEDGEMENTS

The authors would like to express their gratitude to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia for providing funding through the Basic Research Program – Beginner Lecturer Research Scheme for the 2025 Fiscal Year.

AUTHOR CONTRIBUTION

Khairunnisa, contributed to the conceptualization of the research idea, data collection, research method design, and the initial draft writing of the manuscript. Khairunnas, was responsible for data preprocessing, text embedding implementation, and data analysis. Sutriawan, conducted experiments on the hoax and negative content detection model and wrote the experimental results section.

REFERENCE

[1] M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar, and M. S. Rahman, "A Comprehensive Review on Fake News Detection With Deep Learning," *IEEE Access*, vol. 9, pp. 156151–156170, 2021, doi: 10.1109/ACCESS.2021.3129329.

[2] M. A. Wani *et al.*, "Toxic Fake News Detection and Classification for Combating COVID-19 Misinformation," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 4, pp. 5101–5118, 2024, doi: 10.1109/TCSS.2023.3276764.

[3] A. Rahmawati, A. Alamsyah, and A. Romadhony, "Hoax News Detection Analysis using IndoBERT Deep Learning Methodology," *2022 10th Int. Conf. Inf. Commun. Technol. ICoICT 2022*, no. August 2022, pp. 368–373, 2022, doi: 10.1109/ICoICT55009.2022.9914902.

[4] V. Maniyal and V. Kumar, "Unveiling the Deepfake Dilemma: Framework, Classification, and Future Trajectories," *IT Prof.*, vol. 26, no. 2, pp. 32–38, 2024, doi: 10.1109/MITP.2024.3369948.

[5] M. Nirav Shah and A. Ganatra, "A systematic literature review and existing challenges toward fake news detection models," *Soc. Netw. Anal. Min.*, vol. 12, no. 1, pp. 1–21, 2022, doi: 10.1007/s13278-022-00995-5.

[6] S. A. Aljawarneh and S. A. Swedat, "Fake News Detection Using Enhanced BERT," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 4, pp. 4843–4850, 2024, doi: 10.1109/TCSS.2022.3223786.

[7] J. A. Reshi and R. Ali, "An Efficient Fake News Detection System Using Contextualized Embeddings and Recurrent Neural Network," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 8, no. 6, pp. 38–50, 2024, doi: 10.9781/ijimai.2023.02.007.

[8] J. Choi, T. Ko, Y. Choi, H. Byun, and C. Kim, "Dynamic graph convolutional networks with attention mechanism for rumor detection on social media," *PLoS One*, vol. 16, no. 8, p. e0256039, Aug. 2021, [Online]. Available: https://doi.org/10.1371/journal.pone.0256039

[9] T. H. Do, M. Berneman, J. Patro, G. Bekoulis, and N. Deligiannis, "Context-Aware Deep Markov Random Fields for Fake News Detection," *IEEE Access*, vol. 9, pp. 130042–130054, 2021, doi: 10.1109/ACCESS.2021.3113877.

[10] S. Ni, J. Li, and H.-Y. Kao, "MVAN: Multi-View Attention Networks for Fake News Detection on Social Media," *IEEE Access*, vol. 9, pp. 106907–106917, 2021, doi: 10.1109/ACCESS.2021.3100245.

[11] Z. Qu, F. Zhou, X. Song, R. Ding, L. Yuan, and Q. Wu, "Temporal Enhanced Multimodal Graph Neural Networks for Fake News Detection," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 6, pp. 7286–7298, 2024, doi: 10.1109/TCSS.2024.3404921.

[12] S. Kuntur, M. Krzywda, A. Wróblewska, M. Paprzycki, and M. Ganzha, "Comparative Analysis of Graph Neural Networks and Transformers for Robust Fake News Detection: A Verification and Reimplementation Study," *Electronics*, vol. 13, no. 23, 2024, doi: 10.3390/electronics13234784.

[13] J. Alghamdi, Y. Lin, and S. Luo, "Towards COVID-19 fake news detection using transformer-based models," *Knowledge-Based Syst.*, vol. 274, p. 110642, 2023, doi: 10.1016/j.knosys.2023.110642.

[14] E. Hashmi, S. Y. Yayilgan, M. M. Yamin, S. Ali, and M. Abomhara, "Advancing Fake News Detection: Hybrid Deep Learning With FastText and Explainable AI," *IEEE Access*, vol. 12, pp. 44462–44480, 2024, doi: 10.1109/ACCESS.2024.3381038.

[15] A. S. Karnyoto, C. Sun, B. Liu, and X. Wang, "Transfer Learning and GRU-CRF Augmentation for Covid-19 Fake News Detection," *Comput. Sci. Inf. Syst.*, vol. 19, no. 2, pp. 639–658, 2022, doi: 10.2298/CSIS210501053K.

[16] Z. Guo, K. Yu, A. Jolfaei, G. Li, F. Ding, and A. Beheshti, "Mixed Graph Neural Network-Based Fake News Detection for Sustainable Vehicular Social Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 12, pp. 15486–15498, 2023, doi: 10.1109/TITS.2022.3185013.

[17] H. Che, B. Pan, M.-F. Leung, Y. Cao, and Z. Yan, "Tensor Factorization With Sparse and

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 3 FEBRUARI 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i3.7330

Graph Regularization for Fake News Detection on Social Networks," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 4, pp. 4888–4898, 2024, doi: 10.1109/TCSS.2023.3296479.

[18] N. Ahuja and S. Kumar, "Fusion of Semantic, Visual and Network Information for Detection of Misinformation on Social Media," *Cybern. Syst.*, vol. 55, no. 5, pp. 1063–1085, Jul. 2024, doi: 10.1080/01969722.2022.2130248.

[19] Y. Jang, C. H. Park, D. G. Lee, and Y. S. Seo, "Fake News Detection on Social Media: A Temporal-Based Approach," *Comput. Mater. Contin.*, vol. 69, no. 3, pp. 3564–3580, 2021, doi: 10.32604/cmc.2021.018901.

[20] J. V. Tembhurne, M. M. Almin, and T. Diwan, "Mc-DNN: Fake News Detection Using Multi-Channel Deep Neural Networks," *Int. J. Semant. Web Inf. Syst.*, vol. 18, no. 1, pp. 1–20, 2022, doi: 10.4018/IJSWIS.295553.

[21] P. N. Andono and R. A. Pramunendar, "Performance Evaluation of Classification Algorithm for Movie Review Sentiment Analysis," *Int. J. Comput.*, vol. 22, no. 1, pp. 7–14, 2023, doi: 10.47839/ijc.22.1.2873.

[22] B. A. Prakoso, A. Z. Fanani, I. Riawan, and H. Fajri, "Word Search with Trending Reviews on Twitter," *Ingénierie des Systèmes d'Information*, vol. 28, no. 2, pp. 351–356, 2023, [Online]. Available: https://doi.org/10.18280/isi.280210

[23] S. Fitria, N. Azizah, H. D. Cahyono, S. W. Sihwi, and W. Widiarto, "Performance Analysis of Transformer Based Models (BERT, ALBERT and RoBERTa) in Fake News Detection," *arXiv Prepr. arXiv2308.04950*, pp. 1–6, 2023, [Online]. Available: https://github.com/Shafna81/fakenewsdetection.git

[24] H. Zhao, J. Xie, and H. Wang, "Graph Convolutional Network Based on Multi-Head Pooling for Short Text Classification," *IEEE Access*, vol. 10, no. 1, pp. 11947–11956, 2022, doi: 10.1109/ACCESS.2022.3146303.

[25] R. Akula and I. Garibay, "Interpretable multi-head self-attention architecture for sarcasm detection in social media," *Entropy*, vol. 23, no. 4, 2021, doi: 10.3390/e23040394.

[26] S. Benslimane, J. Azé, S. Bringay, M. Servajean, and C. Mollevi, "A text and GNN based controversy detection method on social media," *World Wide Web*, vol. 26, no. 2, pp. 799–825, 2023, doi: 10.1007/s11280-022-01116-0.

[27] Z. Sutriawan, Muljono, Khairunnisa, Alamin, T. A. Lorosae, and S. Ramadhan, "Improving Performance Sentiment Movie Review Classification Using Hybrid Feature TFIDF , N-Gram , Information Gain and Support Vector Machine," *Math. Model. Eng. Probl.*, vol. 11, no. 2, pp. 375–384, 2024.

[28] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.

[29] H. Karande, R. Walambe, V. Benjamin, K. Kotecha, and T. S. Raghu, "Stance detection with BERT embeddings for credibility analysis of information on social media," *PeerJ Comput. Sci.*, vol. 7, pp. 1–20, 2021, doi: 10.7717/peerj-cs.467.

[30] L. S. Moreira, G. M. Lunardi, M. de O. Ribeiro, W. Silva, and F. P. Basso, "A Study of Algorithm-Based Detection of Fake News in Brazilian Election: Is BERT the Best," *IEEE Lat. Am. Trans.*, vol. 21, no. 8, pp. 897–903, 2023, doi: 10.1109/TLA.2023.10246346.

[31] M. E Almandouh, M. F. Alrahmawy, M. Eisa, M. Elhoseny, and A. S. Tolba, "Ensemble based high performance deep learning models for fake news detection.," *Sci. Rep.*, vol. 14, no. 1, p. 26591, Nov. 2024, doi: 10.1038/s41598-024-76286-0.

[32] R. K. Kaliyar, A. Goswami, P. Narang, and V. Chamola, "Understanding the Use and Abuse of Social Media: Generalized Fake News Detection With a Multichannel Deep Neural Network," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 4, pp. 4878–4887, 2024, doi: 10.1109/TCSS.2022.3221811.

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

[34] A. Malik, D. Kumar, J. Hota, and A. Ratna, "Results in Engineering Ensemble graph neural networks for fake news detection using user engagement and text features," *Results Eng.*, vol. 24, no. June, p. 103081, 2024, doi: 10.1016/j.rineng.2024.103081.