



ARTICLE

A Transformer-Based Deep Learning Framework with Semantic Encoding and Syntax-Aware LSTM for Fake Electronic News Detection

Hamza Murad Khan¹, Shakila Basheer², Mohammad Tabrez Quasim³, Raja'a Al-Naimi⁴,
Vijaykumar Varadarajan⁵ and Anwar Khan^{1,*}

¹Department of Electronics, University of Peshawar, Peshawar, 25120, Pakistan

²Department of Information Systems, College of Computer and Information Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

³Department of Computer Science and Artificial Intelligence, College of Computing and Information Technology, University of Bisha, P.O. Box 551, Bisha, 61922, Saudi Arabia

⁴Department of Mathematics, University of Petra, Amman, 1199, Jordan

⁵School of Computer Science, University of Technology, Sydney, 2007, Australia

*Corresponding Author: Anwar Khan. Email: arkhan@uop.edu.pk

Received: 20 June 2025; Accepted: 19 August 2025; Published: 10 November 2025

ABSTRACT: With the increasing growth of online news, fake electronic news detection has become one of the most important paradigms of modern research. Traditional electronic news detection techniques are generally based on contextual understanding, sequential dependencies, and/or data imbalance. This makes distinction between genuine and fabricated news a challenging task. To address this problem, we propose a novel hybrid architecture, T5-SA-LSTM, which synergistically integrates the T5 Transformer for semantically rich contextual embedding with the Self-Attention-enhanced (SA) Long Short-Term Memory (LSTM). The LSTM is trained using the Adam optimizer, which provides faster and more stable convergence compared to the Stochastic Gradient Descent (SGD) and Root Mean Square Propagation (RMSProp). The WELFake and FakeNewsPrediction datasets are used, which consist of labeled news articles having fake and real news samples. Tokenization and Synthetic Minority Over-sampling Technique (SMOTE) methods are used for data preprocessing to ensure linguistic normalization and class imbalance. The incorporation of the Self-Attention (SA) mechanism enables the model to highlight critical words and phrases, thereby enhancing predictive accuracy. The proposed model is evaluated using accuracy, precision, recall (sensitivity), and F1-score as performance metrics. The model achieved 99% accuracy on the WELFake dataset and 96.5% accuracy on the FakeNewsPrediction dataset. It outperformed the competitive schemes such as T5-SA-LSTM (RMSProp), T5-SA-LSTM (SGD) and some other models.

KEYWORDS: Fake news detection; tokenization; SMOTE; text-to-text transfer transformer (T5); long short-term memory (LSTM); self-attention mechanism (SA); T5-SA-LSTM; WELFake dataset; FakeNewsPrediction dataset

1 Introduction

The increasing growth of digital media platforms has changed the information-sharing paradigm, allowing news to immediately reach to global audience. It has also facilitated the widespread dissemination of fake news that are presented as real news [1,2]. False news has the ability to shift public opinion, influence election outcomes, incite violence and create mistrust in legitimate news sources [3,4]. These challenges stem



from the growing use of malicious tools, highlighting the urgent need for reliable and trustworthy methods to globally detect fake news [5–7].

Fake news detection is a cumbersome problem and can be approached from the perspective of text analysis, as one of the strategies. The use of multimodal data, such as images, videos and user metadata has shown promising results in enhancing detection performance of fake news [8]. Additionally, social contexts, such as propagation patterns and user engagement, help differentiate fake news from real content. Therefore, an effective fake news detection system should include comprehensive frameworks that integrate diverse data and contextual information [9,10].

An enhanced Bi-directional Encoder Representations from Transformers (BERT)-based model, incorporating linear, dropout, and activation layers was trained on the self-developed PolitiTweet and benchmark BuzzFeed datasets in [11]. It achieved respective accuracies of 89% and 85% under varied conditions and outperformed the traditional BERT model. However, this approach is prone to overfitting. The framework in [12] used deep Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), LSTM and pre-trained BERT for zero-shot fake news detection without domain-specific labeled data, achieving 88.39% accuracy. Yet, its limitation in handling sarcasm, context-specific nuances and subtle manipulative language indicates the need for further improvement. The inter- and intra-modal relationships for fake news detection with positive and unlabeled data are analyzed in [13] at the cost of high computational. The study in [14] detected multimodal fake news by filtering noisy text, images, and comments using a masked Transformer and a selective comment-context encoder. Although it showed good performance, it relied on pre-defined comment selection rules, limiting adaptability across diverse datasets.

To cope with the aforementioned challenges, this work proposes the T5-SA-LSTM. It is a transformer-based deep learning architecture that combines the T5 Transformer's generative text-to-text capabilities, the global semantic focus of Self-Attention and LSTM for capturing long-range patterns within a unified framework. It applies tokenization and the Synthetic Minority Over-sampling Technique (SMOTE) as data preprocessing techniques to balance class distribution and improve learning. It is designed to address the shortcomings found in the existing BERT-based [11], CNN-RNN hybrid [12] and multimodal Transformer models [13,14]. It ensures robustness, domain adaptability and practical computational efficiency. The model's effectiveness is assessed using the WELFake and FakeNewsPrediction datasets, focusing on improving fake news detection while maintaining generalization across datasets. The key contributions of this research are as follows:

- We propose the T5-SA-LSTM framework that combines the Text-to-Text Transfer Transformer (T5 Transformer) with the Self-Attention mechanism and the LSTM network to enhance the accuracy and robustness of false and real news classification and detection.
- First, we apply tokenization to segment the raw text into meaningful tokens (subwords). Next, the SMOTE technique is applied to balance the dataset and address class imbalance. The tokenized and balanced data is then fed to the T5 model, which serves as a pre-trained text-to-text framework that provides deep contextual understanding and advanced Natural Language Processing (NLP) capabilities for fake news classification.
- Finally, by integrating the Self-Attention mechanism with LSTM's sequential learning capability, the framework effectively captures long-term dependencies and subtle contextual cues. This synergy enables the model to distinguish manipulated content from legitimate news with significant precision and improved interpretability.
- A set of performance metrics is used to evaluate the performance of the proposed approach using different datasets. It demonstrates superior performance over the competitive models.

The rest of this paper is organized as follows: [Section 2](#) presents the literature review. [Section 3](#) describes the classification algorithms used in this study and the proposed model. [Section 4](#) provides details about the datasets used and discusses the experimental results. [Section 5](#) concludes the paper and outlines directions for future research.

2 Literature Review

This section reviews some novel machine and deep learning approaches used in fake news detection. The authors in [\[15\]](#) present a machine learning framework for fake news detection. They compare traditional classifiers (Naïve Bayes, Support Vector Machine, Random Forest) with deep learning models (CNN, LSTM, BERT) using various embedding techniques. Although BERT achieves good accuracy, its high computational cost and sensitivity to biased training data pose challenges to scalable and real-time deployment. In contrast, the proposed model integrates the T5's powerful text-to-text transfer capabilities with a lightweight Self-Attention layer and LSTM, providing better contextual representation with reduced computational cost and better adaptability to new and unseen contexts.

The study in [\[16\]](#) presents an automated fake news detection method using machine learning, deep learning and three-word vectorization techniques. A weighted voting ensemble achieves the respective 88.76% and 87.67% accuracies on the PolitiFact and BuzzFeed datasets. However, its dependence on simple voting ensembles and traditional word vectorization limits its ability to capture deep semantic dependencies and contextual nuances. The proposed technique overcomes this by leveraging T5's deep sequence-to-sequence encoding for rich semantic understanding while the Self-Attention mechanism captures long-term dependencies more effectively than standard LSTM or Gated Recurrent Unit (GRU) layers alone. A comparative study in [\[17\]](#) evaluates various deep learning models for fake news detection and concludes that these models improve accuracy but suffer from high computational cost and reduced zero-shot performance in real-time applications. The proposed T5-SA-LSTM balances the performance-cost trade-off by reusing the T5's encoder capabilities for efficient data representation. In [\[18\]](#), the authors propose the Efficient Transformer-Based Multilevel Attention (ETMA) framework for multimodal fake news detection, combining visual, textual and joint attention to capture modality-specific features. While effective for multimodal misinformation, its performance degrades when visual content is weakly correlated. Since our study focuses on pure textual misinformation, it avoids unnecessary visual overhead and enhances textual representation depth with Self-Attention, making it robust even to subtle or contextually ambiguous fake news.

The SA-Bi-LSTM model combines Self-Attention and bidirectional LSTM to effectively capture long-range dependencies and textual features, achieving 89.98% accuracy on the Information Security and Object Technology (ISOT) fake news dataset. Cross-validation and ablation studies highlight the attention mechanism's role in enhancing fake news detection [\[19\]](#). However, these models typically depend on static embeddings or shallow tokenization. The proposed model extends this by integrating T5's sophisticated text-to-text encoder to generate rich embeddings, which are further refined by the Self-Attention layer before being sequentially processed by the LSTM, ensuring superior capture of nuanced semantics. The CNN-based multichannel deep CNN in [\[20\]](#) utilize multiple kernel sizes to extract varying n-grams for feature diversity. While effective for local patterns, it struggles with sequential dependencies and context beyond fixed window sizes and the multichannel structures used increase computational complexity. The T5-SA-LSTM eliminates the need for heavy convolution by directly learning deep semantic and sequential relations through the T5 and LSTM, with self-attention enhancing token interactions across long texts. The studies in [\[21,22\]](#) enhance multimodal fake news detection by integrating diverse feature modalities, including visual, textual, frequency-domain and emotion-driven features, achieving notable accuracy gains over prior

methods. However, their evaluation on limited datasets, absence of cross-domain and real-time validation, lack of ablation analysis and high computational costs restrict their applications to resource-constrained or time-sensitive settings.

Finally, the algorithms proposed in [23,24] advance multimodal fake news detection by integrating textual and visual features through early, joint and late fusion. They use a multiscale semantic enhancement–progressive multimodal fusion network (SEPM) for improved semantic representation. While achieving notable performance gains—up to 90% accuracy on Gossipcop and Fakeddit, and 3.5% and 2% improvements on Twitter and Weibo—both are limited by small evaluation datasets, lack of comparison with advanced multimodal models and incomplete resolution of the semantic gap. Despite notable progress in machine learning and transformer-based methods for fake news detection [11–24], existing approaches still have challenges in fake electronic news detection. Challenges such as limited contextual understanding of information, bias in learned representations and poor robustness on unseen data persist, showcasing the need for more efficient and reliable frameworks. To bridge these gaps, this study proposes an enhanced transformer-based architecture using the LSTM and Self-Attention to improve fake news detection.

3 Operational Framework

3.1 Tokenization

The tokenization is the process of breaking down a sequence of text into smaller units called tokens [25]. These tokens can be sentences, words, subwords, or even characters, depending on the application [26], as shown in Fig. 1.

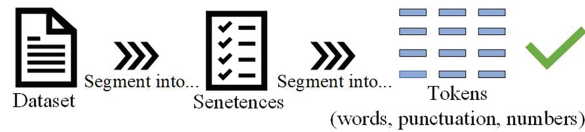


Figure 1: A structured approach to analyze fake news through dataset segmentation and tokenization

Mathematically, tokenization can be seen as a function that maps a sequence of characters to a sequence of tokens [27] as $T: S \rightarrow \{t_1, t_2, \dots, t_n\}$, where S is the input text sequence and $T(S)$ is the tokenized output. Each t_i represents an individual i -th token in the sequence [28].

3.2 Synthetic Minority Over-Sampling Technique (SMOTE)

The SMOTE is a data augmentation and oversampling method used to handle classification imbalances [29].

The SMOTE is a data augmentation and oversampling method used to handle class imbalances. It works by generating synthetic samples of the minority class rather than merely duplicating existing ones [30]. Addressing this class imbalance improves the performance of classifiers in tasks like fake news detection, where real news articles typically outnumber fake news samples [31,32]. To evaluate the effectiveness of these synthetic samples, we analyzed the class distribution before and after applying the SMOTE and observed a substantial increase in the representation of the minority class across all training subsets. The balanced datasets led to noticeable improvements in the detection of the minority class, demonstrating that the generated samples helped the model to reliably detect fake news. In imbalanced datasets, where the minority class has significantly fewer instances than majority class, standard classification algorithms often become biased toward the majority class. To address this challenge, SMOTE generates synthetic fake news samples to balance the dataset and reduce majority class bias as shown in Fig. 2. Visualizing the data with and without

SMOTE confirms that this augmentation enriches the features space, reduces misclassification and enhances the model's generalization and detection accuracy. The SMOTE starts by selecting a random minority class sample and one of its k -nearest neighbors [33] such that $x_i, x_j \in \mathbb{R}^d, i \neq j$, where x_i denotes a randomly chosen feature vector from the fake news class after tokenization and vectorization, while x_j represents one of its k -nearest neighbors within the same class. A new synthetic sample is created as follows [34]:

$$x_{new} = x_i + \lambda \cdot (x_j - x_i) \quad (1)$$

where λ is a random number drawn from a uniform distribution between 0 and 1 ($\lambda \sim U(0,1)$). The term $x_j - x_i$ represents the vector difference between the selected minority sample i and its k -nearest neighbor j of the selected minority sample. The x_{new} is a newly generated synthetic fake news sample on the line segment between x_i and x_j . This process is repeated until the desired number of synthetic samples is generated, ensuring a more balanced dataset.

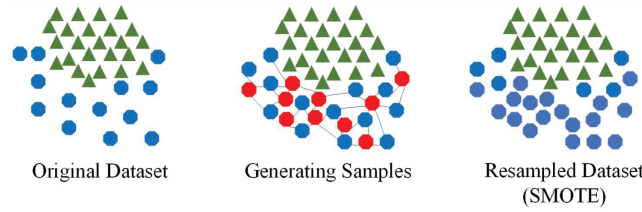


Figure 2: Addressing class imbalance with the SMOTE by generating realistic synthetic data samples

3.3 Text-To-Text Transfer Transformer (T5)

The T5 model adopts the original Transformer's encoder-decoder architecture [35], recasting text classification as a text-to-text task that maps an input, such as a news article or claim, to an output label like "real" or "fake." Its encoder extracts context-rich representations, which the decoder uses to generate the final output [36] as shown in Fig. 3. Given an input sequence $x = (x_1, x_2, \dots, x_n)$, the T5 encoder maps it to a hidden representation h as $h = \text{Encoder}(x)$. The encoder and decoder are composite modules consisting of multiple sub-layers, typically implemented as a stack of N identical transformer layers (e.g., $N = 12$ for T5-Base). They include multi-head self-attention, layer normalization and position-wise feed-forward networks. Within the feed-forward layers, the encoder uses the Gaussian Error Linear Unit (GeLU) activation function to introduce non-linearity. To ensure computational efficiency, we adopt the commonly used sigmoid-based approximation of the GeLU: $\text{GeLU}(x) \approx x \cdot \sigma(1.702x)$, where the factor 1.702 optimizes the sigmoid approximation to closely match the original GeLU's behavior. Importantly, this GeLU formulation is a smooth and non-periodic activation function. Here, $\sigma(x)$ denotes the standard sigmoid function:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

The T5 decoder then predicts the output token sequence $y = (y_1, y_2, \dots, y_m)$ autoregressively, meaning each token y_t (which signifies the single individual token that the T5 decoder is currently trying to predict at a specific timestep t within the output sequence) is generated conditionally on the encoder representation and the previously generated tokens $y < t$ (here y representing the entire sequence of tokens that have already been generated by the decoder before the current timestep t , acting as the history of the output sequence) [37]:

$$P(y|x) = \prod_{t=1}^m P(y_t | y < t, h) \quad (3)$$

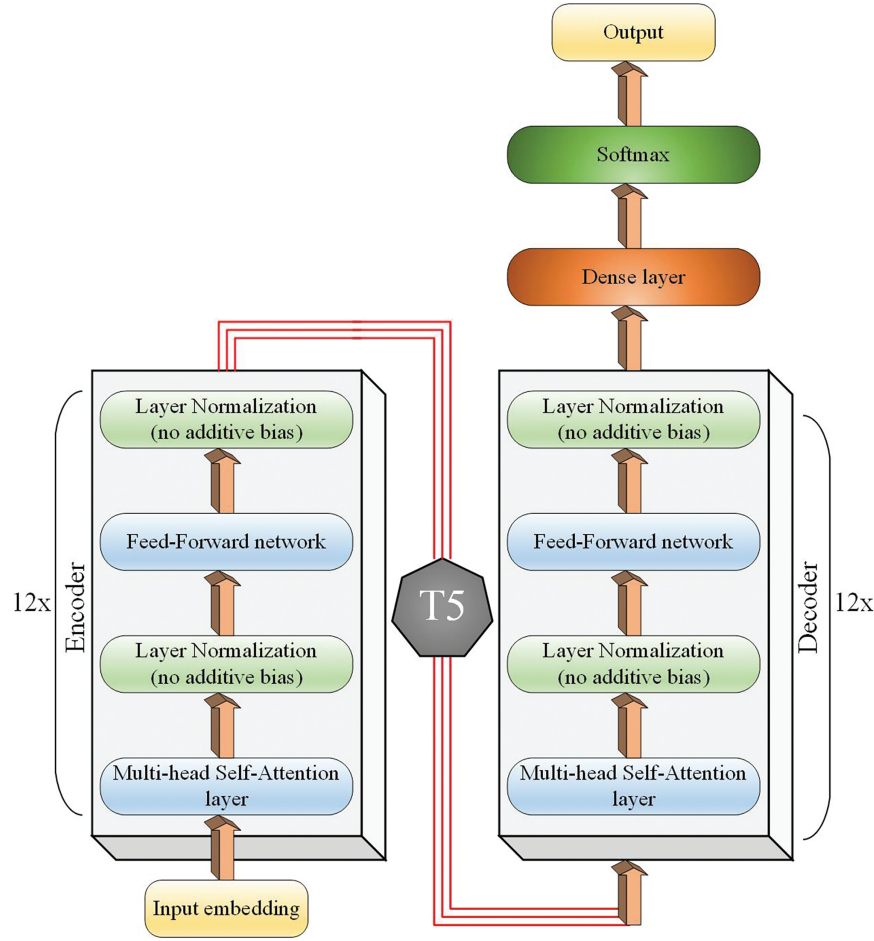


Figure 3: The working mechanism of the transformer-based T5 model

This factorization assumes that each token y_t is conditionally independent given the prior tokens and the encoder output, which is a standard assumption for autoregressive sequence generation. To obtain the vocabulary probability distribution, the decoder's hidden state z_t is projected through a learnable linear layer to generate logits, which are then transformed by the softmax function into a valid probability distribution [38] as:

$$P(y_t | y < t, h) = \text{softmax}(W \cdot z_t) \quad (4)$$

where W is the weight matrix of the dense layer and z_t is the decoder's hidden state at time step t . The T5 is trained by minimizing the cross-entropy loss between the predicted probability distribution and the true target labels $y = (y_1, y_2, \dots, y_m)$ (here y denotes the true label sequence) [39]:

$$\mathcal{L} = - \sum_{t=1}^m y_t \log P(y_t | y < t, h) \quad (5)$$

where y_t donates the true label for the t -th token in the target sequence. Here, y_t represents both the predicted token during generation and the ground truth token during training. For fake news detection, we fine-tune the T5 model on a labeled dataset containing real and fake news articles. The input format is “*classify: [News Text]*” and the output is fake or real.

3.4 Self-Attention (SA)

Self-Attention allows a model to weigh the importance of different parts of an input sequence when making predictions, essentially enabling it to pay attention to relevant information across the sequence [40]. The key idea is that each word in a sequence attends to all the other words, dynamically assigning importance scores [41]. Given an input sequence X of word embeddings: $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{n \times d}$, where n is the sequence length (number of words in a news article) and d is the embedding dimension. The self-attention mechanism transforms X into three learned representations: Query matrix (Q), Key matrix (K), and Value matrix (V). These matrices are computed as $Q = XW_Q$, $K = XW_K$, and $V = XW_V$ [42], where W_Q , W_K , and $W_V \in \mathbb{R}^{d \times d_k}$ are learned weight of the corresponding respective matrices. Then the attention scores are computed using the Scaled Dot-Product Attention [43] as:

$$A = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) \quad (6)$$

where $A \in \mathbb{R}^{n \times n}$ contains attention weights. Here, Q (Queries) and K (Keys) are projected representations of the input and K^T is the transpose of the keys matrix, allowing each query vector to be compared against all key vectors by dot product. The scaling factor d_k (dimension of keys) is used for scaling to prevent large values. Each attention score represents the extent of the focus each word should have on the others.

After computing the attention scores, we generate the output $Z = AV$, where $Z \in \mathbb{R}^{n \times d_k}$ is the new representation of the sequence. The final output of the self-attention mechanism is ZW_o , where $W_o \in \mathbb{R}^{d_k \times d}$ is another trainable matrix [44]. This improves the model's ability to detect fake news by capturing different contextual clues such as misleading phrases, sentiment shifts and deceptive structures.

3.5 Long-Short Term Memory (LSTM)

The LSTM overcomes long-term dependency and vanishing gradient in RNN using memory cells and forget, input, and output gates [45,46] as illustrated in Fig. 4. In fake news detection, it models the sequential and contextual relations in text [47].

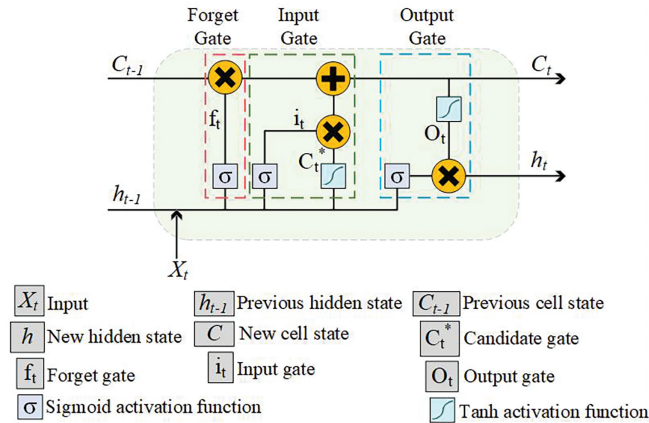


Figure 4: The LSTM's core mechanism for handling long and short-term dependencies in sequences

In the forget gate, f_t is the output of the sigmoid function that controls the flow of information from the previous cell state. This mechanism is essential to prevent the LSTM from retaining outdated or irrelevant

information. A value of 0 forgets the information and a value of 1 retains the information as:

$$f_t = \sigma (W_f h_{t-1} + U_f X_t + b_f) \quad (7)$$

The input gate i_t determines the information to be stored in the cell state. It works together with the candidate gate C_t^* to update the memory as:

$$i_t = \sigma (W_i h_{t-1} + U_i X_t + b_i) \quad (8)$$

The candidate gate C_t^* generates a set of potential values to add to the cell state C . These are “suggested memories” based on the current input X_t and the previous hidden state h_{t-1} , passed through the tanh activation. They are then combined with the decisions of the input gate and the output of the forget gate, as follows:

$$C_t^* = \tanh (W_c h_{t-1} + U_c X_t + b_c) \quad (9)$$

$$C_t = f_t * C_{t-1} + i_t * C_t^* \quad (10)$$

Within the output gate, the parameter O_t determines which parts of the cell state are transferred to the hidden state h_t , forming the output at the current time step t , as formulated in (19). It regulates the exposure of the internal memory to the next layer or future time steps, which is essential for controlled information flow.

$$O_t = \sigma (W_o h_{t-1} + U_o X_t + b_o) \quad (11)$$

$$h_t = O_t * \tanh (C_t) \quad (12)$$

where f_t , i_t , O_t , and C_t^* represent the forget gate, input gate, output gate and candidate gate at time step t , respectively, while X_t denotes the current input word embedding at timestep t . C_{t-1} and h_{t-1} refer to the previous cell state and hidden state of the LSTM network at timestep $t-1$, respectively. C_t and h_t indicate the updated cell state and hidden state at timestep t . W_f , W_i , W_c , W_o , U_f , U_i , U_c , U_o are the weight matrices for all gates while b_f , b_i , b_c , b_o represent the bias terms, with the indicated subscripts as mentioned and used before. The σ and $\tanh$ denotes the sigmoid and hyperbolic tangent activation functions, respectively. Additionally, the symbol $*$ indicates element-wise multiplication in computing the new cell state.

$$\tanh (x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (13)$$

The final hidden state is passed through a dense (fully connected) layer followed by a softmax activation function to obtain class probabilities [48] as:

$$p (y|X) = \text{Softmax} (W_c h_T + b_c) \quad (14)$$

where h_T is a final hidden state representation for the entire news article and W_c and b_c are the respective weight matrix and bias of the classification layer. The softmax function is defined as:

$$\text{Softmax} (z_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (15)$$

where z_i represents the logit or raw score for the i th class generated by the final dense layer from the LSTM's output hidden state and I denotes a specific target class label (for example, ‘real news’ or ‘fake news’). The

term e^{z_i} is the exponential of the logit for class i , which ensures that all output scores become positive and comparable. $\sum_j e^{z_j}$ represents the sum of the exponentials of the logits for all possible classes j in the classification task.

3.6 T5-SA-LSTM (Proposed Model)

The T5, Self-Attention, and LSTM models are integrated to develop the T5-SA-LSTM framework for fake news detection, as shown in Fig. 5. During data preprocessing, tokenization is first applied to break down the input text into meaningful units (subwords tokens), which are then converted into numerical representations through word embeddings. Once the text has been tokenized and embedded, the SMOTE technique is employed to address class imbalance by generating synthetic samples for the minority class (fake news). By applying SMOTE, the model works with a balanced representation of real and fake news articles, which prevents prediction bias, improves classification performance and enhances the model's robustness and generalization capability.

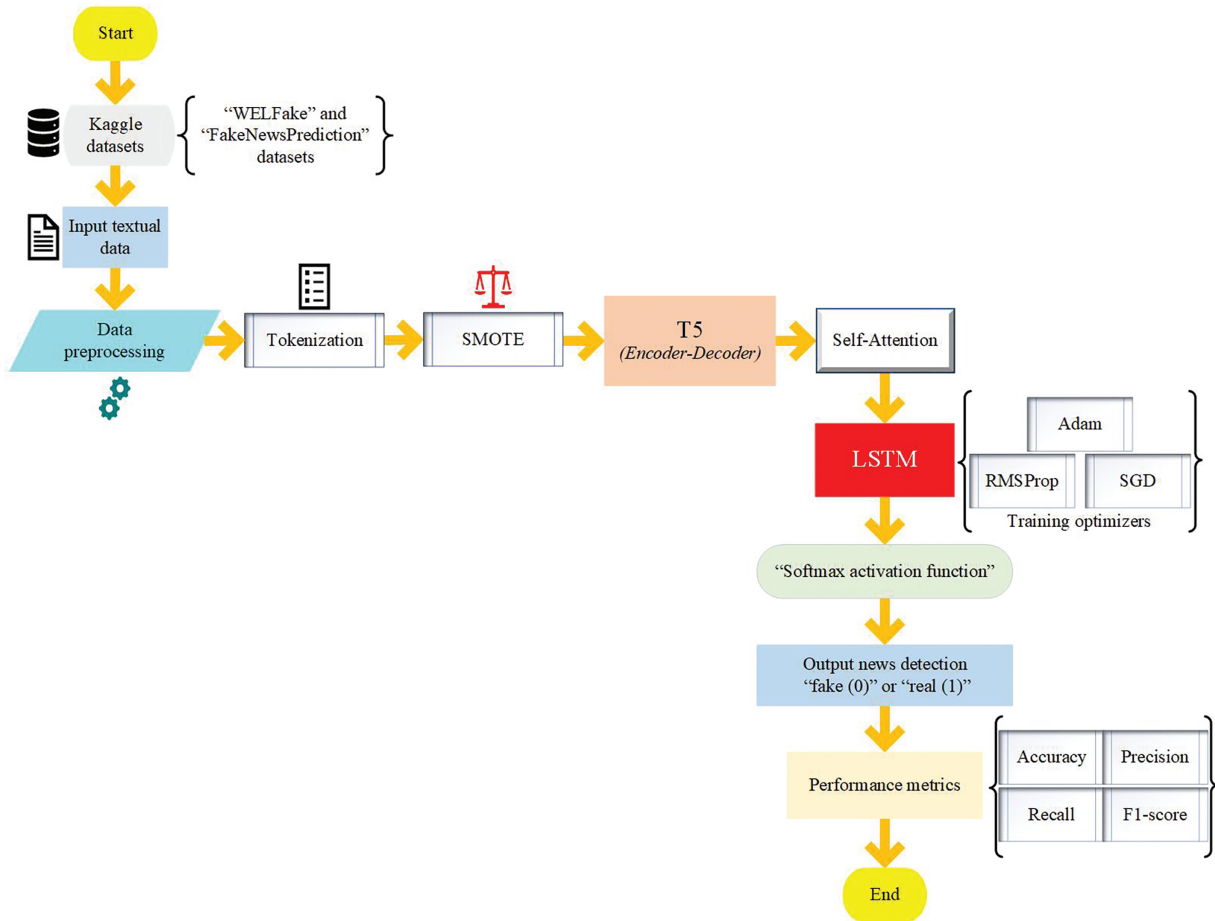


Figure 5: Flowchart of the proposed fake news detection model

The T5 model serves as the primary feature extractor. Its encoder-decoder architecture transforms input text into a meaningful latent representation, capturing contextual dependencies across words. The Self-Attention mechanism refines these representations by dynamically weighing the importance of a word, ensuring that the model focuses on critical textual patterns and deceptive linguistic cues. Subsequently,

the LSTM processes the self-attended feature sequences, capturing long-term dependencies to enhance classification accuracy. To effectively train the LSTM component, we select the Adam optimizer after empirical comparisons with SGD and RMSProp, as Adam provides adaptive learning rates for each parameter and incorporates momentum, which helps the model converge faster and with more accuracy. This makes Adam particularly well-suited for training our deep sequence model on complex textual news data, where gradients can be sparse or noisy. In our experiments, Adam consistently outperforms SGD and RMSProp by achieving higher accuracy and lower validation loss, demonstrating its positive impact on the model's performance and generalization capability. The LSTM's final hidden state passes through a dense layer with a softmax activation function to predict whether the news is real or fake.

The pseudocode of the proposed T5-SA-LSTM framework is shown in Algorithm 1. First, the news dataset is loaded, tokenized, converted to word embeddings and balanced with SMOTE. The data is labeled (fake = 0, real = 1), split into training and testing sets and passed through the T5 transformer to extract latent features. The Self-Attention mechanism refines these features, which are then fed into the LSTM trained using multiple optimizers (Adam, SGD, RMSProp) to select the best configuration. The final model is retrained with the optimal optimizer and the softmax output layer produces the fake or real prediction. Performance is evaluated using accuracy, precision, recall, and F1-score, yielding a robust trained model for reliable fake news detection.

Algorithm 1: Pseudocode for the T5-SA-LSTM framework

Input: Textual News Dataset (Fake and Real)

Output: FinalModel → Trained Model for Fake News Detection (Binary Classification “Fake = 0 or Real = 1”)

Xbest → Optimal Model Configuration

Begin

Step 1: Load Dataset

file → CSV i/p Load

Step 2: Preprocess Data

- Tokenization:

TextTokens → Tokenize(inputText)

- Word Embedding:

EmbeddedText → ConvertToNumerical(TextTokens)

- Handle Class Imbalance using SMOTE:

If ImbalanceDetected → Apply_SMOTE(EmbeddedText)

Else → Continue

Step 3: Classification Labels

- Fake News → Label = 0

- Real News → Label = 1

Step 4: Split Dataset

Dtrain, Dtest → Split(EmbeddedText, Labels, ratio = 80:20)

Step 5: Extract Features using T5 Transformer

- Initialize T5 Model:

T5output → Encode_Decode(Dtrain)

- Generate latent Representations:

(Continued)

Algorithm 1 (continued)

LatentFeatures → Extract(T5output)

Step 6: Apply Self-Attention Mechanism

- Compute Attention Weights:

AttentionWeights → ComputeWeights(LatentFeatures)

- Weighted Feature Representation:

SelfAttendedFeatures → ApplyAttention(LatentFeatures, AttentionWeights)

Step 7: Train LSTM Network with Multiple Optimizers

- Initialize LSTM:

LSTMoutput → Pass_via_LSTM(SelfAttendedFeatures)

- Define Optimizers:

Optimizers = {Adam, SGD, RMSProp}

- Select Best Optimizer based on Validation Performance:

For optimizer in Optimizers:

Train model using optimizer

ValLoss → Evaluate

If ValLoss < BestLoss:

BestOptimizer → optimizer

BestLoss → ValLoss

Step 8: Retrain Model with Best Optimizer

- Fine-tune LSTM with BestOptimizer:

LSTMoutput → Train_LSTM(SelfAttendedFeatures, BestOptimizer)

Step 9: Classification of Dense Layer

- Final Prediction (Fake = 0, Real = 1):

OutputClass → Softmax(LSTMoutput)

If OutputClass ≥ 0.5 → PredictedLabel = **1** (Real News)

Else → PredictedLabel = **0** (Fake News)

Step 10: Model Evaluation

- Compute Performance Metrics:

Accuracy, Precision, Recall, F1-score → Evaluate(PredictedLabel, Dtest)

Step 11: Output Results

Return FinalModel, BestOptimizer, Accuracy, Precision, Recall, F1-score

End

3.7 Performance Metrics

For performance comparison, first a binary classification is made where a True Positive (TP) indicates correctly identifying fake news, a True Negative (TN) means accurate identification of true news as true, a False Positive (FP) occurs when true news is mistakenly labeled as fake and a False Negative (FN) happens when fake news is incorrectly categorized as true. The following performance metrics are used:

- Accuracy (Acc): measures overall correctness. It is the ratio of correctly classified instances to the total instances.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

- Precision (Pr): represents how many of the detected fake news instances are genuinely fake.

$$Pr = \frac{TP}{TP + FP} \quad (17)$$

- Recall (Sensitivity): shows how effectively the model detects actual fake news.

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

- F1-score: balances precision and recall. It is the harmonic mean:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (19)$$

4 Simulation and Experimental Analysis

4.1 Dataset Description and Preprocessing

We used two fake news detection datasets, WELFake and FakeNewsPrediction, to implement the proposed model. These datasets were obtained from [Kaggle.com](https://www.kaggle.com), a reputable and well-known platform that hosts a diverse collection of datasets across various categories [49,50] and are freely available online. Both datasets contained a wide range of news articles, with an emphasis on political topics (e.g., election fraud, fake statements by leaders) and international topics (e.g., diplomatic conflicts, cross-border disputes). The WELFake and FakeNewsPrediction datasets, from the years 2021 and 2024, respectively, offer diverse, high-quality text for fake news detection and are publicly accessible. WELFake contains 62,470 articles (34,486 real, 27,984 fake) curated from Kaggle, McIntire, Reuters and BuzzFeed Political. Its original 78,098 entries were filtered to exclude entries with missing Label values, resulting in four columns: Serial Number, Title, Text and Label. The Fake labels in the WELFake dataset were determined automatically by using web scraping and weak supervision. The collection of fake news articles from sources known for publishing fabricated or satirical contents and real news from reputable mainstream outlets was made with labeling based on source credibility rather than manual fact-checking of each article. Similarly, FakeNewsPrediction includes 20,091 articles (11,812 real, 8279 fake) drawn from various sources to prevent overfitting. From an initial 25,116 entries, rows lacking valid labels were removed. The Fake labels are assigned according to the original sources and fact-checking procedures described by the dataset authors. This dataset features five columns: Serial Number, Title, Author, Text, and Label, ensuring both datasets are well-structured and suitable for binary classification tasks.

[Table 1](#) outlines the system resources used during simulations: an AMD Ryzen 5 5600X CPU, AMD Radeon RX 5600XT GPU, 16 GB RAM, Windows 11 (x64) and Jupyter Notebook (Python 3.11.0) are. This setup supports efficient development and computation. We used data preprocessing techniques such as tokenization and SMOTE.

Table 1: PC setup for executing the proposed research work

Category	Resource	Details
Hardware	CPU	Multi-core processor (AMD Ryzen 5 5600x)
	GPU	AMD radeon (RX 5600xt)
	RAM	16 GB (3200 MHz)

(Continued)

Table 1 (continued)

Category	Resource	Details
Software	Operating system Development environment	Windows 11 (×64 bit) Jupyter notebook (Python 3.11.0)

The precise subword token counts in fake news articles (Fig. 6a,b) reveal distinct linguistic patterns and differences in writing style.

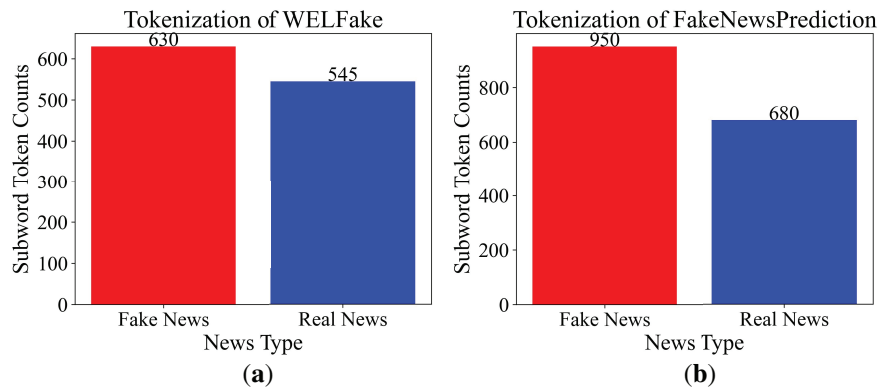


Figure 6: (a) Tokenization of the ‘WELFake’ and (b) ‘FakeNewsPrediction’ textual news datasets

In this study, tokenization was performed at the subword level, which enables a more granular representation of textual information by capturing morphological variations and rare word fragments that traditional word-level tokenization usually overlook. Using subword tokens, the exact token count for fake news articles in the WELFake dataset is 630, compared to 545 for real news articles. Similarly, in the FakeNewsPrediction dataset, fake news articles contain exactly 950 tokens, while real news articles have 680 tokens, as shown in the respective Fig. 6a,b. This observation indicates that fake news articles tend to exhibit higher token counts. It is because fake news often relies on exaggerated claims, emotionally charged narratives, redundant phrasing, speculative statements and clickbait elements designed to persuade and engage readers, all of which may contribute to increase token length [51,52]. In contrast, real news articles adhere to journalistic standards that prioritize clarity and conciseness, which, when tokenized at the subword level, results in a lower token count. After applying the tokenization technique to the original false news datasets, the textual news data is then forwarded to SMOTE for data balancing.

Fig. 7a,b shows the respective actual and SMOTE balanced WELFake dataset while Fig. 8a,b shows the respective operation on the FakeNewsPrediction dataset. Initially, as shown in the original class distribution plots, both datasets exhibit a noticeable imbalance, where real news articles (Class 1) significantly outnumber fake news articles (Class 0).

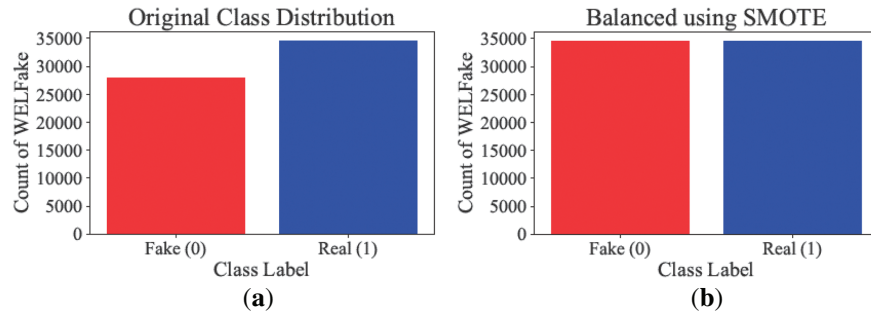


Figure 7: (a,b) The original classification and the SMOTE-generated synthetic samples balancing WELFake

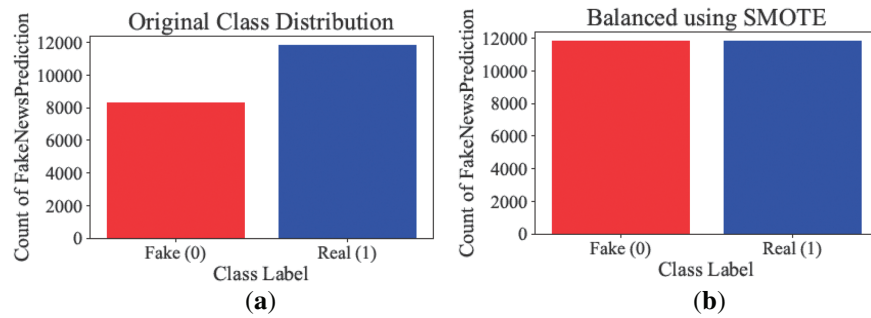


Figure 8: (a,b) The actual and the SMOTE-generated synthetic samples balancing FakeNewsPrediction

4.2 Simulation Results

Table 2 presents the configurations of various models used for fake news detection on the WELFake and FakeNewsPrediction datasets.

Table 2: Hyperparameter tuning details for the proposed and counterpart models for detecting false news

Platform		Models on “WELFake” and “FakeNewsPrediction” Datasets				
	Parameter	T5-SA-LSTM	FakeBERT	BERT + LightGBM	SA-Bi-LSTM	LR-SVM-GRU
Jupyter Notebook (Python 3.11.0)	Neurons	128	128	128	128	128
	Lr.	0.001	0.001	0.001	0.001	0.001
	Epochs	100	100	100	100	100
	Dense layers	32	32	32	32	32
	Mini-batch size	32	32	32	32	32
	Iterations	1950 and 750	1950 and 750	1950 and 750	1950 and 750	1950 and 750
	Optimizer	Adam/SGD /RMSProp	–	–	Adam/SGD /RMSProp	Adam/SGD /RMSProp
	Activation function	GeLU/ σ /Softmax	GeLU/tanh /Softmax	GeLU/tanh /Softmax	σ /tanh /Softmax	σ /tanh /Softmax

Figs. 9 and 10 present a comprehensive analysis of the T5-SA-LSTM model's performance on the WELFake and FakeNewsPrediction datasets, respectively, over 100 training epochs using three optimizers: SGD, RMSProp, and Adam. For both datasets, the accuracy graphs show rapid improvement within the initial 20–30 epochs, after which the curves flatten, indicating convergence and stable learning.

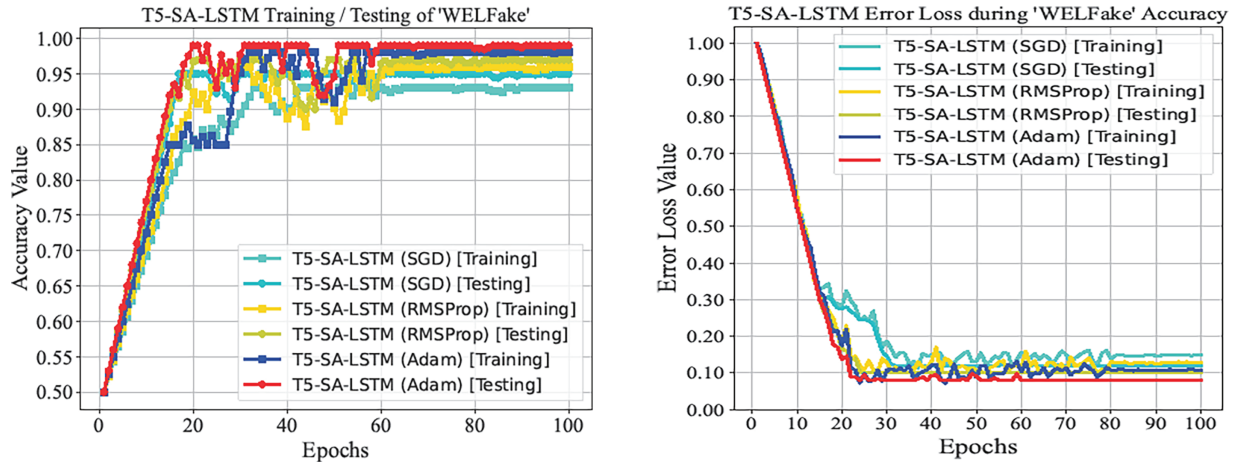


Figure 9: The proposed model accuracy and error loss during training/testing on WELFake

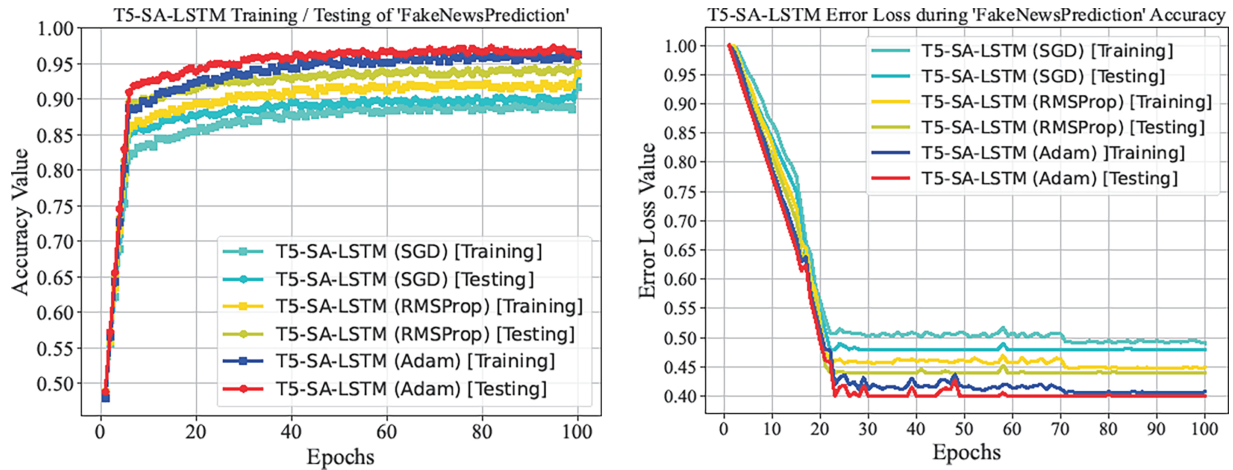


Figure 10: The proposed model accuracy and error loss during training/testing on FakeNewsPrediction

On WELFake (Fig. 9), the T5-SA-LSTM with Adam achieves the highest accuracy, reaching about 98% (training) and 99% (testing), with RMSProp stabilizing at around 96% and 97%, and SGD at 93% and 95%, respectively. The corresponding error loss graph shows Adam maintaining the minimal loss of approximately 0.03 (training) and 0.01 (testing), while RMSProp and SGD have losses in between 0.07–0.04.

A similar trend appears on FakeNewsPrediction (Fig. 10), where Adam achieves around 96% (training) and 96.5% (testing) accuracy, RMSProp reaches 94% and 95%, and SGD about 92% and 93%, with respective error losses of about 0.04–0.03 for Adam and 0.06–0.04 for RMSProp and SGD. The close alignment between training and testing curves for both accuracy and loss indicates excellent generalization and effective overfitting prevention, demonstrating the robustness of the T5-SA-LSTM architecture and its effective convergence achieved with the Adam optimizer.

Fig. 11 provides a comparative analysis of the T5-SA-LSTM model against several benchmark models on the WELFake dataset. The accuracy graph clearly indicates that T5-SA-LSTM with Adam is the most accurate model, achieving approximately 99% accuracy on testing data. The other T5-SA-LSTM variants, RMSProp and SGD, reached around 97% and 95% accuracies, respectively.

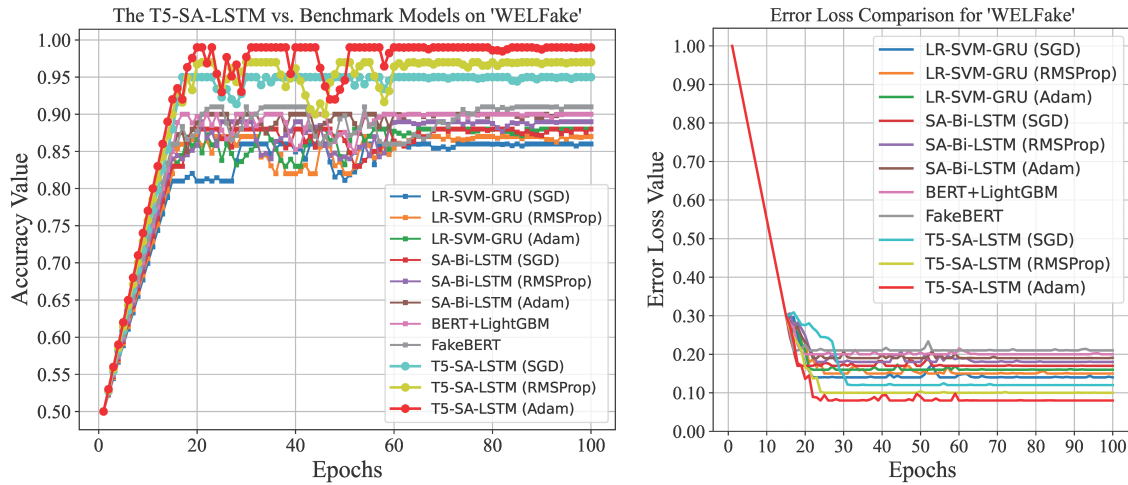


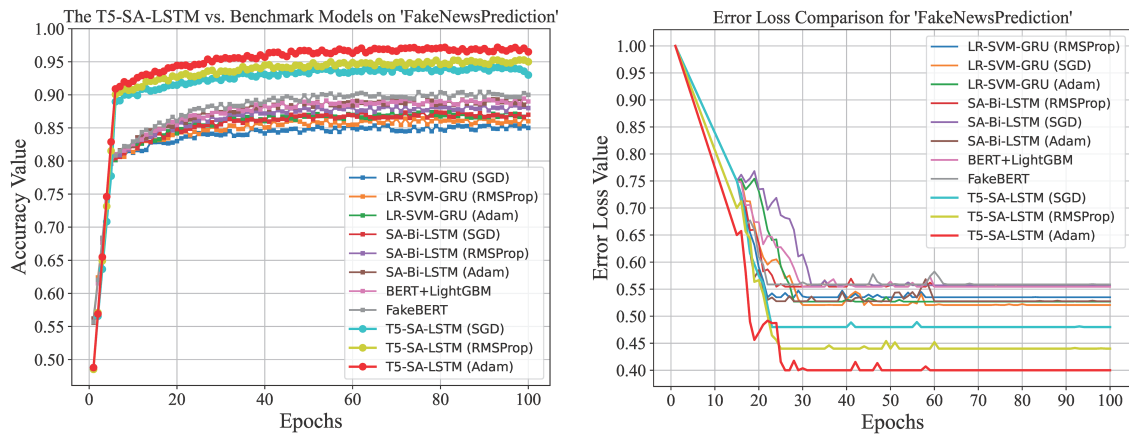
Figure 11: Accuracy benchmarking and error loss convergence of the proposed and counterpart models for fake news detection on the WELFake dataset

The counterpart models, FakeBERT, SA-Bi-LSTM (Adam, RMSProp and SGD) achieved 91%, 90%, 89%, and 88% accuracies, respectively. The LR-SVM-GRU has an accuracy of 88% (Adam), 87% (RMSProp) and 86% (SGD) as shown in Table 3. The error loss graph shows that T5-SA-LSTM with Adam exhibited the lowest error loss, with a value of approximately 0.08, followed by RMSProp with a loss value of about 0.11 and SGD with a loss value of 0.13. BERT + LightGBM recorded an error loss of about 0.15 while FakeBERT showed around 0.16 error loss. The SA-Bi-LSTM variants recorded error losses of 0.17 (Adam), 0.19 (RMSProp), and 0.21 (SGD), while the LR-SVM-GRU variants recorded losses of 0.18 (Adam), 0.20 (RMSProp), and 0.22 (SGD). Fig. 12 provides a comparative analysis of the T5-SA-LSTM model vs. several benchmark models on the FakeNewsPrediction dataset. The accuracy graph clearly indicates that the T5-SA-LSTM with Adam is the most accurate, having an accuracy of 96.5% on testing data. The other T5-SA-LSTM variants, RMSProp and SGD, has accuracies of 95% and 93%, respectively. Among the counterpart models, FakeBERT has 90% accuracy, followed by an accuracy of 89% (Adam), 88% (RMSProp) and 87% (SGD). The LR-SVM-GRU models achieved accuracies of 87% (Adam), 86% (RMSProp), and 85% (SGD) as shown in Table 3. The error loss graph shows that the T5-SA-LSTM with Adam exhibited the lowest error loss of approximately 0.41, followed by RMSProp with a lossof about 0.44 and SGD with a loss of 0.50. BERT + LightGBM recorded an error loss of about 0.55, while FakeBERT showed a loss of about 0.54. The SA-Bi-LSTM variants had loss values of 0.55 (Adam), 0.56 (RMSProp), and 0.58 (SGD). The LR-SVM-GRU variants showed loss values of 0.56 (Adam), 0.57 (RMSProp), and 0.58 (SGD).

Table 3: Comparative analysis of fake news detection models: accuracy, loss, precision, recall, and F1-score across WELFake and FakeNewsPrediction datasets

Model	Acc % (Error loss)		Pr %		Recall %		F1-Score %	
	WEL Fake	FakeNews Prediction	WEL Fake	FakeNews Prediction	WEL Fake	FakeNews Prediction	WEL Fake	FakeNews Prediction
LR-SVM-GRU (SGD)	86%	85% (0.58)	92.2%	88.5%	90.2%	86.6%	91.2%	87.5%
LR-SVM-GRU (RMSProp)	87%	86% (0.57)	94.4%	93.3%	95.0%	93.8%	94.7%	93.5%
LR-SVM-GRU (Adam)	88%	87% (0.56)	96.6%	93.5%	97.1%	94.6%	96.9%	94.0%
SA-Bi-LSTM (SGD)	88%	87% (0.58)	88.3%	89.3%	80.9%	82.3%	84.4%	85.6%
SA-Bi-LSTM (RMSProp)	89%	88% (0.56)	91.4%	92.4%	91.6%	93.5%	91.5%	92.9%
SA-Bi-LSTM (Adam)	90%	89% (0.55)	93.2%	91.7%	91.0%	90.0%	92.0%	90.8%
BERT + LightGBM	90%	89% (0.55)	93.0%	90.2%	88.5%	94.0%	90.5%	92.1%
FakeBERT	91%	90% (0.54)	95.6%	92.6%	97.4%	95.6%	96.4%	94.0%
T5-SA-LSTM (SGD)	95%	93% (0.50)	91.3%	86.1%	88.7%	84.0%	90.0%	85.0%
T5-SA-LSTM (RMSProp)	97%	95% (0.44)	96.5%	91.0%	95.2%	90.2%	95.9%	90.6%
T5-SA-LSTM (Adam)	99%	96.5%	99.6%	97.8%	99.8%	98.0%	99.6%	97.8%
	(0.08)	(0.41)						

Note: The bold entries in Table 3 represent the best-performing models across the respective evaluation metrics. Notably, the T5-SA-LSTM (Adam) achieved the highest accuracy of 99% with the lowest error loss (0.08), along with superior precision (99.6%), recall (99.8%), and F1-score (99.6%) on the WELFake dataset. Similarly, this configuration is also produced promising results on the FakeNews-Prediction dataset, with an accuracy of 96.5%, precision of 97.8%, recall of 98.0%, and an F1-score of 97.8%. These results confirm that the proposed approach, optimized using Adam, provides a more robust feature representation and more effective learning compared to the other models.

**Figure 12:** Accuracy benchmarking and error loss convergence of the proposed and baseline models on the FakeNewsPrediction dataset for fake news detection

Traditional hybrid models, such as LR-SVM-GRU and SA-Bi-LSTM, provide only moderate predictive performance because their architectures rely heavily on manually engineered features and limited sequence modeling capabilities. LR-SVM-GRU combines simple linear and kernel-based classifiers with a GRU, which is effective for capturing short-term sequential patterns but struggles to retain long-range dependencies due to GRU's relatively shallow gating mechanism. Similarly, SA-Bi-LSTM improves upon basic LSTM by adding bidirectionality and self-attention, yet its contextual representations remain constrained by its shallow embedding layers and lack of the deep pre-training that modern transformer encoders provide. Transformer-based models like BERT + LightGBM and FakeBERT address some of these limitations by leveraging pre-trained language representations that capture bidirectional context and nuanced semantic relationships within the text. However, these architectures typically rely solely on the transformer's final embeddings for classification, often neglecting local sequential dependencies that may still be important for distinguishing subtle linguistic patterns typical of deceptive content. Furthermore, when used with additional learners like LightGBM, they require significant computational resources for gradient boosting, which adds complexity to training and inference without always guaranteeing complementary sequence-level refinements.

In contrast, the proposed T5-SA-LSTM framework uses the T5 transformer to perform advanced sequence-to-sequence encoding and decoding, which captures complex contextual relationships by treating the text as a generative problem rather than an embedding extraction task. The output from T5 undergoes an explicit self-attention mechanism that recalibrates the relative importance of different tokens within the encoded latent representation. This ensures that deceptive cues such as misleading qualifiers, contradictory statements or sudden sentiment shifts are dynamically weighted before entering the recurrent layer. Unlike standalone transformers that stop at contextual embeddings, the T5-SA-LSTM feeds these refined, self-attended features into a dedicated LSTM network. This LSTM excels at modeling sequential time dependencies, enabling the system to preserve the order of important context cues and handle longer input sequences more robustly. By stacking self-attention before the LSTM, the framework compensates for the LSTM's known limitations in direct long-range attention by providing it with a weighted, focused input sequence, enhancing its memory retention.

Additionally, the proposed model is optimized using the Adam optimizer, which adaptively adjusts individual learning rates for each parameter based on first and second moment estimates of the gradients. This optimization strategy not only accelerates convergence but also mitigates the risks of vanishing or exploding gradients that can degrade the performance of deep sequential architectures like LSTM. This synergy of advanced sequence-to-sequence contextualization, adaptive attention, and gated recurrence—all tuned with a robust optimizer—enables the proposed T5-SA-LSTM to outperform traditional and transformer-only baselines. It consistently achieves higher accuracy, precision, recall, and F1-scores on both the datasets by effectively detecting subtle manipulative patterns, misleading narratives and context inconsistencies in fake news that simpler hybrid models and vanilla transformers often overlook.

Figs. 13–15 present a comprehensive comparative analysis of the precision, recall, and F1-score results for the proposed T5-SA-LSTM model variants against several baseline models on the WELFake and FakeNewsPrediction datasets, respectively. Fig. 13 shows that on the WELFake dataset, the T5-SA-LSTM with Adam optimizer achieves the highest precision of 99.6%, outperforming its RMSProp (96.5%) and SGD (91.3%) variants. On the other hand, the competitive models: FakeBERT, BERT + LightGBM, SA-Bi-LSTM and LR-SVM-GRU achieved precision values of 95.6%, 93.0%, 88.3%–93.2% and 92.2%–96.6%, respectively. Similarly, on the FakeNewsPrediction dataset, T5-SA-LSTM (Adam) achieved a precision of 97.8%, outperforming RMSProp (91.0%) and SGD (86.1%), FakeBERT (92.6%), BERT + LightGBM (90.2%), SA-Bi-LSTM (89.3%–92.4%) and LR-SVM-GRU (88.5%–93.5%) as shown in Table 3. Fig. 14 further shows that for recall on WELFake dataset, T5-SA-LSTM (Adam) achieves the higher score of 99.8% than RMSProp

(95.2%), SGD (88.7%) and FakeBERT, BERT + LightGBM, SA-Bi-LSTM, and LR-SVM-GRU tat achieved scores in the 80.9%–97.4% range. On the other hand, for FakeNewsPrediction dataset, T5-SA-LSTM (Adam) has a recalls core of 98.0%, greater than RMSProp (90.2%) and SGD (84.0%), and other models that achieved scores in the 82.3%–95.6% range, as detailed in [Table 3](#).

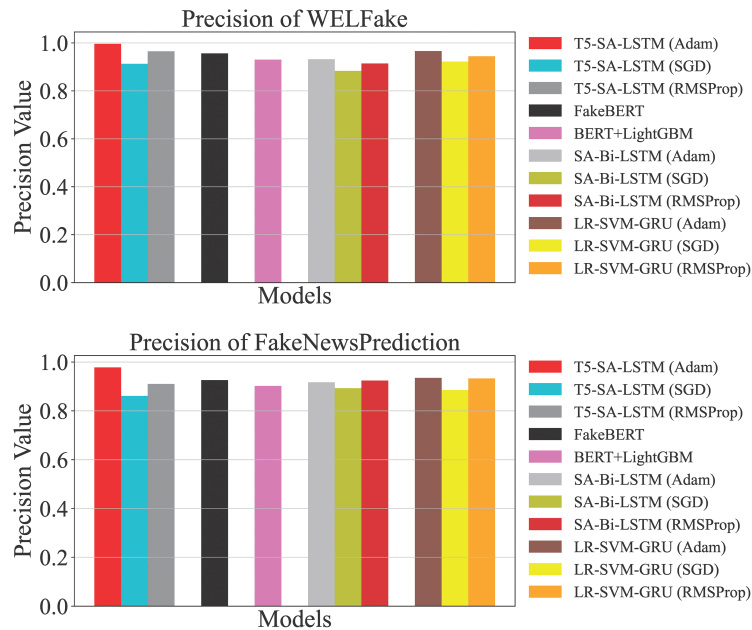


Figure 13: Comparative precision results of the T5-SA-LSTM and other models on WELFake and FakeNewsPrediction datasets

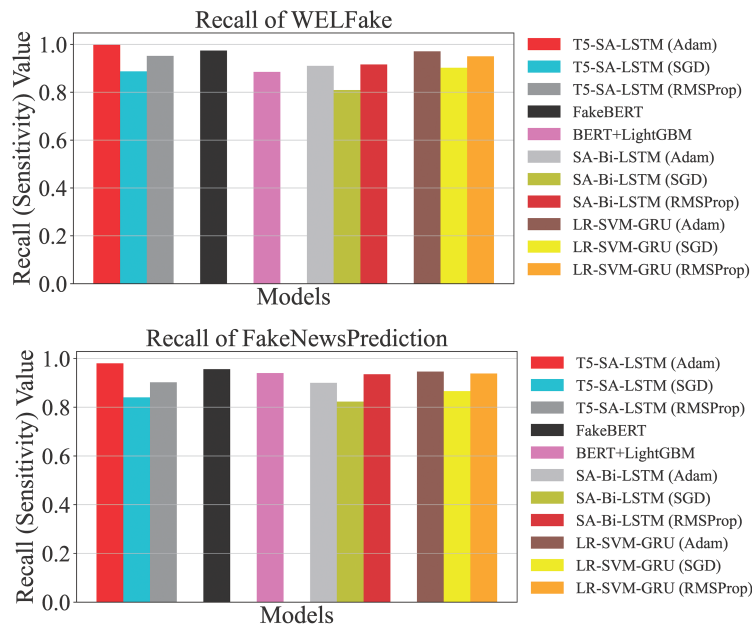


Figure 14: Comparative recall (sensitivity) results of the T5-SA-LSTM and other models on WELFake and FakeNewsPrediction datasets

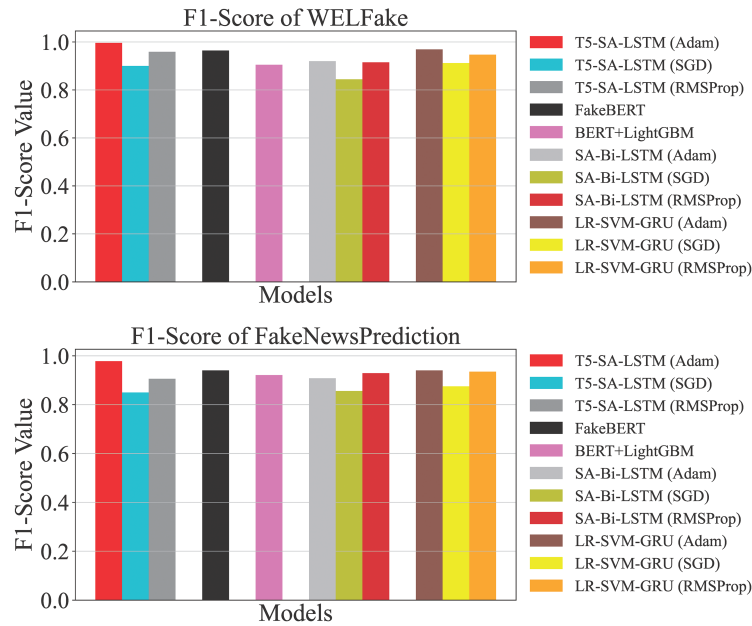


Figure 15: Comparative F1-Score results of the T5-SA-LSTM and other models on WELFake and FakeNewsPrediction datasets

Table 3 presents a detailed comparative analysis of compared approaches, highlighting their performance in terms of accuracy, error loss, precision, recall, and F1-score for detecting fake news. These evaluations have been conducted across two benchmark datasets, namely WELFake and FakeNewsPrediction, to demonstrate the robustness and generalizability of each method.

Finally, Fig. 15 illustrates the F1-scores for the WELFake dataset, where the T5-SA-LSTM (Adam) achieves the highest score of 99.6%, followed by RMSProp (95.9%) and SGD (90.0%). FakeBERT, BERT + LightGBM and SA-Bi-LSTM variants (Adam, RMSProp, SGD) achieved respective scores of 96.4%, 90.5%, 92%, 91.5%, 84.4% and 84.4%. On the FakeNewsPrediction dataset, the T5-SA-LSTM (Adam) again secures the highest F1-score of 97.8%, compared to RMSProp (90.6%) and SGD (85.0%), while FakeBERT and BERT + LightGBM attain 94.0% and 92.1% respective scores. The SA-Bi-LSTM variants scores are 90.8% (Adam), 92.9% (RMSProp), and 85.6% (SGD). LR-SVM-GRU variants achieve 94.0% (Adam), 93.5% (RMSProp), and 87.5% (SGD), scores as shown in Table 3.

These comprehensive results consistently demonstrate that the T5-SA-LSTM model optimized with Adam outperforms the counterpart models in terms of precision, recall, and F1-score as the performance metrics across the evaluated datasets.

4.3 Discussion

Based on the experimental analysis and results of this study, the proposed T5-SA-LSTM model demonstrated substantial efficacy in detecting fake news by combining transformer-based semantic abstraction with LSTM-driven temporal modeling.

The model demonstrated robust performance across both the WELFake and FakeNewsPrediction datasets as shown in Fig. 16a and b, respectively. On WELFake, it achieved 99% testing accuracy, correctly classifying 6200 fake and 6100 real articles, with minimal misclassifications (100 fake, 94 real). It attained a precision of 99.6%, recall of 99.8% and an F1-score of 99.6%. This suggests that the model effectively captures nuanced linguistic and contextual cues due to the semantic strength of the T5 encoder and the temporal

awareness of the LSTM. Similarly, on the FakeNewsPrediction dataset, the model achieved 96.5% accuracy, with 3000 fake and 2900 real samples correctly identified. With a precision value of 97.8%, recall value of 98.0% and F1-score of 97.8, it maintained promising generalization, indicating adaptability to varying data quality and linguistic patterns.

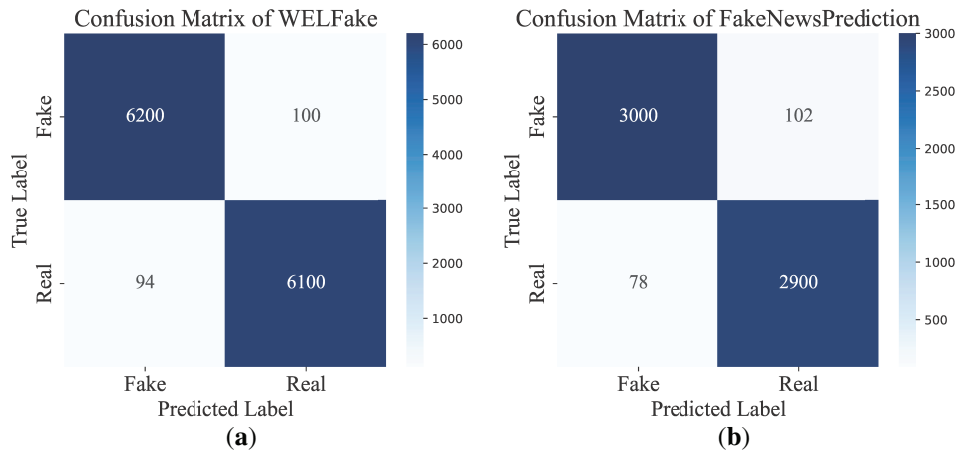


Figure 16: (a,b) Confusion matrices on the WELFake and FakeNewsPrediction datasets demonstrating strong classification performance with minimal misclassifications

While the T5-SA-LSTM (Adam) framework achieves high accuracy, its recall is slightly higher than its precision. This indicates that the model is highly effective at capturing the majority of fake news articles but may occasionally misclassify some true news as fake. This trade-off reduces the risk of fake news going undetected but could generate false alarms that require additional human review. To address this issue, adaptive threshold tuning, post-processing filters or human-in-the-loop systems may help refine predictions and minimize the impact of false positives. This ensures that the model remains reliable and practical for real-time applications such as live news feed moderation, automated fact-checking plugins for social media platforms, misinformation filtering in messaging apps or real-time content screening for online publishing. By combining high recall with appropriate safeguards, the model can support timely and trustworthy news verification further.

5 Conclusion and Future Work

Using T5 Transformer and Self-Attention-enhanced LSTM, the proposed T5-SA-LSTM framework demonstrates robustness in tackling fake electronic news detection problem. To ensure data quality and class balance, data preprocessing techniques such as the Tokenization and SMOTE were applied. These steps enhanced the representational capability of the model by addressing class imbalance and preparing the textual data for effective sequence modeling. The proposed framework was evaluated on two widely recognized benchmark datasets: WELFake and FakeNewsPrediction. On the WELFake dataset, the model achieved an accuracy of 99%, while on the FakeNewsPrediction dataset, it attained an accuracy of 96.5%. These results indicate the superior performance and reliability of the proposed model in detecting fake news across diverse sources over some benchmark and competitive models. Moreover, the model promising performance across two distinct datasets highlights its generalization ability, making it suitable for real-world deployment in news verification systems, social media platforms and digital journalism tools.

To further enhance the capability of the proposed model, future work may prioritize its expansion to a comprehensive multimodal system. This necessitates the integration of specialized modules for extracting

salient features from diverse contents, specifically visual and auditory data. Convolutional Neural Networks (CNNs) or Vision Transformers (ViTs) may be used for images features extraction, while a combination of 3D CNNs and video Transformers will process temporal and spatial information from videos. These extracted multimodal features may then be synergistically used with textual representations, leveraging advanced techniques such as cross-modal attention. This will enable the model to identify discrepancies and inconsistencies across a diverse set of modalities. This will result in a framework capable of detecting misinformation embedded within complex multimedia narratives, thereby significantly advancing the state-of-the-art in automated fake news detection.

Acknowledgement: This research is supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R195), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors are thankful to the Deanship of Graduate Studies and Scientific Research at University of Bisha for supporting this work through the Fast-Track Research Support Program.

Funding Statement: This research is supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R195), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author Contributions: Hamza Murad Khan performed investigation of the research. Anwar Khan provided validation and supervision. Shakila Basheer and Mohammad Tabrez Quasim refined the draft of the manuscript and its overall flow and also provided resources. Raja`a Al-Naimi and Vijaykumar Varadarajan contributed in responding to reviewers' comments. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Both the WELFake and FakeNewsPrediction datasets were selected from reliable sources on the Kaggle database. These datasets are freely accessible online, as referenced in source [52].

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Jouhar J, Pratap A, Tijo N, Mony M. Fake news detection using Python and machine learning. *Procedia Comput Sci.* 2024;233(1):763–71. doi:10.1016/j.procs.2024.03.265.
2. Qu Z, Zhou F, Song X, Ding R, Yuan L, Wu Q. Temporal enhanced multimodal graph neural networks for fake news detection. *IEEE Trans Comput Soc Syst.* 2024;11(6):7286–98. doi:10.1109/TCSS.2024.3404921.
3. Lian Z, Zhang C, Su C, Ali Dharejo F, Almutiq M, Memon MH. FIND: privacy-enhanced federated learning for intelligent fake news detection. *IEEE Trans Comput Soc Syst.* 2024;11(4):5005–14. doi:10.1109/TCSS.2023.3304649.
4. Lin ZH, Wang Z, Zhao M, Song Y, Lan L. An AI-based system to assist human fact-checkers for labeling Cantonese fake news on social media. In: 2022 IEEE International Conference on Big Data (Big Data); 2022 Dec 17–20. Osaka, Japan: IEEE; 2022. p. 6766–20. doi:10.1109/BigData55660.2022.10020949.
5. Li Z, Liu J. Fake and untrue news dataset (FUND): an expanded dataset for fake news classification. In: 2023 IEEE 3rd International Conference on Computer Systems (ICCS); 2023 Sep 22–24; Qingdao, China. doi:10.1109/ICCS59700.2023.10335496.
6. Akhtar P, Ghouri AM, Khan HUR, Amin Ul Haq M, Awan U, Zahoor N, et al. Detecting fake news and disinformation using artificial intelligence and machine learning to avoid supply chain disruptions. *Ann Oper Res.* 2022;327(2):1–25. doi:10.1007/s10479-022-05015-5.
7. Souto Moreira L, Machado Lunardi G, de Oliveira Ribeiro M, Silva W, Paulo Basso F. A study of algorithm-based detection of fake news in Brazilian election: is BERT the best. *IEEE Lat Am Trans.* 2023;21(8):897–903. doi:10.1109/TLA.2023.10246346.
8. Balshetwar SV, Rs A, Dani JR. Fake news detection in social media based on sentiment analysis using classifier techniques. *Multimed Tools Appl.* 2023;2023(23):1–31. doi:10.1007/s11042-023-14883-3.

9. Papadopoulos P, Spythouris D, Markatos EP, Kourtellis N. FNDaaS: content-agnostic detection of websites distributing fake news. In: 2023 IEEE International Conference on Big Data (BigData); 2023 Dec 15–18; Sorrento, Italy. doi:10.1109/BigData59044.2023.10386830.
10. Hu B, Mao Z, Zhang Y. An overview of fake news detection: from a new perspective. *Fundam Res.* 2024;5(1):332–46. doi:10.1016/j.fmre.2024.01.017.
11. Oad A, Hamza Farooq M, Zafar A, Ayesha Akram B, Zhou R, Dong F. Fake news classification methodology with enhanced BERT. *IEEE Access.* 2024;12(1):164491–502. doi:10.1109/access.2024.3491376.
12. Baashirah R. Zero-shot automated detection of fake news: an innovative approach (ZS-FND). *IEEE Access.* 2024;12(6):182828–40. doi:10.1109/access.2024.3462151.
13. Wang J, Qian S, Hu J, Hong R. Positive unlabeled fake news detection via multi-modal masked transformer network. *IEEE Trans Multimed.* 2023;26(3):234–44. doi:10.1109/TMM.2023.3263552.
14. Wang J, Qian S, Hu J, Hong R. Comment-context dual collaborative masked transformer network for fake news detection. *IEEE Trans Multimed.* 2023;26:5170–80. doi:10.1109/TMM.2023.3330074.
15. Mouratidis D, Kanavos A, Kermanidis K. From misinformation to insight: machine learning strategies for fake news detection. *Information.* 2025;16(3):189. doi:10.3390/info16030189.
16. Toor MS, Shahbaz H, Yasin M, Ali A, Fitriyani NL, Kim C, et al. An optimized weighted-voting-based ensemble learning approach for fake news classification. *Mathematics.* 2025;13(3):449. doi:10.3390/math13030449.
17. Roumeliotis KI, Tselikas ND, Nasiopoulos DK. Fake news detection and classification: a comparative study of convolutional neural networks, large language models, and natural language processing models. *Future Internet.* 2025;17(1):28. doi:10.3390/fi17010028.
18. Yadav A, Gaba S, Khan H, Budhiraja I, Singh A, Singh KK. ETMA: efficient transformer-based multilevel attention framework for multimodal fake news detection. *IEEE Trans Comput Soc Syst.* 2024;11(4):5015–27. doi:10.1109/TCSS.2023.3255242.
19. Jian W, Li JP, Akbar MA, Haq AU, Khan S, Alotaibi RM, et al. SA-Bi-LSTM: self attention with bi-directional LSTM-based intelligent model for accurate fake news detection to ensured information integrity on social media platforms. *IEEE Access.* 2024;12:48436–52. doi:10.1109/access.2024.3382832.
20. Kaliyar RK, Goswami A, Narang P, Chamola V. Understanding the use and abuse of social media: generalized fake news detection with a multichannel deep neural network. *IEEE Trans Comput Soc Syst.* 2024;11(4):4878–87. doi:10.1109/TCSS.2022.3221811.
21. Jing J, Wu H, Sun J, Fang X, Zhang H. Multimodal fake news detection via progressive fusion networks. *Inf Process Manag.* 2023;60(1):103120. doi:10.1016/j.ipm.2022.103120.
22. Yadav A, Gupta A. An emotion-driven, transformer-based network for multimodal fake news detection. *Int J Multimed Inf Retr.* 2024;13(1):7. doi:10.1007/s13735-023-00315-3.
23. Lin SY, Chen YC, Chang YH, Lo SH, Chao KM. Text-image multimodal fusion model for enhanced fake news detection. *Sci Prog.* 2024;107(4):368504241292685. doi:10.1177/00368504241292685.
24. Wang H, Guo J, Liu S, Chen P, Li X. SEPM: multiscale semantic enhancement-progressive multimodal fusion network for fake news detection. *Expert Syst Appl.* 2025;283(1):127741. doi:10.1016/j.eswa.2025.127741.
25. Pillai SEVS, Avacharmal R, Reddy RA, Kumar Pareek P, Zanke P. Transductive-long short-term memory network for the fake news detection. In: Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE); 2024 Apr 26–27; Ballari, India. doi:10.1109/ICDCECE60827.2024.10548223.
26. Jadhav M, Patil M, Giri P, Hande Y. Advancing fake news detection: a comparative analysis of SVM and LR models and prospects for dynamic model updating. In: 2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI); 2024 Nov 29–30; Wardha, India. doi:10.1109/IDICAIEI61867.2024.10842732.
27. Dev DG, Bhatnagar V. Hybrid RFSVM: hybridization of SVM and random forest models for detection of fake news. *Algorithms.* 2024;17(10):459. doi:10.3390/a17100459.
28. Zhao J, Zhao Z, Shi L, Kuang Z, Liu Y. Collaborative mixture-of-experts model for multi-domain fake news detection. *Electronics.* 2023;12(16):3440. doi:10.3390/electronics12163440.

29. Rahim LAA, Abu Samah KAF, Aminuddin R. A conceptual hybrid model for fake review detection using implicit ABSA and imbalanced data. In: 2024 IEEE 12th Conference on Systems, Process & Control (ICSPC); 2024 Dec 7; Malacca, Malaysia. doi:10.1109/ICSPC63060.2024.10862280.
30. Umah B, Gupta D, Venugopalan M. Investigating fake job descriptions with TF-IDF and word embeddings. In: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT); 2024 Jun 24–28; Kamand, India. doi:10.1109/ICCCNT61001.2024.10725356.
31. Kathiravan P, Shanmugavadivu P, Saranya R. Mitigating imbalanced data in online social networks using stratified K-means sampling. In: 2023 8th International Conference on Business and Industrial Research (ICBIR); 2023 May 18–19; Bangkok, Thailand. doi:10.1109/ICBIR57571.2023.10147677.
32. Mujahid M, Kina E, Rustam F, Villar MG, Alvarado ES, De La Torre Diez I, et al. Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering. *J Big Data*. 2024;11(1):87. doi:10.1186/s40537-024-00943-4.
33. Azhar NA, Pozi MSM, Din AM, Jatowt A. An investigation of SMOTE based methods for imbalanced datasets with data complexity analysis. *IEEE Trans Knowl Data Eng*. 2023;35(7):6651–72. doi:10.1109/TKDE.2022.3179381.
34. Chen W, Yang K, Yu Z, Shi Y, Philip Chen CL. A survey on imbalanced learning: latest research, applications and future directions. *Artif Intell Rev*. 2024;57(6):137. doi:10.1007/s10462-024-10759-6.
35. Kavuri K, Reddy TR, Gelli A, Priyanka BHDD. A hybrid approach for language variety prediction using BERT and T5 embeddings. In: 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT); 2025 Feb 5–7; Bengaluru, India. doi:10.1109/IDCIOT64235.2025.10914742.
36. Habu R, Ratnaparkhi R, Askhedkar A, Kulkarni S. A hybrid extractive-abstractive framework with pre & post-processing techniques to enhance text summarization. In: 2023 13th International Conference on Advanced Computer Information Technologies (ACIT); 2023 Sep 21–23; Wrocław, Poland. doi:10.1109/ACIT58437.2023.10275584.
37. Sahin F, Amasyali MF. Sentence detailing and its applications. In: 2023 Innovations in Intelligent Systems and Applications Conference (ASYU); 2023 Oct 11–13; Sivas, Turkiye. doi:10.1109/ASYU58738.2023.10296831.
38. Shakil H, Farooq A, Kalita J. Abstractive text summarization: state of the art, challenges, and improvements. *Neurocomputing*. 2024;603(13):128255. doi:10.1016/j.neucom.2024.128255.
39. Masih S, Hassan M, Fahad LG, Hassan B. Transformer-based abstractive summarization of legal texts in low-resource languages. *Electronics*. 2025;14(12):2320. doi:10.3390/electronics14122320.
40. Jamshidi B, Hakak S, Lu R. A self-attention mechanism-based model for early detection of fake news. *IEEE Trans Comput Soc Syst*. 2024;11(4):5241–52. doi:10.1109/TCSS.2023.3322160.
41. Chen J, Zhang T, Yan Z, Zheng Z, Zhang W, Zhang J. Attention-based BiLSTM with positional embeddings for fake review detection. *J Big Data*. 2025;12(1):83. doi:10.1186/s40537-025-01130-9.
42. Kamal A, Abulaish M. Contextualized satire detection in short texts using deep learning techniques. *J Web Eng*. 2024;23(1):27–52. doi:10.13052/jwe1540-9589.2312.
43. Sun P, Bi W, Zhang Y, Wang Q, Kou F, Lu T, et al. Fake review detection model based on comment content and review behavior. *Electronics*. 2024;13(21):4322. doi:10.3390/electronics13214322.
44. Shaeri P, Katanforoush A. A semi-supervised fake news detection using sentiment encoding and LSTM with self-attention. In: 2023 13th International Conference on Computer and Knowledge Engineering (ICCKE); 2023 Nov 1–2; Mashhad, Iran. doi:10.1109/ICCKE60553.2023.10326287.
45. Keya AJ, Shajeeb HH, Rahman MS, Mridha MF. FakeStack: hierarchical Tri-BERT-CNN-LSTM stacked model for effective fake news detection. *PLoS One*. 2023;18(12):e0294701. doi:10.1371/journal.pone.0294701.
46. Aljrees T, Cheng X, Ahmed MM, Umer M, Majeed R, Alnowaiser K, et al. Fake news stance detection using selective features and FakeNET. *PLoS One*. 2023;18(7):e0287298. doi:10.1371/journal.pone.0287298.
47. Rupapara V, Rustam F, Amaar A, Washington PB, Lee E, Ashraf I. Deepfake tweets classification using stacked Bi-LSTM and words embedding. *PeerJ Comput Sci*. 2021;7(4):e745. doi:10.7717/peerj-cs.745.
48. Babar M, Ahmad A, Tariq MU, Kaleem S. Real-time fake news detection using big data analytics and deep neural network. *IEEE Trans Comput Soc Syst*. 2024;11(4):5189–98. doi:10.1109/TCSS.2023.3309704.

49. Fake news detection datasets: WELFake [Internet]. [cited 2025 Aug 18]. Available from: <https://www.kaggle.com/datasets/studymart/welfake-dataset-for-fake-news>.
50. FakeNewsPrediction [Internet]. [cited 2025 Aug 18]. Available from: <https://www.kaggle.com/datasets/lokeshmendake/fake-news-prediction-datasets>.
51. Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science*. 2018;359(6380):1146–51. doi:10.1126/science.aap9559.
52. Zhou X, Zafarani R. A survey of fake news: fundamental theories, detection methods, and opportunities. *ACM Comput Surv*. 2020;53(5):1–40. doi:10.1145/3395046.