Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.076156

ARTICLE

Deep-Learning Approaches to Text-Based Verification for Digital and Fake News Detection

Raed Alotaibi[1,*], Muhammad Atta Othman Ahmed[2], Omar Reyad[3,4,*] and Nahla Fathy Omran[5]

[1]Applied College, Shaqra University, Shaqra, Saudi Arabia
[2]Faculty of Computers and Information, Luxor University, Luxor, Egypt
[3]College of Computing and Information Technology, Shaqra University, Shaqra, Saudi Arabia
[4]Faculty of Computers and Artificial Intelligence, Sohag University, Sohag, Egypt
[5]Faculty of Computers and Information, Qena University, Qena, Egypt

*Corresponding Authors: Raed Alotaibi. Email: alhafi@su.edu.sa; Omar Reyad. Email: oreyad@su.edu.sa

Received: 15 November 2025; Accepted: 05 February 2026; Published: 09 April 2026

ABSTRACT: The widespread use of social media has made assessing users' tastes and preferences increasingly complex and important. At the same time, the rapid dissemination of misinformation on these platforms poses a critical challenge, driving significant efforts to develop effective detection methods. This study offers a comprehensive analysis leveraging advanced Machine Learning (ML) techniques to classify news articles as fake or true, contributing to discourse on media integrity and combating misinformation. The suggested method employed a diverse dataset encompassing a wide range of topics. The method evaluates the performance of five ML models: Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), Long Short-Term Memory networks (LSTMs), Decision Trees (DTs), and Support Vector Machines with Radial Basis Function (SVM-RBF) kernels. The presented methodology included thorough data preprocessing, detailed parameter tuning during model training, and robust statistical analyses to ensure fair and accurate performance comparisons. The results demonstrate that the combination of Term Frequency-Inverse Document Frequency (TF-IDF) with ANN and CNN achieved the highest accuracy of 99.13%, showcasing the effectiveness of these approaches in text-based news classification. The LSTM model followed closely with an accuracy of 98.59%, while the DT and SVM-RBF models achieved accuracies of 85.67% and 90.22%, respectively. These findings highlight the superior performance of deep learning (DL) models when combined with effective feature extraction techniques such as TF-IDF. The models offer practical utility and show promising potential for integration into editorial workflows to facilitate pre-publication news verification. Furthermore, statistical test methods such as Analysis of Variance (ANOVA) and Tukey's Honestly Significant Difference (HSD) tests are also performed. The obtained results clarify significant performance differences among the evaluated models, highlighting their unique capabilities and comparative strengths in the context of fake news detection. Hence, the presented study reinforces the importance of artificial intelligence based tools in promoting media reliability and provides a foundation for future advancements in automated misinformation detection systems.

KEYWORDS: Text verification; fake news detection; machine learning; deep learning; CNN; LSTM; news classification systems; misinformation detection

1 Introduction

In the digital age, the news media landscape has undergone a transformative expansion, offering a wide array of sources that serve as chroniclers of people times, shape public opinion, and support the democratic process and crucial pillars supporting the democratic process [1]. This proliferation brings significant challenges for verifying the authenticity of disseminated information [2]. Misinformation is widespread, often fueled by clickbait headlines and polarizing content shared on social media and other platforms [3]. The significant impact of news on our lives is evident in recent global events [4]. Whether it's staying informed about the pandemic or following the stock market rally, people heavily depend on the news. Studies have found that much of the news on social media is intentionally misleading and is generally categorized under the broader term "fake news" [5]. Fake news is false information presented as news with the intent to mislead. It has been confirmed that spreading fake news impacts sociopolitical domains, human behavior, and the sovereignty of a country [6]. It has also been noted that Artificial Intelligence (AI) is a new tool for detecting and verifying fake news on social media platforms in a short time [7]. In the same line, sophisticated technologies should be used to stop the spread of fake news [8]. The increased use of AI in journalism marks a pivotal shift toward addressing these challenges, making a new era of news creation, dissemination, and consumption [9]. AI now parses large news streams, detects emerging patterns, and can even draft stories, expanding the scale and speed of reporting while raising new risks to accuracy and reliability. Unlike many existing studies that rely on single-run performance metrics, this research establishes a statistically significant benchmark by validating model efficacy across 20 independent runs. The robustness of our results is confirmed through rigorous ANOVA testing, yielding a substantial F-statistic of 2981.91, thereby offering a level of validation and reproducibility often missing in standard deep learning evaluations. The presented study addresses this risk—the spread of fake news—by applying multiple machine learning models to distinguish authentic content from fabrications. Although LSTM models are generally effective for sequential data, their sequence-modeling capability is not leveraged in this specific experimental setting due to the use of TF-IDF features. The results show high-confidence classification of genuine articles, underscoring AI's practical role in news verification and support of journalistic integrity. Integrating the model into practices is a key step in automating pre-publication checks and enhancing media integrity.

The next sections of this paper is structured as follows. Section 2 delves into existing literature and bridging past research with the present inquiry into AI's application in news validation. Section 3 outlines computational approach and analytical design, detailing the strategies employed to assess the efficacy of ML models in news classification. The discussion on the experimental hypothesis and modeling setup is prepared in Section 4, which outlines the research hypotheses and the structured framework created. Section 5 presents the empirical results. Finally, Section 6 summarizes key findings and suggests directions for future research in the critical and evolving field.

2 Related Work

A recent survey in [10] consolidates empirical fake-news detection studies across major Natural Language Processing (NLP) and applied-AI venues, contrasting classical ML with deep models and focusing on reported algorithms and metrics; it exposes gaps in modality coverage and robustness evaluation, robustness-aware design. ZoFia is a two-stage zero-shot framework that (i) uses Hierarchical Salience with a Soft Cosine with Maximal Marginal Relevance (SC-MMR) selector to retrieve fresh external evidence and (ii) conducts a multi large language model, role-based debate to deliver interpretable judgments; it outperforms zero-shot and many few-shot baselines on two public datasets [11]. The co-attention mechanism in a Combined Graph neural network model (CMCG) fuses two Graph Neural Network (GNN) streams—user profiles and user preferences (news content, user history, and sharing cascades)—linked via a co-attention module to model who shares and what is shared, achieving 98.53% on GossipCop and 96.77% on PolitiFact, underscoring the benefits of profile–preference–propagation fusion over content-only baselines [12]. Tajrian et al. present a structured review of fake-news research, organizing it along two axes: how fake news is analyzed (knowledge, style, propagation, and source) and how it is detected (manual vs. AI-driven automatic methods) [13]. They also examine political news through sentiment analysis, showing how media framing shapes public perception. Furthermore, deep learning has introduced novel approaches to news classification. Zhou and Zafarani [14] used deep neural networks to distinguish between real and fake news stories. Their model demonstrates remarkable accuracy, highlighting the potential of deep learning in enhancing the reliability of news shared on social media platforms. Another pivotal area of investigation within the realm of news classification is the challenge posed by imbalanced datasets, particularly in non-English languages. For Bangla news classification, Hasib et al. balance a 437,948-item corpus using Random Under-Sampling (RUS) and Synthetic Minority Oversampling Technique (SMOTE), then benchmark classical ML (logistic regression, decision tree, Stochastic Gradient Descent (SGD)) against ANN, CNN, and Bidirectional Encoder Representations from Transformers (BERT) [15]. BERT tops performance with 99.04% accuracy on the balanced set vs. 72.23% on the imbalanced set, underscoring both the benefit of rebalancing and the advantage of transformer models [15]. Authors in [16] presented a study focusing on the classification, detection, and sentiment analysis of digital news using ML techniques. Their research addressed several challenges associated with digital news. Their findings show the Fake News Detection Model, with an accuracy of 87%, and a sentiment analysis model, with 89% accuracy. Kumar et al. benchmark classic and neural models (Naive Bayes, SVM, Passive-Aggressive, Random Forest, Logistic Regression, LSTM, BERT) on the ISOT dataset (44,898 samples) for fake-news detection, finding SVM achieves the best precision (99.88%), narrowly ahead of Random Forest and Passive-Aggressive [17]. Mishra and Sadia deliver a comprehensive survey of fake-news detection, synthesizing ML and DL approaches across major social platforms while cataloging datasets, features, and model families. They distill strengths and limitations—e.g., generalization gaps, evolving topics, and modality/propagation cues—providing a clear roadmap of effective techniques and open challenges [18].

Recent developments in the field have been significantly shaped by the advent of Transformer-based architectures. Vaswani et al. introduced the attention mechanism [19], which paved the way for models like BERT [20]. Their study utilized several ML and deep learning techniques to categorize news articles accurately. Their findings demonstrated remarkable results, achieving an accuracy of up to 94%.

3 Methodology

The presented work explores various machine learning models to address the nuanced task of classifying text data into fake and true news articles. The suggested approach highlights the use of multiple computational techniques as shown in Fig. 1. Initially, introducing the foundational concepts of ANN, which mimic the human brain's neural structure to capture complex relationships within the data. Moving forward, elucidating the CNN as a powerhouse in processing grid-like topologies of data, or in the case of this study, vector-based features extracted from text, by learning spatial hierarchies of features. The methodology further extends to LSTMs. Showcasing their exceptional ability to retain information across extended sequences, these models are especially well-suited for tasks involving time-series analysis or the processing of sequential text data.

Additionally, the DT model is explored, which provides a transparent decision-making path through a tree-like structure of decisions and outcomes. Lastly, the SVM-RBF is investigated, renowned for its effectiveness in high-dimensional spaces and non-linear data separability. Each model offers unique perspectives and mechanisms for feature extraction and classification that allow for the assesment of their efficacy individually and collectively. This diversified methodology not only enhances the generalization of this study but also contributes novel insights into the robustness of ML applications to discern the integrity of news articles. As depicted in Fig. 1, the process begins with data input and splitting into training, validation, and testing sets. TF-IDF embedding is used for feature extraction. Five models—ANN, CNN, LSTM, DT, and SVM-RBF—perform classification. Outputs are labeled as real or fake news based on evaluation metrics.

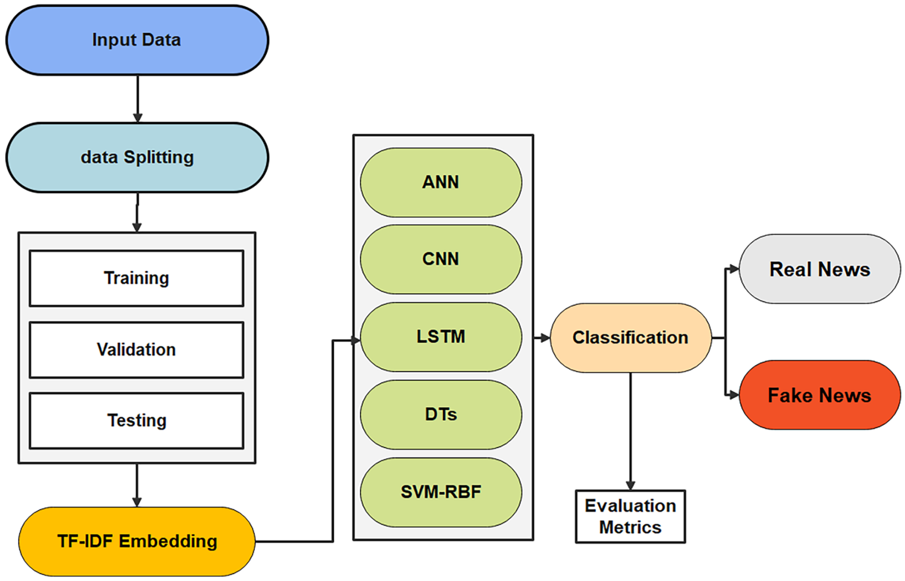


Figure 1: Experimental pipeline overview.

Artificial intelligence tools such as ChatGPT are used in a limited and responsible manner during the writing process of this work. The tools usage is used solely for enhancing the academic style and refining sentence structures. These tools did not contribute to the model design, data analysis, or interpretation of the obtained results. All the methodological decisions, experiments and scientific conclusions were made exclusively by the manuscript authors.

3.1 Dataset Understanding and Preprocessing

The Fake and Real News dataset [21] provides the necessary information to assess the integrity of news articles. The dataset is organized in .csv format and comprises two main classes: Fake and True datasets. The fake data contains 17,903 different values for the title class, while the true data has 20,826 unique values for the title class and 21,192 for the text class. Fig. 2 illustrates the frequency distribution of the fake and true class labels, indicating minimal differences between the two classes.

The overall stability of the dataset is confirmed, with 52.3% of the data labeled as true and 47.7% as fake. Additionally, Fig. 3 presents the distribution of fake news across various subjects, showing that political news has the highest count of fake articles. It demonstrates that fake news articles tend to be longer in length compared to true articles. It is mainly in terms of word count for article titles. Differentiating between real and fake articles can be more challenging considering the main text field.

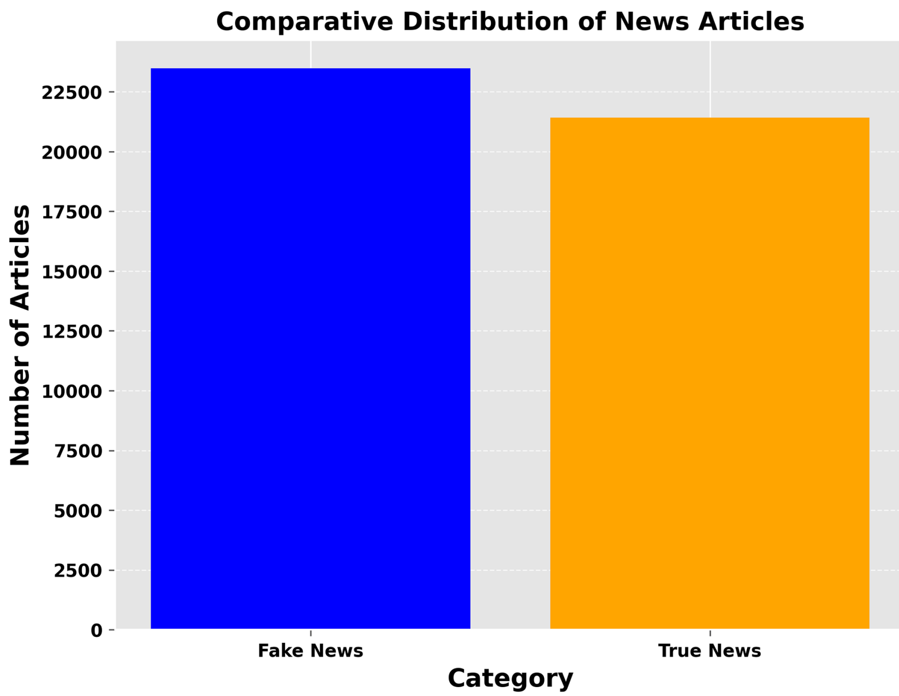


Figure 2: Category-wise distribution of news articles indicating a roughly balanced dataset for classification.

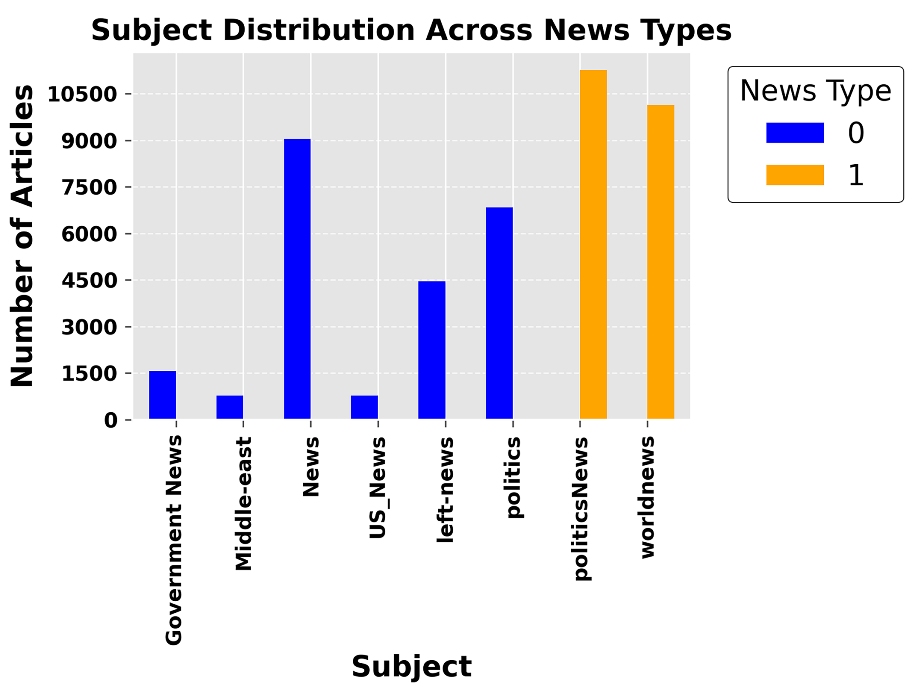
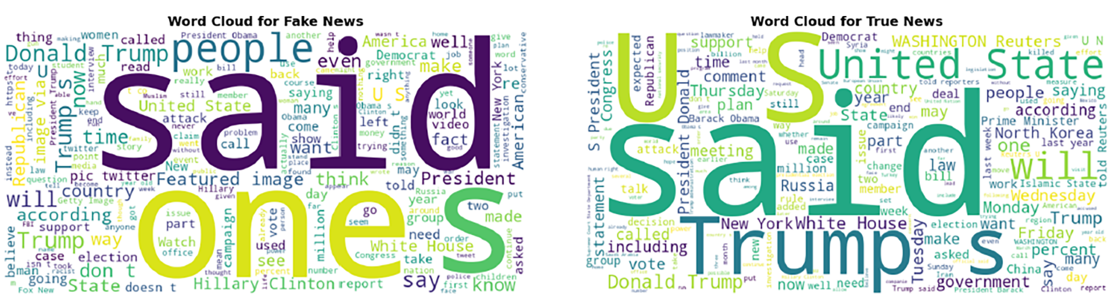


Figure 3: Subject distribution across news categories. The chart compares the number of fake (blue) and true (orange) articles across various subjects. Politics and world news dominate, while government and regional news have lower representation.

3.1.1 Exploratory Data Analysis and Visualization

Exploratory data analysis (EDA) is a critical first step in the data science workflow, bridging data collection and more sophisticated inferential or predictive analytics. In news classification, EDA is paramount as it unveils underlying structures, detects outliers, and extracts insightful patterns that can inform the direction of further analytical endeavors. Leveraging the vast informational expanse of the used dataset, which comprises 44,898 articles from critical periods in contemporary history, it is embarked on a journey to illuminate the multifaceted nature of news narratives. The mathematical underpinnings of the performed analysis are rooted in probability and statistics. By visualizing the frequency distribution of fake vs. true news, this not only quantify the prevalence of each category but also apply a visual lens to discern the symmetry, skewness, and kurtosis of data distribution. This forms the bedrock for hypothesis testing and ML models that follow, imbuing an analysis with robustness derived from mathematical rigour. A bar plot illustrates the dataset's overall fake and true news article distribution (Fig. 2). This visualization is key to understanding each news type's relative balance and prevalence inside the corpus. The subject distribution among the articles was also analyzed. A bar plot captures the frequency of subjects for both fake and true news articles, providing insights into the topics that are more prevalent in one category over the other is shown in Fig. 3. Fig. 4 presents side-by-side word clouds for fake and true news, where subfigure 4a emphasizes sensationalism, while subfigure 4b highlights authenticity anchors in true reporting. These clouds give a quick qualitative sense of the rhetoric and language patterns that characterize both types of news. As demonstrated in Fig. 5a, the distribution of title lengths across the corpus of news articles reveals a central tendency with a peak around 50–60 characters, suggesting a common editorial preference.

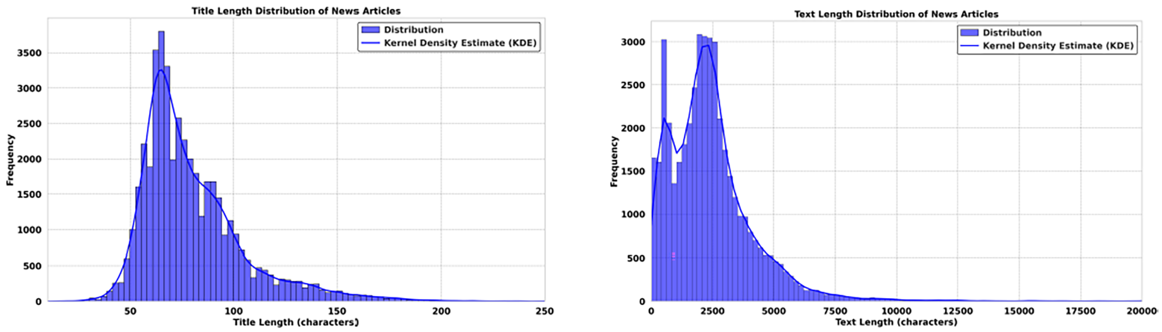
The distribution's tail, extending towards longer title lengths, illustrates the diversity in title construction within the dataset. Understanding these patterns is crucial, as title length can influence reader engagement and reflect on the stylistic choices differentiating genuine news from disinformation. Fig. 5b presents the distribution of text lengths within the used dataset, highlighting a solid preference for conciseness in news articles. The peak in the lower range of the histogram indicates that shorter articles are more prevalent, potentially reflecting editorial standards or reader engagement metrics that favour brevity. The long tail toward the higher length values suggests less frequent but more in-depth reporting in the dataset. Analyzing the length distribution helps in understanding the dataset's composition and can also serve as an indicator of content depth and complexity. Bar chart 2 showcases the balance between fake and true news articles within the dataset, providing a stark visual representation of their relative frequencies. Bar chart 3 illustrates the prevalence of subjects across fake and true news, highlighting thematic disparities and their correlation with news authenticity. It offers a concise overview of topic susceptibility to misinformation, enhancing the grasp of news dissemination dynamics.



(a) Word cloud reveals the linguistic fabric of fake news data.

(b) Word cloud showcases the terms most common in true news data.

Figure 4: Comparative word clouds illustrating linguistic patterns in (**a**) fake and (**b**) true news datasets.



(a) Distribution of news title lengths (with KDE). **(b)** Distribution of article text lengths (with KDE).

Figure 5: Length distributions across the dataset: (**a**) titles and (**b**) full texts. Shorter lengths dominate, yet long-form tails persist, reflecting a balance between brevity and comprehensive coverage.

Fig. 5 presents side-by-side length distributions for titles and full texts. Subfigure 5a shows a pronounced tendency toward succinct titles designed to capture attention quickly; the Kernel Density Estimation (KDE) curve traces a smooth frequency profile that underscores this brevity while allowing for occasional longer titles to convey complex narratives. Subfigure 5b illustrates the broader variability in article body lengths. The distribution peaks at shorter texts and decays gradually, with the KDE highlighting concentration around median lengths and a long tail of in-depth reports. Together, these patterns reflect editorial strategies that balance readability and engagement with the need for comprehensive coverage.

Fig. 6 summarizes corpus-level sentiment patterns for fake vs. true news (scores –1 to +1). The histogram shows fake articles clustering around mildly negative sentiment, while true news exhibits a wider, more balanced spread—consistent with varied reporting tones. The boxplot reinforces this: fake news skews slightly negative, suggesting emotionally charged or sensational language, whereas true news spans neutral to positive and negative. These contrasts point to systematic emotional manipulation in fake content and offer cues for authenticity assessment. The analyses provide a compact baseline for subsequent pattern mining, leveraging conventional statistics with modern visualization to surface structure in the data.

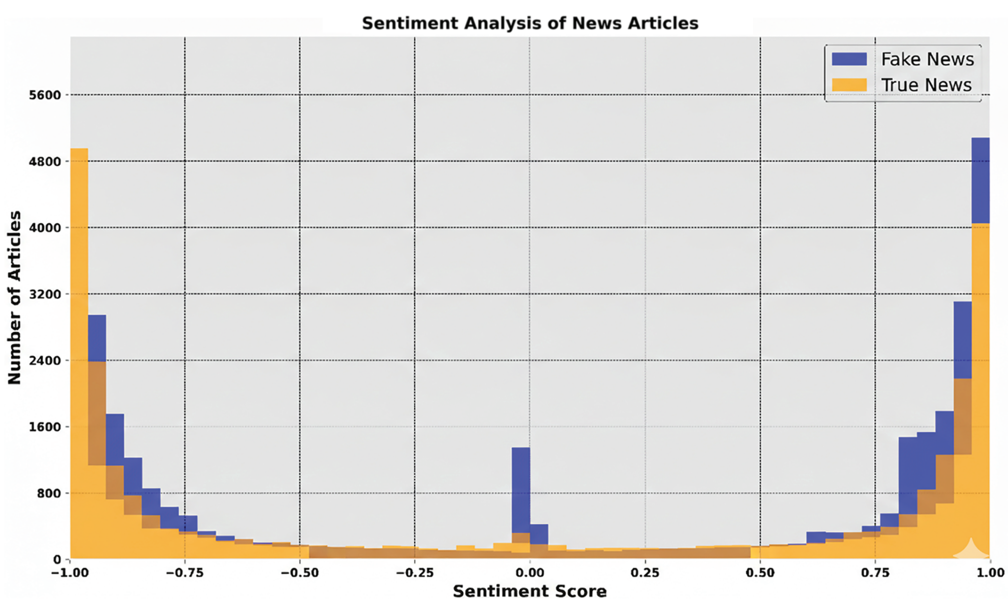


Figure 6: Sentiment analysis of news articles.

3.1.2 Tokenization

The dataset central to this investigation consists of a meticulously curated collection of news articles, each tagged with a 'fake' or 'true' label. These articles encompass a broad spectrum of topics meticulously selected to ensure equitable representation across categories. The preprocessing phase for model training involved steps such as tokenization, elimination of stop words, and data vectorization. This included applying the TF-IDF technique for feature representation. The authors in [22] capture the meaning of each term across the corpus, aiming to optimize the textual data for ML applications. The dataset was used in its complete form, comprising a distribution of 'fake' and 'true' articles, totalling 44,898 entries. While not perfectly equal, the dataset is sufficiently balanced to mitigate significant bias in model training and evaluation.

3.2 TF-IDF Feature Extraction

This study employs a methodical method to explore detecting fake news on digital platforms using NLP techniques. News is collected from publicly available datasets, ensuring a balanced representation of fake and real news. Text preprocessing removes stopwords, punctuations, and irrelevant characters, followed by tokenization and stemming to standardize terms. TF-IDF is utilized to transform text data into numerical feature vectors, reflecting the importance of each term within the overall corpus. Term frequency emphasizes how often a word appears in a specific news article, while inverse document frequency downplays the impact of frequently occurring terms. The TF-IDF method quantifies a term's importance in a document relative to a corpus. Term Frequency (TF) measures the ratio of a term's count in a document to the total terms in that document. Inverse Document Frequency (IDF) is computed as $\log\left(\frac{N}{1+DF(t)}\right)$, where N is the total number of documents, and $DF(t)$ is the document frequency of the term. The TF-IDF score is calculated by multiplying TF with IDF, giving higher importance to words that appear frequently in a document but less commonly across the entire corpus. This generates a feature matrix that serves as input for training and evaluating classification models. To enhance model performance and minimize overfitting, hyperparameter tuning is applied. The study highlights the effectiveness of TF-IDF in uncovering unique patterns associated with the spread of fake news on social media platforms.

3.3 Classification and Taxonomy Modelling

An ANN is structured with layers of interconnected nodes resembling the neural networks in the human brain [23]. It consists of input (I), hidden (H), and output (O) layers. Activation functions introduce non-linearity, which is essential for learning complex patterns.

$$s = \sum_k W_k \times I_k + B \tag{1}$$

Eq. (1) represents the weighted sum, where W_k denotes the weight associated with the k^{th} input node I_k, and B is the bias term. The loss function calculates the Mean Squared Error (MSE) over n instances, comparing actual Y_i and predicted $\hat{Y}_i$ outputs, guiding the network's training phase.

CNNs are built to handle data with a grid-like structure—such as images or sequential text—by learning spatial feature hierarchies through a series of layers, including convolutional, pooling (subsampling), activation, normalization, and fully connected layers [24,25]. Each layer plays a crucial role in the feature extraction and classification process. The convolutional layer is the core building block of a CNN. It applies a set of learnable filters to the input, activating certain features at certain spatial positions. In this experiment, CNNs are used to learn local patterns in the feature space via convolution over TF-IDF dimensions. Also, LSTMs operate on TF-IDF vectors as generic numerical inputs, meaning their gating and memory mechanisms do not exploit temporal dynamics, which explains their comparatively reduced advantage in this experimental configuration.

$$G(x, y) = \sum_{m=-M}^{M} \sum_{n=-N}^{N} F(m, n) . X(x - m, y - n) \tag{2}$$

Eq. (2) details the convolution operation, where $F(m, n)$ represents the filter kernel of size $(2M + 1) \times (2N + 1)$, and X is the input feature map. The result is a feature map G that emphasizes the presence of detected features in the input. Pooling layers reduce the dimensions of the data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer.

$$P(x, y) = \max_{m \in [0, M], n \in [0, N]} X(x.M + m, y.N + n) \tag{3}$$

Eq. (3) shows max pooling, where the maximum value over a $(M + 1) \times (N + 1)$ region is taken as the pooled output, effectively downsampling the feature map X to reduce its dimensions and allow the network to focus on the most prominent features. After the convolution operation, an activation function is applied to introduce non-linearity into the model, enabling it to capture and learn more intricate patterns. The Rectified Linear Unit (ReLU) activation function, which is commonly used due to its simplicity and efficiency. It replaces all negative pixel values in the feature map with zero. Normalization layers adjust and scale activations to speed up the training process and improve model generalization. Batch normalization is a widely used method.

$$B(x) = \gamma \left(\frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} \right) + \beta \tag{4}$$

Eq. (4) describes batch normalization, where μ and σ^2 are the mean and variance of a feature across a mini-batch, γ and β are learnable parameters of scale and shift, and ε is a small constant to avoid division by zero. After several convolutional and pooling layers, the fully connected layer processes the high-level extracted features by flattening them into a vector format, enabling classification into distinct labels. In a fully connected layer, where W is the weight matrix, b is the bias vector, and x is the input vector. This layer combines all the learned features to make a final classification. Within a CNN, these layers operate collaboratively to carry out both feature extraction and classification, which makes CNNs particularly effective for applications like image recognition, natural language processing, and—as applied in this study—differentiating fake from real news articles by analyzing their content and stylistic elements.

LSTMs process sequences by retaining memories of previous inputs using gates [26]. These gates control the flow of information, making LSTMs adept at understanding time-series or sequential data.

$$g_f = \sigma(W_f.[H_{t-1}, I_t] + B_f) \tag{5}$$

The forget gate Eq. (5) decides the extent to which previous state H_{t-1} influences the current, with W_f, I_t, and B_f denoting the weight matrix, current input, and bias, respectively.

$$g_i = \sigma(W_i.[H_{t-1}, I_t] + B_i) \tag{6}$$

Eq. (6) represents the input gate, determining new information's incorporation into the cell state, with parameters analogous to Eq. (5).

$$C_t = g_f \times C_{t-1} + g_i \times tanh(W_C.[H_{t-1}, I_t] + B_C) \tag{7}$$

The cell state update Eq. (7) combines past state C_{t-1} and new information weighted by the forget g_f and input g_i gates' outputs, enabling long-term dependency learning.

$$g_o = \sigma\left(W_o.\left[H_{t-1}, I_t\right] + B_o\right) \tag{8}$$

Eq. (8) details the output gate, which filters the information to be passed as the current output from the cell state, guided by weights W_o and bias B_o.

DTs implement a tree-like model of decisions, using branches to represent the decision paths and leaves to represent outcomes [27]. They classify instances by navigating through the branches based on feature values until reaching a leaf node corresponding to a decision outcome.

$$H(S) = -\sum_{i=1}^{c} p_i \log_2(p_i) \tag{9}$$

Eq. (9) defines the entropy of a set S, where p_i represents the proportion of the samples belonging to class i within S. Entropy quantifies the level of impurity or randomness in a dataset, serving as a key factor in directing the decision-making steps during tree construction.

$$IG(S, A) = H(S) - \sum_{t \in T} \frac{|S_t|}{|S|} H(S_t) \tag{10}$$

Information gain Eq. (10) calculates the reduction in entropy or impurity due to splitting the set S on attribute A. Here, T represents the subsets formed from splitting S by A, and $|S_t|$ is the size of subset t. Decision trees offer a transparent classification approach, allowing the decision path from root to leaf to be clearly followed and interpreted, which makes them well-suited for scenarios that demand clarity in decision-making [28].

SVMs with the RBF kernel are powerful tools for classifying data that is not linearly separable in the input space [29]. They transform the data into a higher-dimensional space in which it becomes linearly separable.

$$K(x, z) = exp\left(-\gamma \|x - z\|^2\right) \tag{11}$$

The RBF kernel Eq. (11) facilitates this projection, where γ is a parameter that determines the spread of the kernel and hence the decision boundary's complexity, x and z are input feature vectors.

$$\min\left(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i\right) \tag{12}$$

Eq. (12) outlines the objective function of SVM optimization, aiming to minimize the margin's width $\|w\|^2$ and the penalty term ξ_i for misclassified points, controlled by the parameter C. SVM with RBF kernel is highly effective in high-dimensional spaces, making it a robust classifier for text and image data where linear separability cannot be assumed.

4 Experimental Hypothesis and Setup

Based on the assumption that advanced ML models—particularly CNNs and LSTMs—are capable of accurately distinguishing between fake and real news articles, this study investigates the subtle textual differences that define each category. It is hypothesized that these models can effectively identify and leverage the differential patterns, stylistic features, and thematic disparities between fake and true news articles for classification purposes. Methodical experimental approach is done by using a diverse set of ML models, specifically ANN, CNN, LSTM, DT, and SVM-RBF, which were trained on the refined data set. Each model's detailed parameters and architectural specifications are elaborated in Table 1, with selections informed by extensive preliminary tuning to optimize performance. Then employing a standard hold-out strategy, the dataset was partitioned into training (80%) and testing (20%) subsets. This partition facilitates an assessment of the models' capacity for generalization to novel data. Finally the effectiveness of each model in correctly classifying news articles was measured using evaluation metrics including accuracy, precision, recall, and F1 score, providing a comprehensive perspective on overall performance.

Table 1: Detailed parameters and architecture specifics of the models utilized in the experiment. For deep learning models (ANN, CNN, LSTM), the Adam optimizer was used with a learning rate of 0.001 and categorical cross-entropy loss.

Model	Parameters	Architecture or Layers
ANN	Epochs: 100, Batch: 32, Dropout: 0.5	Input(TF-IDF Vector), Dense(128, ReLU), Dense(64, ReLU), Output(Sigmoid)
CNN	Filter: 128, Kernel: 5, Pool: 2	Conv1D(128), MaxPool(2), Conv1D(64), GlobalMaxPool, Dense(1)
LSTM	Units: 100, Dropout: 0.2	LSTM(100), Dense(50, ReLU), Dense(1, Sigmoid)
DT	Criterion: Gini, Splitter: Best	Max Depth: None, Min Samples Split: 2
SVM-RBF	C: 1.0, Gamma: Scale	Kernel: RBF, Tolerance: 0.001

Hypotheses Testing and Statistical Significance Analysis

This section outlines the suggested approach to statistically analyze the classification accuracies of the used ML models (ANN, CNN, LSTM, DT, SVM-RBF) over multiple runs. The aim is to identify whether significant differences exist not only between individual pairs of models but also in a one-to-many comparison fashion, assessing each model against the collective performance of others. To evaluate the performance differences between each pair of models, a series of paired t-tests [30] is employed. This method allows us to compare the mean accuracy scores of two models over 20 runs, assuming paired observations are normally distributed. A significant t-statistic, determined against the critical t-value from the t-distribution table at $\alpha = 0.05$, suggests a significant difference in the performance of the two compared models. For a broader comparison of each model's performance against the collective performance of others, the Bonferroni-adjusted ANOVA test [31] is utilized. This adjustment is crucial to maintain the overall Type I error rate while conducting multiple comparisons. The ANOVA test was adjusted for multiple comparisons, with the Bonferroni correction factor applied to the significance level α. The F-statistic is calculated then compared against the critical F-value from the F-distribution table at a predetermined significance level (usually $\alpha = 0.05$). A significant F-statistic (exceeding the critical F-value) suggests that at least one model's mean accuracy is significantly different from the others, warranting rejection of H_0. Given a significant ANOVA result, a post-hoc analysis is conducted to pinpoint which specific models' means differ significantly. The Tukey's HSD test [32] is employed for this purpose, offering a pairwise comparison among the models' accuracy means while controlling for the Type I error rate. Models with a mean difference exceeding the HSD value are considered to have significantly different accuracies, confirming H_1 for those models. The application of the ANOVA test, followed by Tukey's HSD post-hoc analysis, facilitates a comprehensive understanding of the comparative performance landscape of these ML models in the task of news classification. This statistical framework underscores a commitment to rigour and precision in validating advanced ML techniques for enhancing media integrity and information verification processes. Following identifying significant differences via the adjusted ANOVA test, proceeding with a post-hoc analysis using Tukey's HSD test to ascertain specific models that differ in performance. The combination of paired t-tests for individual model comparisons, alongside a Bonferroni-adjusted ANOVA for a comprehensive assessment, ensures a thorough statistical examination of the used ML models' efficacy in classifying news articles. By systematically analyzing performance disparities, this multifaceted statistical approach aims to validate the effectiveness of advanced ML techniques in distinguishing between fake and true news, thereby offering potent tools for media verification and integrity. The ANOVA test is chosen due to its effectiveness in identifying differences in group means in scenarios involving multiple groups.

As detailed in Table 1, the model parameters and architectural nuances are pivotal in contextualizing the experimental findings. The systematic selection of model configurations directly influences the robustness of the classification outcomes. The experimental setup and model training were conducted within a robust computational environment designed to support the intensive processing demands of ML workflows. This section outlines the technical specifications and software stack to facilitate the research. The implementation and testing of ML models were supported by comprehensive hardware, software tools, and libraries. Where the machine learning models were trained using a batch size of 32 and an initial learning rate of 0.001, which was dynamically adjusted via a scheduler based on validation loss. Training proceeded for a maximum of 100 epochs, with early stopping applied based on validation accuracy to mitigate overfitting. Training time varied by model complexity, with CNN and LSTM architectures requiring roughly 2–3 h on the designated GPU hardware. The dataset was randomly partitioned into 80% for training and 20% for testing, maintaining a stratified label distribution across both subsets. Model performance was assessed through a hold-out validation strategy, using test set metrics to evaluate generalizability. All experiments were executed in a controlled environment to ensure consistent and reproducible outcomes.

5 Experimental Findings Analysis and Discussion

In this research, an in-depth evaluation is performed to assess the effectiveness of five distinct models: ANN, CNN, LSTMs, DTs, and SVM-RBF. Each model underwent 20 independent training and evaluation runs to ensure a robust assessment of its performance. The accuracy metric was recorded for each run, providing a basis for the subsequent statistical analysis. The obtained findings reveal distinct performance characteristics across the models, with ANN and CNN models demonstrating exceptional capability in classifying news articles. The LSTM model, while slightly less accurate, showcased its proficiency in handling feature-rich data, a critical attribute for text-based tasks. In contrast, the classical ML models, DT and SVM-RBF, lagged in performance, underscoring the challenges these models face with complex, high-dimensional datasets. The experimental runs yielded the following mean accuracies with their corresponding standard deviations (SD) over 20 iterations as summarized in Fig. 7. The bar chart reports mean ±SD for each model (ANN, CNN, LSTM, DT, and SVM–RBF), highlighting the comparative performance across architectures.

The analysis was started with an ANOVA test to confirm whether the observed model performance differences were statistically significant. This test returned an F-statistic of 2981.91 and a p-value less than 0.001. It demonstrates significant differences in model accuracies. The result necessitated additional exploration through post-hoc analysis to identify the specific models among differences that occurred. Using Tukey's HSD test for pairwise comparison of model accuracies. Differences were significant, as shown in Table 2.

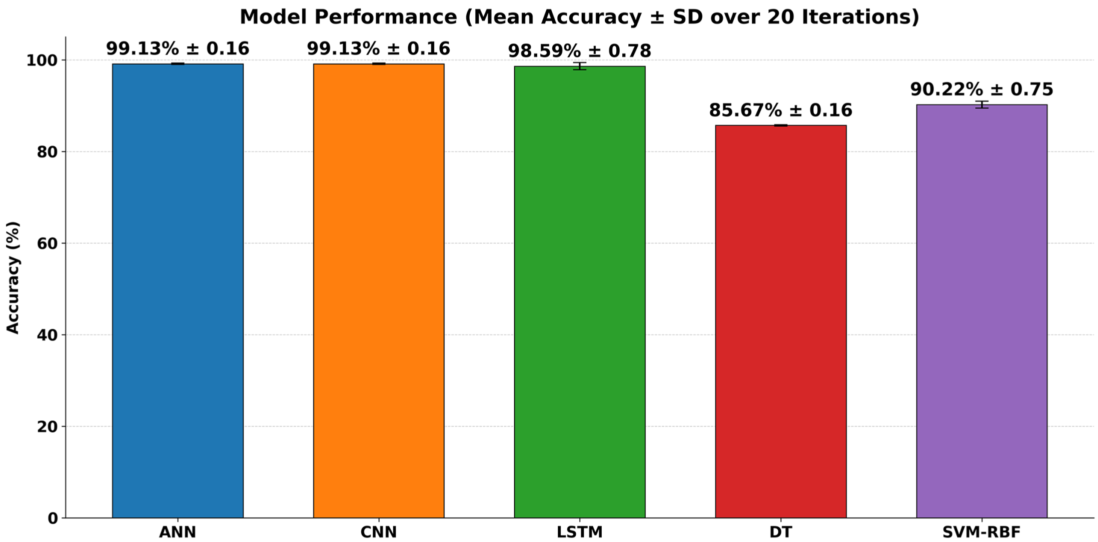


Figure 7: Model performance over 20 runs: mean accuracy (%) with error bars showing standard deviation for ANN, CNN, LSTM, DT, and SVM–RBF.

Table 2: Summary of Tukey's HSD test for pairwise model comparison.

Comparison	Mean Diff.	*p*-adj	Confidence Interval	Reject Null?
ANN vs. CNN	0.0	1.000	[−0.0045, 0.0045]	No
ANN vs. DT	−0.1345	0.000	[−0.139, −0.13]	Yes
ANN vs. LSTM	−0.0054	0.0106	[−0.0099, −0.0009]	Yes
ANN vs. SVM-RBF	−0.089	0.000	[−0.0935, −0.0846]	Yes
CNN vs. DT	−0.1345	0.000	[−0.139, −0.13]	Yes
CNN vs. LSTM	−0.0054	0.0106	[−0.0099, −0.0009]	Yes
CNN vs. SVM-RBF	−0.089	0.000	[−0.0935, −0.0846]	Yes
DT vs. LSTM	0.1292	0.000	[0.1247, 0.1336]	Yes
DT vs. SVM-RBF	0.0455	0.000	[0.041, 0.05]	Yes
LSTM vs. SVM-RBF	−0.0837	0.000	[−0.0882, −0.0792]	Yes

There is no significant performance difference between ANN and CNNs. In contrast, the notable performance differences between these previous models and LSTM, DT, and SVM-RBF indicate that ANN and CNN are more efficient. Importantly, the decision tree model exhibited the weakest performance, and SVM-RBF also underperformed compared to LSTM. It highlights LSTM's robustness in capturing data dependencies, even though its performance is slightly lower than that of ANN and CNN when utilizing TF-IDF features, which lack explicit sequential time-steps. The comparable effectiveness of both ANN and CNN models suggests that tasks involving complex pattern recognition or high-dimensional data processing could greatly benefit from these architectures. Furthermore, the low performance of LSTM, relative to ANN and CNN, may indicate a trade-off between capturing temporal dependencies and overall predictive accuracy when the input representation is non-sequential. The trade-off is important, particularly in applications where model complexity and training time are significant factors. The classical models, such as DT and SVM-RBF, may not perform as well as CNN and ANN models. The variation in model performance highlights the necessity of choosing a model that aligns with the specific requirements. Table 3 compares with other related studies with various data.

Table 3: Comparison of proposed framework accuracy vs. state of the art.

Ref.	Dataset	Methods	Benchmark (%)
[33]	WELFAKE	TF–IDF with BNB	90
[34]	ISOT	ML Algorithms	96.36
[35]	WELFAKE	BERT + Bi-LSTM	98.1
[36]	Kaggle	SVM	>90
[37]	Kaggle	BERT with MBIC	98.69
Ours	**Kaggle**	**TF–IDF with ANN, CNN**	**99.13**

6 Conclusion and Future Work

This study set out to tackle a contemporary challenge in digital news verification by employing various machine learning models to distinguish between genuine and fabricated news headlines. Through careful experimentation and thorough data analysis, the presented study demonstrates that AI can effectively enhance the media's credibility. The results indicate that CNN and ANN achieve the highest accuracy of 99.13%. Crucially, this study validates these findings through extensive statistical testing, including ANOVA and Tukey's HSD analysis across 20 independent runs. This rigorous evaluation confirms that deep learning architectures significantly outperform classical baselines (such as DT and SVM-RBF) in capturing intricate textual patterns, proving their efficacy even when utilizing standard high-dimensional features like TF-IDF. The superior performance of the ANN and CNN models highlights the increasing significance of neural networks in text classification tasks where robust feature mapping is required. The outcomes of this research hold value beyond the academic realm, offering tangible benefits for journalists, media outlets, and providers of IT solutions. With the predictive capabilities of CNN and LSTM, media stakeholders can implement automated systems to efficiently scan large volumes of news content and identify any spam articles. Technological advancement is crucial for ensuring the prevalence of factual reporting worldwide and fostering informed public discourse. However, while this work represents a significant advancement, it is important to recognize its limitations. The exceptionally high accuracy observed in this study may be partially attributed to the specific linguistic characteristics of the dataset, where fake news often exhibits distinct lexical patterns distinguishable by TF-IDF. Misinformation tactics are becoming increasingly sophisticated, and language constantly evolves, necessitating regular updates to machine learning models and validation across more diverse datasets to prevent overfitting. In the future, the features of news content could be enhanced by incorporating more multimodal data, such as images and videos, to improve the effectiveness of fake news detection systems. Additionally, exploring transfer and unsupervised learning approaches offers significant potential for advancing the automated understanding and classification of news content. The present research can be extended into a critical area of digital journalism that focuses on optimizing machine learning in the increasingly complex landscape of journalistic practice. As a society situated at the intersection of media and technology, the study highlights the advancements made in this domain while emphasizing the need for greater collaboration to push the boundaries of AI while preserving the credibility of news.

Acknowledgement: The authors would like to thank the Deanship of Scientific Research at Shaqra University for supporting this work. Also, the authors acknowledge the use of AI language assistance tools for improving the academic presentation and readability of the manuscript.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Raed Alotaibi, writing and visualization; Muhammad Atta Othman Ahmed, conceptualization, software, writing and preparation; Omar Reyad, idea proposal, review and editing and supervision; Nahla Fathy Omran, data curation and investigation. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The dataset file that supports the findings of this work is publicly available as indicated in the references.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alghamdi J, Luo S, Lin Y. A comprehensive survey on machine learning approaches for fake news detection. Multimed Tools Appl. 2024;83(17):51009–67. doi:10.1007/s11042-023-17470-8.
2. Bertini F, Benetton A, Montesi D. Ensuring news integrity against online information disorder through text watermarking and blockchain. Blockchain Res Appl. 2025:100414. doi:10.1016/j.bcra.2025.100414.
3. Ali AA, Latif S, Ghauri SA, Song OY, Abbasi AA, Malik AJ. Linguistic features and Bi-LSTM for identification of fake news. Electronics. 2023;12(13):2942. doi:10.3390/electronics12132942.
4. Olan F, Jayawickrama U, Arakpogun EO, Suklan J, Liu S. Fake news on social media: the impact on society. Inf Syst Front. 2024;26(2):443–58. doi:10.1007/s10796-022-10242-z.
5. Lv J, Gao Y, Li L, Shi L, Li S. Multi-modal fake news detection: a comprehensive survey on deep learning technology, advances, and challenges. J King Saud Univ Comput Inf Sci. 2025;37(9):306. doi:10.1007/s44443-025-00317-7.
6. Harris S, Hadi HJ, Ahmad N, Ali Alshara M. Multi-domain Urdu fake news detection using pre-trained ensemble model. Sci Rep. 2025;15(1):8705. doi:10.1038/s41598-025-91054-4.
7. Vural NE, Kalaman S. Using artificial intelligence systems in news verification: an application on X. İletişim Kuram Ve Araştırma Dergisi. 2024;67:127–41. doi:10.47998/ikad.1466830.
8. Jouhar J, Pratap A, Tijo N, Mony M. Fake news detection using python and machine learning. In: Proceedings of the 5th International Conference on Innovative Data Communication Technologies and Application (ICIDCA 2024); 2024 Jan 10–11; Coimbatore, India.
9. Sonni AF, Hafied H, Irwanto I, Latuheru R. Digital newsroom transformation: a systematic review of the impact of artificial intelligence on journalistic practices, news narratives, and ethical challenges. J Medium. 2024;5(4):1554–70. doi:10.3390/journalmedia5040097.
10. Alshuwaier FA, Alsulaiman FA. Fake news detection using machine learning and deep learning algorithms: a comprehensive review and future perspectives. Computers. 2025;14(9):394. doi:10.3390/computers14090394.
11. Wu L, Jiang X, Sun S, Wen T, Wang Y, Liu M. ZoFia: zero-shot fake news detection with entity-guided retrieval and multi-LLM interaction. arXiv:2511.01188. 2025.
12. Khedairia S, Bennour A, Nahas M, Chefrour A, Marie RR, Al-Sarem M. A co-attention mechanism into a combined GNN-based model for fake news detection. Comput Mater Contin. 2025;85(1):1267–85. doi:10.32604/cmc.2025.066601.
13. Tajrian M, Rahman A, Kabir MA, Islam MR. A review of methodologies for fake news analysis. IEEE Access. 2023;11:73879–93. doi:10.1109/ACCESS.2023.3294989.
14. Zhou X, Zafarani R. A survey of fake news: fundamental theories, detection methods, and opportunities. arXiv:1812.00315. 2018.
15. Hasib KM, Towhid NA, Faruk KO, Al Mahmud J, Mridha MF. Strategies for enhancing the performance of news article classification in Bangla: handling imbalance and interpretation. Eng Appl Artif Intell. 2023;125:106688. doi:10.1016/j.engappai.2023.106688.

16. Ahmed J, Ahmed M. Classification, detection and sentiment analysis using machine learning over next generation communication platforms. Microprocess Microsyst. 2023;98:104795. doi:10.1016/j.micpro.2023.104795.
17. Kumar A, Ansari MIH, Singh K. A fake news classification and identification model based on machine learning approach. In: Proceedings of the Information and Communication Technology for Competitive Strategies (ICTCS 2022); 2022 Dec 9–10; Chandigarh, India. Singapore: Springer Nature; 2023. p. 473–84.
18. Mishra A, Sadia H. A comprehensive analysis of fake news detection models: a systematic literature review and current challenges. Eng Proc. 2023;59(1):28. doi:10.3390/engproc2023059028.
19. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. In: Advances in neural information processing systems. Red Hook, NY, USA: Curran Associates, Inc.; 2017. p. 5998–6008. doi:10.65215/2q58a426.
20. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers); 2019 Jun 2–7; Minneapolis, MN, USA. p. 4171–86.
21. Bisaillon C. Fake and real news dataset. 2019[cited 2025 Jan 1]. Available from: https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset.
22. Huang CH, Yin J, Hou F. A text similarity measurement combining word semantic information with TF-IDF method: a text similarity measurement combining word semantic information with TF-IDF method. Chin J Comput. 2011;34(5):856–64. doi:10.3724/sp.j.1016.2011.00856.
23. Okwu MO, Tartibu LK. Metaheuristic optimization: nature-inspired algorithms swarm and computational intelligence, theory and applications. Cham, Switzerland: Springer International Publishing; 2021. p. 133–45.
24. El-Rahiem BA, Ahmed MAO, Reyad O, El-Rahaman HA, Amin M, El-Samie FA. An efficient deep convolutional neural network for visual image classification. In: Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019); 2019 Mar 28–30; Cairo, Egypt. Cham, Switzerland: Springer International Publishing; 2019. p. 23–31.
25. Khan A, Sohail A, Zahoora U, Qureshi AS. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev. 2020;53(8):5455–516. doi:10.1007/s10462-020-09825-6.
26. Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. 2019;31(7):1235–70. doi:10.1162/neco_a_01199.
27. Costa VG, Pedreira CE. Recent advances in decision trees: an updated survey. Artif Intell Rev. 2023;56(5):4765–800. doi:10.1007/s10462-022-10275-5.
28. Krishna NLSR, Adimoolam M. Fake news detection system using decision tree algorithm and compare textual property with support vector machine algorithm. In: Proceedings of the 2022 International Conference on Business Analytics for Technology and Security (ICBATS); 2022 Feb 16–17; Dubai, United Arab Emirates. p. 1–6.
29. Hu L, Wei S, Zhao Z, Wu B. Deep learning for fake news detection: a comprehensive survey. AI Open. 2022;3:133–55. doi:10.1016/j.aiopen.2022.09.001.
30. Al-Mutair H, Berri J. MSDSI-FND: multi-stage detection model of influential users' fake news in online social networks. Computers. 2025;14(12):517. doi:10.3390/computers14120517.
31. Cabin RJ, Mitchell RJ. To bonferroni or not to bonferroni: when and how are the questions. Bull Ecol Soc Am. 2000;81(3):246–8. doi:10.1016/0197-2456(88)90095-5.
32. Abdi H, Williams LJ. Tukey's honestly significant difference (HSD) test. Encycl Res Des. 2010;3(1):1–5. doi:10.4135/9781412961288.n181.
33. Mohsen F, Chaushi B, Abdelhaq H, Karastoyanova D, Wang K. Automated detection of misinformation: a hybrid approach for fake news detection. Future Internet. 2024;16(10):352. doi:10.3390/fi16100352.
34. Ribeiro Bezerra JF. Content-based fake news classification through modified voting ensemble. J Inf Telecommun. 2021;5(4):499–513. doi:10.1080/24751839.2021.1963912.
35. Al-Quayed F, Javed D, Jhanjhi NZ, Humayun M, Alnusairi TS. A hybrid transformer-based model for optimizing fake news detection. IEEE Access. 2024;12:160822–34. doi:10.1109/ACCESS.2024.3476432.

36. Wierzbicki A, Shupta A, Barmak O. Synthesis of model features for fake news detection using large language models. In: Proceedings of the Computational Linguistics Workshop at 8th International Conference on Computational Linguistics and Intelligent Systems (CoLInS-2024); 2024 Apr 12–13; Lviv, Ukraine.
37. Kasampalis A, Chatzakou D, Tsikrika T, Vrochidis S, Kompatsiaris I. Bias detection and mitigation in textual data: a study on fake news and hate speech detection. In: Advances in information retrieval. Cham, Switzerland: Springer Nature; 2024. p. 374–83. doi:10.1007/978-3-031-56063-7_29.