

Article

Deep Dive into Fake News Detection: Feature-Centric Classification with Ensemble and Deep Learning Methods

Fawaz Khaled Alarfaj ^{1,*}  and Jawad Abbas Khan ²¹ Department of Management Information Systems, School of Business King Faisal University (KFU), Al-Ahsa 31982, Saudi Arabia² Department of Computer Science, COMSATS University, Wah Campus, Rawalpindi 47040, Pakistan; jawadkhan31849@gmail.com

* Correspondence: falarfaj@kfu.edu.sa

Abstract: The online spread of fake news on various platforms has emerged as a significant concern, posing threats to public opinion, political stability, and the dissemination of reliable information. Researchers have turned to advanced technologies, including machine learning (ML) and deep learning (DL) techniques, to detect and classify fake news to address this issue. This research study explores fake news classification using diverse ML and DL approaches. We utilized a well-known “Fake News” dataset sourced from Kaggle, encompassing a labelled news collection. We implemented diverse ML models, including multinomial naïve bayes (MNB), gaussian naïve bayes (GNB), Bernoulli naïve Bayes (BNB), logistic regression (LR), and passive aggressive classifier (PAC). Additionally, we explored DL models, such as long short-term memory (LSTM), convolutional neural networks (CNN), and CNN-LSTM. We compared the performance of these models based on key evaluation metrics, such as accuracy, precision, recall, and the F1 score. Additionally, we conducted cross-validation and hyperparameter tuning to ensure optimal performance. The results provide valuable insights into the strengths and weaknesses of each model in classifying fake news. We observed that DL models, particularly LSTM and CNN-LSTM, showed better performance compared to traditional ML models. These models achieved higher accuracy and demonstrated robustness in classification tasks. These findings emphasize the potential of DL models to tackle the spread of fake news effectively and highlight the importance of utilizing advanced techniques to address this challenging problem.



Citation: Alarfaj, F.K.; Khan, J.A. Deep Dive into Fake News Detection: Feature-Centric Classification with Ensemble and Deep Learning Methods. *Algorithms* **2023**, *16*, 507. <https://doi.org/10.3390/a16110507>

Academic Editor: Takeshi Yamada

Received: 23 September 2023

Revised: 27 October 2023

Accepted: 30 October 2023

Published: 3 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: fake news detection; deep features; ensemble-based methods; machine learning; deep learning

1. Introduction

In today’s digital age, online media has become integral to our lives. With the increasing accessibility and diversity of online interaction and information-sharing platforms, text, photos, and multimedia content can reach larger audiences faster than ever before [1]. According to the American Press Institute, a significant percentage of individuals, 59% of those aged 18 to 34 and 56% of those aged 35 to 49, rely on digital media for news updates and access to the latest information [2]. However, this heavy dependance on online media for news consumption establishes inherent risks. It exposes individuals to the danger of encountering misleading, manipulative, or even violence-inciting fake news [3].

Consuming news through social media has become gradually common [4]. Social networks have become popular as platforms for news utilization due to their diverse multimedia formats, affordability, and ability to facilitate rapid news broadcasting [5]. However, fake news producers often develop these very features, rapidly spreading false information [6]. Consequently, the wide usage of social media has led to the broader dissemination of misinformation, making news consumed from online feeds unreliable [7]. The spreading of fake information can severely impact society. For instance, the circulation of fake news can manipulate the results of significant public events [8]. During the 2016

United States presidential election, the impact of fake news on shaping public opinion was evident [9]. False stories circulated widely on social media platforms, leading to a significant influence on people's perceptions. Examples include claims such as Pope Francis endorsing Donald Trump for president and Hillary Clinton's alleged participation in a child trafficking ring. These fabricated stories not only contributed to political polarization but also misled voters, potentially influencing the election's outcome [10].

Additionally, children's exposure to fake news poses a critical concern, as they are more sensitive to believing fake information than adults [11]. Children are still developing critical thinking skills, so they require assistance distinguishing between reliable and unreliable information. Moreover, fake news can propagate negativity among people when consumed by young users [12]. Recent research has shown that users between 18 and 28 are the most active on social media [13]. Consequently, the early identification of fake news on social media has gained popularity and garnered significant attention.

Several fake news detection methods, including conventional machine learning (ML) and deep learning (DL) models, have been developed [14]. In traditional approaches, features from the news must be extracted before they are classified. Conversely, DL models can automatically extract significant features from text or images in the news [15]. Due to the time-consuming process of manual feature extraction, DL approaches are often preferred over traditional approaches [16]. Ensemble methods are widely used to merge multiple models into a composite model that achieves better performance than individual models. The accuracy and diversity of each model significantly impacts the effectiveness of an ensemble model. In general, the utilization of ML and DL models to detect fake news stems from their limitations in comprehending and interpreting the intricacies of human language [17].

The alignment prediction of headline and article body pairs was conducted using the term frequency-inverse document frequency (TF-IDF) and a deep neural network [18]. Although this approach shows promise in detecting fake news, distinguishing between real news and articles that employ humor or satire may face challenges, as they often rely on linguistic subtleties and context-specific meanings. To address this issue, ref. [8] introduced an ensemble-based ML technique incorporating a deep neural network model for classifying fake news. However, many existing methods, including [19], rely on a single prototype embedding model. Such models often overlook polysemous terms, which pose multiple interpretations or meanings depending on the context [15]. The lack of multiple prototype embeddings for each can result in difficulties in accurately processing and understanding polysemous sentences [16]. The primary goal of this study was to identify the most effective approach for accurately classifying false information. The driving force behind this endeavor is the need to combat the widespread dissemination of misleading information and to maintain international peace by identifying and rectifying incorrect information. The major contributions of this research are as follows:

- We propose a comprehensive fake news detection classifier for news article, considering the platform's unique characteristics and challenges.
- We conduct experiments and analyze multiple ML and DL techniques to evaluate their performance on the fake news dataset, aiming to address the problem effectively.
- Exploring various deep features, such as TF-IDF, n-gram, Word2Vec, and global vectors for word representation (GloVe), to find the optimal combination that enhances the detection process and improves classifier accuracy.
- Comparing diverse ML and DL techniques provides insights into their strengths and weaknesses for fake news detection, aiding researchers in making informed decisions about their applications.

The remainder of the study is organized as follows. Section 2 discusses recent research studies relevant to fake news detection methods. Section 3 presents the proposed research methodology, deep features, diverse ML and DL methods, and their architecture to resolve the challenges of fake news detection. Section 4 presents a detailed empirical analysis of

the model with comprehensive analysis. Finally, in Section 5, conclusions, implications, and future work are discussed.

2. Related Work

This section provides an overview of the existing research and approaches developed for fake news detection, focusing on various features, ML techniques, and deep learning algorithms.

2.1. ML Algorithms

In existing studies, diverse ML algorithms have been applied to resolve the issues and challenges of fake news detection. A fake news detection approach [20] utilized a random forest (RF) classifier by incorporating twenty-three textual features. The authors applied different feature selection techniques to identify the most significant features. They compared results with benchmark techniques like GBM, XGBoost, and the Ada Boost regression model. A comparative analysis of machine learning classifiers presented in [21] includes support vector machine (SVM), naïve Bayes (NB), random forest (RF), and logistic regression (LR). The study evaluated the performance of these classifiers on diverse datasets. SVM achieved the highest accuracy on the liar, fake job posting, and fake news datasets.

Moreover, SVM and RF attained the highest accuracy for the fake job posting dataset, followed by LR and NB. Similarly, another fake news detection model [22] introduced a novel genetic algorithm (GA) and ML classifiers as the fitness function, outperforming traditional methods. Subsequently, ref. [23] focused on the automated classification of news articles into real and fake news by extracting different textual properties using an ML approach. They evaluated the performance of various ML algorithms using benchmark datasets. According to the investigation, RF ensemble-based method outperformed other models and attained 90.40% accuracy for news article classification.

Kaliyar et al. [24] focused on detecting and classifying fake news distributed on social networks to enhance trust and transparency in social network recommendation systems. The authors discussed various traditional ML models for fake news detection and multiclassification using unlabeled data. The proposed approach improved the accuracy of trust-aware recommender systems through a semi-supervised strategy. LR achieved the highest accuracy in classifying fake news. Rathod et al. [25] proposed a model for detecting fake news in online resources and social media platforms. The authors used Natural Language Processing (NLP) and ML techniques to classify news articles as fake and real news based on source authenticity. Rezaei et al. [26] applied ML algorithms, including passive-aggressive (PA), NB, and SVM, to identify fake news. However, relying solely on simple classification methods may not yield optimal results. Integrating ML techniques with text-based processing techniques has significantly improved fake news detection. Challenges arise due to the limited availability of diverse corpora for training and differentiation purposes. Their approach showed better results through experimental analysis using publicly available datasets, achieving accuracy levels of up to 93%.

Faust et al. [27] focused on leveraging spreading networks and graph features to construct a model for detecting fake news. Fourteen graph features were extracted and evaluated across 13 ML models. Notably, propensity and centrality emerged as key factors influencing the classification process. The best-performing models achieved better accuracy using modified SVM on the Twitter15 and Twitter16 datasets. Chouliara et al. [28] introduced a robust framework for detecting Thai fake news, encompassing three modules: information retrieval, NLP, and ML. A web-crawler information retrieval technique was employed to collect data from Thai online news websites, and relevant features were extracted using NLP. They compared multiple conventional ML models with ensemble-based ML models, and Adda Boost emerged as the best-performing model. Table 1 presents the performance detail of each ML model for fake news classification.

Table 1. ML algorithms for classifying fake news.

Ref	Year	Model	Performance Evaluation Matrix (%)
[20]	2020	RF + XG Boost	F1 score: 92.00
[21]	2021	SVM	Accuracy: 91.70
[22]	2022	GA	Accuracy: 80.50
[23]	2022	RF	Accuracy: 90.40
[24]	2020	Semi-Supervised LR	Accuracy: 87.91
[25]	2022	Random subspace (SVM)	Accuracy: 89.76
[26]	2022	PA + NB	Accuracy 93.00
[27]	2023	Bagging + SVM	Accuracy: 91.40
[28]	2023	Boosting + Add Boost	Accuracy: 94.92

2.2. DL Algorithms

ML models rely on predefined features and statistical algorithms to detect fake news. In contrast, DL models automatically learn features and hierarchies from raw data, enabling these models to capture more intricate patterns and representations. Ahmed et al. [29] proposed a novel approach for detecting fake news related to COVID-19 using biomedical information extraction (BioIE) combined with ML models. The authors analyzed COVID-19 news articles, extracted novel features using BioIE algorithms, and trained ML classifiers. The study demonstrated that incorporating biomedical information improves the performance of fake news detection models. Rai et al. [30] proposed techniques that rely on natural language analysis and DL models. They applied a hybrid approach to combine CNN with an attention-based mechanism to highlight the significant part of textual data. The utilization of attention in CNN models enables them to concentrate on specific segments of input data during prediction and output generation. This process assigns different levels of significance to individual elements within a sequence, facilitating the model in grasping pertinent context for enhanced comprehension and generation capabilities [31]. However, CNN-based models cannot learn past and future dependencies from a text but have become popular due to their convergence to near-optimal solutions with low computational complexity. Luvembe et al. [32] proposed a multimodal transformer using two-level visual features (MTTV), which leverages transformer models to process text and image data uniformly. Two-level visual features, global and entity levels, are utilized to enhance the utilization of news images. The extended transformer model enables full interaction between multimodal data, capturing semantic relationships. The authors also introduced a scalable classifier to improve the balance of fine-grained fake news detection. Experimental results demonstrated the effectiveness of MTTV compared to existing methods.

Prabhakar Kaila et al. [2] proposed a dual-channel CNN (DC-CNN) model with attention-pooling to address challenges such as noisy data, redundancy, and long-distance dependencies in fake news detection. DC-CNN utilizes skip-gram and fast-Text deep embeddings to reduce noisy data and enhance the model's learning ability for non-derived words. A parallel dual-channel pooling layer replaces the traditional CNN pooling layer, combining the advantages of max-pooling for local information and attention-pooling for context semantics and global dependencies. Experimental results on COVID-19 fake news datasets showed that the proposed model performed best in handling noisy data and balancing local–global feature correlations. Onan et al. [33] introduced a method for assessing the credibility of knowledge-based facts using a multi-layer perceptron (MLP) model to address the limitations of existing models in classifying facts as fake or genuine. Their approach utilizes state-of-the-art word embedding models for feature embedding, such as Word2Vec, GloVe, and TF-IDF. The experimental results showed that the MLP trained on triples vectorized with GloVe and counted vectorize outperformed the baseline ML models in terms of accuracy. The authors also proposed an algorithm for the joint

extraction of triples and associated named entity tags, contributing as an additional feature for training the models.

To detect multimodal fake news, Sahoo et al. [34] focused on the fusion of abundant information across different modalities. To address this challenge, they proposed a mutual attention neural network (MANN) model that learns the relationships between different modalities in the news. The MANN model consists of four components: multimodal feature extractor, mutual attention fusion, fake news detector, and irrelevant event discriminator. The model was evaluated on the Weibo dataset, where it outperformed other state-of-the-art methods. Trueman et al. [35] proposed a bidirectional long short-term memory (Bi-LSTM) DL approach with the addition of self-attention, enhancing clarity, a crucial aspect of deep learning. The proposed hybrid DL model aims to improve the accuracy of false news detection. The research emphasized the need for enhanced mechanisms and presented the approach as a potential solution. By combining advanced DL techniques, these studies have contributed to developing more effective and accurate models for fake news detection across various modalities and contexts. Furthermore, with the advent of transformers and bi-directional DL model such as BERT and its different variations boost the capabilities to NLP models to understand the context and long-range dependencies within natural language text [36]. Additionally, their self-attention mechanisms enable them to efficiently process input sequences of varying lengths, revolutionizing the field of NLP, and enabling advancements in a wide range of applications [37]. Table 2 shows the performance of each DL model for fake news classification.

Table 2. DL algorithms used for fake news classification.

Authors	Year	Model	Performance Evaluation Matrix
[29]	2017	RNN + BioIE	Accuracy: 91.00
[30]	2022	Attention + CNN	Accuracy: 89.00
[32]	2023	MTTV + Transformers	Accuracy: 92.36
[38]	2020	DC-CNN	Accuracy: 90.63
[33]	2021	MLP	F1 Score: 81.01
[34]	2021	MANN	Accuracy: 90.75
[35]	2021	Self-Attention + Bi-LSTM	Accuracy: 92.21

3. Proposed Methodology

This study proposes a hybrid ensemble method and a DL model to classify news articles as ham or spam. The purpose of the model is to assess and monitor the impact of this content. For this study Kaggle fake news dataset has been used. The extracted dataset was initially raw and contained ambiguities, noise, and optional information. Following the data collection stage, the datasets underwent a series of preparation stages, including removing links, stop words, punctuation, tokenizing, lemmatizing, and stemming the data. After removing any stop words, min-max normalization was performed on the dataset. The textual aspects were collected from the news data on top of the retrieved features. These ensemble methods consist of different ML-based algorithms, such as PA, LR, multinomial naïve Bayes (MNB), Bernoulli NB (NN), Gaussian NB (GNN), and LSTM, BERT and RoBERTa. Next, we split the dataset into two sets: training and testing. The PA, LR, MNB, Bernoulli NN, Gaussian NN ML models, and DL-based LSTM models were used to estimate whether the news was ham or spam. The results were assessed using accuracy, precision, recall, and F1 scores as the performance metrics, and the ML and DL models were compared. The collected data were then evaluated considering previously established methods for classifying news as ham or spam. The proposed methodology is shown in Figure 1.

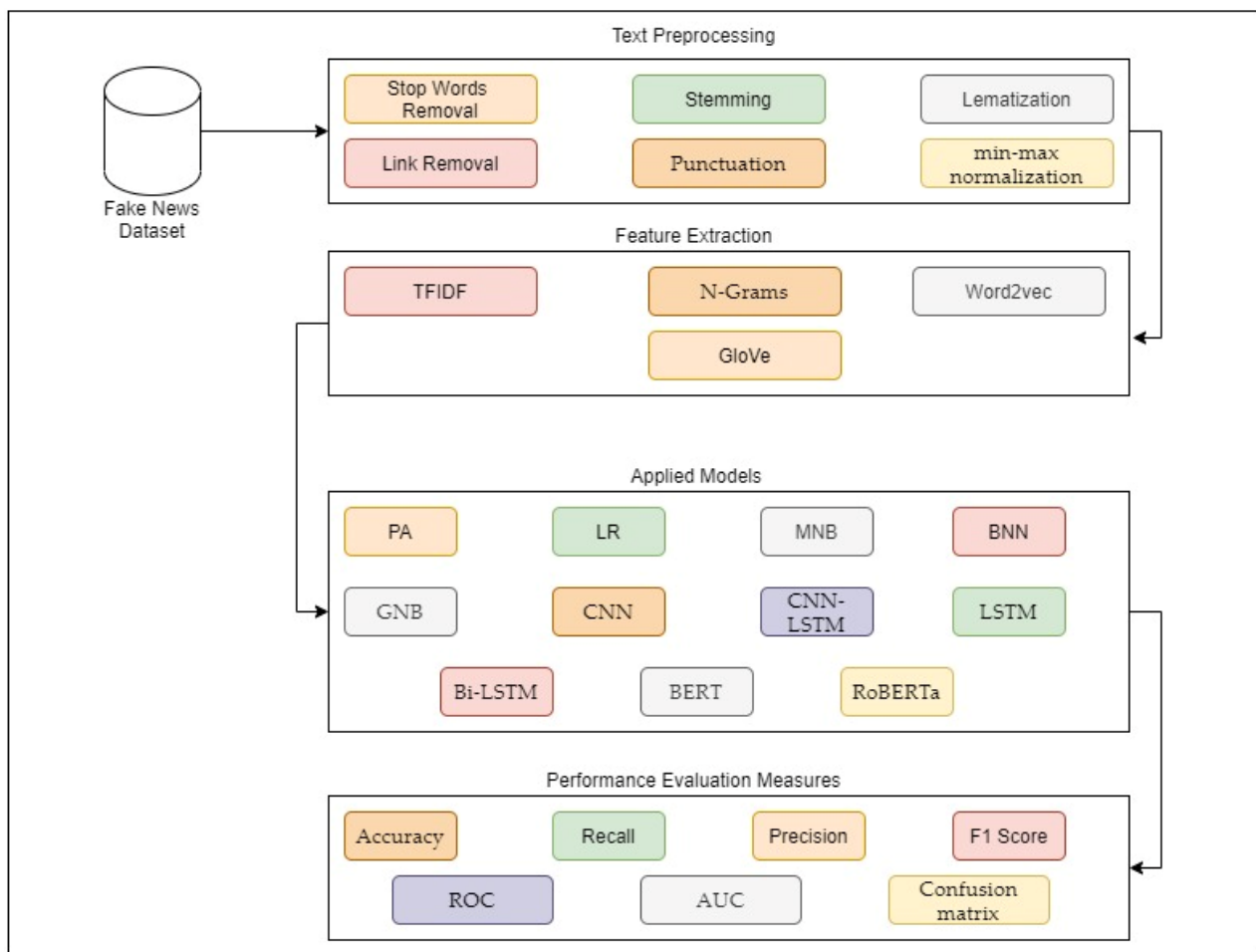


Figure 1. Proposed methodology for detecting fake news.

3.1. Data Acquisition and Pre-Processing

A reliable and labeled dataset is crucial to developing an effective fake news classifier. In this study, we utilized the “Fake News” dataset (<https://www.kaggle.com/c/fake-news/data> accessed on 15 September 2023), which comprises a collection of news articles labeled as either “reliable” (real news) or “unreliable” (fake news). The dataset contains 87,500 articles, distributed into 70,000 and 17,500 classes for reliable and unreliable data, respectively. The dataset provides essential features, such as the article title, text content, publication date, source, and author, enabling us to leverage multiple aspects for classification. The dataset comprises a diverse range of news articles obtained from various sources. Several pre-processing steps were applied to prepare the data for training the fake news classifier. These steps aimed to clean and transform the raw text data into a format suitable for ML algorithms.

The utilization of specific pre-processing techniques in this study is motivated by established best practices and prior research in the NLP and fake news detection. The application of these pre-processing techniques is crucial in our pursuit of developing an effective fake news classifier. In the era of information overload and the rapid dissemination of news through digital platforms, the quality of the data we feed into our models is paramount. By carefully preparing and cleaning the text data, we aim to reduce noise, extract meaningful information, and enhance the discriminative power of our classifier. These techniques are not merely procedural but are essential for uncovering the underlying patterns and distinguishing characteristics that differentiate fake news from real news. The following pre-processing techniques were employed:

Stop Word Removal: Stop word removal is a crucial step in fake news detection, eliminating common, insignificant words like “the” and “is” to focus on meaningful content.

This process reduces noise, improves efficiency, and helps identify relevant keywords and phrases.

Word Tokenization: Word tokenization involves splitting the text into individual words or tokens. We utilized a tokenization algorithm to break down the text content of each article into a sequence of tokens, facilitating further analysis at the word level.

URL Removal: Many news articles contain URLs or web links that may not contribute to the classification task. Therefore, we removed the URLs from the text data to focus solely on the textual content of the articles.

Stemming: Stemming is a technique that reduces words to their base or root form, removing suffixes or prefixes. This process helps reduce the feature space's dimensionality and captures the words' core meaning. We applied a stemming algorithm to transform words into their base form, thus aiding the classification process.

By implementing these pre-processing steps, we aimed to enhance the quality and relevance of the text data, thus improving the performance of the fake news classifier. The pre-processed dataset was then ready for the subsequent feature engineering, model training, and evaluation stages.

3.2. Feature Extraction

Handling a large dataset with many variables can be computationally intensive and challenging. Feature extraction, a dimensionality reduction technique, breaks down large amounts of raw data into more manageable groups. Feature extraction refers to methods for selecting and/or combining variables into features, significantly reducing the amount of data that must be processed while accurately describing the initial dataset.

3.2.1. TF-IDF

After the pre-processing phase, the *TF-IDF* algorithm extracted textual features from the dataset. This algorithm computes the frequency of a word with its sentimental meaning. The algorithm calculates the frequency of terms in a document and multiplies them by the inverse frequency of words that appear regularly in several texts. The frequency of documents in a corpus is then calculated by the *TF-IDF* algorithm [28] using the following mathematical formulation shown in Equation (1):

$$TF - IDF(w, d, D) = TF(w, d) * IDF(w, D) \quad (1)$$

where $TF(w, d)$ represents the term frequency of word w in document d , and $IDF(w, D)$ represents the inverse document frequency of the word w in the set of documents D .

3.2.2. N-Grams

The n-gram method is frequently employed in statistical NLP. N-grams are continuous sequences of n-words from a given text. In this research study, we employ different word sequence models. These include unigrams (1-g), bigrams (2-g), trigrams (3-g), and so on, depending on the word length of the text. In the case of word sequences, consider the example sentence, "He went to the zoo". The trigrams for this sentence can be formed as follows:

"# He went" (adding a blank space # before the first word)

"He went to"

"went to the"

"to the zoo"

"the zoo #" (adding a blank space # after the last word).

represents a blank space denoting the sentence's beginning and end.

3.2.3. Word2Vec

Word2Vec is a technique for encoding semantic information from text documents into vector form. Using extensive English text corpora, we trained a skip-gram model for the English language, resulting in 300-dimensional word vectors. Word2Vec analyses the

semantic perspective of the content using a three-layered deep NN, which then groups words with similar contexts.

3.2.4. GloVe

The GloVe paradigm uses word co-occurrence in a matrix to produce word embeddings. A significant corpus of words was created, and each word was individually examined. GloVe constructs a large co-occurrence matrix that captures the frequency of words appearing together in the text. The algorithm then learns lower-dimensional embeddings by factorizing this matrix while preserving these co-occurrence patterns. A GloVe model with 200 dimensions was trained using text corpora in English to represent words comprehensively [39].

3.3. Ensemble ML Algorithms

ML algorithms, which are built on the fundamental principles of artificial intelligence (AI), can learn from input data, continually improve their performance, and generate predictions based on newly acquired knowledge. ML models can be applied individually or in combination with other ML techniques to achieve better results. This is achieved through ensemble methods, which combine multiple models to reduce prediction errors by adding bias values to the model and decreasing the prediction dispersion and variance. This study used ensemble ML algorithms to determine whether a news article was fake or genuine.

3.3.1. Passive Aggressive Classifier

The passive-aggressive classifier (PAC) is an online learning algorithm that belongs to large-margin classifiers. This algorithm is well-suited for handling extensive datasets and adapts in response to each new instance it encounters. As an online learning algorithm, the PAC updates its weights based on new input. A key feature of the PAC is its regularization parameter, C , which allows for a trade-off between the margin's size (space between the classes) and the number of misclassifications. At each iteration, the PAC looks at a new instance, determines whether it was correctly classified, and adjusts its weights accordingly. If the instance was correctly classified, the weights do not change. If the instance was incorrectly classified, the PAC updates its weights based on the misclassified instance to better classify subsequent instances. The amount to which the PAC adjusts its weight depends on the regularization parameter C and the degree of confidence in classifying that instance. Higher values of C lead to aggressive weight updates, while lower values result in conservative updates.

3.3.2. Logistic Regression

LR has a rich history of application in the biological and social sciences, particularly with categorical dependent variables. The importance of LR can be observed in tasks such as categorizing spam emails. Because linear regression requires a threshold to be set for classification and cannot efficiently handle categorization tasks, misclassification may lead to serious consequences. Logistic regression offers a better way to address this limitation. LR is preferable to linear regression for categorical classification because it ensures that the predicted values fall within a range of 0 to 1. This formulation allows LR to produce probability estimates and reliable categorical predictions. The mathematical equation of the LR algorithm is shown in Equation (2):

$$y = \frac{\hat{e}(b_0 + b_1 * x)}{(1 + \hat{e}(b_0 + b_1 * x))} \quad (2)$$

where the variables x , y , b_0 , and b_1 represent the coefficient of a single input value, the predicted output, the bias or intercept term, and the corresponding input value, respectively.

3.3.3. Multinomial Naïve Bayes Algorithm

Popular Bayesian learning methods, such as the MNB algorithm, are widely used in NLP tasks. This algorithm makes informed predictions about classifying text data, such as emails or news articles, using the guiding principles of the Bayesian theorem. The method computes the probability of each potential class for a given sample and selects the class with the highest probability. A significant feature of the NB classifier is the assumption that the features are independent. This means that the presence or absence of other features does not influence the presence or absence of a particular feature. Due to this independence assumption, the algorithm can effectively combine different approaches and successfully classify a variety of features from the dataset. The mathematical equation of the MNB algorithm, based on the Bayesian theorem, is shown in Equation (3):

$$P(A|B) = \frac{P(B|A)}{P(B)} \quad (3)$$

3.3.4. Bernoulli Naïve Bayes

The Bernoulli NB algorithm is a member of the naïve Bayesian family of algorithms. It is considered a compact model that takes only binary values. The most straightforward illustration is checking whether each word is present in a text. When counting word frequencies is unimportant, the Bernoulli NB algorithm can yield more accurate results.

Simply put, we need to summarize all the binary features indicating whether a word is included in a text. These features are used instead of counting the number of times a word appears in a document. The mathematical equation of the Bernoulli NB algorithm is shown in Equation (4):

$$P(x_i|y) = P(i|y)x_i + (1 - P(x_i|y))(1 - x_i) \quad (4)$$

3.3.5. Gaussian Naïve Bayes

The Gaussian NB algorithm is a probabilistic classification method that uses strong independence assumptions to apply the Bayesian theorem. This algorithm assumes that the presence or absence of one feature value does not affect the presence or absence of another in the context of categorization (although this differs from independence in probability theory). Despite their naïve independence assumption, naïve Bayes classifiers are well-known in ML for their expressiveness, scalability, and respectable accuracy. However, as the training set's size grows, NB classifiers' effectiveness may deteriorate. Various factors influence their performance. NB classifiers have the advantage of not requiring parameter tuning, scaling well with larger training datasets, and effectively handling continuous features. These advantages contribute to their effectiveness.

In the Gaussian NB algorithm (Equation (5)), certain assumptions are made, such as the independence of variance with respect to Y (denoted as σ_i), the independence of variance with respect to x_i (denoted as σ_k), or both (denoted as σ):

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y^2)}{2\sigma_y^2}\right) \quad (5)$$

These assumptions play a crucial role in formulating the Gaussian NB algorithm.

3.4. DL Models

DL models, however, are a subset of ML models specifically designed to mimic the workings of the human brain's NNs. DL models, particularly artificial neural networks, consist of multiple layers of interconnected nodes (neurons) that learn representations of input data at different levels of abstraction. These models can automatically learn complex features and hierarchies from raw data without the need for manual feature engineering. In the context of fake news detection, DL models, such as LSTM, CNNs, and different

combinations of LSTM and CNNs, can process and analyze textual information to make predictions about the authenticity of news articles.

3.4.1. Long Short-Term Memory

LSTM is a recurrent neural network (RNN) frequently used in problems in which predictions are made sequentially. The usefulness of RNNs is limited by long-term dependency issues caused by vanishing gradient concerns due to their operation—information persistence. LSTM comprises input, output, forget gates, and hidden states that keep track of historical and current data timestamps. The input gate accepts new data and assesses their importance. The output gate produces a result based on the learned information, while the forget gate helps separate old knowledge from current information in the model.

3.4.2. Convolutional Neural Networks

CNNs are prominent DL methods that excel at analyzing and classifying images. They can automatically learn unique features and patterns by optimizing learnable weights and biases [40]. CNNs have the advantage of requiring less pre-processing than other classification techniques, which is one of their benefits. CNNs can independently identify and enhance the filters and attributes necessary for accurate classification during training. Simpler approaches, however, often require manual filter construction.

3.4.3. CNN-LSTM

The CNN-LSTM model, which combines the LSTM and CNN layers, enables sequence prediction and feature extraction from input data. These models are frequently used for tasks such as predicting visual time series and generating textual descriptions from collections of images. They benefit from various applications, including activity recognition, image and video annotation, and image tagging. By combining the strengths of the CNN and LSTM models, the CNN-LSTM architecture can handle complex tasks that require spatial and temporal understanding.

3.4.4. BERT

BERT represents a cutting-edge advancement in NLP and has demonstrated exceptional performance across a spectrum of language understanding tasks, including text classification. The foundational architecture of BERT is rooted in the Transformer model, which effectively processes input sequences of varying lengths. This architecture comprises a stack of transformer encoder layers, each consisting of two integral sub-layers: a multi-head self-attention mechanism and a position-wise feedforward neural network. These components synergistically enable BERT to capture context-rich information from both directions, fostering a bidirectional understanding of text. Furthermore, BERT undergoes unsupervised pre-training on a substantial text corpus, where it hones its ability to predict missing words within sentences and discern the sequential coherence of sentences in the original text [41].

3.4.5. RoBERTa

RoBERTa, short for “A Robustly Optimized BERT Pretraining Approach”, is a BERT variant meticulously crafted to enhance the pre-training phase for improved language comprehension. While, like BERT, RoBERTa undergoes pre-training on an extensive text corpus, it distinguishes itself by implementing refinements in this process. These include employing larger batch sizes and extending the training duration for more epochs, thereby exposing the model to a more diverse array of textual data and variations. During pre-training, RoBERTa adheres to the Masked Language Model task, wherein it randomly masks words within sentences and trains the model to predict these masked terms based on contextual cues, thereby facilitating the acquisition of intricate contextual representations. Furthermore, RoBERTa integrates data augmentation techniques such as sentence order shuffling and additional sentence span randomization within input documents. This

strategic augmentation fosters superior generalization capabilities by exposing the model to a broader spectrum of sentence arrangements.

3.5. Performance Evaluation Measures

In this study, we applied widely used evaluation measures such as accuracy, Precision, recall, F1-score, AUC, ROC, and confusion matrix to perform empirical analysis. These measures are discussed in detail in the following sections.

3.5.1. Accuracy

Accuracy is an essential metric for evaluating classification models. It measures the proportion of correct predictions made by the model. The formal definition of accuracy is shown in Equation (6):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

3.5.2. Precision

Precision is crucial for evaluating a model's ability to produce precise predictions. It rates how well the model performs in making accurate predictions. The mathematical definition of Precision is shown in Equation (7):

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

3.5.3. Recall

Recall measures the percentage of accurate positive predictions compared to all potential positive predictions. It assesses a model's capacity to locate and consider all pertinent positive examples. The mathematical definition of recall is shown in Equation (8):

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

3.5.4. F1 Score

The F1 score combines recall and Precision using the harmonic mean. The F1 scoring formula is represented by Equation (9):

$$F1_{Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

3.5.5. Receiver Operating Characteristic Curve

The receiver operating characteristic (ROC) curve depicts the performance of a classification model at various thresholds. It displays the true positive rate (TPR) and the false positive rate (FPR). The ROC curve illustrates the trade-off between these rates for various classification thresholds. Equations (10) and (11) represent the TPR and the FPR, respectively:

$$TPR = \frac{TP}{TP + FN} \quad (10)$$

$$FPR = \frac{FP}{FP + TN} \quad (11)$$

3.5.6. Area under the Curve

The area under the curve (AUC) summarizes the performance across all available classification criteria. The AUC is the probability that the model will value a randomly chosen positive example more highly than a randomly picked negative example.

only after these stop words are removed from the corpus. Figure 3b displays the top 20 unigrams from the cleaned corpus, allowing us to analyze the most common unigrams after the cleaning process. Figure 3c presents the top 20 bigrams of the uncleaned corpus, while Figure 3d shows the top 20 bigrams of the cleaned corpus. Furthermore, Figure 3 illustrates the distribution of n-gram frequencies, with darker colors representing higher frequencies and lighter colors indicating less frequent n-grams.

Table 4. Dataset column description and data type.

Column	Description	Detail
ID	Unique ID of the news article	Object
Title	Title of the news article	Object
Author	Name of article	Object
Text	Complete text of the news article	Object

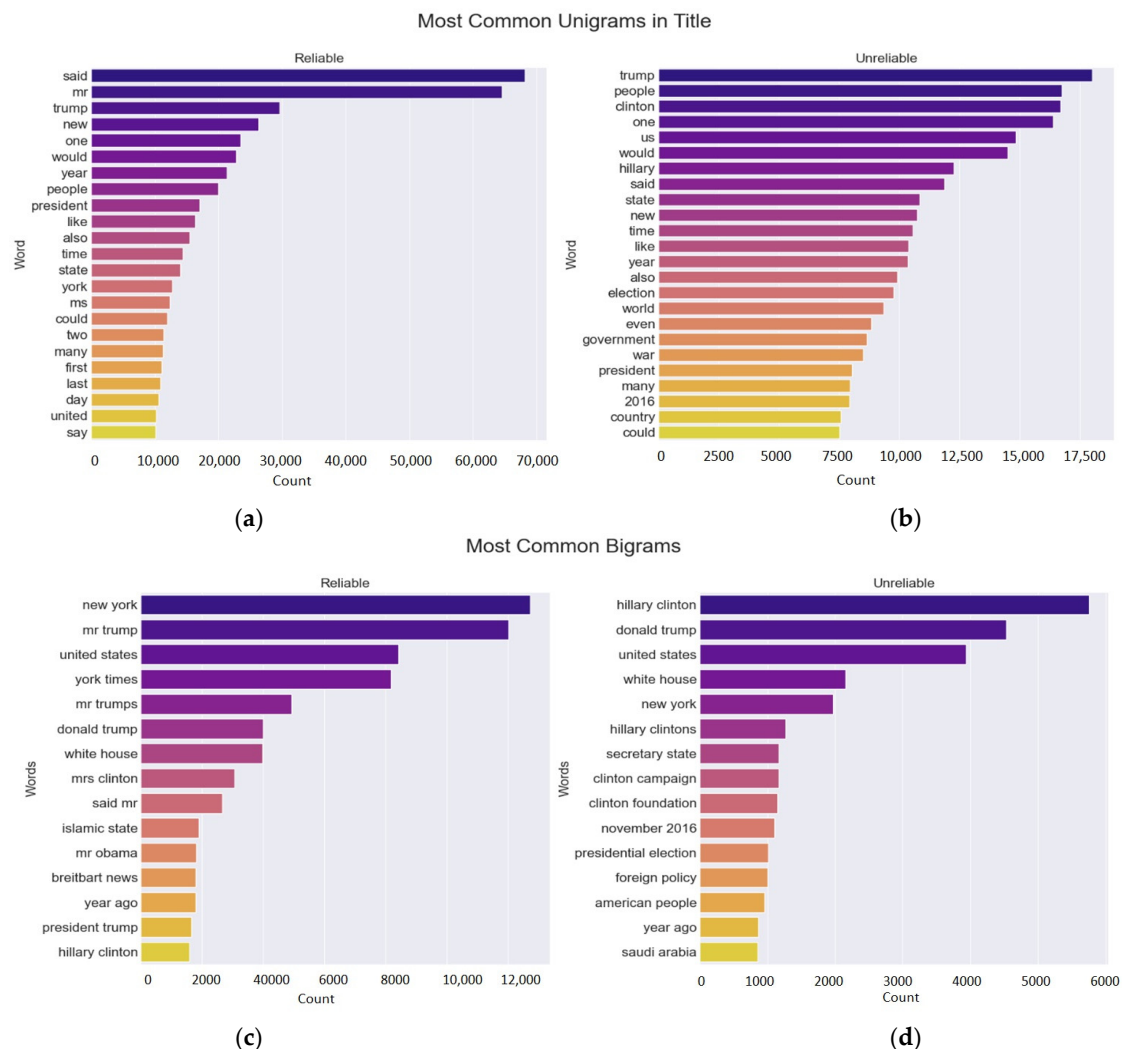


Figure 3. (a–d) top 20 n-grams of the clean and uncleaned corpora.

4.2. Experimentation with the Ensemble Method

During the initial stage of dataset pre-processing, the text is converted to lowercase, tokenized, and lemmatized. Additionally, stop words and links are removed, and words are stemmed. This process yields a pre-processed dataset with 59,589 features, which serves as input for TF-IDF feature extraction. The selected features are then utilised to train

four distinct machine learning classifiers (PA, LR, MNB, Bernoulli NB, and Gaussian NB), along with a deep learning-based LSTM, for the purpose of classifying the text into “ham” or “spam” categories. The model’s performance is evaluated using various metrics, such as accuracy, precision, recall, ROC, AUC, and F1 score, to gain insights into the effectiveness and efficiency of the classification process.

4.2.1. Ensemble ML Algorithms

The computed results of ML algorithms using TF-IDF, and n-grams are shown in Table 5. The results show that LR outperforms the other algorithms, achieving 90.06% accuracy with TF-IDF and 60.74% accuracy with the n-gram. In addition, DT attains 65.93% accuracy using TF-IDF and 62.59% accuracy using n-grams. The ROC curve is shown for text-based features, such as TF-IDF and N-gram, where LR demonstrates better results than the other algorithms.

Table 5. Experimentation with ensemble methods.

Methods		Evaluation Metrics			
		Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
ML Models	PA	97.71	98.00	98.00	98.0
	LR	99.06	98.90	97.80	98.34
	Multinomial NB	96.85	99.41	98.42	98.90
	Bernoulli NB	98.14	98.27	99.14	98.70
	Gaussian NB	89.14	89.09	90.10	89.59
DL Models	LSTM	90.68	90.93	91.41	90.72
	Bi-LSTM	94.72	93.75	94.89	94.75
	BERT	95.50	94.50	95.20	95.10
	RoBERTa	96.10	95.50	95.89	95.12
	CNN	84.50	84.12	84.42	84.21
	CNN-LSTM	85.60	84.98	85.53	85.25

4.2.2. DL Algorithm Results

Table 5 presents the computed results of the DL algorithms using TF-IDF. According to the results, it can be observed that the RoBERTa transformers-based model achieved the highest overall performance, with impressive scores across all metrics, ranging from 95.12.75% to 96.10%. The BERT and Bi-LSTM models also demonstrated strong performance, consistently scoring above 90% across all metrics. On the other hand, the LSTM, CNN, and CNN-LSTM models achieved slightly lower scores but still showed reasonable accuracy and Precision in identifying fake news.

Figure 4 shows the confusion matrix of the PA, LR, MNB, BNN, GNN, LSTM, BI-LSTM, BERT, RoBERTa, CNN, and CNN-LSTM models. Figure 5 shows the ROC curve of the PA, LR, MNB, BNN, GNN, LSTM, BI-LSTM, CNN, and CNN-LSTM models.

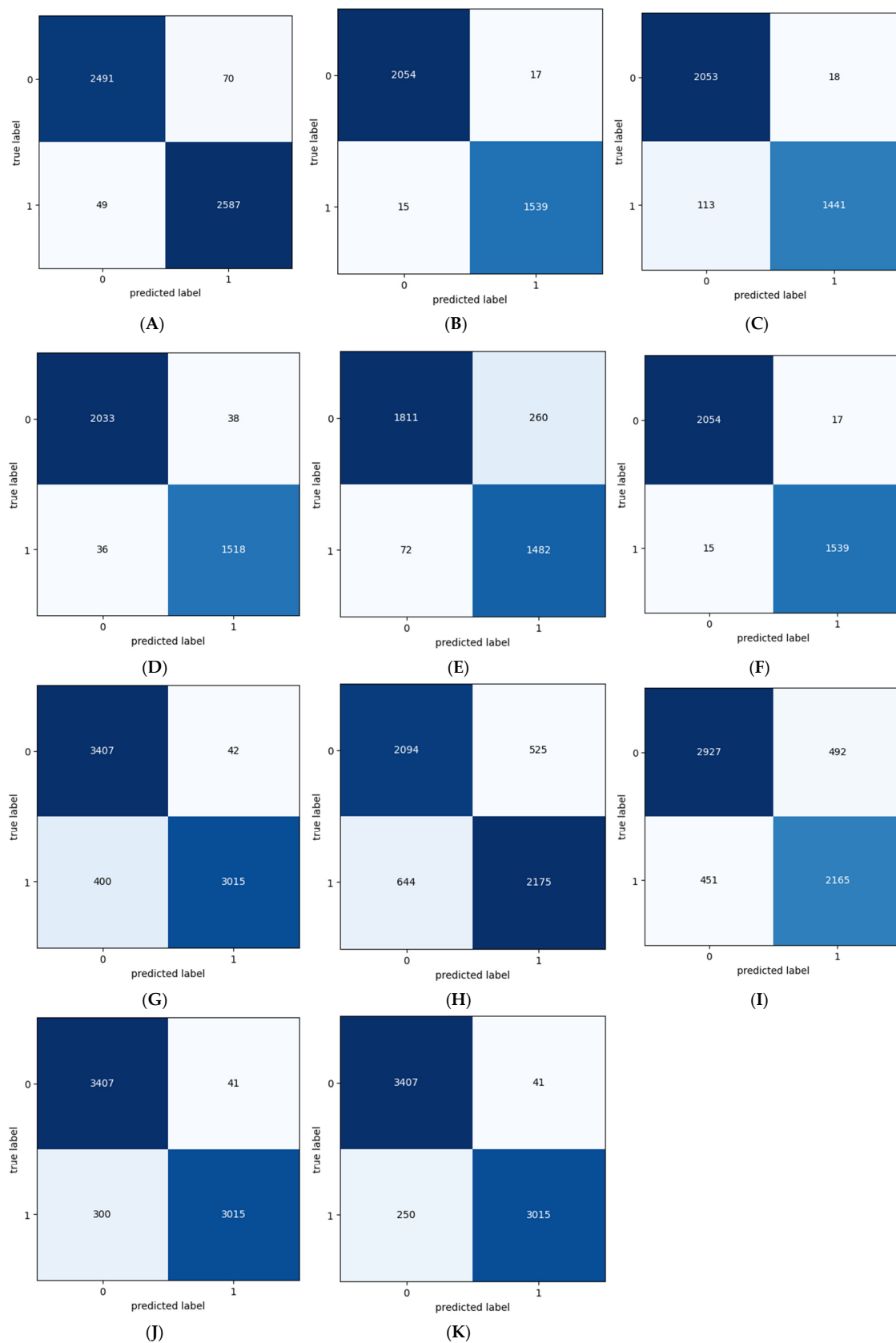


Figure 4. Confusion matrix of ensemble and DL models. (A), PA, (B), LR, (C), MNB, (D), BNB, (E), GNN, (F), LSTM, (G), BI-LSTM, (H), CNN, (I) CNN-LSTM, (J) BERT, (K) RoBERTa.

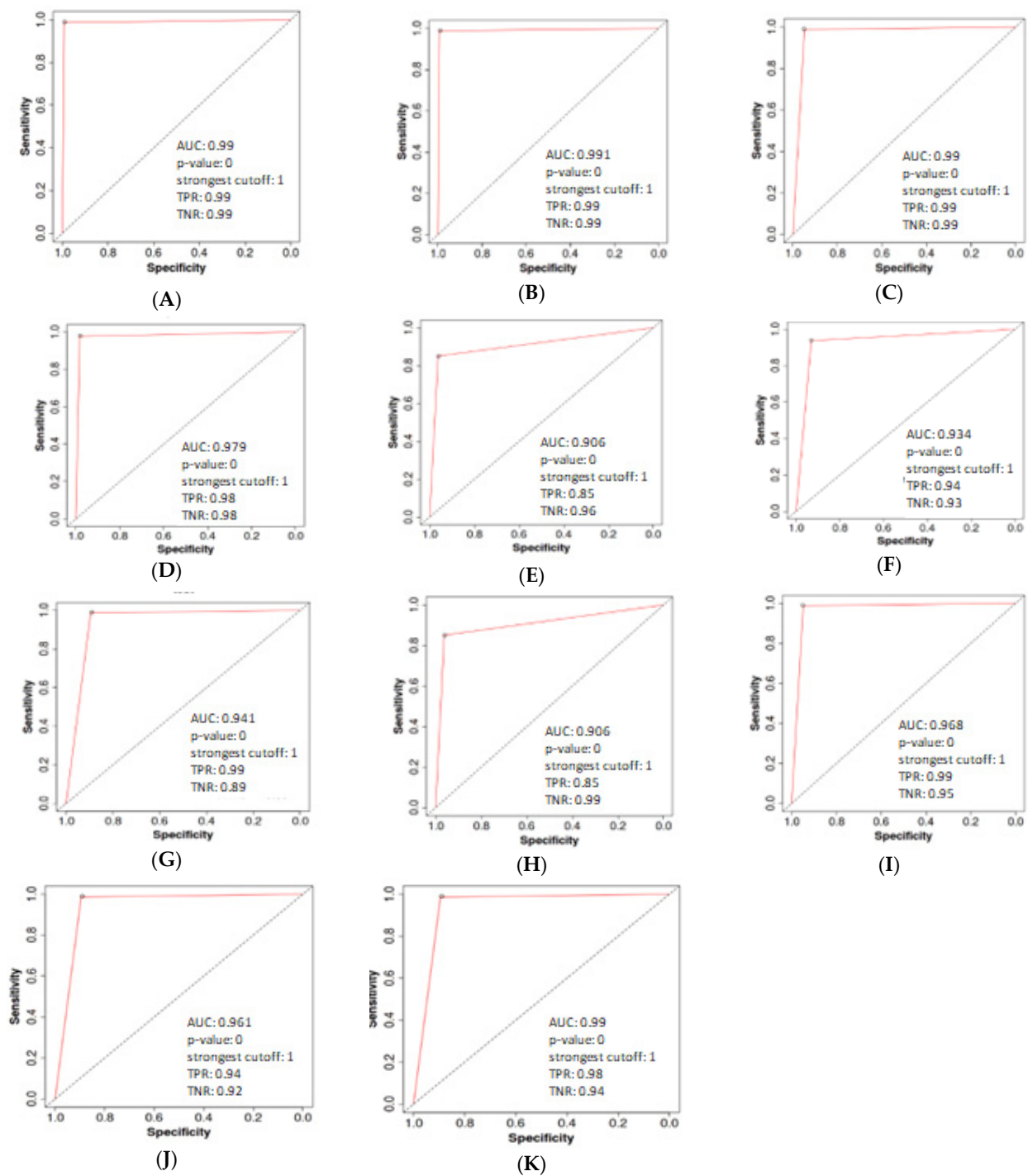


Figure 5. ROC curve of ensemble and DL models. (A), PA, (B), LR, (C), MNB, (D), BNB, (E), GNN, (F), LSTM, (G), BI-LSTM, (H), CNN, (I) CNN-LSTM, (J) BERT, (K) RoBERTa.

Table 6 shows the AUC of the PA, LR, MNB, BNN, GNN, LSTM, BI-LSTM, BERT, RoBERTa, CNN, and CNN-LSTM models. Among the traditional machine learning models, Passive Aggressive and LR exhibit remarkable performance, achieving high AUC values of 0.99 and displaying equally impressive TPR and TNR scores of 0.99. MNB and BNB also demonstrate favorable results, achieving competitive AUC values of 0.968 and 0.979, respectively, and maintaining high TPR and TNR percentages. On the other hand, GNB exhibits a slightly lower AUC value of 0.906, indicating a comparatively less robust performance. Though it maintains a good TNR of 0.96, its TPR score is relatively lower at 0.85.

Table 6. AUC of the ensemble and DL models.

Model	AUC (%)	TPR (%)	TNR (%)
Passive Aggressive	0.99	0.99	0.99
Logistic Regression	0.991	0.99	0.99
Multinomial Naïve Bayes	0.968	0.99	0.95
Bernoulli's Naïve Bayes	0.979	0.98	0.98
Gaussian Naïve Bayes	0.906	0.85	0.96
LSTM	0.934	0.94	0.93
BI-LSTM	0.941	0.99	0.89
BERT	0.961	0.96	0.92
RoBERTa	0.969	0.98	0.94
CNN	0.910	0.84	0.97
CNN-LSTM	0.89	0.86	0.88

Regarding the deep learning models, the RoBERTa model achieves the highest AUC score of 0.967, showcasing an exceptional performance in fake news detection, along with a near-perfect TPR of 0.98. However, its TNR of 0.94 indicates a slightly higher false positive rate. Both BERT and Bi-LSTM models display commendable performance, with AUC values of 0.951 and 0.941, respectively. The LSTM model shows a balanced TPR and TNR of 0.94 and 0.93, while the CNN model achieves a good TNR of 0.97, though with a slightly lower TPR of 0.84. Table 7 presents the comparison of the proposed model against the conventional ML and DL methods. Compared to other models, the proposed methodology stands out, with RoBERTa achieving an accuracy of 96.10% and PA achieving an even higher accuracy of 97.71%. These results demonstrate the promising performance of the proposed RoBERTa and PA models in detecting fake news, showcasing their potential for effective and reliable classification.

Table 7. Results comparison of the proposed methodology with previous studies.

Ref	Year	Model	Performance Evaluation Matrix
[20]	2020	RF + XG Boost	F1 score: 92.00
[21]	2021	SVM	Accuracy: 91.70
[22]	2022	GA	Accuracy: 80.50
[23]	2022	RF	Accuracy: 90.40
[24]	2020	Semi-Supervised LR	Accuracy: 87.91
[25]	2022	Random subspace (SVM)	Accuracy: 89.76
[26]	2022	PA + NB	Accuracy 93.00
[27]	2023	Bagging + SVM	Accuracy: 91.40
[28]	2023	Boosting + Add Boost	Accuracy: 94.92
[29]	2017	RNN + BioIE	Accuracy: 91.00
[30]	2022	Attention + CNN	Accuracy: 89.00
[32]	2023	MTTV + Transformers	Accuracy: 92.36
[38]	2020	DC-CNN	Accuracy: 90.63

Table 7. Cont.

Ref	Year	Model	Performance Evaluation Matrix
[33]	2021	MLP	F1 Score: 81.01
[34]	2021	MANN	Accuracy: 90.75
[35]	2021	Self- Attention + Bi-LSTM	Accuracy: 92.21
Proposed Methodology		RoBERTa	Accuracy: 96.10
		PA	Accuracy: 97.71

5. Conclusions

Fake news classification is a dynamic and continuously evolving research area. In this study, we propose a novel model for classifying fake news using “Fake News” dataset obtained from the Kaggle’s platform. We applied several pre-processing steps to prepare the dataset, including punctuation and link removal, tokenization, and lemmatization. Combining the datasets established the relationship between genuine and fake news. Feature extraction was carried out using TF-IDF on the data, and these extracted features were then utilized in multiple ML models, such as PA, LR, MNB, GNB, and BNB, as well as DL models like CNN, LSTM, and BI-LSTM, to classify fake news. We utilized various evaluation metrics to assess the models’ performance, including accuracy, precision, recall, F1 score, AUC, ROC, and confusion matrix. Among the models tested, LR emerged as the top performer, achieving the highest accuracy, and showcasing its effectiveness in fake news classification. This research contribution lies in presenting a superior-performing LR model compared to existing studies in the field. For future work, we aim to explore the impact of incorporating temporal information from tweets and utilizing more advanced deep learning architectures may further improve the model’s performance.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, F.K.A. and J.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [GRANT4,967].

Data Availability Statement: The dataset used in this research is publicly available and can be accessed at the following link: <https://www.kaggle.com/c/fake-news/data> (accessed on 14 September 2023).

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

References

- Guo, Y. A mutual attention based multimodal fusion for fake news detection on social network. *Appl. Intell.* **2023**, *53*, 15311–15320. [CrossRef]
- Ma, K.; Tang, C.; Zhang, W.; Cui, B.; Ji, K.; Chen, Z.; Abraham, A. DC-CNN: Dual-channel Convolutional Neural Networks with attention-pooling for fake news detection. *Appl. Intell.* **2023**, *53*, 8354–8369. [CrossRef]
- Altheneyan, A.; Alhadlaq, A. Big Data ML-Based Fake News Detection Using Distributed Learning. *IEEE Access* **2023**, *11*, 29447–29463. [CrossRef]
- Helmstetter, S.; Paulheim, H. Weakly Supervised Learning for Fake News Detection on Twitter. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 28–31 August 2018; pp. 274–277. [CrossRef]
- Hammouchi, H.; Ghogho, M. Evidence-Aware Multilingual Fake News Detection. *IEEE Access* **2022**, *10*, 116808–116818. [CrossRef]
- Guo, Y.; Song, W. A Temporal-and-Spatial Flow Based Multimodal Fake News Detection by Pooling and Attention Blocks. *IEEE Access* **2022**, *10*, 131498–131508. [CrossRef]
- Raza, S.; Ding, C. Fake news detection based on news content and social contexts: A transformer-based approach. *Int. J. Data Sci. Anal.* **2022**, *13*, 335–362. [CrossRef]
- Phan, H.T.; Nguyen, N.T.; Hwang, D. Content-Context-Based Graph Convolutional Network for Fake News Detection. In *Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence*; Fujita, H., Fournier-Viger, P., Ali, M., Wang, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 571–582.

9. Segura-Bedmar, I.; Alonso-Bartolome, S. Multimodal Fake News Detection. *Information* **2022**, *13*, 284. [\[CrossRef\]](#)
10. Siino, M.; Di Nuovo, E.; Tinnirello, I.; La Cascia, M. Fake News Spreaders Detection: Sometimes Attention Is Not All You Need. *Information* **2022**, *13*, 426. [\[CrossRef\]](#)
11. Galli, A.; Masciari, E.; Moscato, V.; Sperli, G. A comprehensive Benchmark for fake news detection. *J. Intell. Inf. Syst.* **2022**, *59*, 237–261. [\[CrossRef\]](#)
12. Shao, Y.; Sun, J.; Zhang, T.; Jiang, Y.; Ma, J.; Li, J. Fake News Detection Based on Multi-Modal Classifier Ensemble. In Proceedings of the 1st International Workshop on Multimedia AI against Disinformation, in MAD '22, Newark, NJ, USA, 27–30 June 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 78–86. [\[CrossRef\]](#)
13. Barbosa, V.N.; Neto, F.M.M.; Filho, S.A.; Silva, L. A Comparative Study of Machine Learning Algorithms for the Detection of Fake News on the Internet. In Proceedings of the XVIII Brazilian Symposium on Information Systems, in SBSI '22, Curitiba, Brazil, 16–19 May 2022; Association for Computing Machinery: New York, NY, USA, 2022. [\[CrossRef\]](#)
14. Shushkevich, E.; Alexandrov, M.; Cardiff, J. BERT-based Classifiers for Fake News Detection on Short and Long Texts with Noisy Data: A Comparative Analysis. In *Text, Speech, and Dialogue*; Sojka, P., Horák, A., Kopeček, I., Pala, K., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 263–274.
15. Goldani, M.H.; Safabakhsh, R.; Momtazi, S. Convolutional neural network with margin loss for fake news detection. *Inf. Process. Manag.* **2021**, *58*, 102418. [\[CrossRef\]](#)
16. Ying, L.; Yu, H.; Wang, J.; Ji, Y.; Qian, S. Multi-Level Multi-Modal Cross-Attention Network for Fake News Detection. *IEEE Access* **2021**, *9*, 132363–132373. [\[CrossRef\]](#)
17. Do, T.H.; Berneman, M.; Patro, J.; Bekoulis, G.; Deligiannis, N. Context-Aware Deep Markov Random Fields for Fake News Detection. *IEEE Access* **2021**, *9*, 130042–130054. [\[CrossRef\]](#)
18. Tseng, Y.-W.; Yang, H.-K.; Wang, W.-Y.; Peng, W.-C. KAHAN: Knowledge-Aware Hierarchical Attention Network for Fake News Detection on Social Media. In Proceedings of the Companion Proceedings of the Web Conference 2022, WWW'22, Lyon, France, 25–29 April 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 868–875. [\[CrossRef\]](#)
19. Mosallanezhad, A.; Karami, M.; Shu, K.; Mancenido, M.V.; Liu, H. Domain Adaptive Fake News Detection via Reinforcement Learning. In Proceedings of the ACM Web Conference 2022, WWW'22, Lyon, France, 25–29 April 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 3632–3640. [\[CrossRef\]](#)
20. Silva, R.M.; Santos, R.L.S.; Almeida, T.A.; Pardo, T.A.S. Towards Automatically Filtering Fake News in Portuguese. *Expert Syst. Appl.* **2020**, *146*, 113199. [\[CrossRef\]](#)
21. Felber, T. Constraint 2021: Machine Learning Models for COVID-19 Fake News Detection Shared Task. *arXiv* **2021**, arXiv:2101.03717. [\[CrossRef\]](#)
22. Farooq, M.S.; Naseem, A.; Rustam, F.; Ashraf, I. Fake news detection in Urdu language using machine learning. *PeerJ Comput. Sci.* **2023**, *9*, e1353. [\[CrossRef\]](#)
23. Onan, A.; Korukoğlu, S. Exploring Performance of Instance Selection Methods in Text Sentiment Classification. In *Artificial Intelligence Perspectives in Intelligent Systems*; Silhavy, R., Senkerik, R., Oplatkova, Z., Silhavy, P., Prokopova, Z., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 167–179.
24. Kaliyar, R.K.; Goswami, A.; Narang, P.; Sinha, S. FNDNet—A deep convolutional neural network for fake news detection. *Cogn. Syst. Res.* **2020**, *61*, 32–44. [\[CrossRef\]](#)
25. Rathod, S. Exploring Author Profiling for Fake News Detection. In Proceedings of the 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), Los Alamitos, CA, USA, 27 June–1 July 2022; pp. 1614–1619. [\[CrossRef\]](#)
26. Rezaei, S.; Kahani, M.; Behkamal, B.; Jalayer, A. Early multi-class ensemble-based fake news detection using content features. *Soc. Netw. Anal. Min.* **2022**, *13*, 16. [\[CrossRef\]](#)
27. Faust, O.; Hagiwara, Y.; Hong, T.J.; Lih, O.S.; Acharya, U.R. Deep learning for healthcare applications based on physiological signals: A review. *Comput. Methods Programs Biomed.* **2018**, *161*, 1–13. [\[CrossRef\]](#)
28. Chouliara, V.; Koukaras, P.; Tjortjis, C. Fake News Detection Utilizing Textual Cues. In *Artificial Intelligence Applications and Innovations*; Maglogiannis, I., Iliadis, L., MacIntyre, J., Dominguez, M., Eds.; Springer Nature Switzerland: Cham, Switzerland, 2023; pp. 393–403.
29. Ahmed, H.; Traore, I.; Saad, S. Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. In *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*; Traore, I., Woungang, I., Awad, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 127–138.
30. Rai, N.; Kumar, D.; Kaushik, N.; Raj, C.; Ali, A. Fake News Classification using transformer based enhanced LSTM and BERT. *Int. J. Cogn. Comput. Eng.* **2022**, *3*, 98–105. [\[CrossRef\]](#)
31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–11.
32. Luvembe, A.M.; Li, W.; Li, S.; Liu, F.; Xu, G. Dual emotion based fake news detection: A deep attention-weight update approach. *Inf. Process. Manag.* **2023**, *60*, 103354. [\[CrossRef\]](#)
33. Onan, A.; Toçoğlu, M.A. A Term Weighted Neural Language Model and Stacked Bidirectional LSTM Based Framework for Sarcasm Identification. *IEEE Access* **2021**, *9*, 7701–7722. [\[CrossRef\]](#)
34. Sahoo, S.R.; Gupta, B.B. Multiple features based approach for automatic fake news detection on social networks using deep learning. *Appl. Soft Comput.* **2021**, *100*, 106983. [\[CrossRef\]](#)

35. Trueman, T.E.; Kumar, A.; Narayanasamy, P.; Vidya, J. Attention-based C-BiLSTM for fake news detection. *Appl. Soft Comput.* **2021**, *110*, 107600. [[CrossRef](#)]
36. Siino, M.; Tinnirello, I.; La Cascia, M. T100: A modern classic ensemble to profile irony and stereotype spreaders. In Proceedings of the CLEF 2022: Conference and Labs of the Evaluation Forum, Bologna, Italy, 5–8 September 2022; pp. 2666–2674.
37. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2019**, arXiv:1810.04805.
38. Kaila, D.R.P.; Prasad, D.A.V. Informational flow on Twitter–Corona virus outbreak–topic modelling approach. *Int. J. Adv. Res. Eng. Technol. (IJARET)* **2020**, *11*, 7.
39. Mohapatra, A.; Thota, N.; Prakasam, P. Fake news detection and classification using hybrid BiLSTM and self-attention model. *Multimedia Tools Appl.* **2022**, *81*, 18503–18519. [[CrossRef](#)]
40. Mangione, S.; Siino, M.; Garbo, G. Improving Irony and Stereotype Spreaders Detection using Data Augmentation and Convolutional Neural Network. In Proceedings of the CLEF 2022—Conference and Labs of the Evaluation Forum, Bologna, Italy, 5–8 September 2022; pp. 2585–2593.
41. Tinn, R.; Cheng, H.; Gu, Y.; Usuyama, N.; Liu, X.; Naumann, T.; Gao, J.; Poon, H. Fine-tuning large neural language models for biomedical natural language processing. *Patterns* **2023**, *4*, 100729. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.