Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 10, 1515-1540
2025
Publisher: Learning Gate
DOI: 10.55214/2576-8484.v9i10.10684

A Transfer learning-based model for fake news detection

Oluwakemi Christaina ABIKOYE[1], Taofeek Akintunde ABDULSALAM[2], Enesi Femi AMINU[3], Joseph Bamidele AWOTUNDE[4], Abidemi Emmanuel ADENIYI[5]
[1]Department of Computer Science, University of Ilorin, 240003, Nigeria; abikoye.o@unilorin.edu.ng (O.C.A.).
[2]Department of Computer Science, Lens Polytechnic Offa, Kwara, Nigeria; binarycode1995@gmail.com (T.A.A.).
[3]Department of Computer Science, Federal University of Technology, Minna, Niger, Nigeria; binarycode1995@gmail.com (E.F.A.).
[4]Department of Computer Science, University of Ilorin, 24003, Nigeria; awotunde.jb@unilorin.edu.ng (J.B.A.).
[5]Department of Computer Science, College of Computing and Communication Studies, Bowen University, Iwo, Nigeria; abidemi.adeniyi@bowen.edu.ng (A.E.A.).
[5]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India.

Abstract: The world we know today has vast information, events, data, and news spreading rapidly across the globe at exponential rates. The invention of modern telecommunication technology facilitates this exponential growth of fake news. With the exponential dissemination of news, fake news has also gained the potential to spread fast across social media and blog post spaces. Fake news is primarily aimed at spreading false information that is capable of influencing the opinion of the masses negatively. Previous researchers have proposed many techniques, including machine learning and deep learning, for fake news detection. Extensive studies show that past researchers have not yet established satisfactory performance, while some suffer from model overfitting on training sets with poor performance on testing data. Achieving optimum performance using a machine learning approach is highly difficult, especially when a statistical approach is adopted for the Natural Language Processing task. Hence, this study proposed a Transfer Learning-based Bidirectional Long Short-Term Memory to Machine Learning (Bi-LSTM-2-ML) model for the detection of fake news. The Bi-LSTM architecture is pre-trained on the phoney news dataset for extracting highly contextualized feature embedding, which is further used to enhance the training process of various machine learning algorithms, including SVM, LR, NB, DT, and KNN. The experimental results show that the Bi-LSTM-2-ML fake news detection model yielded better performance than previous approaches, and the Bi-LSTM-2-ML model enhanced the performance of machine learning algorithms with an 8.54% average increase in accuracy. In conclusion, the Bi-LSTM-2-ML model has demonstrated better performance by enhancing machine learning algorithms for detecting fake news.
Keywords: *Bi-LSTM, Deep learning, internet, Fake news, Machine learning, Social media, Transfer learning.*

1. Introduction

In recent years, social media has emerged as the primary platform for news and information dissemination, facilitated by the internet, enabling an environment for the easy posting of media content [1]. Recent research indicates a shift in news consumption, with many users preferring digital-based news over traditional newspapers. However, the authenticity of news on social media (digital news) is a challenging process, unlike content from radio and television, which are subjected to critical review and supervision before broadcasting [2]. The detection of fake news remains essential due to the widespread dissemination of misinformation, rumors, and fake news on platforms such as Facebook, Instagram, Twitter, and WhatsApp [3]. Training intelligent systems to detect fake news online is challenging due to the diverse forms and sources of interdisciplinary news. To effectively detect phony news, various

History: Received: 25 August 2025; Revised: 6 October 2025; Accepted: 10 October 2025; Published: 24 October 2025
* Correspondence: abidemi.adeniyi@bowen.edu.ng

Natural Language Processing (NLP) tools, such as part-of-speech tagging, sentiment analysis, and vectorization techniques, need to be adopted [4]. Additionally, other resources and tools are available to help individuals identify and spot fake news on their social media feeds by analyzing essential facts and critically evaluating news content. The primary objective of fake news detection is to prevent the quick dissemination of false information across various platforms, including social media and messaging applications. Fake news detection systems can result in harmful consequences, such as mob violence, if not properly addressed [3]. Additionally, Hossain et al. [5] revealed that fake news detection aims to identify misleading information and prevent unwarranted acts of violence, which could ultimately safeguard societal peace. Researchers have proposed machine learning algorithms and deep learning techniques in the past for fake news detection, including Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbour (KNN), Logistic Regression (LR), Decision Tree (DT), Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), and Bidirectional Long Short-Term Memory (Bi-LSTM) [6].

Artificial intelligence (AI) is a broad term denoting the field of study where machines are trained to perform intelligent tasks without human intervention [7]. The AI domain is divided into subfields such as knowledge processing, pattern recognition, natural language processing, deep learning, and machine learning for building intelligent systems. 2021 reveals that machine learning technology greatly reduces human work by using statistical and mathematical algorithms to achieve a task more efficiently. Natural language processing (NLP) provides the tools and technology to build communication interfaces between humans and computers using natural languages and computational algorithms [8]. NLP has wide applications in medicine, language translation, sentiment analysis, and false news detection Khurana et al. [9]. Patel and Patel [10] confirmed that deep learning models have been successfully deployed to address natural language processing tasks with high-performance records; commonly used algorithms include Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN), and Bi-LSTM. The RNN architecture is the basis of all recurrent-based network architectures, which are primarily used for natural language processing tasks and processing of sequential-based data (text data) in a feed-forward manner [10]. Long Short-Term Memory (LSTM) architecture comprehensively addresses short-term memory limitations in a recurrent neural network. The architecture incorporates internal mechanisms known as gates, which can control the flow of information [5]. LSTMs are widely adopted by researchers to address text-based classification problems [11] and Bidirectional LSTM (Bi-LSTM) models have demonstrated high effectiveness in gathering sequential information from both forward and backward directions of the text data [12].

1.1. Motivation

Pandey et al. [6] revealed that the drastic change in news dissemination from traditional newspapers to electronic-based news forecasting has contributed to the widespread issues of fake news in today's world. This resulted from a limited or no supervision policy scheme for monitoring news being published online [5]. Many researchers in the past have proposed various techniques to address this issue, but with less satisfactory performance [6, 13].

Achieving optimal performance in the domain of NLP is highly challenging. Hence, concepts such as transfer learning and hybridization techniques are suggested by researchers in this field as Kaliyar et al. [14], Jain et al. [3], Kumar et al. [11], and Jaybhaye [13] to improve the performance of a real-time fake news detection model. Additionally, Pandey et al. [6] and Jaybhaye [13] utilize various machine learning algorithms for building fake news detection models, including Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Logistic Regression (LR), Decision Tree (DT), and Naïve Bayes (NB) algorithms. The result presentation of Pandey, et al. [6] shows that the highest accuracy attained in their study is 90.04%, which creates room for further improvement, and Jaybhaye [13] suffers from overfitting during training with low performance on testing data. Hence, this study proposes an enhanced machine learning-based algorithm using transfer learning and hybridization techniques.

The primary aim is to enhance the performance of a machine learning-based fake news detection model using the Transfer Knowledge approach from Bi-directional Long Short-Term Memory. Furthermore, an extensive experimental study will be conducted on the effect of Pre-Trained Bi-LSTM weights on the performance of various Machine Learning Algorithms (LR, SVM, DT, NB, and DT). Finally, the proposed Bi-LSTM-2-ML model will be evaluated against existing machine learning models. The research work intends to solidify the hypothetical claim and suggestion made by previous researchers on the possibilities of improving the performance of fake news detection models, along with their accuracy, using transfer learning techniques. Hence, proper study and comparative experiments will be conducted between benchmark studies, Pandey et al. [6] and Jaybhaye [13] the proposed Bi-LSMT-2-ML transfer learning fake news detection model for proper verification.

The following are the key contributions of the proposed model:

- The study enhances the performance of an ML-based fake news detection model using the transfer knowledge approach called Bi-LSTM.
- An extensive experimental study will analyze the impact of pre-trained Bi-LSTM weights on the performance of different machine learning algorithms, including LR, SVM, DT, and NB.
- The Bi-LSTM-2-ML model was evaluated and compared with existing ML-based models to establish its effectiveness.
- A comprehensive comparative study was conducted with the benchmark studies to ensure proper evaluation of the proposed model.

The remainder of the study is as follows: Section 2 discusses related work, Section 3 discusses the methodology used in this study, Section 4 discusses the results and discussion, and Section 5 concludes the study.

2. Related Work

The rapid proliferation of misinformation on digital platforms has raised significant concerns about its impact on society, necessitating effective methods for identifying and mitigating fake news [15, 16]. Recent studies indicate that ML and DL models significantly outperform traditional methods in detecting fake news, showcasing their robustness across various datasets. These advanced models leverage sophisticated algorithms and embeddings to enhance classification accuracy, making them essential tools in combating misinformation. Models like CNN, BERT, and GPT have shown exceptional capabilities in identifying fake news by understanding contextual language and detecting patterns in text [17]. Combining DL with traditional ML methods has led to improved performance metrics, including accuracy and F1 scores [18]. Techniques such as Bi-GRU and Bi-LSTM have demonstrated superior results in Arabic fake news detection, achieving F1 scores above 0.98 [19]. Research has explored multimodal approaches that analyze text, images, and videos, enhancing detection capabilities beyond text alone [20]. The use of embeddings such as TF-IDF and FastText has been crucial in improving model performance, with TF-IDF yielding the best results for certain classifiers [21]. While these advanced models show promise, challenges remain in ensuring their adaptability across different languages and media formats, highlighting the need for ongoing research in this area.

Apostol and Truică [22] discovered that the transformation and digitization of social media come with many advantages but also open the doors to widespread disinformation and malformation via fake news Pandey et al. [6] and Kaliyar et al. [14]. Apostol and Truică [22] used a document embedding approach to build multiple models and present an architecture based on multi-labelled classification. The results obtained show that the approach performs better than the complex state-of-the-art deep neural network model. Hossain et al. [5] developed a Bi-LSTM model with GloVe, FastText word embedding, and applied cross-fold validation to detect Bangla fake news. A Gated Recurrent Unit (GRU) deep learning model is introduced to detect fake news. Based on the experimental analysis conducted, an accuracy of 96% was achieved using the Bi-LSTM deep learning model, and 77% accuracy was achieved using the GRU model for detecting and classifying fake news from real news.

Pandey et al. [6] and Jaybhaye [13] introduce machine learning techniques such as K-Nearest Neighbour, Support Vector Machine, Decision Tree, and Logistic Regression classifiers to predict and identify fake news from real news accurately. Performance comparison of Pandey et al. [6] shows that logistic regression performs best with 90.48% accuracy. However, Jaybhaye [13] face challenges of data overfitting during training, which resulted in poor performance on testing data. The researcher obtained 86% accuracy using the LSTM model, and it was suggested that the researcher build a fake news model due to its better performance than the machine learning model. However, Kaliyar et al. [14] revealed that the use of bidirectional encoding for fake news detection performs better than the unidirectional-based model. Hence, the researcher proposed a Bidirectional Encoder Representation from a transformer (FakeBERT) model integrated with a single-layer deep CNN network. As a result, the developed model achieved an accuracy of 98.90%, which outperformed the existing study.

7Furthermore, Nasir et al. [23] adopt the hybridization of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) for detecting fake news. The proposed CNN-RNN model was trained on both ISOT and FAKE datasets, with approximately 100% accuracy achieved using the ISOT dataset and 60% accuracy using the FAKE dataset. Thota et al. [24] also address the same issues using the hybridization of Bidirectional Encoder Representations from Transformers (BERT), Alite-BERT, and Transformer-XL Style Segmented Memory (XLNet). The model performs efficiently with 98% precision, recall, accuracy, and F1-score.

In today's world, social media platforms are ranked as the primary internet tools used to spread news across the globe. Choudhary and Arora [25] claim that social media is an essential tool for disseminating information and creating awareness among the public, but it has no authentication schemes for validating the origin and content of all published news. A linguistic model for extracting essential syntactic and semantic information describing the features of fake news from real news is proposed by the researchers to address the widespread spread of phony news uploaded on the internet. The developed model was able to achieve an average accuracy of 86% for detecting and classifying phony news. The comparative analysis shows that the developed model performs more efficiently than the existing models. Vijjali et al. [26] explore the alarming rate of fake news occurrence during the COVID-19 pandemic, which results in people finding it difficult to differentiate false information from real COVID-19 news. Hence, the researcher proposed a state-of-the-art machine learning model based on a two-level automated pipeline for detecting COVID-19 fake news. The study reveals that the Bidirectional Encoder Representations from Transformers (BERT) (92%) and the ALBERT (94%) models perform satisfactorily.

Furthermore, most fake information created on the internet today is a result of manipulated image, audio, text, and video data being uploaded online. Singhal et al. [27] revealed that manipulated data could easily spread via social media due to the fast-growing nature of social media users. The researcher proposed a multimodal-based fake news detection system, which internally utilized a Bidirectional Encoder from Transformer (BERT) architecture to extract features from text and utilized the VGG19 pre-trained model for image feature extraction. As a result, the developed model outperformed the existing state-of-the-art model trained on Twitter and Weibo datasets with an accuracy increase of 3.27% and 6.83%.

Kumari and Singh [28] presented a DL multimodal framework for fake news detection that significantly outperforms existing state-of-the-art techniques. It achieves an accuracy of 99.22% for text analysis and 93.12% when combining text and images. This demonstrates that advanced ML and DL models can effectively address the challenges of fake news detection, surpassing traditional methods by leveraging NLP and image analysis through LSTM networks and the CLIP model. Gul et al. [29] introduced an innovative approach that combines GNNs with advanced DL techniques to enhance the detection of fake news, moving beyond traditional methods that focus mainly on text analysis and

DOI: 10.55214/2576-8484.v9i10.10684

source credibility assessment. This proposed model leverages the structural information of news propagation networks to better understand the connections and patterns indicative of misinformation.

Sari et al. [30] developed a fake news detection model utilizing an enhanced CNN-BiLSTM architecture combined with GloVe word embedding techniques, achieving a high accuracy of 96% in distinguishing fake news from real news. This demonstrates the effectiveness of the proposed model in addressing the challenges posed by the spread of fake news in the digital age. The research highlights the importance of training data ratios, showing that a larger training data ratio leads to better model performance. Brinda et al. [31] proposed a heterogeneous deep learning model that enhances fake news detection by incorporating dissemination structure features from social networks, such as user interactions, sharing patterns, and network topology, which provide deeper insights into how information spreads beyond just text analysis. Deepthi and Shastry [32] proposed an unsupervised misinformation detection technique (UMD) that achieves an overall accuracy of 92.82%, outperforming existing models such as transformer and hybrid CNN-RNN models by approximately 3%, demonstrating the effectiveness of deep learning in fake news detection.

Zhu [33] highlighted that DL models, such as OPCNN-FAKE, DC-CNN, and CNN-LSTM, show significant accuracy advantages over traditional machine learning methods in fake news detection, addressing the limitations of conventional approaches in handling expanding data sets and streams. Ladouceur et al. [34] highlight that DL algorithms, specifically BERT and LSTM, have demonstrated superior performance in fake news detection compared to traditional methods, particularly when incorporating context features, achieving better accuracy in classifying news from the COVID-19 dataset. Porya et al. [35] investigated various models, including deep learning algorithms such as CNN and LSTM, alongside conventional methods like RF and GB. It highlights challenges in achieving effective fake news detection, indicating that robust performance remains elusive despite advancements.

Traditional ML-based models often require extensive labeled datasets, which can be costly and time-consuming to attain [36]. To address this, transfer learning has emerged as a promising approach, leveraging pre-trained models to adapt to new tasks with minimal data requirements. By utilizing pre-trained language models like BERT, RoBERTa, or XLNet, transfer learning enables the extraction of deep semantic features, enhancing the accuracy and efficiency of fake news detection systems [19, 37]. Recent studies demonstrate that transfer learning models can outperform conventional methods, achieving robust performance across diverse fake news datasets [38-42]. This section explores the latest advancements in transfer learning for fake news detection, with a focus on model architectures, dataset utilization, and cross-domain adaptability, drawing insights from recent research in this dynamic field.

Table 1.
Summarised the related work in this study

Authors	Methods	Dataset	Contributions	Research Gap
Roumeliotis et al. [17]	CNN, BERT, and GPT	Self-dataset	The research addresses the growing need for scalable and unbiased tools to identify fake news across diverse platforms and languages, which could assist journalists, social media platforms, and policymakers in mitigating the spread of misinformation.	The study does not explicitly address the limitations or challenges faced by each ML model (CNN, BERT, and GPT) in the context of fake news.
Almandouh et al. [19]	CNN-LSTM, RNN-CNN, RNN-LSTM, and Bi-GRU-Bi-LSTM	AFND and ARABICFAKETWEETS	The research integrates FastText word embeddings with various ML and DL methods, and employs advanced transformer-based models, optimizing their performance through hyperparameter tuning for effective Arabic fake news detection.	The research primarily focuses on Arabic fake news detection, indicating a gap in the application of the developed models to other languages. Future research could expand the methodologies and models to accommodate multilingual fake news detection, addressing the global nature of misinformation.
Al-Alshaqi et al. [20]	ML-based model, BERT+CNN	ISOT & MediaEval 2016	The study proposes a comprehensive framework for fake news detection that integrates multiple modalities, including text, images, and videos, utilizing both ML and DL techniques. This approach addresses the complexities of modern information dissemination and manipulation.	While the Random Forest classifier achieved high accuracy for textual data, the research does not explore the limitations or potential biases of the classifiers used, nor does it provide insights into the interpretability of the models, which are critical for understanding the decision-making process in fake news detection.
Al-Tarawneh et al. [21]	SVMs, MLPs, & CNNs	TruthSeeker	The study emphasizes the importance of selecting appropriate embedding techniques based on model architecture to maximize fake news detection performance.	The study suggests future research directions, such as integrating contextual embeddings and exploring hybrid model architectures to further improve detection capabilities against misinformation.
Kumari and Singh [28]	NLP, LSTM, & CLIP	Fakeddit	The study introduces a comprehensive multimodal framework for fake news detection that integrates advanced techniques to analyze both multilingual text and visual data, addressing the limitations of existing methods that primarily focus on linguistic characteristics.	The paper highlights the need for a more efficient model that integrates both text and visual data analysis, suggesting that current approaches may lack a comprehensive multimodal framework that leverages advanced techniques for improved detection accuracy.
Gul et al. [29]	GNNs		The model highlights the effectiveness of GNNs in addressing the challenges posed by the spread of false information.	The paper does not address the limitations or potential biases in the data used for training the GNN model, which could affect the generalizability and reliability of the fake news

				detection results.
Sari et al. [30]	CNN-BiLSTM	WELFake	The study developed a fake news detection model using an optimized CNN-BiLSTM architecture, achieving an accuracy of 96% on the WELFake dataset.	The study recommends further research to explore different preprocessing techniques, indicating that the current methods may not be exhaustive and that alternative approaches could enhance the model's performance in fake news detection.
Brinda et al. [31]	BERT+ ULMFiT	Quint Dataset, Boom Live	The model achieves high accuracy rates, demonstrating its effectiveness in combating evolving fake news tactics, with a reported accuracy of 95.27% on The Boom Live dataset and 92.1% on The Quint dataset, showcasing its generalizability and robustness in detecting fake news.	The paper highlights the limitations of traditional machine learning methods that rely solely on semantic text analysis, indicating a research gap in developing more robust detection methods that incorporate additional features beyond text, such as dissemination structure features from social networks.
Deepthi and Shastry [32]	Bi-GRU	not provided	The study highlights the effectiveness of deep learning in detecting fake news.	Subjectivity and bias in news authenticity comparison, and challenges in contextualising news content effectively.
Zhu [33]	OPCNN-FAKE, DC-CNN, and CNN-LSTM	not provided	These models effectively address the limitations of conventional approaches in managing growing data sets and continuous data streams.	Poor scalability of DL models for fake news and slow training speed of existing DL models.
Ladouceur, et al. [34]	BERT+LSTM	5,000 news articles extracted from the COVID-19 dataset	The model demonstrated superior performance in fake news detection compared to traditional methods.	Problem of feature selection
Porya, et al. [35]	RF, GB, CNN, & LSTM	Two widely used datasets	The paper explores preprocessing techniques and model architectures, tested on two datasets, contributing to fake news detection research and guiding future advancements for newcomers.	Model implementation issues hinder effective fake news detection, and a lack of clean, unbiased data poses a challenge in research.

3. Research Methodology

The study proposed a Transfer learning-based Bidirectional Long-Short-Term Memory (Bi-LSTM) for a machine learning (ML) fake news detection model (Bi-LSTM-2-ML model). Technically, the Bi-LSTM architecture is used to pre-train the fake news detection model to generate highly contextualized bidirectional feature encodings or vectors, which are later transferred to a machine learning algorithm for further training.

3.1. Methodological Process

Figure 1 illustrates the conceptual framework of how the Bidirectional Long Short-Term Memory (Bi-LSTM) to Machine Learning (ML) transfer learning-based model (Bi-LSMT-2-ML) for detecting fake news will be designed and developed. The various technologies, entities, and tools that will be utilized in designing and developing the Bi-LSTM-2-ML fake news detection model are also described in Figure 1. The fake news dataset is downloaded from the Kaggle repository in a Comma Separated Values (CSV) format, and the data scientist uses the Jupyter Notebook environment to load the downloaded fake news dataset into the IDE using the Python 'load_csv' function. Furthermore, the loaded fake news data will be forwarded to stage (1), or the preprocessing stage, which includes data exploration, data cleaning, and data visualization.

The exploration helps to gain more insights into the dataset and identify possible correlations, relationships, or features that can assist during machine learning prediction. The X and Y variables (independent and dependent variables) of the fake news dataset are extracted in stage 2. Furthermore, a transformation (Word2Vec) will be conducted on the raw text to convert the text data into numerical features or word embeddings. After feature extraction, the next step is to split the dataset into training and testing samples. The training sample will be utilized to train the Bi-LSTM deep learning model to extract highly contextualized information about the fake news data, then use the pre-trained Bi-LSTM weights to train the machine learning algorithm in a transfer learning approach. The developed model will then be evaluated using various performance metrics such as accuracy, precision, recall, and F1-score.

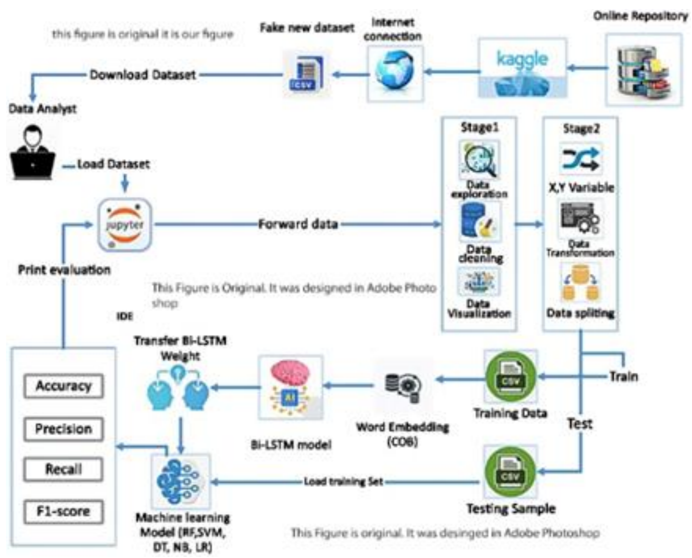

Figure 1.
Transfer Learning Bi-LSTM-2-ML Proposed Framework.

Figure 1 illustrates the methodological process involved in developing the proposed Transfer Learning Bi-LSTM-2-ML Fake News Detection Model using transfer learning. The proposed stages include the following:

3.1.1. Fake News Data Collection

This first stage includes data collection and gathering from the Kaggle repository. The dataset is downloaded in a comma-separated value (CSV) format into the local system drive for further processing. However, the Kaggle repository provides data scientists with a variety of datasets for addressing different tasks, a development environment (Jupyter Notebook), and virtual hardware processing units (GPU, TPU, and CPU) for solving complex tasks. The fake news dataset link on the Kaggle repository is provided on the website.
https://github.com/GeorgeMcIntire/fake_real_news_dataset

3.1.2. Data Exploration and Visualisation

This stage is also known as data exploratory analysis (EDA); it involves conducting an in-depth study of the fake news dataset to extract important insights and explore the dataset structure. Additionally, this step helps uncover hidden patterns, empty strings, duplicate samples, stop words, and special symbols present in the dataset. Data visualization refers to a graphical representation of information and exploration conducted on the dataset. This study will use various charts from the Python library (Matplotlib.pyplot) to perform visual exploration. Firstly, the pandas 'read_csv' function is used to load the dataset into a table of rows and columns, which is generally referred to as the Pandas DataFrame. The Pandas DataFrame's rows and columns can be accessed via indexing, and various columns can be indexed using column names. The sample fake news dataset is indexed using the column name 'text', and with a numeric index of '0', the first row of data in that column can be accessed. The first square bracket denotes the column name, while the second bracket denotes the row number. Based on the sample text explored, it was revealed that the fake news text samples contain symbols, punctuations, escape characters, and stop words that have no meaning or contextual contribution to the fake news text samples. Moreover, it is essential to explore the shape of the dataset, which indicates the number of rows (samples) and columns (features) in the dataset. Figure 2 exploration revealed that the dataset contains 4,594 news-labeled samples and 2 features (text, label). The labels denote the response variable of either fake or real news.

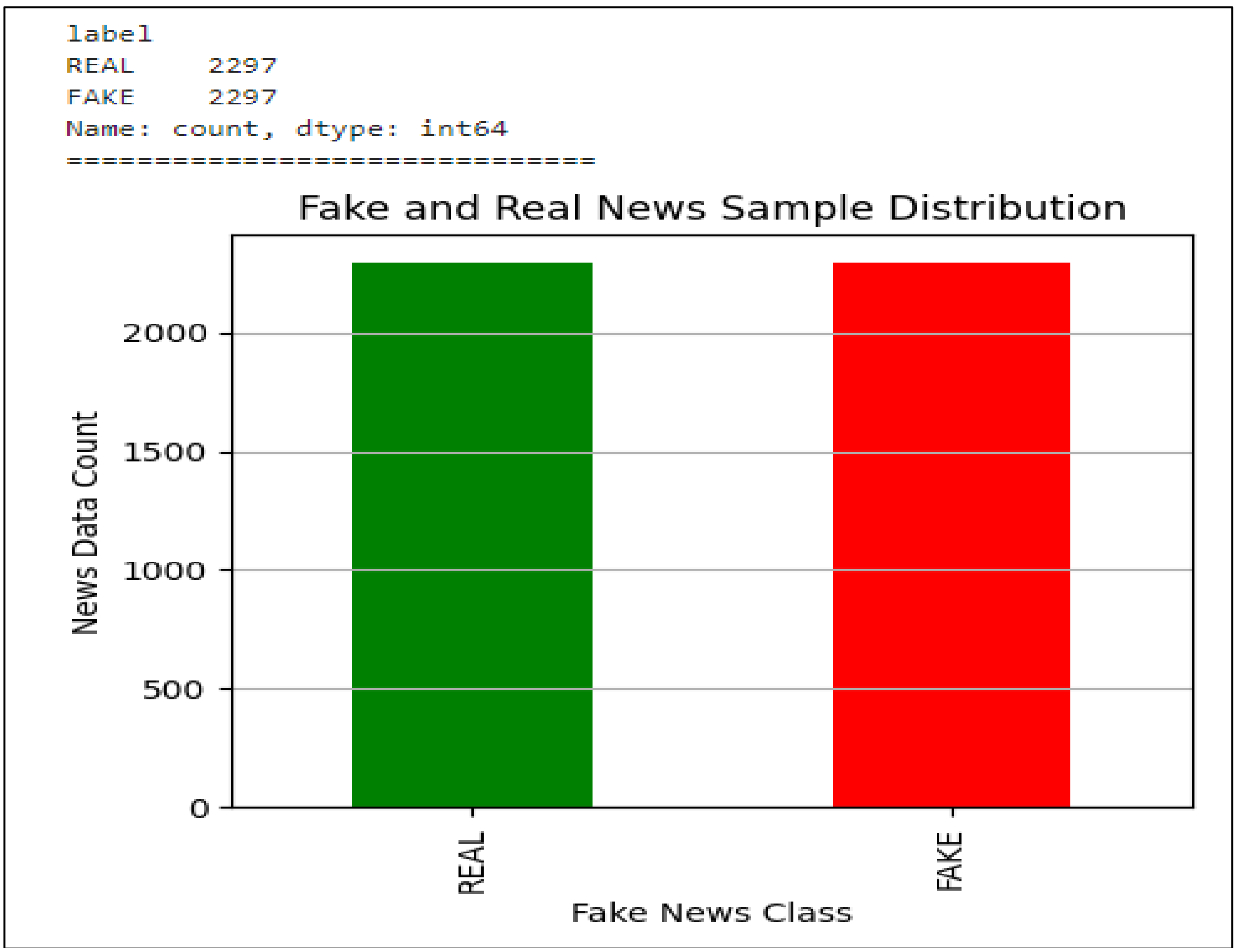

Figure 2.
Data Distribution based on Label Class.

ISSN: 2576-8484
Vol. 9, No. 10: 1515-1540, 2025
DOI: 10.55214/2576-8484.v9i10.10684

Figure 2 shows a bar chart revealing the label class distribution (fake sample count and real sample count). This reveals that the real and fake classes contain equal numbers of data class groups (FAKE and REAL). Each class contains 2,297 samples. The exploration conducted shows that the dataset is balanced.

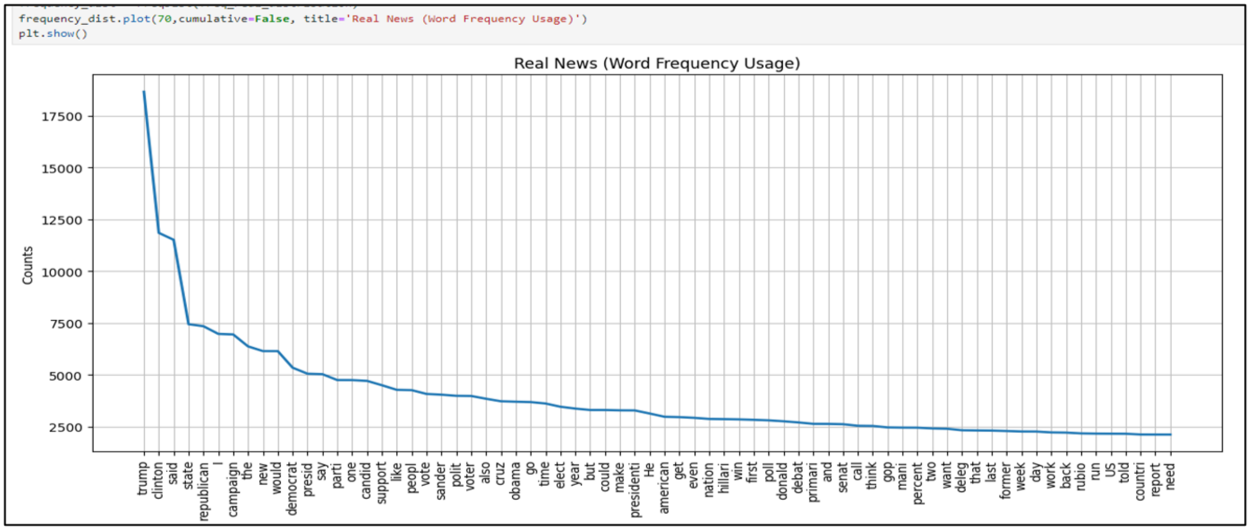

Figure 3.
Word Frequency Usage (Real News Class).

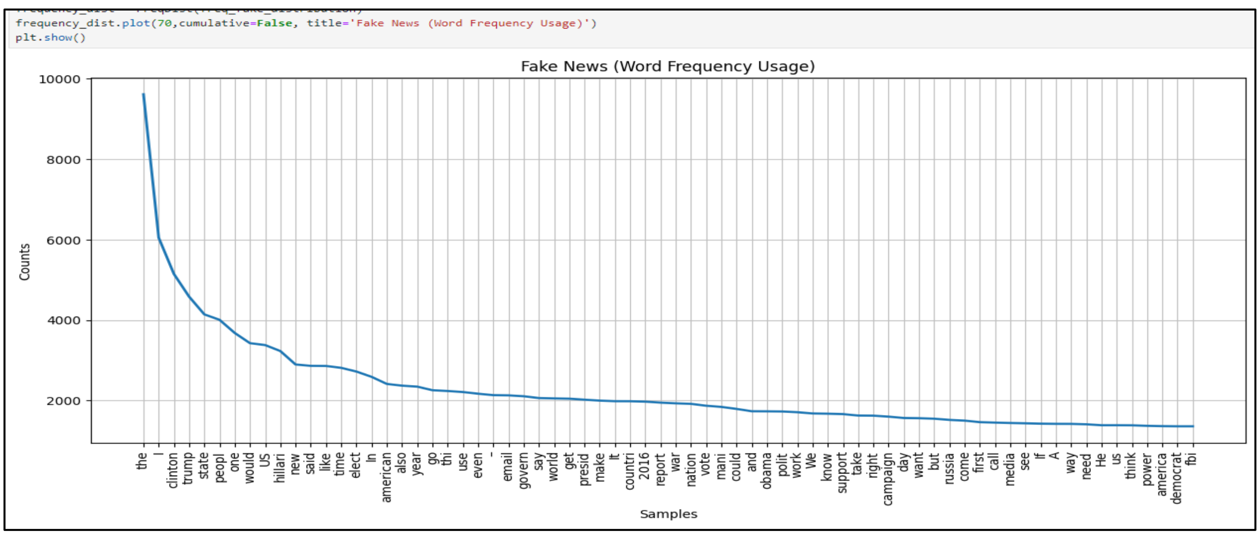

Figure 4.
Word Frequency Usage (Fake News Class).

The frequency count analysis provides a statistical distribution of vocabulary usage. With the analysis, the most frequently used terms in fake news samples and real news samples can be seen using the frequency line plot in Figures 3 and 4. The line chart represents the word count or frequency. Figure 3 revealed that the word 'Trump' is the most frequently used term in real news samples, followed by 'Clinton', 'said', 'state', 'Republican', and 'I', with a frequency of above 7,500. While 'The' is the most frequently used term in fake news samples, as shown in Figure 4, followed by 'I', 'Clinton', 'Trump', 'state', and 'people,' with a frequency of above 4,000.

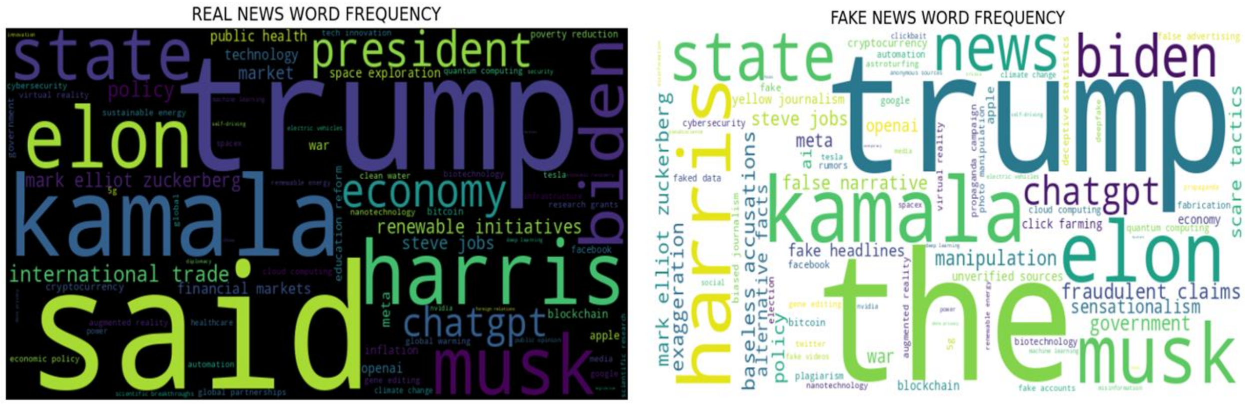

Figure 5.
WordCloud Representation of Fake News and Real News.

Figure 5 shows the WordCloud representation of word frequency usage in both real and fake news vocabulary. The word cloud uses text size to describe or explore how words are frequently used. The larger the text, the more frequently the vocabulary appears. Based on Figure 5, it shows that the words 'Trump,' 'state,' 'Clinton,' and 'said' are bigger in the text, revealing that they are mostly used in real news vocabulary. However, the same figure reveals that the terms 'the,' 'I,' 'Clinton,' 'state,' 'people,' 'us,' and the like are mostly used in fake news vocabulary.

3.1.3. Data Preprocessing and Cleaning

The Natural Language Tool Kit (NLTK) is an essential tool for solving natural language processing tasks [8]. Data preprocessing and cleaning are conducted using modules (stopwords, PorterStemmer, and tokenizer) from the NLTK library. The cleaning process is based on the exploration previously conducted on the dataset. The cleaning steps include removing stopwords, punctuation marks, special symbols, and special characters, while the preprocessing activities include stemming, tokenizing, and text case conversion.

Furthermore, to perform the preprocessing operation, three helper functions (remove_punctuation, remove_stopwords, and stem_and_lower_text) are created to remove punctuations, special characters, stop words, and reduce similar vocabulary to their stem words using PorterStemmer from the NLTK library. Moreover, the label encoding involves the transformation of the fake news label (FAKE, TRUE) into a numeric representation for machine learning processing. In this study, the FAKE label entry is transformed into zeros (0), while the TRUE entries are transformed into ones (1). The resulting data frame after transformation is further processed during feature extraction.

3.1.4. Feature extraction (WordEmbeddings)

The process of transforming text data into numeric vectors is known as vectorization. However, Word Embedding refers to feature vectors generated using deep learning-based architectures. Standard techniques used include Word2Vec, FastText, and GloVe [43]. This study utilises the Word2Vec (CBOW) technique for generating 2-dimensional features from the fake news dataset.

3.1.4.1. Word2Vec (CBOW)

It is essential to extract numerical features from textual data for proper machine learning computation. This study utilized the Word2Vec (CBOW) model from the Gensim library to extract feature vectors for each vocabulary with 300 vectors (dimensions). However, the feature vector of 'trump' is explored, and it is revealed that each feature vector is stored in an array of length 300. With a 2-D feature vector matrix table format of columns and rows, a matrix can be generated for each news sample sentence. The columns represent the numeric features, and the rows contain a 300-dimensional long feature vector representing that particular sample. Figure 14 shows the feature matrix of 300 features by 4,594 samples that was generated for the fake news dataset.

3.1.4.2. Sequencing & Embedding Matrix

Sequencing and generation of the embedding matrix is a technique used to create a special feature that can be passed to the embedding layer of the Bi-LSTM architecture. The TensorFlow library will be used to perform various preprocessing operations and extraction of numeric sequences for the embedding layer. The study uses the TensorFlow library for preprocessing and converting fake news text into a numeric sequence. This preprocessing and transformation are essential to have a suitable format for proper loading into the embedding layer of the Bi-LSTM architecture. A maximum vocabulary of 10,000 is generated, and each row (news samples) is tokenized and padded to 500 long; this ensures that all fake news data samples are in an equal sequence (word length) for each text sample. Finally, the texts are converted to numeric representation using the *text_to_sequence* function of the TensorFlow library.

The sequence generated is used to create the embedding matrix. Therefore, generating the embedding matrix also requires the word2vec representation and utilizes each word from the generated vocabulary. As a result of this process, an 80,859 by 500 (2-dimensional) embedding matrix is produced. The embedding matrix can now be passed as an embedding parameter to the Bi-LSTM input layer. Algorithm 1 illustrates the pre-processing steps.

3.1.5. Data Splitting

Data splitting entails dividing the dataset into training and testing samples, which results in having separate data for training the fake news detection model and evaluating the model's performance using the remaining samples (testing samples). 70% of the fake news dataset is allocated for training, and the remaining 30% is assigned to model evaluation. This study uses the *model_selection.train_test_split* function from the scikit-learn module to split the fake news dataset.

Algorithm 1: Pre-processing

Input: The Dataset
Output: Pre-processing to be fed to the model
Step 1: Text Cleaning
For each data point, do
Remove URLs, special characters, numerical values, emoticons, emojis, mentions, and hashtags.
Remove non-alphanumeric characters
Remove extra spaces and trim the text.
Convert text to lowercase.
end for
Step 2: Tokenisation
For each cleaned data *set*, do
Split *the text* into individual words (tokens)
end for
Step 3: Stop Word Removal
For each token $w \in$ *data* do
If w is a stop word, then
Remove w from *data*
end if

```
end for
Step 4: Train-Test Split
Split the dataset into a training set and a testing set.
Step 5: Word2Vec Embedding
Train a Word2Vec model on a dataset.
Vectorise the data using Word2Vec
Step 6: FastText Embedding
Train the FastText model on the data.
Vectorise training and test data using FastText
Step 7: Handle NaNs in Embeddings
For each vector v ∈ T_train, T_test do
If any value in v is NaN, then
Replace v with a zero vector.
end if
end for
Return Preprocessed Data
```

3.1.6. Bi-LSTM Training

The Bidirectional Long Short-Term Memory (Bi-LSTM) architecture is a recurrent neural network that processes text input data in both directions. This architecture will be used to pre-train the fake news prediction model to extract highly contextualized word embedding vectors, which will be used as input vectors for various machine learning algorithms during training in a transfer learning approach. The Bi-LSTM architecture layers are created using TensorFlow and the Keras module. A sequential Bi-LSTM architecture is created with seven layers to train the fake news detection models. The architecture consists of one input layer, one embedding layer, two Bidirectional LSTM layers, one LSTM layer, one fully connected dense layer, and finally one output layer.

Embedding layer processes 500 vectors of numeric values as input, while the Bi-LSTM, LSTM, and Dense layers consist of 256 neurons, 64 neurons, and 32 neurons for weight, bias, and loss computation. Finally, the output layer has only one neuron denoting either fake (0) or real (1) news output.

3.1.7. Loading Pre-Trained Bi-LSTM Weight

The highly contextualized weights or features from the Bi-LSTM model are saved and reloaded (reused) as pre-trained weights. The weights are used as feature vectors to train various machine learning algorithms.

3.1.8. Machine learning Algorithms

Five machine learning algorithms, including Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Random Forest (RF), Naïve Bayes (NB), and Decision Tree (DT), will be trained to detect fake news. The pre-trained fake news weight (word embeddings from Bi-LSTM) will be used to enhance the training process of the listed machine learning algorithms.

3.1.8.1. Naïve Bayes (NB) Algorithm

Naïve Bayes is widely used to compute conditional probability, which originates from Bayes' theorem by expressing the likelihood of an event occurring based on the occurrence of another event [44]. Naïve Bayes functions as a classifier within the domain of supervised learning algorithms, and it performs predictions based on probability scores from different classes [45].

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$
(1)

$P(A|B)$: Probability of event A such that event B has already occurred.

3.1.8.2. Support Vector Machine (SVM)

SVM Algorithm is highly efficient for addressing binary classification problems [3]. SVM is a supervised machine learning algorithm that can be used for solving both classification and regression problems; it is based on the idea of finding a hyperplane that best separates the dataset into two classes. Hyperplanes are decision boundaries that help machine learning algorithms classify data points accurately [3]. The algorithm provided in Figure X is based on feature set N (0-N) and is sorted using information gained in decreasing order of accuracy.

3.1.8.3. K-Nearest Neighbours (KNN)

The K-NN algorithm operates on the concept of similarity measures between new samples and existing samples. The algorithm groups the new sample point into the category that is most likely to be one of the available categories. It stores all accessible data and classifies new data points using similarity measures.

3.1.8.4. Decision Tree Algorithm

Khanam et al. [46] introduce a decision tree as a vital technique used for solving classification problems, which operates like a flow chart. Each internal node of a decision tree sets a condition or "test" on an attribute, guiding the branching based on these conditions. Ultimately, the leaf node holds a class label determined by evaluating all attributes. The distance from the root to a leaf node defines the classification rule.

3.1.8.5. Logistic Regression

Logistic Regression is a frequently employed classification method specifically designed for assigning labels to observations within a distinct set of classes. In addressing binary classification tasks, Logistic Regression has proven to be a successful choice [47]. The main strength of logistic regression techniques lies in its sigmoid function, which generates a probability value that is subsequently linked to a class within a discrete set for comparing two or more classes [47].

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

1 + e-x

If $S(x) > 0.5$, *assign label = class A, else assign label = class B, where* $0 < S(x) < 1$ *and* (A, B) *are sets of classes* [47].

3.1.9. Model Evaluation

It's essential to evaluate the performance of the fake news prediction models using the test sample data and measure their performance using standard performance metrics such as accuracy, precision, recall, and F-score.

3.1.9.1. Accuracy

The accuracy metric is used when the dependent or output variable is categorical or discrete. It represents the fraction of time a model correctly predicts the total predictions it makes [48].

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

3.1.9.2. Precision

The precision of text classification indicates the level of exact prediction carried out by the model [48]. Technically, given all the positive cases (thus the relevant classes), to how many classes are correctly classified by the model [48].

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

3.1.9.3. Recall

The Recall and Precision metrics are complementary to each other. Recall indicates how well a model can identify positive instances in a problem statement. It is calculated based on all the positive predictions made by the model and the proportion of these predictions that are truly positive [48].

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

3.1.9.4. F1-Score

The F1-score metric combines precision and recall into a single measure; it also captures the trade-off between these metrics, representing both completeness and accuracy [48].

$$\text{F1-score} = \frac{2\times P_i \times R_i}{P_i + R_i} \quad (6)$$

3.2. System Analysis and Design

This subsection introduces various conceptual diagrams to illustrate how the proposed fake news detection model will be designed and developed. The system framework and flow chart diagram are designed using the Microsoft Visio application. Finally, the Bi-LSTM-2-ML algorithm will also be illustrated in this subsection.

Algorithm 2: Bi-LSTM-2-ML transfer learning-based Model
Input: Fake news dataset
Output: Evaluation results R

$Start$
$Load\ the\ labeled\ fake\ news\ dataset$
$Data\ Cleaning$
  3.1 $Tokenize\ the\ text$
  3.2 $Remove\ punctuation\ marks$
  3.3 $Remove\ stop\ words$
  3.4 Apply stemming using *PorterStemmer*
  3.5 Convert all text to lowercase.
Replace label '*fake*' with 0 and '*real*' with 1
Apply word embeddings using *word2vec()*
$Apply\ tensor_{tokenization}(\quad) for\ feature\ extraction$
$Use\ padding_{sequence}(\quad) to\ ensure\ eual\ length\ pad\ sequences$
$Split\ dataset\ with\ train_{size} = 70\ and\ test_{size} = 30$
Build and train a Bi-LSTM model using learning.train()
Use the trained Bi-LSTM model as a feature extractor
Extract Bi-LSTM features.
Train multiple ML classifiers on extracted features (KNN, GNB, SVM and DT)
Evaluate each ML model

DOI: 10.55214/2576-8484.v9i10.10684

```
    Store results in R
    Return R
END
```

Algorithm 2 illustrates the step-by-step approach for developing the proposed Bi-LSTM-2-ML fake news detection model. The algorithm is divided into nine major steps, each with associated subtasks. These steps include data loading, data cleaning, data transformation, feature extraction, data splitting, Bi-LSTM deep learning training, machine learning using transfer learning, and model evaluation.

4. Implementation

The pre-trained Bi-LSTM model implementation, along with the Bi-LSTM transfer learning for machine learning training, will be comprehensively discussed in this sub-section. Marjory Python libraries such as sklearn, TensorFlow, and Keras will be used for training the deep/machine learning models. Firstly, a sequential-based Bi-LSTM model is built, compiled, and summarized before the fake news training process is started.

4.1. Deep learning Implementation (Bi-LSTM)

```
Model: "sequential_20"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding_20 (Embedding)    (None, 500, 500)          40429500

 bidirectional_27 (Bidirect  (None, 500, 256)          644096
 ional)

 bidirectional_28 (Bidirect  (None, 500, 256)          394240
 ional)

 lstm_50 (LSTM)              (None, 64)                82176

 dense_28 (Dense)            (None, 32)                2080

 dense_29 (Dense)            (None, 1)                 33

=================================================================
Total params: 41552125 (158.51 MB)
Trainable params: 1122625 (4.28 MB)
Non-trainable params: 40429500 (154.23 MB)
```

Figure 6.
Bi-LSTM Sequential Model Architecture.

Based on Figure 6, the Bi-LSTM sequential model instance is created, along with other essential layers, using the TensorFlow Keras API. Additional layers are added to the Bi-LSTM model instance using the *model.add* function; the layers have been added, including one (1) Embedding layer, two (2) Bidirectional LSTM layers, one (1) LSTM layer, one (1) fully connected Dense layer, and lastly, one (1) output layer. The embedding layer contains parameters such as the vocabulary size (vocabulary_size), output dimension (output_dim), sequence length (input_length), feature matrix (weights), and finally, the trainable parameter has been set to false (trainable=False). The input sequence (embedding size) is set to 500 features per sequence, and the two layers of Bidirectional LSTM contain 128 neurons with a 20% dropout rate. The LSTM layer contains only 64 neurons with a 10% dropout rate for each training iteration, while the dense layer contains 32 neurons and a ReLU activation function to introduce non-linearity. The final layer contains a single neuron outputting the value of either '0' or '1' with a sigmoid activation function, which is used for binary classification tasks.

Furthermore, the Bi-LSTM model architecture is compiled and summarized before training. The Bi-LSTM architecture is compiled using *Adam* as the optimization algorithm, *binary_crossentropy* as the loss function, and *accuracy* as the performance metric. The training of the Bi-LSTM model is initiated using the *fit* function of the Bi-LSTM model object. The *fit* function uses parameters that define how the training process is conducted; this includes *X_train* (predictor features), *y_train* (response features), *validation_split* of 20% (data sample for training validation), and an *epoch* set to 12. The training process summarizes training loss, accuracy, *val_loss*, and *val_accuracy* for each iteration. Finally, the last accuracy and loss on the 12th iteration represent the highest accuracy and minimum loss achieved by the model.

4.2. Machine Learning Implementation

This subsection details the training process of various machine learning models, including Support Vector Machine (SVM), Decision Tree (DT), Naïve Bayes (NB), Logistic Regression (LR), and K-Nearest Neighbour (KNN). Additionally, transfer learning-based models are retrained using Bi-LSTM pre-trained weights to enhance performance. To train these machine learning models, the fit function is called, passing the 'x_train' and 'y_train' data samples as parameters. Once trained, the models for fake news detection can be evaluated and used for predictions.

5. Result and Discussion

5.1. Deep Learning Performance

Figure 7 shows the training performance measure of the Bi-LSTM fake news detection model using line chart and a confusion matrix. The line chart illustrates the change in performance based on accuracy and loss for every epoch. The plot shows a smooth trend in accuracy increase for both training and validation samples, but initially, the training accuracy was low for the first epoch, and later performed efficiently on the training sample. For every epoch, the model significantly drops in loss, which shows a smooth learning curve.

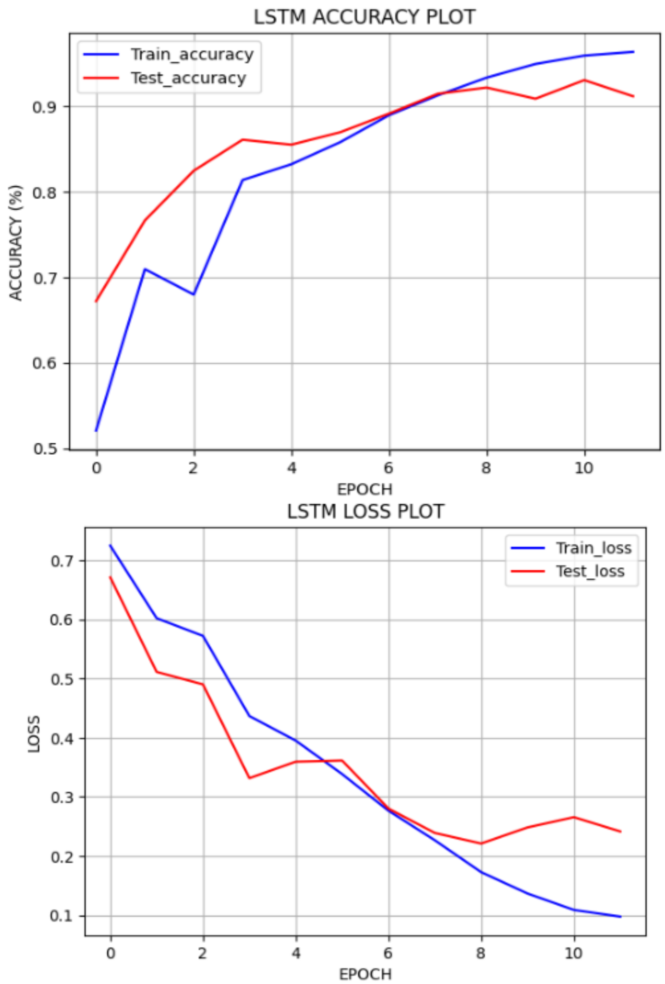

Figure 7.
Bi-LSTM Accuracy & Loss.

Figure 8 shows the confusion matrix illustrating the interception of the actual output to be predicted (Actual Value) and the predicted value (Prediction Value). The confusion matrix indicates that real news was misclassified 44 times as fake news, and fake news was misclassified 46 times as real news.

ISSN: 2576-8484
Vol. 9, No. 10: 1515-1540, 2025
DOI: 10.55214/2576-8484.v9i10.10684

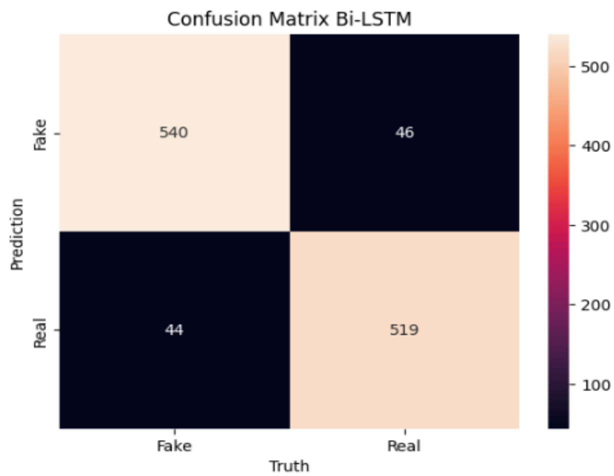

Figure 8.
Bi-LSTM Confusion Matrix Plot.

Figure 9 shows the Bi-LSTM model's performance metrics, which achieved an accuracy of 92.16%, a precision of 92.18%, a recall of 91.85%, and an F1-score of 92.02%.

```
view_collective_result(y_test, norm_prediction, 'Bi-LSTM Model',

Rusult Summary for :Bi-LSTM Model
=============================================
Overall Accuracy :  0.9216710182767625
Overall Precision : 0.9218472468916519
Overall Recall :    0.9185840707964602
Overall F1-score :  0.9202127659574469

ml_result_table = pd.DataFrame(dl_results)
ml_result_table.set_index('model')
```

model	accuracy	precision	recall	f1score
Bi-LSTM Model	0.921671	0.921847	0.918584	0.920213

Figure 9.
Bi-LSTM Performance Summary.

5.2. MaChinese Learning Performance (Bi-LSMT-2-ML)

This subsection comprehensively discusses the result summary of the machine learning training process using the Bi-LSTM trained weights. Table 2 shows the accuracy, precision, recall, and F1-score for each benchmark machine learning algorithm trained without the Bi-LSTM pre-trained weight. Based on Table 2, the Support Vector Machine (SVM) achieved 88.1% accuracy, which is the best accuracy observed. Logistic Regression follows the SVM performance with 87.5% accuracy, followed by the K-Nearest Neighbour with 85.8%, the Decision Tree with 81.8%, and finally, the Naïve Bayes Algorithm with the lowest performance of 79.4% accuracy. The result summary in Table 2 serves as a benchmark, which will be used later for comparing and validating the impact of transfer learning on the machine learning algorithms. Table 2 is graphically represented using the bar chart visual displayed in Figure 10.

Table 2.
Machine Learning Performance (Without Bi-LSTM).

S/N	Model	Accuracy	Precision	Recall	F1-score
0	Logistic Regression	0.875272	0.904412	0.851801	0.877318
1	K Nearest Neighbour	0.858593	0.861454	0.869806	0.86561
2	Naive Bayes	0.794054	0.834862	0.756233	0.793605
3	SVM	0.881073	0.918919	0.847645	0.881844
4	Decision Tree	0.818709	0.834278	0.815789	0.82493

The bar chart in Figure 10 Graphically represents the performance of various machine learning algorithm models without the Bi-LSTM weight. The accuracy, precision, recall, and F1-score are displayed in bars for every model. The bar chart shows a graphical representation of Table 2, and it can be seen that SVM performs best, and the Decision Tree has the lowest performance (below 80%).

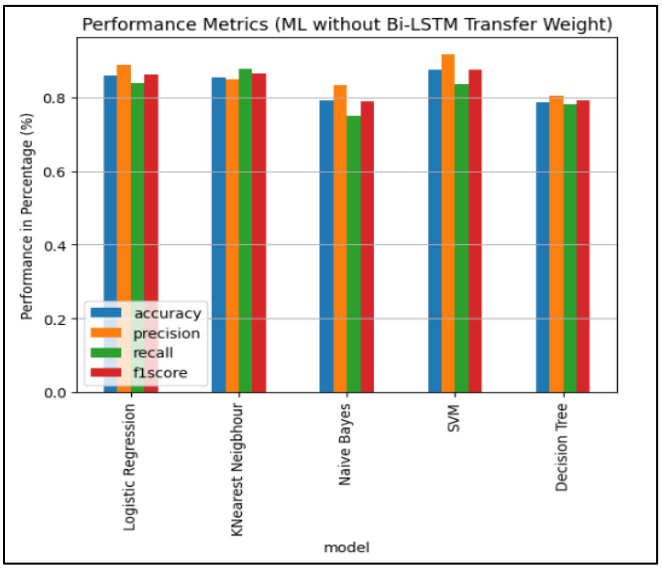

Figure 10.
Machine learning Algorithm performance comparison.

The bar chart in Figure 11 illustrates the performance of various machine learning algorithms in terms of accuracy, precision, and recall. Figure 13 presents the performance of these algorithms in terms of accuracy, precision, recall, and F1-score, utilizing Bi-LSTM transfer learning to enhance model performance. The bar chart provides a graphical representation of Table 2, indicating that Logistic Regression (LR) performs the best, while the Decision Tree algorithm has the lowest accuracy performance at 91.2%.

ISSN: 2576-8484
Vol. 9, No. 10: 1515-1540, 2025
DOI: 10.55214/2576-8484.v9i10.10684

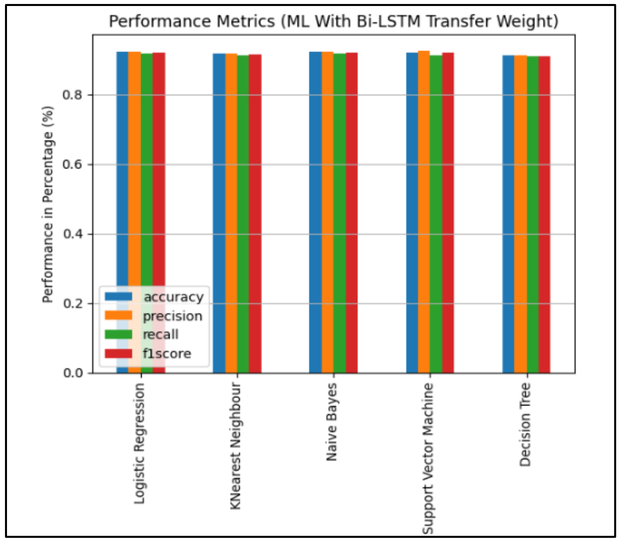

Figure 11.
Machine learning Algorithm with Bi-LSTM Transfer Learning.

Table 3 shows the performance increase in accuracy, precision, recall, and F1-score, using the concept of transfer learning to improve the performance of the trained machine learning algorithm with the Bi-LSTM pre-trained model. Based on Table 3, it is revealed that Logistic Regression, Naïve Bayes, and Support Vector Machine attain accuracies of 92.2%, 92.1%, and 92.0%, respectively, while Nearest Neighbour and Decision Tree obtain accuracies of 91.6% and 91.2%.

Table 3.
Machine Learning Performance (Bi-LSTM-2-ML).

S/N	Model	Accuracy	Precision	Recall	F1-score
1	Logistic Regression	0.921671	0.921847	0.918584	0.920213
2	K-Nearest Neighbor	0.916449	0.916519	0.913274	0.914894
3	Naive Bayes	0.921671	0.921847	0.918584	0.920213
4	Support Vector Machine	0.920801	0.926259	0.911504	0.918822
5	Decision Tree	0.912097	0.911348	0.909735	0.91054

Furthermore, Table 4 shows the performance comparison between the benchmark accuracy of various machine learning fake news detection models and the proposed Bi-LSTM transfer learning-2-ML fake news detection models. It is clearly shown that the machine learning algorithm trained using the Bi-LSTM transfer weight performs significantly better than machine learning fake news detection without the Bi-LSTM enhancement. It is also revealed that Support Vector Machine improves in performance from 87.4% to 92.0%. In comparison, Logistic Regression achieved 92.2% from 86.0%, K-Nearest Neighbour from 85.4% to 91.6%, Naïve Bayes obtained an accuracy increase from 79.0% to 92.1%, and finally, Decision Tree attained 91.2% from 78.5%. This performance difference can be seen in the bar chart in Figure 12.

Table 4.
ML Vs Bi-LSTM-2-ML.

S/N	Model	Proposed Bi-LSTM Transfer learning-2-ML	Machine learning (Benchmark)
1	Logistic Regression	92.16%	86.00%
2	K-Nearest Neighbor	91.64%	85.49%
3	Naive Bayes	92.16%	79.04%
4	SVM	92.08%	87.45%
5	Decision Tree	91.20%	78.53%

ISSN: 2576-8484
Vol. 9, No. 10: 1515-1540, 2025
DOI: 10.55214/2576-8484.v9i10.10684

Figure 12 shows the performance comparison between the Bi-LSTM transfer learning-2-ML and the machine learning (benchmarked) models' accuracy. Regarding the figure, it is revealed that the proposed Bi-LSTM transfer learning-2-ML performs efficiently, with an increase of 8.54% in average accuracy, surpassing the benchmark machine learning algorithms. Table 5 presents a comparison of the proposed algorithm with the state-of-the-art studies.

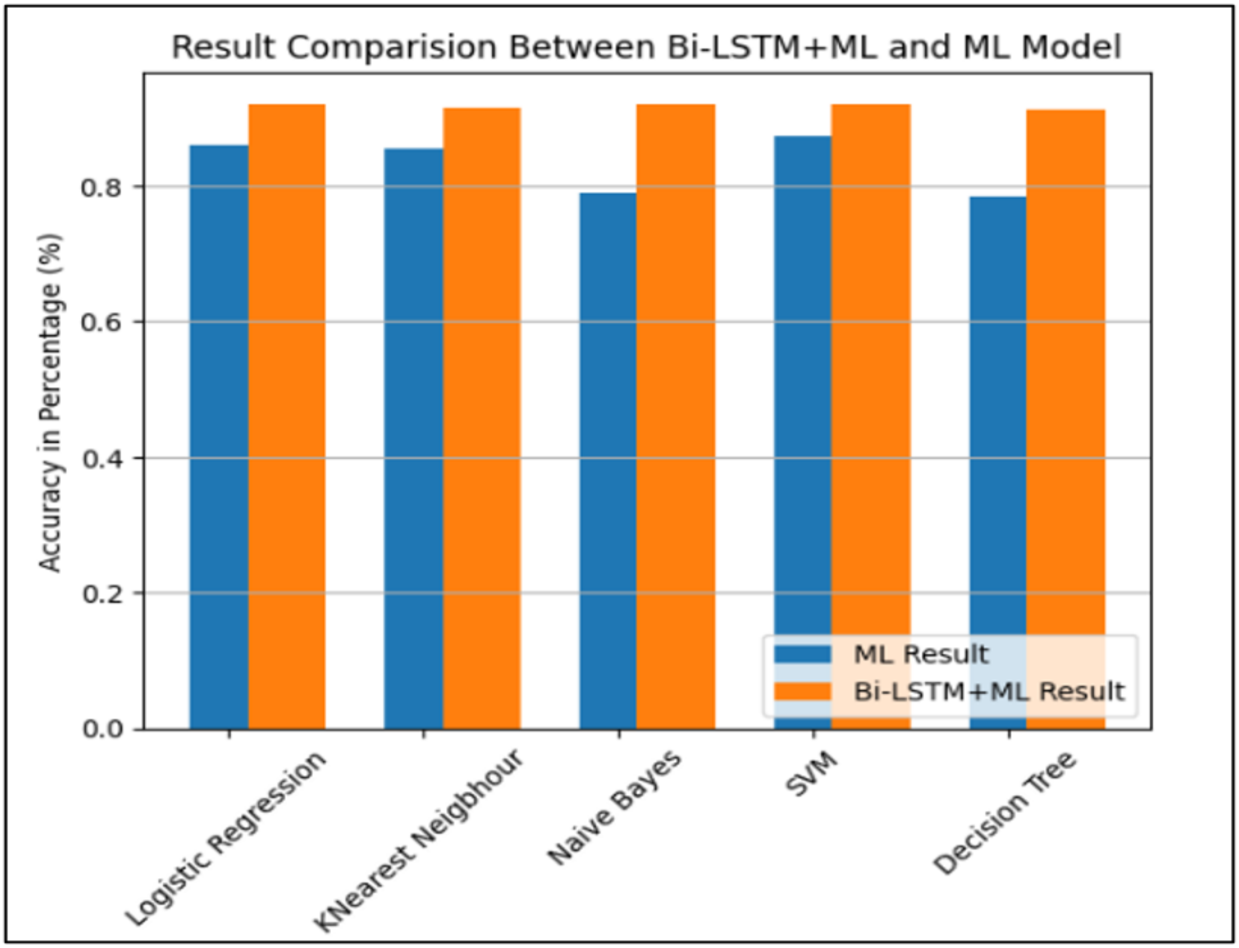

Figure 12.
ML Vs Bi-LSTM-2-ML.

Table 5.
Proposed Bi-LSTM-2-ML Vs Benchmark Models.

S/N	Model	Machine learning (Benchmark)	Pandey et al. [6]	Jaybhaye [13]	Proposed Bi-LSTM Transfer learning-2-ML
1	Logistic Regression	86.00%	90.46%	91.22%	92.16%
2	K-Nearest Neighbor	85.49%	89.98%	74.68%	91.64%
3	Naive Bayes	79.04%	86.89%	79.41%	92.16%
4	SVM	87.45%	73.33%	*****	92.08%
5	Decision Tree	78.53%	89.33%	89.00%	91.20%

Table 5 and Figure 13 show the results and performance comparison between the proposed Bi-LSTM-2-ML fake news detection model, benchmark model, Pandey et al. [6]. Based on the table and figure, it is revealed that the proposed Bi-LSTM-2-ML fake news detection model performs significantly better than the existing models, and the Bi-LSTM pre-trained weights have enhanced the performance of all the machine learning models with better accuracy than the existing models.

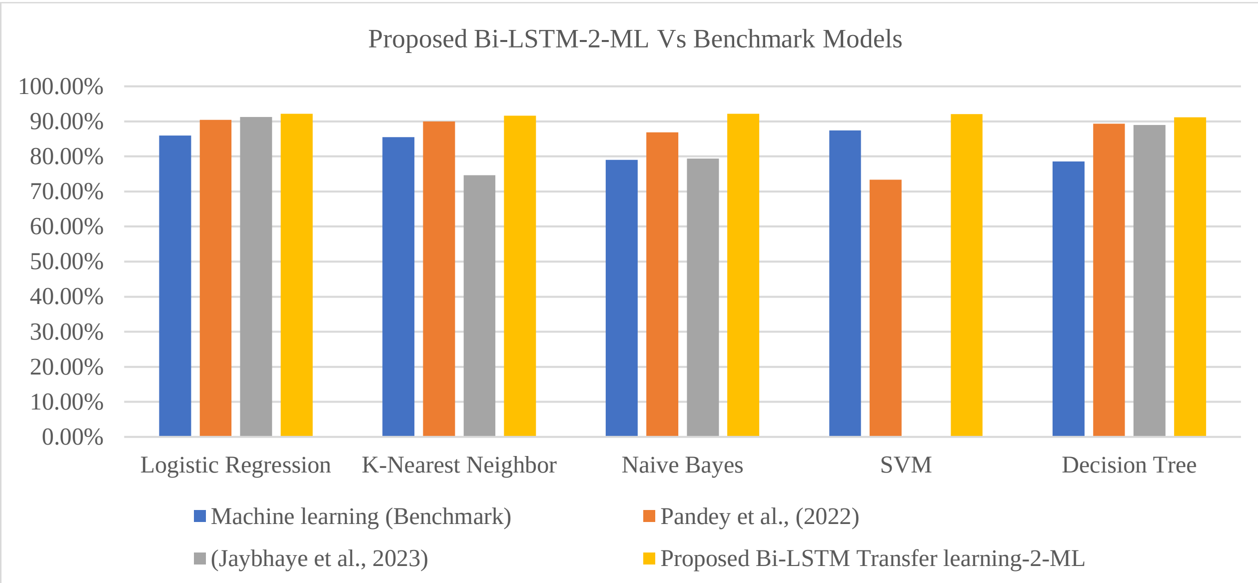

Figure 13.
Proposed Bi-LSTM-2-ML Vs Benchmark Models Performance Comparison.

6. Conclusion and Recommendation

The research study developed a fake news detection model using five machine learning models, including Logistic Regression (LR), K-Nearest Neighbour (KNN), Naïve Bayes (NB), Support Vector Machine (SVM), and Decision Tree (DT). The research utilises the Bi-LSTM deep learning model to enhance the performance of the prior machine learning models mentioned, using the transfer learning approach. Hence, the implemented Bi-LSTM-2-ML fake news detection model has significantly demonstrated great performance with an average increase of 8.54% in accuracy compared to the benchmark approach. Pandey et al. [6] proposed a similar machine learning model for fake news detection, but the model's performance was not satisfactory, while Jaybhaye [13] suffered from model overfitting during training and poor performance on testing samples. This issue is addressed by enhancing the performance of the machine learning model using Bi-LSTM contextualized weights in training the machine learning algorithms, and also addressing the overfitting issue using dropout techniques to penalize the Bi-LSTM training weights. The performance comparison between the proposed Bi-LSTM-2-ML model and the benchmark ML models shows that the Bi-LSTM weights significantly enhance the performance accuracy of the ML model used for building Bi-LSTM-2-ML models. Furthermore, an extensive study can be conducted on other deep learning architectures or transformer-based architectures to improve machine learning using a transfer learning approach. Architectures such as LSTM, GPT, BERT, Albert, GRU, and similar models can be adopted. This will reveal further insights and confirm the impact of deep learning models on machine learning performance through a transfer learning approach.

Data Availability Statement:

The data that support the findings of this study are openly available in Kaggle at https://github.com/GeorgeMcIntire/fake_real_news_dataset.

Declaration of Conflict of Interest:
All the authors wish to declare that no conflict of interest exists in writing and submitting this manuscript. All authors agree to submit this manuscript to a reputable journal outlet.

Declaration Statement:
i. All authors agree to be accountable for the content and conclusions of the article.
ii. No third-party services were involved in the research or manuscript preparation.
iii. All contributing authors of this manuscript are properly listed as authors and have been acknowledged.
iv. The authors declared that no AI software has been used to prepare the manuscript.
v. The authors also declared that all the images used are original, and all the images were obtained from our implementation.

Transparency:
The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Acknowledgements:
All the authors acknowledge that the data used in this study were obtained from an online repository as provided in the data availability statement.

Copyright:

References
[1] C. Song, N. Ning, Y. Zhang, and B. Wu, "A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks," *Information Processing & Management*, vol. 58, no. 1, p. 102437, 2021. https://doi.org/10.1016/j.ipm.2020.102437

[2] G. E. R. Agudelo, O. J. S. Parra, and J. B. Velandia, "Raising a model for fake news detection using machine learning in Python," in *Conference on e-Business, e-Services and e-Society (pp. 596-604). Cham: Springer International Publishing*, 2018.

[3] A. Jain, A. Shakya, H. Khatter, and A. K. Gupta, "A smart system for fake news detection using machine learning," in *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT) (Vol. 1, pp. 1-4). IEEE*, 2019.

[4] A. M. Braşoveanu and R. Andonie, "Semantic fake news detection: A machine learning perspective," in *International Work-Conference on Artificial Neural Networks (pp. 656-667). Cham: Springer International Publishing*, 2019.

[5] E. Hossain, M. Nadim Kaysar, A. Z. M. Jalal Uddin Joy, M. Mizanur Rahman, and W. Rahman, "A study towards Bangla fake news detection using machine learning and deep learning," in *Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021 (pp. 79-95). Singapore: Springer Singapore*, 2021.

[6] S. Pandey, S. Prabhakaran, N. S. Reddy, and D. Acharya, "Fake news detection from online media using machine learning classifiers," in *Journal of Physics: Conference Series (Vol. 2161, No. 1, p. 012027). IOP Publishing*, 2022.

[7] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *Journal of Industrial Information Integration*, vol. 23, p. 100224, 2021.

[8] P. Goyal, S. Pandey, and K. Jain, *Deep learning for natural language processing*. Berkeley, CA: Apress, 2018.

[9] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: State of the art, current trends and challenges," *Multimedia Tools and Applications*, vol. 82, pp. 3713-3744, 2023. https://doi.org/10.1007/s11042-022-13428-4

[10] R. Patel and S. Patel, "Deep learning for natural language processing," presented at the Information and Communication Technology for Competitive Strategies (ICTCS 2020) Intelligent Strategies for ICT (pp. 523-533). Singapore: Springer Nature Singapore, 2021.

[11] S. Kumar, R. Asthana, S. Upadhyay, N. Upreti, and M. Akbar, "Fake news detection using deep learning models: A novel approach," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 2, p. e3767, 2020. https://doi.org/10.1002/ett.3767

[12] K. R. Asish, A. Gupta, A. Kumar, A. Mason, M. K. Enduri, and S. Anamalamudi, "A tool for fake news detection using machine learning techniques," in *2022 2nd International Conference on Intelligent Technologies (CONIT) (pp. 1-6). IEEE*, 2022.

[13] M. Jaybhaye, "Fake news detection using an LSTM-based deep learning approach," presented at the ITM Web of Conferences, 2023.

[14] R. K. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11765-11788, 2021.

[15] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, "Combating fake news: A survey on identification and mitigation techniques," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 3, pp. 1-42, 2019. https://doi.org/10.1145/3305260

[16] H. Thakar and B. Bhatt, "Fake news detection: recent trends and challenges," *Social Network Analysis and Mining*, vol. 14, no. 1, p. 176, 2024.

[17] K. I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "Unmasking misinformation: Leveraging CNN, BERT, and GPT models for robust fake news classification," 2024.

[18] E. Hashmi, S. Y. Yayilgan, M. M. Yamin, S. Ali, and M. Abomhara, "Advancing fake news detection: Hybrid deep learning with fasttext and explainable ai," *IEEE Access*, vol. 12, pp. 44462-44480, 2024.

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

[20] M. Al-Alshaqi, D. B. Rawat, and C. Liu, "Ensemble techniques for robust fake news detection: Integrating transformers, natural language processing, and machine learning," *Sensors*, vol. 24, no. 18, p. 6062, 2024. https://doi.org/10.3390/s24186062

[21] M. A. Al-Tarawneh, O. Al-irr, K. S. Al-Maaitah, H. Kanj, and W. H. F. Aly, "Enhancing fake news detection with word embedding: A machine learning and deep learning approach," *Computers*, vol. 13, no. 9, p. 239, 2024. https://doi.org/10.3390/computers13090239

[22] E.-S. Apostol and C.-O. Truică, "Efficient machine learning ensemble methods for detecting gravitational wave glitches in LIGO time series," in *2023 IEEE 19th International Conference on Intelligent Computer Communication and Processing (ICCP) (pp. 79-86). IEEE*, 2023.

[23] N. Nasir, F. Rehman, and S. F. Omair, "Risk factors for bacterial infections in patients with moderate to severe COVID-19: A case-control study," *Journal of Medical Virology*, vol. 93, no. 7, pp. 4564-4569, 2021.

[24] S. K. Thota, M. Gujjalapudi, C. Anam, R. Chiluka, P. Mamidi, and S. M. Magar, "Comparative Evaluation of the retention of maxillary complete denture bases made by heat cure acrylic resin processed on high expansion stone and type III dental stone: An In-Vivo study," *Int J Cur Res Rev*, vol. 13, no. 15, p. 14, 2021.

[25] A. Choudhary and A. Arora, "Linguistic feature based learning model for fake news detection and classification," *Expert Systems with Applications*, vol. 169, p. 114171, 2021.

[26] R. Vijjali, P. Potluri, S. Kumar, and S. Teki, "Two stage transformer model for COVID-19 fake news detection and fact checking," *arXiv preprint arXiv:2011.13253*, 2020.

[27] S. Singhal, S. Choudhary, and P. C. Biswal, "Return and volatility linkages among International crude oil price, gold price, exchange rate and stock markets: Evidence from Mexico," *Resources Policy*, vol. 60, pp. 255-261, 2019.

[28] S. Kumari and M. P. Singh, "A deep learning multimodal framework for fake news detection," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16527-16533, 2024.

[29] H. Gul, F. Al-Obeidat, M. Wasim, A. Amin, and F. Moreira, "Advancing fake news detection with graph neural network and deep learning," *Journal of Physics: Complexity*, vol. 6, no. 2, p. 025001, 2025.

[30] W. K. Sari, I. S. B. Azhar, Z. Yamani, and Y. Florensia, "Fake news detection using optimized convolutional neural network and bidirectional long short-term memory," *Computer Engineering & Applications Journal*, vol. 13, no. 3, pp. 25-33, 2024.

[31] B. Brinda, C. Rajan, and K. Geetha, "Detecting evolving fake news in social media by leveraging heterogeneous deep learning model," in *2024 Second International Conference on Advances in Information Technology (ICAIT)*, 2024, vol. 1, pp. 1-5.

[32] K. Deepthi and A. K. Shastry, "A deep learning methodology-based unsupervised misinformation detection (umd) technique for identifying fabricated narratives," in *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, 2024: IEEE, pp. 1-6.

[33] H. Zhu, "Application of deep learning models in the identification and screening of fake news," *Applied and Computational Engineering*, vol. 87, pp. 162-166, 2024.

[34] R. Ladouceur, A. Hassan, M. Mejri, and F. Jaafar, "Can deep learning detect fake news better when adding context features?," presented at the 2024 IEEE International Conference on Cyber Security and Resilience (CSR), 2024.

[35] A. Porya, V. Singh, and B. B. Sinha, "Towards robust fake news detection: An emperical investigation of preprocessing and modeling techniques," in *2024 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 2024, vol. 1: IEEE, pp. 80-85.
[36] J. B. Awotunde, R. G. Jimoh, A. L. Imoize, A. T. Abdulrazaq, C.-T. Li, and C.-C. Lee, "An enhanced deep learning-based deepfake video detection and classification system," *Electronics*, vol. 12, no. 1, p. 87, 2022. https://doi.org/10.3390/electronics12010087
[37] J. Alghamdi, Y. Lin, and S. Luo, "Towards COVID-19 fake news detection using transformer-based models," *Knowledge-Based Systems*, vol. 274, p. 110642, 2023. https://doi.org/10.1016/j.knosys.2023.110642
[38] M. Ghayoomi and M. Mousavian, "Deep transfer learning for COVID-19 fake news detection in Persian," *Expert Systems*, vol. 39, no. 8, p. e13008, 2022. https://doi.org/10.1111/exsy.13008
[39] M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar, and M. S. Rahman, "A comprehensive review on fake news detection with deep learning," *IEEE Access*, vol. 9, pp. 156151-156170, 2021.
[40] J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," *Machine Learning with Applications*, vol. 4, p. 100032, 2021. https://doi.org/10.1016/j.mlwa.2021.100032
[41] J. Alghamdi, Y. Lin, and S. Luo, "A comparative study of machine learning and deep learning techniques for fake news detection," *Information*, vol. 13, no. 12, p. 576, 2022. https://doi.org/10.3390/info13120576
[42] M. Q. Alnabhan and P. Branco, "Fake news detection using deep learning: A systematic literature review," *IEEe Access*, vol. 12, pp. 114435 - 114459, 2024. https://doi.org/10.1109/ACCESS.2024.3435497
[43] M. H. Al-Tai, B. M. Nema, and A. Al-Sherbaz, "Deep learning for fake news detection: Literature review," *Al-Mustansiriyah Journal of Science*, vol. 34, no. 2, pp. 70-81, 2023. https://doi.org/10.23851/mjs.v34i2.1292
[44] M. Villagracia Octaviano, "Fake news detection using machine learning," in *2021 5th International Conference on E-Society, E-Education and E-Technology*, 2021, pp. 177-180.
[45] E. M. Sabido *et al.*, "Insights into the variation in bioactivities of closely related Streptomyces strains from marine sediments of the Visayan Sea against ESKAPE and ovarian cancer," *Marine Drugs*, vol. 19, no. 8, p. 441, 2021.
[46] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, "Fake news detection using machine learning approaches," *IOP Conference Series: Materials Science and Engineering*, vol. 1099, no. 1, p. 012040, 2021.
[47] S. Sangamnerkar, R. Srinivasan, M. Christhuraj, and R. Sukumaran, "An ensemble technique to detect fabricated news article using machine learning and natural language processing techniques," in *2020 International Conference for Emerging Technology (INCET), pp. 1-7. IEEE*, 2020.
[48] A. I. Kadhim, "An evaluation of preprocessing techniques for text classification," *International Journal of Computer Science and Information Security*, vol. 16, no. 6, pp. 22-32, 2018.