Journal of Nonlinear and Convex Analysis
Volume 26, Number 4, 2025, 923–940

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

CBIATFAKE: CNN AND BIDIRECTIONAL LSTM WITH ATTENTION TO LINGUISTIC CONTEXT FOR FAKE NEWS CLASSIFICATION

MOHD MERAJ, AHMAD KAMAL*, AND JABBAROV JAMOLIDDIN SINDOROVICH

Abstract. Online Social Media (OSM) is a global platform for rapidly and inexpensively disseminating information. However, it has also made incorrect and misleading information more visible, leading to hostility, aggression, and violence in real life. Existing content modification procedures on OSM platforms are often ineffective in controlling online fake news. This study develops a deep learning-based model named as CBiAtFake for detection of fake news in OSM. The CBiAtFake model comprises five layers: input layer, embedding layer, CBA network, dense layer, and output layer. The input layer represents preprocessed news text that is followed by a word embedding employed for converting the input text into a numeric representation. The CBA network, consisting of CNN, bidirectional LSTM, and attention network effectively learns contextual information about the fake news text in two directions, making it one of the most essential components of the proposed model. The CBiAtFake is further enhanced by a large number of manually crafted linguistic auxiliary features that are concatenated to the CBA vector. The dense layer sets out the model for classification, and finally, an output layer classifies the news as fake or non-fake. To demonstrate and assess the CBiAtFake, experiments are performed on a benchmark dataset. CBiAtFake has demonstrated outstanding efficiency compared to the existing state-of-the-art, achieving a significant accuracy rate of 98.96%. This study will assist researchers in enhancing their knowledge of the application of deep neural networks, notably bidirectional LSTM, and the significance of linguistic factors in the identification of fake news.

1. Introduction

The emergence of Fake News (FN) on the OSM represents a substantial threat to the news media, democratic institutions, and the fundamental right to freely express oneself in contemporary society. Recent surveys[1] have revealed that the rise of OSM platforms has expedited the propagation of intentionally generated misinformation. This widespread distribution of false information, widely recognized as fake news, poses a substantial risk of societal harm, including the potential to incite violence. Although fake news is not a novel phenomenon, its detection remains a formidable challenge, largely due to a tendency among individuals to accept and disseminate misleading information. In recent years, FN detection has garnered more attention, particularly following the US presidential election of 2016[1]. The identification of FN poses a challenge for individuals, as it commonly necessitates a

2020 *Mathematics Subject Classification.* 68T50.
Key words and phrases. Fake news, supervised classification, machine learning, deep learning, bi-LSTM, attention, linguistic features engineering.
*Corresponding author.
[1]https://www.weforum.org/agenda/2019/03/fake-news-what-it-is-and-how-to-spot-it/

high level of expertise in the specific domain addressed by the fabricated information. However, advances in artificial intelligence (AI) offer promising solutions for detecting fake news articles on the web and, more specifically, on OSM platforms. These AI techniques are critical in curbing the dissemination of FN by providing tools that can analyze and discern the veracity of information across vast digital landscapes. According to research by Shu et al.[25], it was found that the writing style of fake news significantly diverges from that of authentic news articles. Fake news is often crafted with the deliberate intent to mislead readers and create societal discord. To tackle this issue, numerous scholars have concentrated on analyzing the writing styles of news content. They have proposed various linguistic features as effective discriminators for identifying fake news[6, 9]. These features help differentiate between genuine and fabricated news by examining the stylistic elements unique to false information, thereby aiding in the detection and mitigation of fake news spread. Linguistic and semantic analyses are instrumental in discerning the patterns and structures inherent to news articles, enabling the extraction of pivotal information from their text. As noted by Shu et al.[26], FN-related linguistic Features in textual data can be analyzed at multiple levels, such as the word, sentence, and overall content. Various methodologies are utilized to analyze text at the word level, including Bag-of-Words, n-grams, Term Frequency (TF), TF-Inverse Document Frequency (TF-IDF), and word embedding techniques[1, 19]. Sentence-level features, such as Parts-of-Speech (POS) tags and sentence length, provide significant insights based on the structure of sentences[4]. Moreover, efforts to delineate content-level attributes from raw documents are essential in distinguishing FN, with sentiment analysis identified by Dickerson et al. [7] as a particularly effective tool for FN detection. These approaches enable a comprehensive examination of textual content. Further, the use of Deep Learning (DL) methodologies has significantly improved the ability to classify and analyze FN articles[13, 14, 31, 33]. Approaches like Recurrent Neural Networks (RNNs) and their extensions, such as Long Short-Term Memory (LSTM) networks, are highly effective for analyzing text. These models are commonly employed to model and replicate the dynamics between long sequences of input and output data, providing enhanced proficiency in managing the intricacies of textual information[10]. In this paper, a novel deep neural network model called CBiAtFake is presented that is specifically intended to detect FN on OSM. The model is built from several complex neural network layers where every layer represents a different characteristic of fake news. First, the text is processed by an input layer, and then embedded into a numerical representation by an embedding layer, and then a CBA network learns the effective contextual information in two directions. Finally, a dense layer incorporates auxiliary features and sets up the model for FN classification. The CBA layer consisting of a Convolutional Neural Network (CNN) layer, stacked Bidirectional LSTM (BiLSTM) and an attention layer followed by the additional layer to incorporate handcrafted FN-related linguistic features is the novel aspect of the proposed model, as it learns contextual information regarding various forms of the input text both forwards and backward. The significant findings that this article presents are as follows: 1) This paper investigates the challenges of FN (Fake News) detection and establishes that deep learning techniques generally outperform conventional machine learning techniques in terms of efficacy. 2) A new deep neural network, named CBiAtFake, is introduced for FN classification. This model incorporates sophisticated layers such as pre-trained embeddings and the CBA network, and it also leverages fake news-specific linguistic features to enhance its predictive capabilities. 3) Through rigorous experimentation, it has been demonstrated that the model exceeds the performance of existing state-of-the-art approaches by attaining an accuracy rate of 98.96% on the benchmark dataset. The structure of the rest of the paper is organized as follows: Section 2 reviews the relevant literature on machine learning and deep learning techniques for FN detection. Section 3 outlines the methodology and functional aspects of the CBiAtFake. The experimental setup and results analysis are detailed in Section 4, followed by the conclusions in Section 5.

2. Related works

The detection of intentionally generated fake, misleading, and deceptive information has emerged as a problem of great difficulty in the field of classical text classification[15]. Early works relied on conventional machine learning techniques, but more recent research works have found that deep learning is more effective. Furthermore, previous literature emphasizes the critical relevance of extracting significant text features from the content of articles for classification. Various researchers[2, 3] have revealed that diverse text characteristics in news articles are crucial in effectively detecting fake news. Castillo et al.[4] studied various feature-based techniques for assessing the trustworthiness of tweets. Study performed by Ott el al.[18] utilized word count and Part-of-Speech (POS) tags as discriminating elements in a dataset of opinion spam, obtaining an accuracy rate of 90%. Additionally, lexical and syntactic attributes were scrutinized by researchers[5] in their efforts to detect fake news and to devise methodologies for automated clickbait detection, ultimately proposing a hybrid approach for the identification of misleading content. In related work, Singh and his colleagues[27] examined the utility of the Linguistic Inquiry Word Count (LIWC) lexicon in identifying fake news, adopting a Support Vector Machine (SVM) classifier and reporting an accuracy of 87%. Moreover, Pe ́rez-Rosas et al. [21] have demonstrated that semantic information contained within the LIWC lexicon effectively enhances the performance of classifiers across different textual domains.

Ahmed et al. [1] developed an n-gram and ML-based method for FN detection, evaluating two distinct feature extraction techniques. Their findings indicate that the use of TF-IDF for text representation coupled with a Linear Support Vector Machine (LSVM) yields effective results. Gravinis et al. [9] conducted comprehensive research on various feature sets recommended for FN detection, employing a range of ML based techniques and reported that the integration of word embeddings and advanced linguistic feature, in conjunction with SVM and ensemble learning strategies, yields a high level of classification accuracy. Sahoo and Gupta[23] investigated the utilization of various characteristics related to Facebook accounts to analyze user behavior, employing both ML and deep learning methodologies. They utilized an LSTM model that integrates features from user profiles and news articles. This model was tested on a dataset comprising over 15,000 news articles sourced from diverse Facebook accounts, yielding enhanced classification accuracy.

Chaudhry and Arora[6] conducted thorough research into content-based features, developing a robust set of features that encompass syntactic, sentiment, grammatical, and readability aspects. The incorporation of these linguistic features into a sequential neural network classification model led to a substantial improvement in accuracy, illustrating the efficacy of their feature selection. Additionally, Jiang et al. [11] explored a variety of machine-learning techniques for representing news text, focusing on TF, TF-IDF, and word embedding methods. They proposed a stacking approach to classify fake news, demonstrating that layered modeling strategies can effectively enhance the accuracy. This approach leverages the strengths of different representation techniques to optimize classification performance.

To construct a more accurate method for FN detection, a variety of researchers have pursued sophisticated deep-learning approaches. Ma et al. [17] designed a model using Recurrent Neural Networks (RNNs) aimed at deciphering concealed representations that encapsulate the dynamic context of pertinent posts over time. They argue that this RNN model is more effective than traditional rumor detection techniques that primarily depend on manually crafted features. In a separate initiative, Yang et al. [33] introduced a novel approach using deep Convolutional Neural Networks (CNNs) that merges both textual and visual features, broadening the scope of FN detection. Additionally, Umer et al. [30] worked on stance detection and suggested a hybrid CNN-LSTM architecture that incorporates dimension reduction technique, Principal Component Analysis (PCA). Further contributions include the application of Capsule Neural Networks (CapsNet) for multitask learning, as discussed in studies [8, 32]. Employing a combined strategy of CapsNet and bidirectional LSTM, Sridhar, and Sanagavarapu [28] managed to achieve an impressive 97.96% accuracy on Kaggle's fake news dataset. Exploring different aspects of FN detection, Ruchansky et al. [22] investigated the impact of user roles and their interactions, alongside article text features. On another front, Kaliyar et al. [13] utilized the pre-trained GloVe embedding method in their CNN-based networks to effectively extract contextual features, achieving a remarkable accuracy of 98.36% in their tests on Kaggle's dataset. These pioneering methods are detailed in Table 1, showcasing the current cutting-edge techniques in FN detection and underscoring the efficacy of various deep-learning methods.

The findings from previous research highlight that deep learning techniques are reasonably effective for detecting fake news on OSM. Several studies have underlined the importance of contextual analysis in distinguishing textual ambiguities to accurately identify fake news. Prominent among these are efforts leveraging CNN and LSTM networks, which have started to effectively utilize contextual data. Despite these advancements, there are notable challenges that persist in current methodologies. One major limitation is the inability of existing models to adequately address the multi-directional nature of context, which is crucial for understanding the full scope of textual interactions. Moreover, there is a lack of sufficient attention on integrating key linguistic features and embedding vectors which could enhance detection capabilities. To address the aforementioned challenges, this article introduces a deep learning-based model, CBiAtFake, for classifying fake news content. The proposed model integrates a multi-layered architecture that includes CNN, BiLSTM, and attention network. This design is aimed at enhancing the detection capabilities for

Author	Features/Embedding	Model	Dataset
Castillo et al. [4]	Linguistic features	ML	Twitter data
Singh et al. [27]	Linguistic features	ML	Kaggle fake news data
Ahmed et al. [1]	n-gram TF-IDF	ML	ISOT
Gravinis et al. [9]	Linguistic features	ML, Ensemble	UN Biased
Sahoo & Gupta [23]	Content, social context, and profile features	ML, LSTM	Facebook data
Choudhary & Arora [6]	Linguistic Features	ML, LSTM	Horne and Adali dataset
Jiang et al. [11]	TF-IDF, word-embedding	ML, Stacking	ISOT
Ma et al. [17]	Word-embedding	RNN, LSTM	Twitter, Weibo
Yang et al. [33]	Linguistic Features	CNN	Kaggle fake news data
Umer et al. [30]	Word embedding	PCA, CNN-LSTM	Facebook dataset
Goldani et al. [8]	n-gram, word embedding	Capsule Network	ISOT, LIAR
Kaliyar et al. [13]	GloVe	Deep learning, CNN	Kaggle fake news data
Sridhar & Sanagavarapu [28]	Word embedding	Deep learning, Capsule Network	Kaggle fake news data

TABLE 1. Summary of different approaches for fake news detection

fake news by harnessing the strengths of each component. Additionally, the model employs both pre-trained embeddings and a set of hand crafted linguistic features to construct a more comprehensive and robust FN detection system.

3. Proposed fake news classification approach

3.1. Business Understanding.

(i) Problem Statement
A collection of text-based news articles posted by users on OSM platforms has been extracted. The objective of the fake news classification problem addressed in this research work is to classify each user post as either Fake News (FN) or Non-Fake News (NFN).

(ii) Problem Solution Overview
Figure 1 illustrates the various components that make up the proposed Fake news detection method. Before introducing the proposed model, we begin by collecting and preprocessing the necessary data. Using the NLTK toolkit, the preprocessing phase cleanses and normalizes the raw news text. Our multilayered CBiAtFake model receives the fine-grained dataset as input. Additionally, it employs a variety of hand-crafted linguistic features.

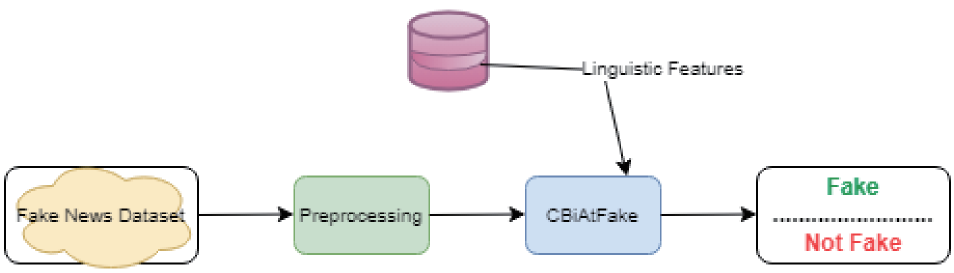

FIGURE 1. Components of the proposed fake news binary classification model

3.2. **Dataset Understanding.** The study utilized Kaggle's Fake and Real News dataset[2], which provides a substantial and balanced repository of news articles posted focusing 2016 presidential of U.S. This labeled dataset contains a total of 20,800 news stories with their titles, of which 10,540 are verified as legitimate and 10,260 as fake. The articles, all written in English, span a broad spectrum of subjects from political and economic debates to sports and entertainment. This dataset is widely used by the academic community to develop, train, and assess the efficacy of various models tailored for fake news classification. Details regarding the dataset are presented in Table 2.

Table 2. Kaggle's Fake and Real News Dataset Description

Dataset Properties	**Value**
Total No of News Instances	20800
NFN	10540
FN	10260
Type	Political, Business, etc.
Attributes	ID, title, author, text, label

3.3. **Data Preparation.** The process of preparing the dataset is considered essential for ensuring the accurateness and excellent performance of the proposed FN detection model since it is necessary to correctly arrange and purify the data used for training and testing. The data preparation phase continues with feature extraction, feature scaling, and the elimination of unnecessary features using correlation analysis techniques.

3.3.1. *Data Cleaning.* Data preprocessing involves cleaning and transforming raw data to make it appropriate for analysis. The first stage is tokenizing the body of each article to create a corpus. Using the corpus, we can extract word characteristics. After tokenizing each article, the NLTK[3] program was used to perform the following operations:

- Separated each article into fragments (removed all whitespace)
- Lowercased all words
- Eliminated punctuation marks and stop words
- Stemmed the text for standardization
- Lemmatized words for converting to their root form

3.3.2. *Linguistic Features Extraction.* Previous research [6,9] has demonstrated that the incorporation of handcrafted linguistic features related to the contextual framework of FN markedly enhances the efficacy of classifiers. This study delineates a range of handcrafted linguistic attributes aimed at constructing an effective fake news classification model. The inclusion of features encapsulating both contextual

[2]https://www.kaggle.com/competitions/fake-news/data

[3]https://www.nltk.org/

and syntactic information notably augments the capabilities of the proposed CBiAtFake model, thereby improving its classification performance

Readability Features: Readability pertains to the ease with which a reader can comprehend text, which is influenced by the complexity of the text's vocabulary and syntactic structure. Various readability indices, such as syllable counts, the automated reliability index, gunfire fog, and Flesch reading ease scores, are employed to extract textual characteristics based on predefined readability criteria.

Grammatical Features: This study collects Part-of-Speech (POS) tags from news articles, which serve as valuable indicators of the grammatical attributes of the text. Fake news creators frequently utilize adjectives and adverbs in conjunction with nouns to enhance the impact of their narratives. For experimental purposes, grammatical features such as the number of nouns, pronouns, adjectives, and conjunctions, along with the count of positive and negative words and average word length, are extracted.

Sentiment Features: The identification of sentiment-driven features is crucial for the development of effective fake news classification strategies. This research focuses on two primary sentiment-related features: the subjectivity and polarity scores of news articles. These parameters are instrumental in differentiating between factual content and opinionated expressions within news articles.

The Linguistic Inquiry and Word Count (LIWC): Observations have indicated that the selection of words by news creators can reveal insights into their personality, thought processes, writing style, and interpersonal relationships. Building on the framework proposed by Tausczik and Pennebaker [29], this paper conducts feature identification through psychological analysis of news articles. LIWC analysis is utilized to categorize words into several psychological categories, such as social, cognitive, personal, and emotional, facilitating a deeper understanding of the linguistic patterns in fake news articles

3.3.3. *Feature Scaling.* It will be necessary to conduct feature scaling on the data to deal with magnitudes that vary widely. Without scaling features, the classification algorithm could be biased towards larger-magnitude features. The current study utilizes the standard scalar approach for regulating the autonomy of data within a predetermined range

3.4. **The Proposed Model: CBiAtFake.** CBiAtFake is a DL-based model designed to address the constraints of existing methods for fake news classification. Figure 2 depicts the comprehensive architecture of the CBiAtFake which comprises five layers of sophisticated deep neural networks. The acquired data set is utilized to train the model once it has been preprocessed. Before being transmitted to the embedding layer, the input texts are tokenized and converted to their respective input vectors. The two-dimensional output matrix generated by the embedding layer using Glove has been taken as input by the CBA network, which consists of a CNN layer, and stacked BiLSTM networks followed by an attention network. CNN layer extracts global features of the news text while the stacked BiLSTM produces contextual representations of FN from distinct directions. The output matrix is then sent to the attention network so that the fake news-related terms receive the most attention. The CBA network finally procedure, CBA vectors. Additionally, a number of shallow and deep auxiliary features are derived from four distinct categories (discussed in section 3.3.2). All of this additional data is compiled into a feature vector. To enrich the model, the language feature vector and the CBA vectors are concatenated and forwarded to the dense layer. The output layer of the neural network employing a sigmoid function classifies the input news as FN or NFN. The subsequent sub-sections describe the function-specifics of CBiAtFake.

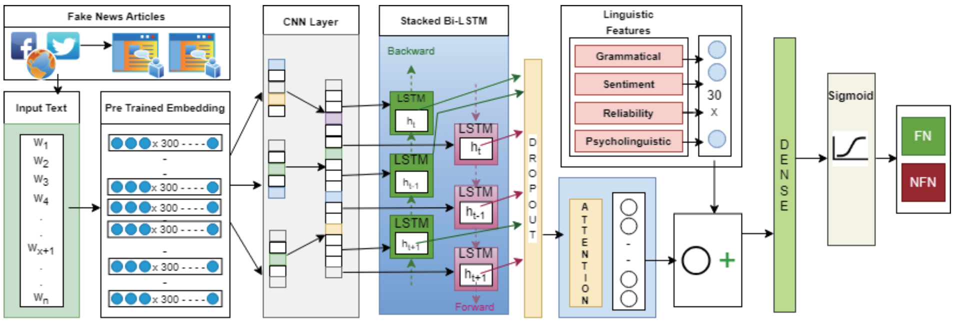

FIGURE 2. Layered Architecture of the CBiAtFake Model

3.4.1. *Input Layer.* The input layer performs tokenization on the preprocessed text of a news article x of length $w.x$ words. If the tokens were not already indexed from the prior news article, they are added to a dictionary for further use. Each token is then replaced with its index in the dictionary, resulting in a numeric vector u representing the tokens. Depending on the length of the news article, the length of u in news articles may also vary. Moreover, u is transformed to a padded vector v of fixed length, $|v| = p \geq |u|$, such that the total length of the input vectors representing all the news articles remains constant at p. In v, u occupies the first $|u|$ positions, and $(p-|u|)$ positions are padded with zeros. The resulting embedding layer receives a vector of fixed length, $v \in R^p$.

3.4.2. *Embedding Layer.* Text embeddings represent words, phrases, or entire documents as dense vectors of real numbers. In text classification tasks, pre-trained embedding techniques are applied to enhance the model's performance on the big dataset [20]. The news text in this study is embedded using *Global Vectors for Word Representation* (GloVe), which performs the transformation by learning from the context in which words appear. The GloVe model is based on the fundamental idea of using statistics to study the relationships between words. It aggregates global word co-occurrence matrices from a given corpus using an unsupervised learning technique to produce word embeddings. Each value in this matrix represents a pair of words that appear together frequently [34]. In order to implement the suggested model, each token's input vector is transformed into a dimension-d distributional vector representation. After that, this input vector is converted into an embedding matrix $I \in \mathbb{R}^{p\times d}$, where a vector representing a token appears in each row.

3.4.3. *CBA Network.* A multilayer network known as the CBA, comprising CNN, BiLSTM, and Attention layers, processes the matrix I and generates contextual information about the fake news in the manner shown below.

CNN Layer: The Convolutional Neural Network (CNN) involves sliding a smaller array known as a "kernel" or "filter" over the input data and conducting many convolution operations on the input data to generate the transformed output known as a feature map. The Convolutional layer primarily performs operations based on a matrix resulting in an output after being subjected to an activation function [35]. When these network filters detect a specific type of feature at a specific spatial point in the input, they exhibit learning behavior. Our proposed model comprises three parallel convolutional layers, each employing a distinct size of kernel. The generation of the global feature vector is performed in an appropriate manner. Rectified Linear Unit (ReLU) is used as the activation function, and the pool size is configured to three. A total of thirty-two filters are used to extract a rich set of features.

The convolution operation applied to the input matrix I by a convolutional layer can be represented using the following equation:

$$F_i = \mathrm{ReLU}(W_i * I + b_i) \tag{3.1}$$

Where:

- F_i represents the feature map generated by the i-th filter.
- W_i represents the weight matrix of the i-th filter of the convolutional layer.
- b_i is the bias term associated with the i-th filter.
- $\mathrm{ReLU}(x) = \max(0, x)$ denotes the Rectified Linear Unit function.
- $*$ denotes the convolution operation.
- $i = 1, 2, \ldots, 32$ since a total of 32 filters are utilized, each potentially with a different kernel size as mentioned for the three parallel layers.

Regarding the three concurrent convolutional layers featuring varying kernel dimensions, output feature maps from each layer are denoted as $F_1^{(k)}$, $F_2^{(k)}$, and $F_3^{(k)}$, where k indicates the kernel size. The max pooling operation with a maximum pool size of 3 can be represented as:

$$P_i^{(k)} = \mathrm{MaxPool}(F_i^{(k)}, 3) \tag{3.2}$$

Stacked BiLSTM: The LSTM is a subtype of RNN. This model is capable of preserving information for an extended time frame, enabling it to effectively manage dependencies over the long term and alleviate the issue of vanishing gradients [10]. However, while LSTM can extract contextual sequences based on fake news from news content, it only operates in one direction. Therefore, the proposed deep neural network implements a BiLSTM [24] layer that is stacked. It is composed of 128 units of LSTM organized into two levels. One layer is responsible for processing the input sequence in the forward direction, while the second layer handles the opposite direction. As a result, it manages long-term dependencies successfully and is able to extract contextual sequences based on fake news from the input news text in both directions [12].

In this work, the output vector G produced by the CNN layer is passed onto each LSTM layer independently. Both LSTM outputs, forward and backward, are produced in the form of two hidden states $\vec{h}_f$ and $\vec{h}_b$, which represent the forward and reverse contextual sequences of the input news text, respectively:

$$\vec{h}_f = \mathrm{LSTM}_f(G, \theta_f), \tag{3.3}$$

$$\vec{h}_b = \mathrm{LSTM}_b(G, \theta_b) \tag{3.4}$$

where LSTM_f and LSTM_b represent the LSTM layers for forward and backward processing, with θ_f and θ_b being their respective parameters.

The preceding outputs are subjected to a dropout layer to prevent overfitting and thereby improve generalization error. The resultant representations V_f produced by this layer are transmitted to the attention layer:

$$V_f = \mathrm{Dropout}(\vec{h}_f \oplus \vec{h}_b, r) \tag{3.5}$$

where r is the dropout rate, determining the proportion of features to randomly drop out to prevent overfitting.

Finally, V_f is transmitted to an attention layer, which is designed to weigh the importance of different words in the input sequence.

Attention Layer: A neural network-based method known as attention makes an effort to derive the most significant information possible from the input data by analyzing it in terms of the data's semantics as a whole [16]. It is accomplished by concentrating on terms in the input text that are essential to the context. An attention mechanism at the word level is considered in this work. The attention layer takes V_f as input and determines the relative importance of fake news-related tokens.

First, a score is assigned to each token in the sequence to establish its significance. This process is typically accomplished using a compact neural network or a basic linear transformation, followed by a nonlinear function. The latent representation y_t is generated by the activation function (tanh) that receives as input the product of T_w (trained weight matrix) and V_f (feature vector) plus b_w (bias vector):

$$y_t = \tanh(T_w V_f + b_w) \tag{3.6}$$

The scores are then normalized across the sequence using a softmax function to obtain the attention weights, denoted as $a_{(t)}$, based on the hidden representation y_t:

$$a_{(t)} = \frac{\exp(y_{(t)})}{\sum_t \exp(y_{(t)})} \tag{3.7}$$

The computation of the CBA network vector s_k for the source (news text) is performed by Equation (3.8), which relies on the values of $a_{(t)}$ and V_f. This context vector represents the input sequence with a focus on the important elements relevant to the task:

$$s_k = \sum_{i=1}^{t} a_{t_i} V_{f_i} \tag{3.8}$$

3.4.4. *Enhancement using Auxiliary Features.* As discussed in section 3.3.2, there are a few crucial fake news-related features that can provide an additional boost to the proposed deep learning model if appropriately combined with the previous layer. Therefore, we have manually crafted a total of 30 additional features that can be categorized into four groups, including readability, grammar, sentiment, and LIWC. Utilizing these auxiliary features yields a feature vector $l \in R^d$ with $d = 30$.

3.4.5. *Dense Layer.* The CBA vector s_k, produced through the CBA layer, is combined with the FN-related extracted feature vector l through concatenation to form the extended feature vector C_v, denoted as $C_v = s_k \oplus l$. The vector C_v is subsequently transmitted to the dense layer to set up the model for classification.

3.4.6. *Output Layer.* The model is trained using binary cross-entropy as the preferred loss function. The utilization of a sigmoid function is implemented in the final layer to classify the result into two classes: fake news (FN) and non-fake news (NFN).

4. Experimental setup and result evaluation

The evaluation study and results of CBiAtFake are presented in this section, which consists of subsections such as evaluation metrics used, hyperparameter settings, result analysis, ablation study, and analysis of the impacts of the various parameters used in this study.

4.1. **Evaluation Metrics.** Four standard metrics are used to assess the efficacy of the CBiAtFake model: F1 score, Accuracy, Recall, and Precision.

Accuracy: The quantification of correctly predicted outcomes relative to the total number of cases being evaluated is known as accuracy.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{4.1}$$

Recall: Recall is the ratio of correctly identified positive cases to the total number of positive instances.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4.2}$$

Precision: The computation involves determining the proportion of correctly predicted positive outcomes by the model relative to its total number of positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4.3}$$

F_1 Score: This coefficient assists in evaluating the model's predictive efficacy based on the results from both recall and precision.

$$F_1 \text{ score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

4.2. **Hyperparameter Setting.** Every deep-learning solution requires the optimization of hyperparameters. We worked diligently on numerous hyperparameters to enhance the efficacy of the model. The optimal list of variables for the CBiAtFake model implementation is shown in Table 3.

TABLE 3. Optimized hyperparameters values used in the CBiAtFake model

Hyperparameter	**Value**
Total Number of Layers	5
Dimension of GloVe embedding	300
Number of Neurons in BiLSTM Layer	128
Loss Function	Binary cross entropy
Activation Function	Sigmoid
Dropout	0.5
Learning Rate	0.1
Optimizer	Adam
Epoch	15
Batch Size	30

4.3. **Result Analysis.** CBiAtFake is implemented in Python utilizing the Keras library . The dataset has been split into two parts: 80% is designated for training, and the remaining 20% is used for testing. In the evaluation of the model, a total of eight baseline methods were taken into account. These methods encompassed four traditional ML approaches, namely Naive Bayes, SVM, Decision Tree (DT), and Random Forest. Additionally, four neural network models were included in the evaluation, specifically Dense Neural Network (DNN), CNN, LSTM, and Gated Recurrent Unit (GRU). The DNN is composed of three layers, each containing 64 neurons, CNN consists of 128 filters and a width of 3. Both the GRU and LSTM networks consist of 128 units of neurons each.

The efficacy of various classifiers utilized in the experiment is summarized in Table 4. Among the ML methods, the Naive Bayes algorithm demonstrated higher efficiency with an accuracy rate of 88.90% while SVM produces the worst outcome with a low accuracy of 66.73%. We have observed that traditional machine learning techniques perform inadequately in FN detection.

Further, to analyze the efficacy of the neural model for FN detection, extensive experiments are performed with the pre-trained word embedding technique GloVe using the identical dataset. The classification accuracy of the GloVe-enabled technique with DNN has marginally increased to 91.20% compared to baseline ML techniques. CNN with 10 epochs yielded a test accuracy of 92.76% while The LSTM model demonstrated a classification accuracy of 97.25% when trained for the same number of epochs. GRU outperformed all other deep-learning-based models in the experiment and achieved 98.20% testing accuracy. It has been observed that derivatives of RNN, such as LSTM and GRU, perform well in fake news classification due to their context-learning properties. Finally, the proposed CBiAtFake model is examined utilizing the GloVe embedding model. To optimize the performance of the model we have performed the hyperparameter tuning and obtained the optimal values for the different parameters as shown in Table 3. The analysis demonstrates a training accuracy of 99.90% and a testing accuracy of 98.96% with 15 epochs. Furthermore, it is important to highlight that the CBiAtFake model demonstrates superior performance compared to other classifiers used in this study. This enhanced performance can be attributed to the model's ability to learn text sequences in a bidirectional manner and its incorporation of hand-crafted language features related to fake news. Moreover, empirical evidence suggests that the CBiAtFake model achieves a stable equilibrium in training, manifesting neither under-fitting nor overfitting. Figures 3 and 4 display the accuracy and model loss over epochs for CBiAtFake model, evaluated using both training and testing data.

Table 4. Classification results

Evaluation parameter	**Accuracy**	**Precision**	**Recall**	**F1 score**
Model				
Naïve Bayes	88.90	88.10	87.50	88.40
Decision Tree	72.40	72.10	72.37	72.50
SVM	66.73	67.20	67.30	66.50
Random Forest	71.50	71.30	71.75	71.20
DNN	91.20	88.20	88.20	88.20
CNN	92.76	90.62	94.78	91.52
LSTM	98.10	97.30	98.71	97.61
GRU	98.20	97.98	98.42	98.18
CBiAtFake	98.96	98.88	99.06	98.85

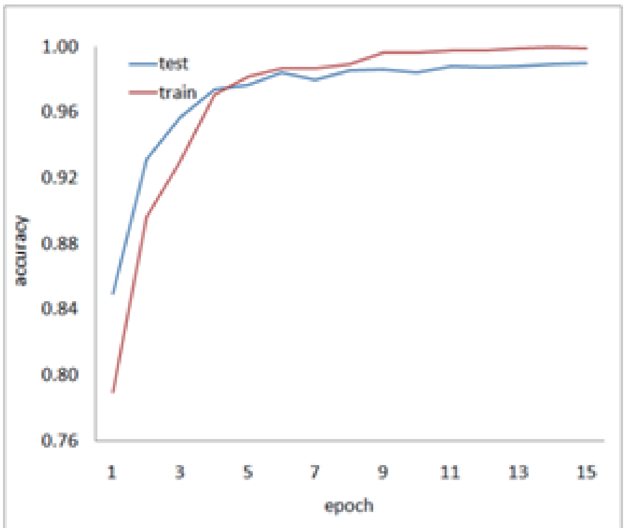

Figure 3. Model accuracy and epoch

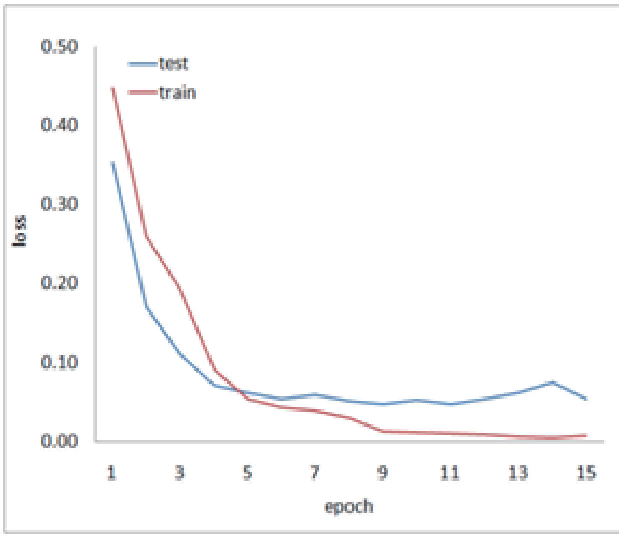

FIGURE 4. Model loss and epoch

4.4. **Ablation Study.** An ablation study is performed to assess the impact of each individual component within the CBiAtFake model on its overall performance. Table 5 presents the findings of the investigation based on the F1 score values. It can be observed that CBiAtFake consistently demonstrates excellent performance compared to alternative options having a reduced number of components. The CBA layer appears as the main component of the model and excluding the same will drastically reduce the F1 score to 93.30%. Thus, the CBA layer serves a vital part in extracting contextual information from the news text in multiple directions. By excluding the auxiliary features from the CBiAtFake, the F1 score reduces to 98.35%. Therefore, the validation of the significance of our well-informed model utilizing external features has been approved. The F1 score is diminished further when the layered BiLSTM network is excluded. The utilization of sequence modeling networks, such as RNN and LSTM, is prevalent in contemporary research. Among these methods, the stacked BiLSTM stands out as one of the most critical elements in the acquisition of contextual sequences in both directions. In the end, usage of the attention layer is accessed, and it has been observed that the exclusion of this network lowers the overall F1 score to 98.20%. Thus, it can be demonstrated that the attention layer is also an essential factor by emphasizing the key fake news-related words.

4.5. **Comparative Analysis of CBiAtFake.** In this study, we examine the following state-of-the-art methods and compare them with the proposed CBiAtFake model. These methods were assessed using an identical dataset.

(i) Ahmad et al. [1]: The method uses n-gram, TF-IDF, and word embedding and uses various ML techniques for FN classification.
(ii) Singh et al. [27]: This paper uses linguistic features including LIWC, and applied ML methods for fake news classification.

Table 5. Ablation study of the CBiAtFake Model

Evaluation parameter	F1 score
Without Attention layer	98.20
Without BiLSTM Stack layer	95.42
Without Aux. Features layer	98.35
Without CBA layer	93.30
With All layers	98.85

(iii) Kaliyar et al. [13]: Their research utilizes GloVe embedding in a deep CNN network to construct a DL-based model for FN detection.
(iv) Chaudhry and Arora [6]: Their study employs various categories of linguistic features for FN detection, including grammatical, reliability, syntactic, and sentiment. These features are examined and evaluated with various ML and DNN techniques that include the LSTM.

Table 6 presents a concise overview of the effectiveness of existing methods in comparison to the suggested CBiAtFake method. At first sight, it appears that deep learning approaches outperform ML-based methods. CNN and RNN-based methodologies are quite useful for classifying fake news. When compared to the ML-based technique [1, 4] demonstrates a considerable improvement in performance accuracy to 95.20%, indicating that linguistic features also serve a significant impact in FN classification. In addition, the comparative analysis revealed that bidirectional learning and attention to the context of fake news classification is more accurate than the CNN-based method Kaliyar et al. [12] and the LSTM-based method Choudhary and Arora [4]. Further, we observed that the GloVe model was beneficial in extracting the context of words related to fake news during training, resulting in greater accuracy. Overall CBiAtFake outperformed the state-of-the-art, achieving an accuracy rate of 98.96%.

Table 6. Comparative analysis of CBiAtFake with existing state-of-the-art

Author	Features/Model	Accuracy
Singh et al. [27]	Linguistic Features, ML	87.00
Ahmed et al. [1]	TF-IDF, word-embedding, ML	92.00
Choudhary & Arora [6]	Linguistic Features, ML, LSTM	95.20
Kaliyar et al. [13]	Pretrained embedding, Deep CNN	98.30
CBiAtFake	GloVe, CNN, BiLSTM, Attention to FN-related words	98.96

5. Conclusion

The present study introduces a novel deep neural newtoek model CBiAtFake, designed to detect FN on the OSM. The model comprises five distinct layers of advanced DNN. The GloVe is employed for the purpose of generating the embedding vector to represent the news text. The CBA network is responsible for the acquisition of linguistic contextual information in both the forward and backward direction of the input sequence. This is accomplished with the assistance of CNN followed by bidirectional LSTM and attention mechanism. Additionally, CBiAtFake is enhanced by the extensive collection of manually crafted linguistic features to make it well-informed. The CBiAtFake outperformed the current state-of-the-art methods in the comprehensive experiments that we ran and achieved an accuracy of 98.96% in classifying FN. Its computation does not necessitate the use of high-end hardware resources. The software can operate on a CPU that possesses a maximum RAM capacity of 8 GB. This finding paves the way for future research paths that will generalize CBiAtFake for use in other text classification problems, especially those that do not have access to high-end hardware resources. Although CBiAtFake is quite good at detecting fake news articles in English text, it does have some limitations. For instance, the CBiAtFake model has not undergone evaluation for the task of detecting fake news across multiple classes. Finally, its applicability to multilingual or code-mixed data is an exciting area of future research.

References

[1] H. Ahmed, I. Traore and S. Saad, *Detection of online fake news using n-gram analysis and machine learning techniques*, in: International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, Springer, Cham, 2017, pp. 127–138.

[2] S. A. Alkhodair, S. H. Ding, B. C. Fung and J. Liu, *Detecting breaking news rumors of emerging topics in social media*, Information Processing & Management **57** (2020): 102018.

[3] S. Bauskar, V. Badole, P. Jain and M. Chawla, *Natural language processing-based hybrid model for detecting fake news using content-based features and social features*, International Journal of Information Engineering and Electronic Business. **11** (2019), 1–10.

[4] C. Castillo, M. Mendoza and B. Poblete, *Information credibility on Twitter*, in: Proceedings of the 20th International Conference on World Wide Web, 2011, pp. 675–684.

[5] Y. Chen, N. J. Conroy and V. L. Rubin, *Misleading online content: Recognizing clickbait as false news*, in: Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, 2015, pp. 15–19.

[6] A. Choudhary and A. Arora, *Linguistic feature-based learning model for fake news detection and classification*, Expert Systems with Application. **169** (2021): 114171.

[7] J. P. Dickerson, V. Kagan and V. Subrahmanian, *Using sentiment to detect bots on Twitter: Are humans more opinionated than bots?*, in: 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), IEEE, 2014, pp. 620–627.

[8] M. H. Goldani, S. Momtazi, and R. Safabakhsh, *Detecting fake news with capsule neural networks*, Applied Soft Computing. **101** (2021): 106991.

[9] G. Gravanis, A. Vakali, K. Diamantaras and P. Karadais, *Behind the cues: A benchmarking study for fake news detection*, Expert Systems with Applications **128** (2019), 201–213.

[10] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, Neural computation. **9** (1997), 1735–1780.

[11] T. A. O. Jiang, J. P. Li, A. U. Haq, A. Saboor and A. Ali, *A novel stacking approach for accurate detection of fake news*, IEEE Access. **9** (2021), 22626–22639.

[12] A. Kamal, T. Anwar, V. K. Sejwal and M. Fazil, *BiCapsHate: Attention to the Linguistic Context of Hate via Bidirectional Capsules and Hatebase*, IEEE Transactions on Computational Social Systems, 2023.

[13] R. K. Kaliyar, A. Goswami, P. Narang, and S. Sinha, *FNDNet: A deep convolutional neural network for fake news detection*, Cognitive Systems Research **61** (2020), 32–44.

[14] S. Kula, M. Choraś, R. Kozik, P. Ksieniewicz and M. Woźniak, *Sentiment analysis for fake news detection by means of neural networks*, in: International Conference on Computational Science, Springer, Cham, 2020, pp. 653–666.

[15] S. Kumar and N. Shah, *False information on web and social media: A survey*, arXiv preprint arXiv:1804.08559, 2018.

[16] T. Luong, H. Pham and C. D. Manning, *Effective approaches to attention-based neural machine translation*, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015, pp. 1412–1421.

[17] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K. F. Wong and M. Cha, *Detecting rumors from microblogs with recurrent neural networks*, in Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16), 2016.

[18] M. Ott, Y. Choi, C. Cardie and J. T. Hancock, *Finding deceptive opinion spam by any stretch of the imagination*, in: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, vol. 1, Association for Computational Linguistics, 2011, pp. 309–319.

[19] F. A. Ozbay and B. Alatas, *Fake news detection within online social media using supervised artificial intelligence algorithms*, Physica A: Statistical Mechanics and its Applications. **540** (2020): 123174.

[20] J. Pennington, R. Socher and C. D. Manning, *Glove: Global vectors for word representation*, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532–1543.

[21] V. Pérez-Rosas, B. Kleinberg, A. Lefevre and R. Mihalcea, *Automatic detection of fake news*, in: Proceedings of the 27th International Conference on Computational Linguistics, 2018, pp. 3391–3401.

[22] N. Ruchansky, S. Seo and Y. Liu, *CSI: A hybrid deep model for fake news detection*, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017, pp. 797–806.

[23] S. R. Sahoo and B. B. Gupta, *Multiple features-based approach for automatic fake news detection on social networks using deep learning*, Applied Soft Computing. **100** (2021): 106983.

[24] M. Schuster and K. K. Paliwal, *Bidirectional recurrent neural networks*, IEEE transactions on Signal Processing **45** (1997), 2673–2681.

[25] K. Shu, A. Sliva, S. Wang, J. Tang and H. Liu, *Fake news detection on social media: A data mining perspective*, ACM SIGKDD Explorations Newsletter **19** (2017), 22–36.

[26] K. Shu, S. Wang, and H. Liu, *Defend: explainable fake news detection*, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2019, pp. 395–405. `https://doi.org/10.1145/3292500.3330935`.

[27] D. V. Singh, R. Dasgupta and I. Ghosh, *Automated fake news detection using linguistic analysis and machine learning*, in: International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS), 2017, pp. 1–3.

[28] S. Sridhar and S. Sanagavarapu,*Fake news detection and analysis using multitask learning with BiLSTMCapsNet model*, in: International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE), IEEE, 2021, pp. 905–911.

[29] Y. R. Tausczik and J. W. Pennebaker, *The psychological meaning of words: LIWC and computerized text analysis methods*, Journal of Language and Social Psychology. **29** (2010), 24–54.

[30] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi and B. W. On, *Fake news stance detection using deep learning architecture (CNN-LSTM)*, IEEE Access **8**(2020), 156695–156706.

[31] Y. Wang, M. Huang, X. Zhu and L. Zhao, *Attention-based LSTM for aspect-level sentiment classification*, in: Proceedings of the 2016 conference on empirical methods in natural language processing, 2016, pp. 606–615.

[32] L. Xiao, H. Zhang, W. Chen, Y. Wang and Y. Jin, *Mcapsnet: Capsule network for text with multi-task learning*, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 4565–4574.

[33] Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li and P. S. Yu, *TI-CNN: Convolutional neural networks for fake news detection*, arXiv preprint arXiv:1806.00749, 2018.

[34] X. Zhang, J. Zhao, and Y. LeCun, *Character-level convolutional networks for text classification*, Advances in Neural Information Processing Systems, 2015, pp. 649–657.

[35] B. Zhong, X. Xing, P. Love, X. Wang and H. Luo, *Convolutional neural network: Deep learning-based classification of building quality problems* Advanced Engineering Informatics **42** (2019), 46–57.

Manuscript received April 24, 2024
revised August 27, 2024

M. Meraj
Department of Mathematics, Jamia Millia Islamia, New Delhi, India
E-mail address: meraj.jmi001@gmail.com

A. Kamal
Department of Mathematics, Jamia Millia Islamia, New Delhi, India
E-mail address: akamal1@jmi.ac.in

J. J. Sindorovich
Department of Computer Science, Samarkand State University, Samarkand, Uzbekistan
E-mail address: jamoliddin.jabbarov@mail.ru