J. Integr. Sci. Technol. 2022, 10(2), 110-119 Article

Journal of Integrated
SCIENCE & TECHNOLOGY

Automated hybrid Deep Neural Network model for fake news identification and classification in social networks

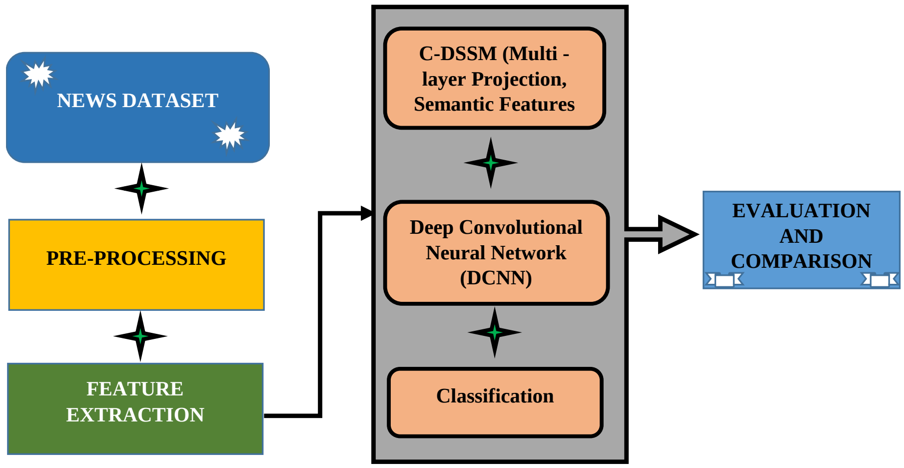
Roshan R. Karwa*, Sunil R. Gupta

Department of Computer Science & Engineering, Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati-444701, Maharashtra, India.

Received on: 29-Apr-2022, Accepted and Published on: 07-July-2022

ABSTRACT

The rapid growth of social media has far-reaching impacts on civilization, traditions, and economics, including both beneficial and unfavourable implications. Since social networking sites have become more frequently utilized for transmitting data, they have also become a gateway for the distribution of fake news for diverse financial and legislative goals. Artificial Intelligence (AI) and Natural Language Processing (NLP) approaches have a lot of ability for academics who wish to design models that can recognize fake news automatically. On the other hand, identifying fake news is a difficult issue because it demands systems that describe the news and then contrast it to the actual news to categorize it as fake. Thus, to overcome this, this paper introduces Hybrid Deep Neural Network Model, in which C-DSSM and Deep CNN models have been utilized. It identifies and classifies fake news using the LIAR dataset. According to experimental results, the proposed model obtained an accuracy of 92.60%, a recall of 92.40%, a precision of 92.50%, and an F1 score of 92.50%. Furthermore, the proposed model is compared to earlier studies for fake news identification using the LIAR dataset, and the proposed model's performance is remarkable. As a result, the proposed hybrid model gives better results in detecting and classifying fake news on social networks.



Keywords: Deep learning, Fake news detection, Rumors, Misinformation, Dis-information, Convolution, Dynamic Semantic Structural Model

INTRODUCTION

Social media is becoming a vital aspect of everyday life due to great improvements in technology and communications.[1,2] It proved to be an excellent method for people to receive data. Individuals have used social media to express their thoughts, views, and feelings. Since it regards the advantages of utilizing social media, consumers prefer it over conventional news resources like magazines and televisions. The origin of false news is a user, who then uploads it on social media sites and spreads it to several other people without verifying the truth of the news or data they are providing. With so much data or news available, the concern of whether that supplied news or data is authentic or fake arose. Fake news is usually broadcast in the attempt to misdirect or influence to acquire a political or financial benefit. Assume the preceding example: During the latest Indian voting, there seemed to be a large debate over the legitimacy of numerous news articles endorsing specific candidates and the political thought processes that underpin them. In light of this heightened attention, detecting fake news is critical to minimizing its detrimental influence on people and society.

Fake news has been largely viewed as among the most severe threats to society, media, and freedom of opinion. It has eroded public faith in government entities. According to researchers, false news via Twitter has been frequently forwarded by plenty of persons and circulates way faster, particularly in the case of political news.[3] Economic and ideological incentives are the two primary motivations for disseminating fake news.[4] For instance,

*Corresponding Author Roshan R. Karwa, Badnera, Amravati, India, 444701.
Email: karwa.roshan@gmail.com
Cite as: *J. Integr. Sci. Technol., 2022, 10(2), 110-119.*

 ISSN: 2321-4635 http://pubs.iscience.in/jist

during the 2016 US Presidential election, several young people in the Macedonian town of Veles got wealthy by spreading sensationalist and frequently false news through penny-per-click advertising.[5] Fake news has been a source of contention for decades. However, it was highlighted by Donald Trump during the 2016 US Presidential Election. As per ref. [6] during the election, the top twenty false electoral reports received 8,711,000 sharing, responses, and comments via Facebook, while the top 20 finest functioning news on 19 major websites earned 7,367,000 shares, reactions, including opinions. Even the stock exchange isn't immune to false news; for example, a fraudulent tweet about an "explosion" wiped out $130 billion in stock value in 2013, and fraudulent news was blended with unlawful stock market influence.[7] Fake news is not restricted to this profession; it also influences other industries, such as product reviews.[8] As evidenced by the rising usage of language like "post-truth," which was chosen as the global word of the year by Oxford Dictionaries in 2016,[9] those events and setbacks have fueled false news studies and discussion.

Fake news has evolved into serious public concern and a big obstacle for all those aiming to battle misinformation. The ease with which knowledge may be disseminated through sharing has resulted in an exponential increase in its fabrication. The legitimacy of social media material is particularly jeopardized in areas where false news is popular.[10] Many attempts have been made and continue to be made to understand, recognize, and decrease, if not eliminate, the propagation of false news. Still, the area has become increasingly sought after with the development of AI that allows machines to learn quickly, detect, and forecast.

The automatic authentication of a text item as genuine or false is a difficult problem.[11] Various social networking services modify news based on personal beliefs or interests. Fake news was disinformation and perhaps even altered news disseminated through social media to tarnish the prestige of an individual, gang, enterprise, or political group. Due to the overall extent of fake news that has recently been propagated in the media, computational approaches for detecting it are in high demand. Fake news identification attempts to aid users in spotting several sorts of fictitious news. Individuals might judge if the news is true or fake predicated on prior experiences with false or authentic news. We can get false news via social media using a variety of approaches. We need to improve our ability to distinguish between phoney and authentic news. A human can't detect all phoney news. Human fact-checking is time-consuming and costly. As a result, there's a pressing necessity to develop a technology that could automatically recognize and categorize news as legitimate or phoney; this may be achieved, for example, through employing artificial intelligence strategies, whereas detection accuracy and feature extraction approaches have considerable limits. Therefore, our research proposes the Hybrid Deep Neural Network Model to automatically recognize and categorize the news.[12]

C-DSSM and Deep CNN models have been utilized in this proposed model, in which additional convolutional layers in the C-DSSM retrieve local contextual information. In contrast, a max-pooling layer forms a global feature vector. Furthermore, the C-DSSM architecture's first tiers depict the feature extraction strategy, whilst the D-CNN layers indicate the categorization process.

Related Works

Fake news has aggressively spread through social media and has become a major issue in modern civilization. To limit the detrimental effects of false news, it is vital to recognize it. This section outlines the work that has previously been completed to detect false news.

Ruchansky et al.[13] Presented a model incorporating three objectives for a more accurate and automated prediction. Based on the response and text, a Recurrent Neural Network is used to preserve the temporal pattern of user activity; nevertheless, the process is slow and hard to train. To increase the processing time, author Karimi et al.[14] proposed a CNN-LSTM network for feature extraction. The authors introduce approaches for incorporating information from multiple sources and discriminating between different degrees of fakeness. The author proposes the Multi-Source Multi-class Fake News Detection framework MMFD, combining automated feature extraction by CNN-LSTM. Still, it requires more memory to train multi-source fusion and automated degrees of fakeness detection into a coherent and interpretable model.

Deligiannis et al.[15] proposed the graph convolutional neural network (GCN). The authors designed a multi-entry neural network architecture (MENET) that can handle diverse types of information to predict the continuous geographical coordinates of Twitter users. Even though Graph Convolutional Neural Network has a limitation of time and space complexity are higher.

Shu et al.[16] proposed a Fake News Ratio to extract the features from real-world datasets. However, it increases unanticipated problems. Wu et al.[17] introduced a novel strategy for classifying false material on social media that gathers contextual message semantic features utilizing the recurrent architecture of an RCNN network and then learns sentiment depiction employing a convNet architecture. In contrast, RCNN is not able to detect fake news in real-time.

Dong et al.[18] proposed an attention-dependent bi-directional Gated Recurrent Units (GRU) prototype to retrieve the attributes from news data as well as a deep framework to excerpt concealed portrayals of side data and afterwards merged the two hidden vectors deriving from the extractions as mentioned earlier into an attention matrix and learned an attention dispersion over the vectors. Ultimately, the circulation is being utilized to help identify fake news. Moreover, Bi-GRU has a limitation of vanishing and exploding gradient.

Wang et al.[19] suggested an Event Adversarial Neural Network (EANN). This end-to-end architecture can create event-invariant features and therefore assist in recognition of misleading info on freshly obtained occurrences. The event discriminator's function might be eliminating event-specific elements while retaining event-specific characteristics. The Event Adversarial Neural Network (EANN) needs the training to operate, which is difficult.

Nyow et al.[20] established procedures for discovering relevant Tweets' features and application architecture for systematically automating the categorization of an online article and proposing enhanced Tweets' qualities. Because of the boundaries imposed by Twitter and its API, this research has limitations. At this point, the crawling time from Twitter necessitates a significant amount of waiting time, and only tweets marked as 'public' may be examined.

Gravanis et al.[21] presented a strategy for detecting false news based on content-based characteristics and Machine Learning (ML) techniques. The authors developed an improved collection of linguistic indicators with great capabilities for distinguishing false news from legitimate news items. Furthermore, the suggested technique might be the foundation for a tool that allows publishers to swiftly determine whether articles require further exploitation based on their validity.

Jadhav et al.[22] proposed utilizing upgraded Recurrent Neural Networks and a Deep Structural Semantic Framework to identify and categorize fraudulent news items. Without pre-existing domain knowledge, the offered technique intuitively identifies critical characteristics linked with fake news and obtains 99 percent accuracy, whereas the RNN computation is slow.

Zhang et al.[23] proposed an innovative analytics-driven method for identifying false news. We begin by describing the architecture for the suggested technique and the foundational computational framework, containing information on execution and validation against a collection of news input. The overall efficacy of the suggested technique is demonstrated by the design and testing of a breakthrough False News Identification scheme. The presented scheme obtains a classification accuracy of 92.49 percent.

Qayyum et al.[24] offered an in-depth review of a block chain-based paradigm for bogus news avoidance, stressing the numerous architectural obstacles and concerns of such a block chain-based paradigm for coping with fake news. Block chain, just a decentralized ledger system, claims to offer openness and assurance to this emerging "post-truth" civilization by facilitating characteristics such as intelligent contracts, decentralized consent, and tamper-proof authentication. The pitfall is acknowledged as a limitation of their proposed system.

When assessing and categorizing news articles and statements, Wang et al.[25] created a systematic automated strategy for recognizing the numerous cases. This strategy is erected on a five-tiered hierarchy of fakeness. It methodically investigates a wide range of social media signals, capturing both the content and language of postings and the sharing and dispersion among users. The logistic regression method has a problem when group proficiency distributions have little overlap.

Ahmad et al.[26] suggested utilizing machine learning ensemble approaches such as KNN and SVM for automatic news article categorization. The authors investigate several textual features that might be utilized to distinguish between true and fraudulent information. The authors use such characteristics to train many machine learning models utilizing distinct ensemble approaches. Therefore, real-time fake news detection in videos could be a promising future approach.

Kesarwani et al.[27] established a unique framework for recognizing false news leveraging a K-Nearest Neighbor classifier. The greatest categorization accuracy attained by this approach was 79 percent. The KNN is very time-consuming in the classification process.

Smitha et al.[28] employed machine learning and natural language processing to show a framework for identifying misleading news among news articles. Numerous features of engineering technologies, including count vector, TF-IDF, and embedding, have been used to generate the feature vector in this presented research whereas TF-IDF is slow for large vocabularies but doesn't use word semantic similarity. Seven different Machine Learning Categorization systems have been trained to detect whether the news is accurate or false.

Ozbay et al.[29] presented a two-stage methodology for identifying bogus information in social media. The first stage of the approach is executing a multitude of pre-processing processes for the data gathering to convert unorganized data sets. It is assumed that the counts of various terms give independent evidence of similarity.

Asghar et al.[30] established a BiLSTM-CNN paradigm, a hybrid of a BiLSTM and a CNN. This BiLSTM, popularly referred to as the sequential tier, retains sequence data in both directions (forward as well as backward), whereas the CNN layer captures characteristics from the rich illustration supplied by the BiLSTM layer, besides the tweet is eventually categorized as a rumour or non-rumour. Other characteristics, such as visuals and social context, can be evaluated in addition to text-based features to achieve more efficient results.

Li et al.[31] established MCNN-TFW, a multi-stage convolutional neural network-dependent fraudulent news identification method wherein MCNN retrieves article depiction. At the same time, WS estimates the weight of hypersensitive terms for every news item. The authors next employed a methodology to quantify the weight of sensitive words (TFW), revealing their larger relevance with bogus or true labels. The performance of our technique in a greater range of applications will be evaluated in future work.

Graph-aware Co Attention Networks (GCAN), a novel neural network-based system created by Lu et al.[32] also anticipates fake news depending on the initial tweet and its propagation-based users. Moreover, GCAN might provide earlier diagnoses of false news with excellent results. The scientists believe that GCAN would be employed for more than merely recognizing phoney news on social media, but also for sentiment identification, hateful speech identification, including tweet trend forecasting. In future work, the author looks into model generalization. Furthermore, since most fake news focuses on specific events, the authors utilized GCAN to investigate how to eliminate event-specific features to improve performance and explainability.

Kaur et al.[33] established a novel multi-level voting ensemble method. The pool of characteristics retrieved from the corpus is fed into the specified machine learning models using three feature extraction approaches. A web-based GUI for the proposed fake news detection system will be developed in the future to categorize content as fake or true on real-time social media platforms such as Facebook, Instagram, and Twitter.

Jiang et al.[34] applied cross-validation to appraise the effectiveness of five machine learning techniques, including three deep learning systems on two datasets of fake & real news of diverse sizes. Researchers also utilized term frequency, term frequency-inverse document frequency, and embedding techniques to gain text descriptions for machine learning and deep learning architectures. These proposed approaches, however, have some accuracy limits.

Kaliyar et al.[35] suggested a new methodology (FakeBERT) relying on BERT (Bidirectional Encoder Representations from

Transformers) by integrating multiple concurrent blocks of a single-layer deep Convolutional Neural Network (CNN) containing dynamic kernel sizes including filters with the BERT. This blend is excellent for dealing with ambiguity.

Nasir et al.[36] developed a novel blended deep learning framework for false news detection that blends convolutional and recurrent neural networks. The current study assists in this area by presenting an approach centered on innovative approaches that illustrate deep neural networks' application for falsified news identification. It employs a mix of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), which enhances the effectiveness of the suggested phoney news identification system. However, the RNN is difficult to process for longer sequences.

However, a significant proportion of the research has been performed to recognize false news in social media by employing Artificial Intelligence approaches. Despite this, detection accuracy and feature extraction approaches have considerable limits. To overcome the limits mentioned above, our research creates a novel system that depends on artificial intelligence technologies to recognize false news in social media. The upcoming section describes this novel hybrid deep neural network approach.

Automated Hybrid Deep Neural Network Model for Fake News Identification and Classification in Social Networks

The advent of the World Wide Web, including the super quick approval of social networks (like Fb and Twitter), primed the route for unprecedented rates of information transfer of data in mankind's life. In contrast, it is difficult to automate the categorization of a written article as misleading or disinformation; Artificial Intelligence techniques are utilized to detect bogus news on social media. Despite this, the accuracy of detection and the methods for extracting features have significant limitations. Thus to overcome this, our research proposes a deep learning strategy for automated news article categorization, namely, a hybrid deep neural network model, in which, C-DSSM and Deep CNN models have been utilized, which identifies and classifies bogus news. Initially, the input strings from specified data input were supplied into the input layer, wherein subsequent convolutional layers in the C-DSSM retrieve local contextual information. In contrast, a max-pooling layer generates a global feature vector. Convolution is a sliding window-dependent feature extraction technique in C-DSSM that collects the contextual features of a word. Those strings have been already pre-processed to produce a bag of words. These word bags are again passed on to the C-DSSM model. The preliminary layers of the C-DSSM design reflect the feature extraction approach, whilst the D-CNN layers indicate the categorization process. This proposed architecture is seen in Figure 1.

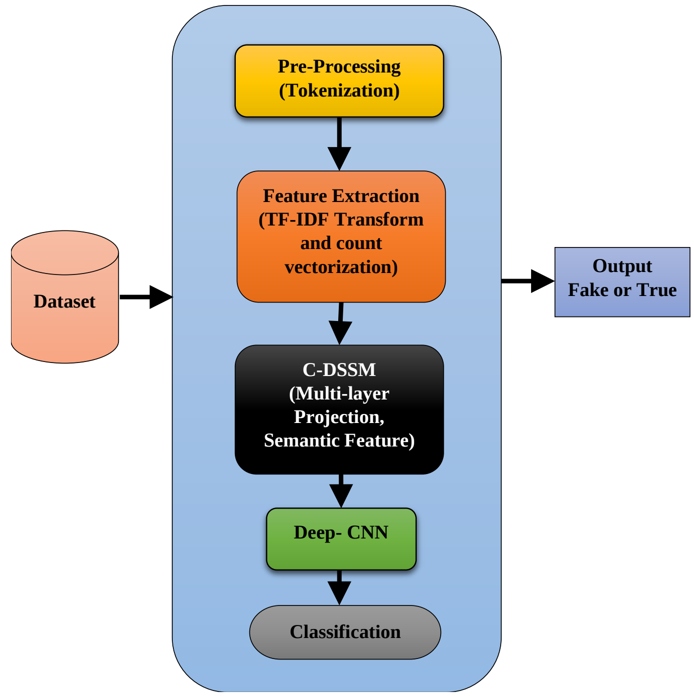


Figure 1. The architecture of the Proposed Method

Extraction of High Semantic Features by Convolution Dynamic Semantic Structural Model (C-DSSM)

The growing usage of social media has repercussions for society, culture, and commerce, including both favourable and negative results. Recognizing misinformation is hard because it demands algorithms to summarise the news and afterward contrast it to genuine news to categorize it as fraudulent. Utilizing the Twitter dataset, our proposed approach identifies and categorizes bogus news leveraging C-DSSM and an upgraded DCNN model. The input strings from the Twitter dataset were supplied into the input layer in this model. Those strings get pre-processed into a word bag. The pre-processed bag of words is then input into the C-DSSM model, wherein convolution is considered a sliding window-dependent feature extraction strategy in C-DSSM, and it gathers a word's contextual features. Furthermore, word hashing has been the first stage in the C-DSSM model, proceeded by multi-layer nonlinear projection and semantic feature generation. The C-DSSM has the following specific structure as shown in figure 2.

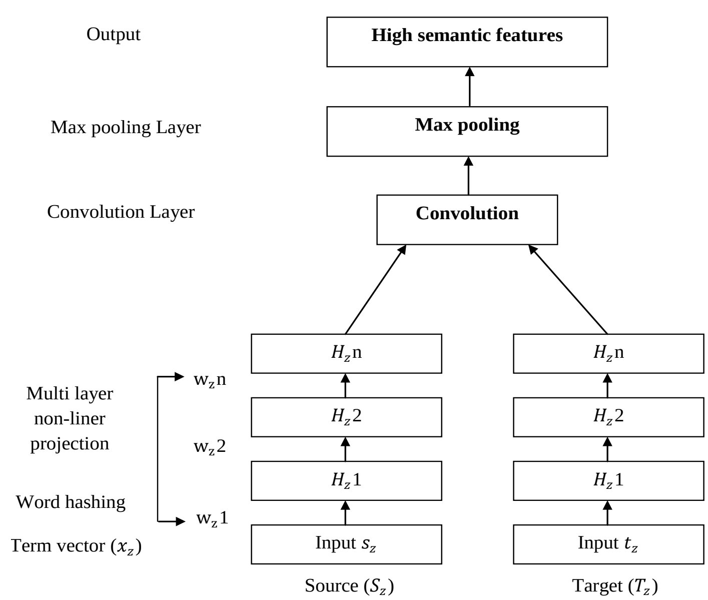


Figure 2. Structure of the C-DSSM

Initially, input from the dataset is provided to the input layer. Then, the input data is pre-processed using tokenization, and additional characteristics are retrieved utilizing the Term Frequency, and Inverse Document Frequency (TF-IDF)[37] transform with count vectorization. Moreover, pre-processed statements are fed into the Convolution Dynamic Semantic Structural Model (C-DSSM). The C- DSSM architecture for translating raw text data into high semantic features has been depicted in Figure 2. The C-DSSM begins with word hashing, lowering the dimensionality of the bag-of-words term vectors employing the word hashing algorithm. In a word (e.g. good), hashing was used for the beginning and ending marks (e.g. #good#). The word is then broken down into letter n-grams (for example, #go, goo, ood, od#). Eventually, the word was expressed as a vector of letter n-grams. The semantic features are generated using a multi-layer nonlinear projection that recognizes keywords/concepts. The C-DSSM then employs additional convolutional layers to retrieve local contextual characteristics, with a max-pooling layer yielding a global feature vector. This convolution procedure was viewed as a sliding window-dependent feature extraction method. Furthermore, the output layer is a semantic layer that offers the high-level semantic feature vector of the input word sequence.

$$R_z(S_z, T_z) = \cos(y_{zs_z}, y_{zt_z}) = \frac{y_{zs_z}^{T_z} \cdot y_{zt_z}}{\|y_{zs_z}\| \|y_{zt_z}\|} \quad (1)$$

Equation (1) above cosine similarity is used to calculate the similarity between the source and target vectors. The similarity score, $R_z(S_z, T_z)$, is calculated using a cosine similarity between two semantic vectors (y_{zs_z} and y_{zt_z}) that describe similarity relations between two statements (S_z and T_z).

Algorithm 1: High Semantic feature extraction by C-DSSM

Start
Step 1: Input data: Source (S_z), Target (T_z)
Step 2: Mapping of raw text
Step 3: Computing word hashing for fake news
Step 4: Generate the word hashing layer (w_z1, w_z2, $w_z3 \ldots w_z n$)
Step 5: Develop a convolution layer and extracts the contextual features from the convolution layer
Step 6: Obtain useful features from the Max pooling layer
Step 7: Perform equation 1
Step 8: Output (High semantic features y_z)
End

Moreover, high semantic features are retrieved in this section, after which the classification is addressed in the next part.

Classification of News (Fake or True) by Deep Convolutional Neural Network (DCNN)

A couple of convolutional layers and a max-pooling layer were replicated in our proposed DCNN architecture, followed by a flattening layer and two fully linked layers. Every convolutional layer has a kernel size of 3*3 with a stride of 1, while all max-pooling layers have a kernel size of 2*2 with a stride of 1. ReLu is a function utilized to activate functions in the hidden layers. Softmax gets utilized as an activation function because the output layer conveys the phoney news.

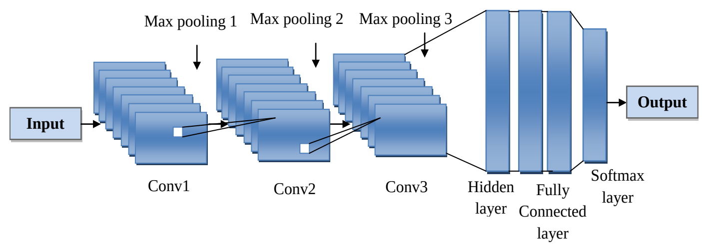


Figure 3. Architecture of DCNN

The C-DSSM output gets loaded into deep convolutional neural networks, which are utilized for categorization. A deep convolutional network's structure comprises convolution, pooling, activation, and fully linked layers, as shown in figure 3. A convolution would be a function that multiplies a collection of weights by the inputs of a neural network. The convolution maps were then treated by a nonlinear activation layer, including Rectified Linear Unit (ReLu) that replaces negative integers in the filtered sentences with zeros. Over time, the pooling layers decrease the assertions, retaining just the most significant facts. The words with the greatest value, for example, are maintained for each batch of statements; this is known as max pooling.

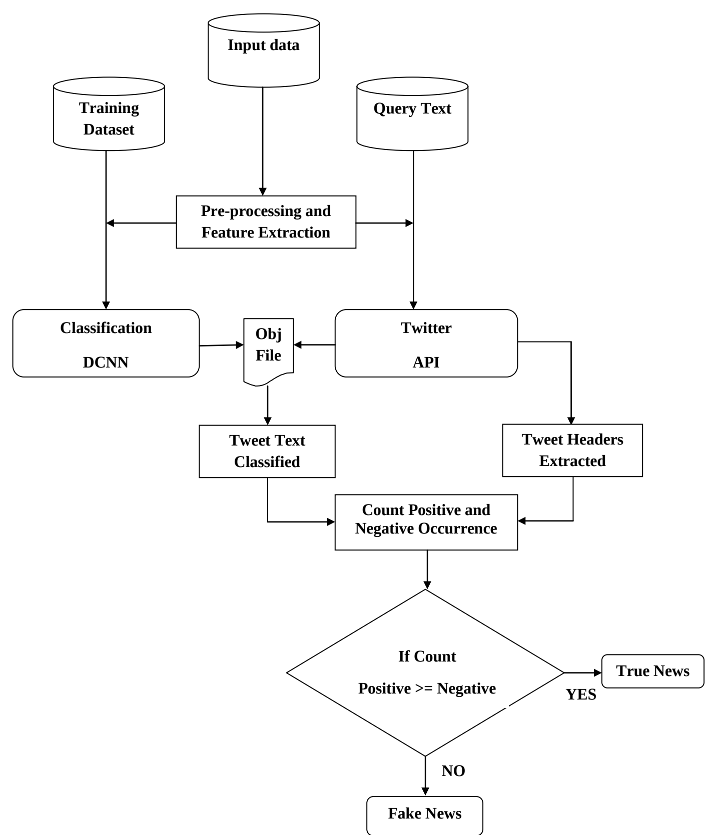


Figure 4. Overall Architecture of the proposed method

After numerous cycles of convolution and pooling layers, the network concludes with a typical multi-layer perceptron or fully connected neural network. Convolution and pooling layers had already filtered, rectified, and compressed the statement's flattened pixels, which have been supplied as an input vector into fully linked layers. The softmax function is utilized for the outputs of the fully linked layers, resulting in the likelihood of the statements being given a class. The output is in binary format as an outcome. The output layer returns either 0 (false news predicted yes) or 1 (false news predicted no) (false news predicted no). As a result, the proposed hybrid model gives better results in detecting and classifying fake news on social networks. Figure 4 shows the overall architecture of the proposed approach.

RESULT AND DISCUSSION

This section covers the implementation outcomes and our proposed system's performance. Moreover, comparison results from the baseline approach are also discussed.

Tool	: PYTHON3
OS	: Windows 7 (64 bit)
Processor	: Intel Premium
RAM	: 8GB RAM

Dataset Description

The LIAR dataset,[38] gathered from politifact.com, contains 12,836 short claims by famous US politicians. True, Mostly True, Half True, Barely True (mostly false), False, or Pants-on-Fire are the different classifications for statements (6 classes). These statements were obtained from various sources, including news releases, TV/radio interviews, campaign speeches, TV advertising, tweets, debates, and Facebook posts, among others. It provides a finer-grained measure of truthfulness than a binary (true or false) label, which is relevant given that statements might comprise both true and incorrect claims. LIAR is significantly larger than previous datasets, facilitating the development of statistical and computational approaches to fake news detection. The research on establishing a broad-coverage fake news detector is also made possible by LIAR's legitimate, real-world short statements from varied contexts with diverse speakers. This dataset was utilized for political NLP research, stance classification, argument mining, issue modeling, rumour detection, and stance classification. In our research, the tweet dataset is pre-processed and turned into a bag of words. The bags of words are then classified using the CDSSM-DCNN model. Figure 5 depicts a sample of news in the LIAR dataset. The proposed system is categorized into two parts: the testing and training phase.

(1) Steps of Training Phase:

- Insert Training dataset
- Tokenisation and stemming procedures are used to pre-process the training data.
- TF-IDF transform is used to extract features.
- Construct a CDSSM-DCNN model and obtain an object file.

(2) Steps of Testing Phase:

- Text query is inserted from the test dataset
- Tokenisation and stemming methods are used to pre-process each text query.
- The TF-IDF transform extracts features and transmits them to the Twitter API.
- Twitter data is generated, and it is passed to the object file.
- Then, count positive and negative occurrences in the tweet header
- Verify the fake and true news based on the positive and negative count occurrences.

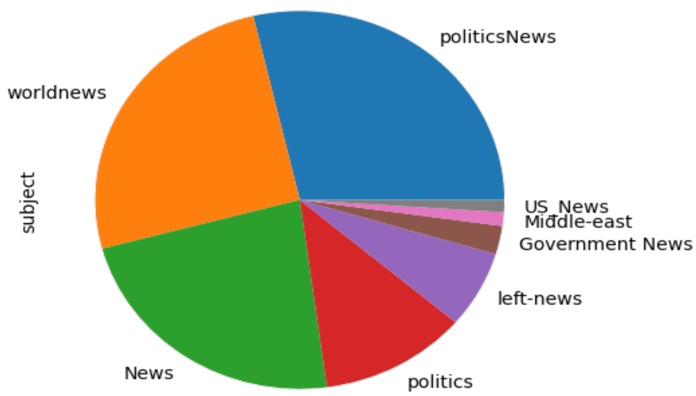


Figure 5. LIAR dataset News

Pre-processing

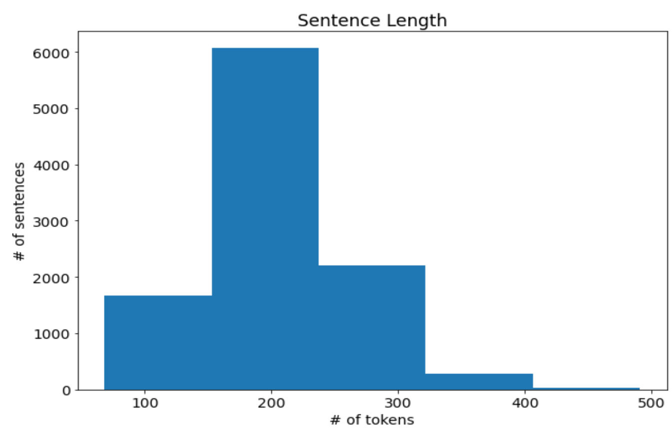


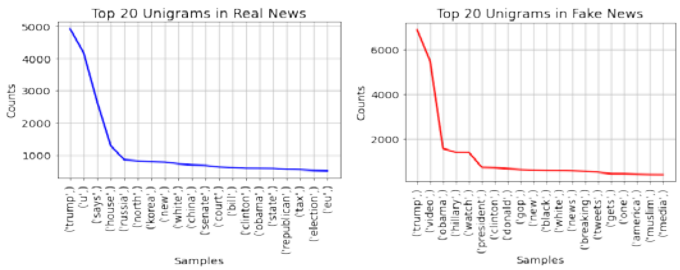
Figure 6. Sentence Length

Figure 6 illustrates the tokenization of a sentence. Separating the text into individual sentences is known as sentence tokenization. Since this sentence tokenizer does not separate individual words, the offensive text is kept in replacement form during the tokenization process. Following the generation of individual sentences, reverse substitutions are performed, resulting in a set of enhanced sentences that replicates the original sentence. In the pre-processing stage, sentence length is obtained in the form of a histogram shown in figure 6.

Feature Extraction

Convolution Dynamic Semantic Structural Model (C-DSSM) is utilized in this feature extraction process. The text is broken down into n-grams (n=1,2,3) such as unigram, bigram, and trigram, which is illustrated as follows:

Unigrams in Real and Fake news

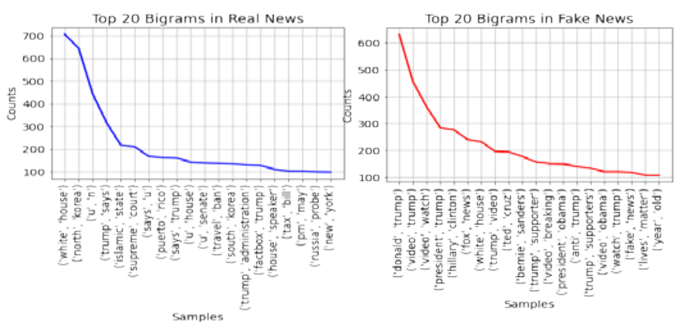


(a) (b)

Figure 7. (a) Top 20 Unigrams in Real News, (b) Top 20 Unigrams in Fake News

Figures 7(a) and 7(b) demonstrate the top 20 most typically utilized unigrams in fake and real news from the dataset. Both utilize similar terminology like 'trump.' However, there are some significant variances. For example, fake news says that John McCain has done nothing to assist veterans, which Donald Trump says, but real news claims that Donald Trump opposes marriage equality.

Bigrams in Real and False News

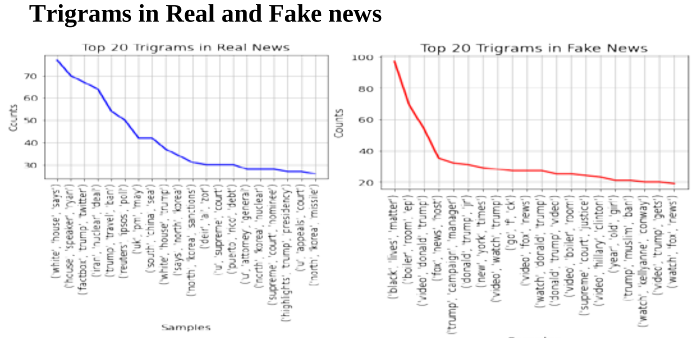


(a) (b)

Figure 8. (a) Top 20 Bigrams in Real News, (b) Top 20 Bigrams in False News

Figures 8(a) and 8(b) demonstrate the dataset's top 20 most frequently utilized unigrams in false and real news. The following process obtains these bigrams. Initially, hashing will be used at the beginning and the ending of each sentence in the dataset. Then the word is broken down into bigrams, and the semantic feature is extracted by novel C-DSSM, shown in figure 8(a), (b).

Trigrams in Real and Fake news



(a) (b)

Figure 9. (a) Top 20 Trigrams in Real News, (b) Top 20 Trigrams in Fake News

Figure 9 illustrates the top 20 trigrams in real and false news. Also, this study states that boosting the n-gram size diminishes the accuracy when applying all of the classifiers. Our research obtains improved accuracy with more feature values. Given that both fake and real news evaluations contained similar words, C-DSSM works better in extracting features. According to our findings, fake news contains more filler/functions and content keywords than real statements. Moreover, they utilize more verbs as well as adverbs than real news. Real news, on the other hand, use more nouns and adjectives.

Classification of Real and Fake News

Predicted values are labelled Positive and Negative in our research, whereas actual values are labelled Real and Fake. In figure 10 True positive value is 4812, the True Negative value is 5869, the False Positive value is 437, and the False Negative value is 421. This confusion matrix is essential for determining performance metrics like Recall, precision, specificity, and accuracy.

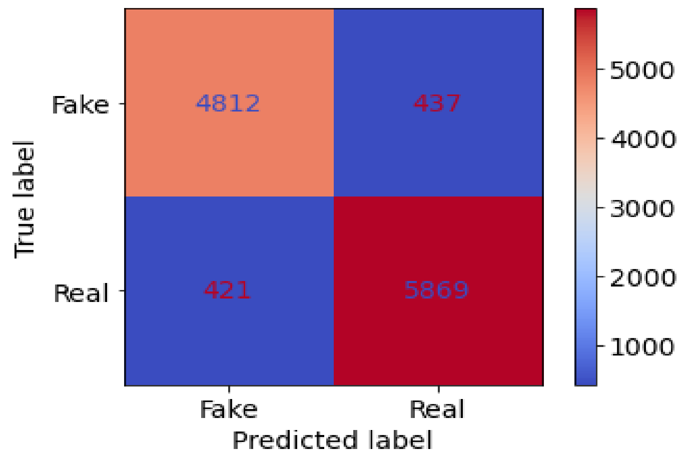


Figure 10. Classification of Real as well as Fake News

Figure 10 classifies the LIAR dataset's real and false news. On the main diagonal, a confusion matrix shows the correctly classified news (top left to bottom right). The incorrect labels are displayed in the other cells, referred to as true negatives or false negatives. Consequently, our proposed hybrid model gives better results in categorizing fake news on social networks.

Evaluation Metrics

This section details the performance of our proposed technique, wherein numerous criteria such as accuracy, F1 score, precision, and Recall have been utilized to assess the usefulness of the unique technique in identifying and categorizing false news. Several performance metrics, including accuracy, precision, Recall, as well as F1 score, were assessed utilizing the following formulas:

$$\text{Accuracy} = \frac{\text{True Positive}+\text{True Negative}}{\text{True Positive}+\text{True Negative}+\text{False Positive}+\text{False Negative}} \tag{2}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive}+\text{False Positive}} \tag{3}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive}+\text{False Negative}} \tag{4}$$

$$\text{F1 Score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \tag{5}$$

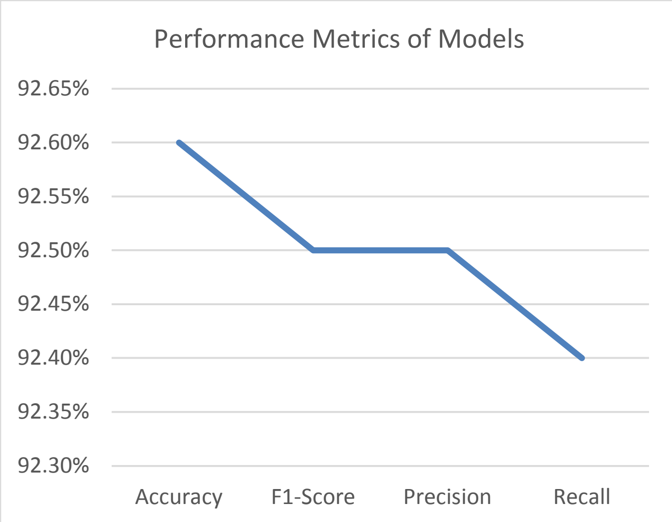


Figure 11. Performance metrics of models

Figure 11 illustrates the performance evaluation metrics of the proposed method. The obtained values of accuracy, F1 Score, Precision, and Recall are 92.60%, 92.50%, 92.50%, and 92.40%. The performance of our proposed attains higher accuracy, F1 score, precision, and Recall by utilizing novel C-DSSM and improved DCNN.

Comparison Metrics

This section deals with the comparison of the existing techniques in which the proposed technique is compared to baseline approaches such as Bi-directional Long Short Term Memory – Convolutional Neural Network (Bi-LSTM-CNN),[30] Hybrid Convolutional Neural Network – Recurrent Neural Network (CNN-RNN),[36] Bi-directional Long Short Term Memory (Bi-LSTM)], Capsule Neural Network (CapsNet) and Deep learning (Ensemble learning).[39]

Table 1. Accuracy Comparison

Methods	Accuracy (%)
CNN-RNN [36]	50
Bi-LSTM [39]	60
CapsNet [39]	64.9
Bi-LSTM-CNN [30]	86.12
Ensemble Learning [39]	89.8
Proposed	92.60

The overall accuracy comparison is shown in Figure 12. The accuracy of the proposed technique improves by using DCNN-based C-DSSM. Our proposed approach attains higher accuracy when compared to the baseline as Hybrid Convolutional Neural Network – Recurrent Neural Network (CNN-RNN).[36] Bi-directional Long Short-Term Memory (Bi-LSTM) [39], Capsule Neural Network (CapsNet),[39] Bi-directional Long Short-Term Memory – Convolutional Neural Network (Bi-LSTM-CNN)[30] and Ensemble learning[39] such as 50%, 60%, 64.9%, 86.12%, and 89.8%. As a result, our unique novel technique has a 92.60% accuracy, which is higher than baseline approaches.

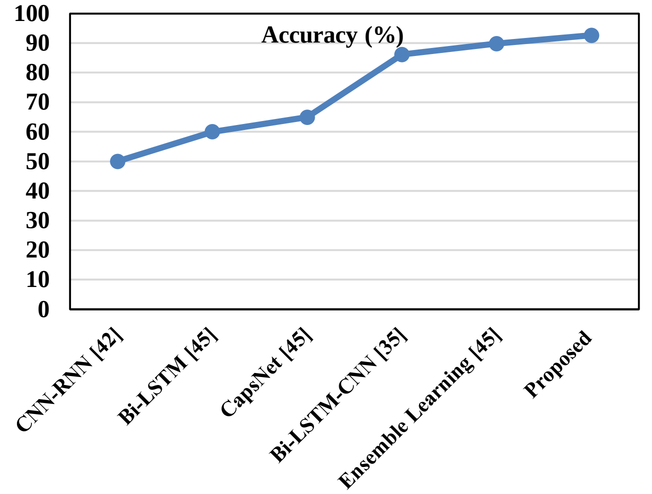


Figure 12. Comparison of methods for fake news detection accuracy

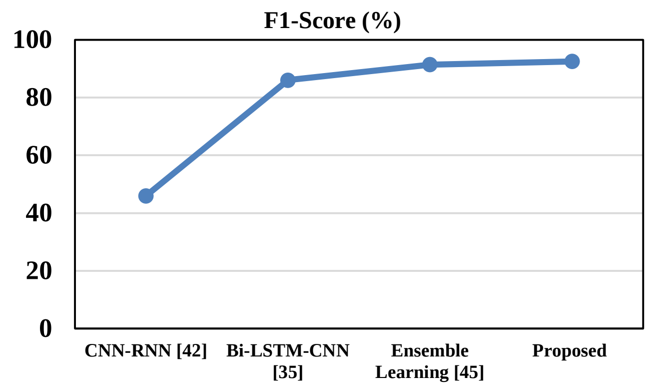


Figure 13. Comparison of methods for fake news detection F1-score

The overall F1-Score comparison is shown in Figure 13. The F1-Score of the proposed technique improves by using DCNN-based C-DSSM. Our proposed approach attains a higher F1-Score when compared to the baseline as Hybrid Convolutional Neural Network – Recurrent Neural Network (CNN-RNN),[36] Bi-directional Long Short-Term Memory – Convolutional Neural Network (Bi-LSTM-CNN)[30] and Bi-directional Long Short-Term Memory – Gated Recurrent Unit [39] such as 46%, 86%, and 91.4%. As a result, our unique novel technique has a 92.50% F1-Score, higher than baseline approaches.

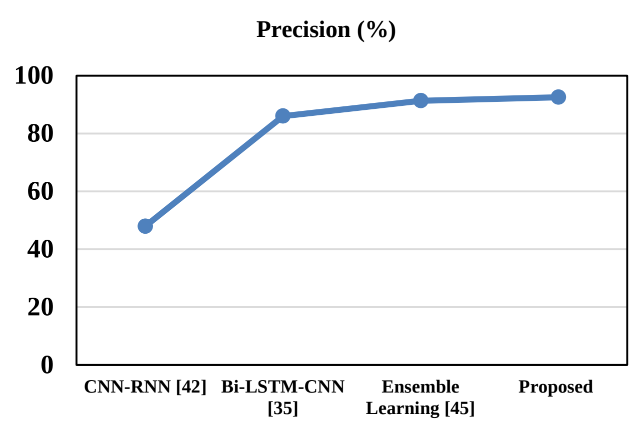


Figure 14. Comparison of methods for fake news detection Precision

The overall precision comparison is shown in Figure 14. The precision of the proposed technique improves by using DCNN-based C-DSSM. Our proposed approach attains a higher precision when compared to the baseline as Hybrid Convolutional Neural Network – Recurrent Neural Network (CNN-RNN)[36], Bi-directional Long Short-Term Memory – Convolutional Neural Network (Bi-LSTM-CNN)[30], and Bi-directional Long Short-Term Memory – Gated Recurrent Unit[39] such as 48%, 86%, and 91.3%. As a result, our unique novel technique has a 92.50% precision, which is higher than baseline approaches.

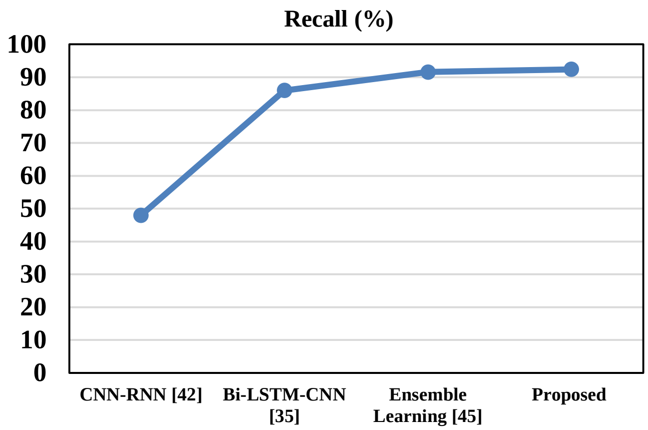


Figure 15. Comparison of methods for fake news detection Recall

The overall recall comparison is shown in Figure 15. The Recall of the proposed technique improves by using DCNN-based C-DSSM. Our proposed approach attains a higher precision when compared to the baseline as Hybrid Convolutional Neural Network – Recurrent Neural Network (CNN-RNN),[36] Bi-directional Long Short-Term Memory – Convolutional Neural Network (Bi-LSTM-CNN),[30] and Bi-directional Long Short-Term Memory – Gated Recurrent Unit[39] such as 48%, 86%, and 91.6%. As a result, our unique novel technique has a 92.40% recall, which is higher than baseline approaches.

Conclusion

Many individuals acquire daily news from social networking rather than conventional media. Furthermore, fake news has been disseminated via social media, with harmful implications for individuals and society. In this research, a detailed description of a novel C-DSSM, as well as an enhanced D-CNN model for detecting fake news, is described. The results indicate the accuracy of the hybrid C-DSSM and D-CNN performance is better than the other classifiers. Furthermore, our research contrasts the proposed methodology with the standard tactic since the proposed approach offers the finest depiction and boosts system performance. Moreover, our research compares the proposed approach to the baseline approach since the proposed approach provides the finest demonstration and enhances the system's performance. Our proposed research yielded remarkable results, with an accuracy of 92.60 percent while employing the LIAR dataset, which is exceptionally effective in identifying fake news. As a result, the proposed work is extremely desirable in detecting fake news and increasing accuracy.

Reference

1. Y. Liu, T. Bakici. Enterprise social media usage: The motives and the moderating role of public social media experience. *Computers in Human Behavior* **2019**, 101, 163–172.
2. M. Pivec, A. Maček. Employment background influence on social media usage in the field of European project management and communication. *J. Business Res.* **2019**, 94, 280–289.
3. S. Vosoughi, D. Roy, S. Aral. The spread of true and false news online. *Science* **2018**, 359(6380), 1146–1151.
4. H. Allcott, M. Gentzkow. Social Media and Fake News in the 2016 Election. *J. Economic Perspectives* **2017**, 31(2), 211–236.
5. A. Smith, V. Banic. Fake News: How a partying Macedonian teen earns thousands publishing lies. In: NBC News; Accessed: March 20, 2022.
6. H. Li, G. Fei, S. Wang, B. Liu, W. Shao, A. Mukherjee, J. Shao. Bimodal distribution and co-bursting in review spam detection. In: Proceedings of the 26th International Conference on World Wide Web **2017**.
7. V. Sandulescu, M. Ester. Detecting singleton review spammers using semantic similarity. In: Proceedings of the 24th international conference on World Wide Web(ACM) **2015**.
8. Amy B. Wang. Post-truth 'named 2016 word of the year by Oxford Dictionaries. Washington Post **2016**.
9. S.I. Manzoor. Fake News Detection Using Machine Learning approaches A Systematic Review. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), ICOEI **2019**, 230–234.
10. A. Bondielli, F. Marcelloni. A survey on fake news and rumour detection techniques. *Inform. Sci.* **2019**, 497, 38–55.
11. F. Cardoso, A. Cristina & B. Garcia. Can Machines Learn to Detect Fake News? A Survey Focused on Social Media. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, **2019**, 6, 2763–2770.
12. 1. K. Bhosle, B. Ahirwadkar. Deep learning Convolutional Neural Network (CNN) for Cotton, Mulberry and Sugarcane Classification using Hyperspectral Remote Sensing Data. *J. Integr. Sci. Technol.* **2021**, 9 (2), 70–74.
13. R. Natali, S. Seo, 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**, 797-806.
14. K. Hamid, P. Roy, S. Saba-Sadiya, J. Tang. Multi-source multi-class fake news detection. In Proceedings of the 27th international conference on computational linguistics **2018**, 1546-1557.
15. D. Nikos, T. Huu, D.M. Nguyen, X. Luo. Deep learning for geolocating social media users and detecting fake news. In NATO Workshop **2018**.
16. K. Shu, S. Wang, H. Liu. Understanding user profiles on social media for fake news detection. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), IEEE **2018**, 430-435.
17. L. Wu, Y. Rao, H. Yu, Y. Wang, A. Nazir. False information detection on social media via a hybrid deep model. In International Conference on Social Informatics, Springer, Cham. **2018**, 323-333.
18. M. Dong, L. Yao, X. Wang, B. Benatallah, Q.Z. Sheng, H. Huang. Dual: A deep unified attention model with latent relation representations for fake news detection. In International conference on web information systems engineering, Springer, Cham. **2018**, 199-209.
19. Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, J. Gao. Eann: Event adversarial neural networks for multi-modal fake news detection. In Proceedings of the 24th acmsigkdd international conference on knowledge discovery & data mining **2018**, 849-857.
20. N. Nyow, H.N. Chua. Detecting fake news with tweets' properties. In 2019 IEEE Conference on Application, Information and Network Security (AINS), IEEE **2019**, 24-29.
21. G. Gravanis, A. Vakali, K. Diamantaras, P. Karadais. Behind the cues: A benchmarking study for fake news detection. *Expert Systems with Applications* **2019**, 128, 201-213.
22. S.S. Jadhav, S.D. Thepade. Fake news identification and classification using DSSM and improved recurrent neural network classifier. *Applied Artificial Intelligence* **2019**, 33(12), 1058-1068.
23. C. Zhang, A. Gupta, C. Kauten, A.V. Deokar, X. Qin. Detecting fake news for reducing misinformation risks using analytics approaches. *Eur. J. Operational Res.* **2019**, 279(3), 1036-1052.

24. A. Qayyum, J. Qadir, M.U. Janjua, F. Sher. Using blockchain to rein in the new post-truth world and check the spread of fake news. *IT Professional* **2019**, 21(4), 16-24.
25. L. Wang, Y. Wang, G. Melo, G. Weikum. Understanding archetypes of fake news via fine-grained classification. *Social Network Analysis and Mining* **2019**, 9(1), 1-17.
26. I. Ahmad, M. Yousaf, S. Yousaf, M.O. Ahmad. Fake news detection using machine learning ensemble methods. *Complexity* 2020 **2020**.
27. A. Kesarwani, S.S. Chauhan, A.R. Nair. Fake News Detection on Social Media using K-Nearest Neighbor Classifier. In 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), IEEE **2020**, 1-4.
28. N. Smitha, R. Bharath. Performance Comparison of Machine Learning Classifiers for Fake News Detection. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE **2020**, 696-700.
29. F. Ozbay, B. Alatas. Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics Appl.* **2020**, 540, 123174.
30. M.Z. Asghar, A. Habib, A. Habib, A. Khan, R. Ali, A. Khattak. Exploring deep neural networks for rumor detection. *J. Ambient Intelligence Humanized Computing* **2021**, 12(4), 4315-4333.
31. Q. Li, Q. Hu, Y. Lu, Y. Yang, J. Cheng. Multi-level word features based on CNN for fake news detection in cultural communication. *Personal and Ubiquitous Computing* **2020**, 24(2), 259-272.
32. Y.J. Lu, C.T. Li. GCAN: Graph-aware co-attention networks for explainable fake news detection on social media. arXiv preprint arXiv:2004.11648 **2020**.
33. S. Kaur, P. Kumar, P. Kumaraguru. Automating fake news detection system using multi-level voting model. *Soft Computing* **2020**, 24(12), 9049-9069.
34. T. Jiang, J.P. Li, A. UlHaq, A. Saboor, A. Ali. A Novel Stacking Approach for Accurate Detection of Fake News. IEEE Access **2021**, 9, 22626-22639.
35. R.K. Kaliyar, A. Goswami, P. Narang. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimedia Tools and Applications* **2021**, 80(8), 11765-11788.
36. J.A. Nasir, O.S. Khan, I. Varlamis. Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights* **2021**, 1(1), 100007.
37. C. Buntain and J. Golbeck. Automatically identifying fake news in popular Twitter threads. 2017 IEEE International Conference on Smart Cloud, SmartCloud 2017, New York City, NY, USA, November 3-5, **2017**.
38. W.Y. Wang. liar, liar pants on fire: A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648 **2017**.
39. N. Aslam, I.U. Khan, F.S. Alotaibi, L.A. Aldaej, A.K. Aldubaikil. Fake detect: A deep learning ensemble model for fake news detection. *Complexity* **2021**, 2021.