

# Mc-DNN: Fake News Detection Using Multi- Channel Deep Neural Networks

Jitendra Vikram Tembhurne, Indian Institute of Information Technology, Nagpur, India\*

Md. Moin Almin, Tezpur University, Tezpur, India & Zaloni Technologies India Pvt. Ltd., Assam, India

Tausif Diwan, Indian Institute of Information Technology, Nagpur, India

## ABSTRACT

With the advancement of technology, social media has become a major source of digital news due to its global exposure. This has led to an increase in spreading fake news and misinformation online. Humans cannot differentiate fake news from real news because they can be easily influenced. A lot of research work has been conducted for detecting fake news using artificial intelligence and machine learning. A large number of deep learning models and their architectural variants have been investigated, and many websites are utilizing these models directly or indirectly to detect fake news. However, state-of-the-arts demonstrate the limited accuracy in distinguishing fake news from the original news. The authors propose a multi-channel deep learning model, namely Mc-DNN, leveraging and processing the news headlines and news articles along different channels for differentiating fake or real news. They achieve the highest accuracy of 99.23% on ISOT Fake News Dataset and 94.68% on Fake News Data for Mc-DNN. Thus, they highly recommend the use of Mc-DNN for fake news detection.

## KEYWORDS

Convolutional Neural Networks, Deep Neural Networks, Ensemble Architectures, Fake News Detection, Multi-Channel Model

## INTRODUCTION

As of 2019, 86.6% of the world population in the developed world has an access to the internet (Statista, 2019) and a person spends an average time of 2:22 (hh:mm) daily on various social media accounts (Review42, 2016). Due to the ease of internet access, it's natural for people to search and follow news from various social media platforms instead of traditional news media such as newspapers and televisions. The main reasons behind this spike in online news reading are the ease of availability of timely news, nice presentation, better recommendations, and less expensive related to the traditional means. Moreover, social media allows its users to share, comment, and discuss the news with an individual, followers, groups, or in public domain. For example, around 62% news from social media read by the adults of U.S. in 2016, a considerable enhancement from 49% as compared to 2012 (Gottfried & Shearer, 2016). Nevertheless, social media news are outperforming the news appearing on the television (BBC News, 2020). However, like the two sides of a coin, there is a major flaw in the availability and authenticity of news posted on social media. The quality of the news articles

DOI: 10.4018/IJSWIS.295553

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

posted on these online portals is relatively lower than those published by reputed news agencies and their online prints. Moreover, the credibility of the news is little bit uncertain that might produce the possibilities of spreading fake news in the community.

Fake news is the spread of disinformation through news channels, print media or online platforms i.e. social media, mostly fake news are deliberated to misguide the people in large. Due to the digital platforms, we witness the rapid increase in fake news. The dissemination of fake news is often targeted to damage the image of a person, group, or community. Furthermore, these news are nicely fabricated to catch the headlines by reaching more people. Moreover, advertising agencies earn revenue from such fake news posted as a headlines and clickbait stories in online (Wikipedia: Fake News, 2020). Nearly, half of the readers report that they see fake news stories on their social media at least once a day. Our automated model shall help individuals, industries and firms to detect and identify the authenticity of the news. There is a huge impact on the shares of the companies on the basis of these fraudulent news. As we incorporate both news headlines and news articles to detect the authenticity, higher confidence is provided to its stakeholders in terms of fake news detection accuracy.

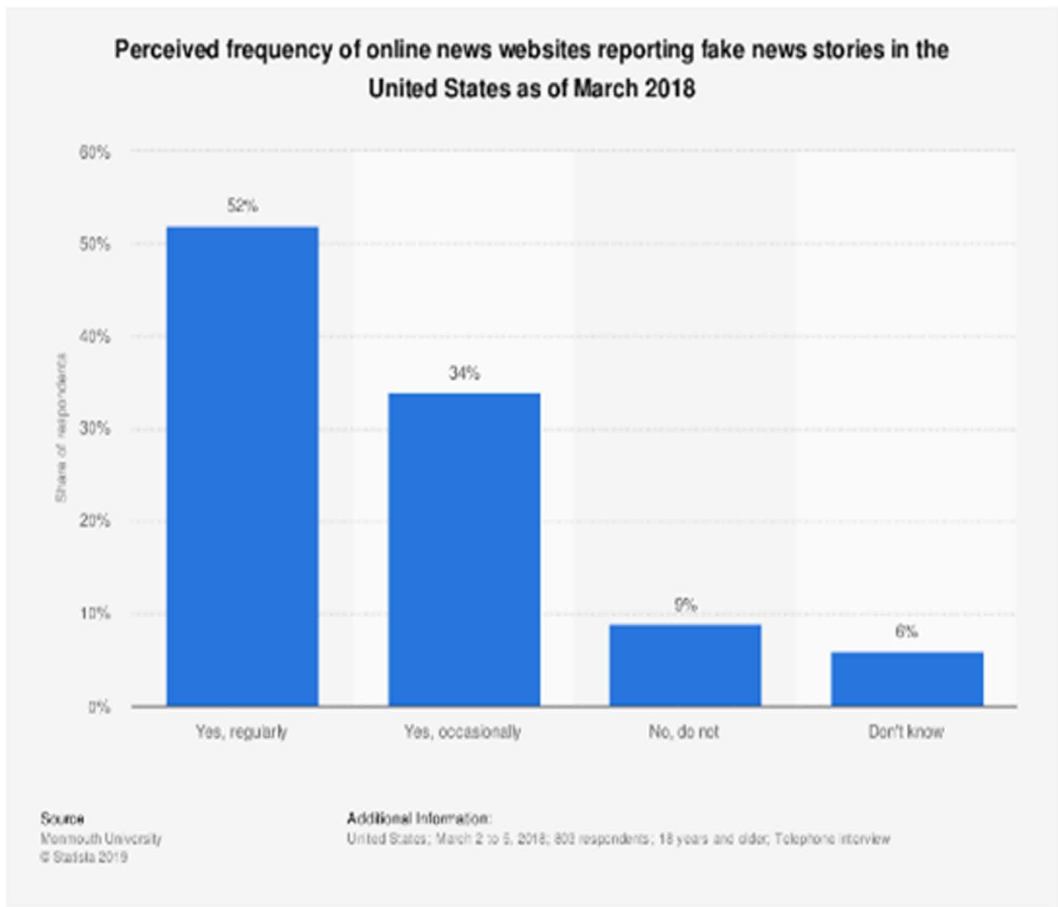
In Figure 1 (a), we can see the perceived frequency of online news websites reporting fake news stories in the U.S. (March 2018). In U.S., fake news was reported by the news websites is about 52%, fake news was occasionally reported by the news website is about 34%, and only 9% confirmed that no fake news being reported online (Statista, 2020). A Survey conducted by the American Trends Panel in March-April 2014, shows that Facebook was the highest communicator of fake news (Pew Research Center, 2015). The result of the survey is shown in Figure 1 (b). Overall, the spread of disinformation could incite political violence, sabotage elections, and unsettle diplomatic relations leading to the deterioration of conflicts. As people are more inclined to believe what they see and read, fake news is an especially dangerous tool to spread of unverified information - an issue made worse by how quickly and easily social media platforms can be used to facilitate this spread (BIC, 2020).

To mitigate the impact of fake news towards to benefit the society and healthy public news ecosystem, we need a tool or model for automatic detection of fake news on social media and to flag them unauthentic as early as possible. Detecting fake news is a complicated task and might require a lot of information to detect the fallacy in them. For example, the articles' headlines also possess a lot of fake information and can also be used to detect fake news. Researchers in the past have not investigated this correlation between article headlines and the article's contents. The research conducted in this paper solve the problem of fake news detection by experimenting with several deep learning models on the news headlines and the fake or real news data, to achieve an effective solution that can flag fake news quickly and accurately.

The challenges in the fake news detection comprise the lack of accuracy reported by the various models which is based on the inadequate context-specific data relevant to the news, selection of effective features in fake news classification, mismatching between the news headlines and news articles, and deciding the role of a deep model in efficient features' extraction in the process of fake news detection. Traditional Machine Learning (ML) algorithms rely on handcrafted features extraction and selection, however, deep learning models especially convolutional neural network (CNN) is designed for automatic features extraction. Moreover, features' extraction using deep learning models is much reliable and robust, suited for the problem under consideration i.e., fake news detection.

Our experimentations on the news articles are based on the understanding that the relation between different tokens present in the article. Each model extracts a different information and identifies the correlation between article headlines and the article's contents, and then predict the authenticity of the news. Previously, researchers have explored either the news article or the news headlines separately for the possible detection of its authenticity using various ML/DL models, as reflected in related work section. It can be demonstrated that utilization of both of these i.e., headlines and article' contents certainly improve the performance of detection of news authenticity. To the best of our knowledge, this is the first work that incorporates headlines and article' contents both on separate channels using latest deep models for the news authenticity detection. We propose the novel deep

Figure 1a. Fake news reported by news website in U.S. (March 2018)



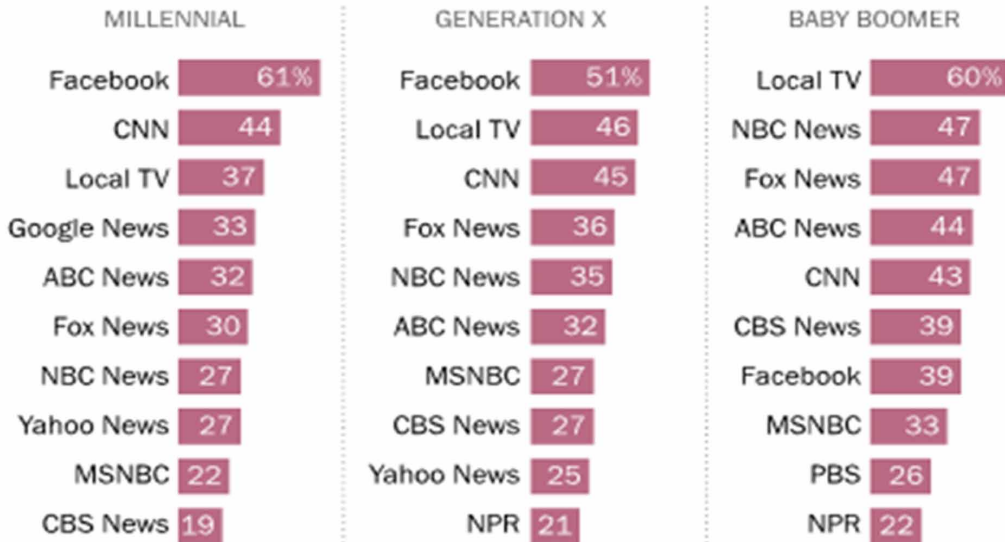
learning model which is the ensemble of different models and also propose the optimized utilization of deep architecture or layers. These layers provide an insight that how the words of a news article are interrelated and it can help in deciding the authenticity of a news article.

In this research, we developed a multi-channel model (Mc-DNN) to detect the fake news with the help of deep convolutional layers without hand-crafted feature engineering for feature extraction. Mc-DNN learns most promising features automatically by the deep network at various layers. We have been motivated from the literature – optimum use of deep layers in model design (Kaliyar *et al.*, 2020), use of features' extracted in sentence level for sentiment analysis (SA) (Fu *et al.*, 2017), adoption of content in SA (Pan *et al.*, 2018), and to develop model based on ensemble to achieve more accuracy (Reddy *et al.*, 2020), to design the new model for fake news detection. In addition, proposed model outperforms the existing models on the large datasets of fake news. The popularity of CNNs in text classification, has found its applicability in the industry (Roy *et al.*, 2018; Zhou & Zafarani, 2018), thus, makes it more suitable to be used in Mc-DNN. Therefore, we suggest the application of Mc-DNN model in insurance, retails and business to analyze the news. We also demonstrate the optimal utilization of CNNs for the detection of fake news, the result analysis shows that the Mc-DNN outperforms the state-of-the-art ML / deep learning (DL) techniques for fake news detection. The accuracy achieved by the Mc-DNN is 99.23% which is significantly higher in comparison with state-of-the-arts, hence we suggest the proposed model for accurate detection of fake news.

Figure 1b. % of people received news about politics and government from different sources.

## Among Millennials, Facebook Far Exceeds Any Other Source for Political News

*% who got news about politics and government in the previous week from...*



American Trends Panel (wave 1). Survey conducted March 19-April 29, 2014. Q22, Q24A.  
Based on online adults.

**PEW RESEARCH CENTER**

The major contributions in our research approach are threefold and presented as follows:

1. To propose a novel deep learning model (Mc-DNN) utilizing the efficient features' extraction from news headlines and news articles separately for fake news detection with the higher accuracy.
2. To investigate the role and applicability of different models along these two channels for robust features' extraction.
3. To present the performance of the proposed model against state-of-the-arts deep models for fake news detection.

The remainder of the paper is structured as follows. In Section 2, we present the literature survey based on fake news detection. Section 3 highlights the methodology adopted and datasets used. The results and discussions are demonstrated in Section 4. Finally, we conclude the paper in Section 5.

### RELATED WORKS

Fake news spreading is one of the important problems that can be solved using DL, and researchers have spent a considerable amount of time trying to find an efficient solution to this epidemic. In (Mahabub, 2020), an ensemble voting classifier based on an intelligent detection system to deal with news classification is proposed. Authors applied various ML algorithms for fake news detection

such as Naïve Bayes, k-nearest neighbors (k-NN), support vector machine (SVM), random forest (RF), ANN, logistic regression, gradient boosting, Ada boosting, etc. In addition, best three ML algorithms were utilized in ensemble voting classifier after cross-validation and maximum accuracy of 94.5% is reported. Further, the two effective deep learning models for solving fake news detection problem in online news contents of multiple domains is proposed by (Saikh, 2020). The systems yield encouraging performance by outperforming the current handcrafted feature engineering-based systems with a significant margin of 3.08% and 9.3% by the two models, respectively on FakeNews AMT and Celebrity datasets. Borges *et al.* (2019) developed a ML-based model for the detection of stance sentences consist of news headline such as discusses, unrelated, agrees, and disagrees. A deep model with string similarity features were adopted to detect stance sentences. The model featuring – representation of text efficiently, document classification, and inference of natural languages. Specifically, neural attention and bi-directional recurrent neural networks (BiRNNs) are employed on temporal dimension with max-pooling layer. Furthermore, supervised ML is employed to offer the solution for aspect-based SA for reviews on Arabic Hotels' (Al-Smadi *et al.*, 2018). The deep RNN and SVM are adopted to process the various features such as syntactic, semantic, lexical, word, and morphological obtained from the reviews.

Semi supervised learning is useful when less amount of labelled dataset and large amount of unlabelled dataset is available for training and evaluation of the model, an investigation for detection of the fake news is presented by (Mansouri, 2020). Subsequently, (Dong *et al.*, 2020) developed a novel framework of two-path deep semi-supervised learning - one path is for supervised learning and the other for unsupervised learning. The supervised learning path learns on the limited amount of labelled data whereas the unsupervised learning path can learn on a huge amount of unlabelled data. Furthermore, these two paths implemented with CNNs were jointly optimized to complete semi-supervised learning. Moreover, a shared CNN is built to extract the low-level features from both labelled as well as unlabelled data to feed them into these two paths. To verify this framework, a Word CNN based semi-supervised learning model was implemented and tested on LIAR and PHEME datasets.

Eventually, Vehicular ad-hoc network (VANET) is investigated by (Gaurav *et al.*, 2020) to perform vehicles communication to exchange various messages for efficient travel intended to the passenger. However, sometimes intruder propagates the false information such as traffic jam, accidents, etc., creates adverse environment for the security of vehicle. Thus, an entropy-based technique is developed to detect the false information. Subsequently, multimodal inputs such as text, audio and video are utilized to extract sentiments. The 3D segmentation is presented by (AlZu'bi *et al.*, 2020) using Fuzzy C-Means algorithm for medical images to detect the real or fake between simulated / actual medical data.

A theory-driven model was proposed for fake news detection by (Zhou *et al.*, 2020). This model investigates news content at different levels: syntax-, semantic-, lexicon-, and discourse-level. The news is represented at each level, depends on the theory of forensic psychology and social relevance, supervised ML model is adapted to detect fake news. Authors explored the features to detect fake news such as patterns of fake news, fake news interpretability for feature engineering, relationships amongst fake news, disinformation in fake news, etc. Progressively, (Roy *et al.*, 2018) developed various DL models to detect fake news and categorized into the pre-defined fine-grained classes. Bi-directional Long Short Term Memory (Bi-LSTM) and CNN is used to represent the input and the combined result were fed into a Multi-layer Perceptron (MLP) model and classification of fake news is performed lastly. In 2018, (Ajao *et al.*, 2018) tried to identify fake news on Twitter using hybrid CNN and RNN models. The proposed framework detects and classifies fake news messages from Twitter posts and achieved an accuracy of 82%. This approach intuitively identifies relevant features associated with fake news stories without background knowledge of the domain. Further, many researchers tried to create datasets for training fake news detection models. One of the prominent

dataset is ISOT Fake News Dataset (Ahmed *et al.*, 2018; Ahmed *et al.*, 2017), we utilized this dataset for the experimentation in this paper.

Previous researchers not worked extensively with the ensemble deep learning models for fake news detection, however, we propose new deep learning models with optimized layers and utilize the features extracted from the ensemble models for efficient prediction of real or fake news. Traditional ML methods have been applied to identify fake news online, but suffered from the limited accuracy. Moreover, news headlines and news articles are the two major pillars based on we can decide the authenticity of the news. Unfortunately, to the best of our knowledge, none of the study focused on these two aspects separately.

Moreover, the way we utilized the datasets for training our model is novel in nature. Researchers have either taken only the news headlines or the news article as the input to the ML/DL models for predicting the fallacy in the news articles, however, some of them have combined both the news headlines and articles into one blob of data. This blobbed data is then processed using a single channel in the ensemble approaches. We propose dual channels deep model wherein one channel extracts feature from news headlines whereas other channel processes news articles. The extracted features from these channels are merged and further processed to predict the authenticity of the news.

## METHODS AND MATERIALS

In this paper, we present the novel architecture for the detection of fake news. Our models act as a binary classifier that predicts whether a news article is fake or real. Here, we proposed five multi-channel ensemble models using various deep models.

### Datasets

We used two popular datasets for fake news detection namely Fake News Data (FND) (available on Kaggle) (Fake News Dataset, 2020) and ISOT Fake News Dataset (Ahmed *et al.*, 2018; Ahmed *et al.*, 2017). These datasets contains the well annotated real and fake news, collected from real-world print and online media sources. The authentic articles were extracted by crawling the articles from several news providing websites like *Reuters.com*. The datasets consist of files containing large amounts of collected news articles. The collection of fake news articles are performed from unreliable websites i.e. flagged by PolitiFact and Wikipedia, and equal number of news articles were stored for real and fake classes.

Each news article in the datasets contains the information such as *article title*, *text*, *type*, and *publication date*. The redundant article type and date fields are dropped, however title and news field are retained. In addition, the news articles are shuffled and a flag is associated as label (0 or 1) for real or fake news indicator with each of these article. The dataset is divided into training and testing set in the ratio of 80:20. Both the training and testing set consists of the news headlines and their corresponding news articles.

### Fake News Dataset

In this section, we discuss on the key features of FND developed by UTK Machine Learning Club and utilized for a competition hosted to find whether a news article is fake or not. The dataset consist of following attributes:

1. *id*: unique id for a news article
2. *title*: the title of a news article
3. *author*: author of the news article
4. *text*: the text of the article; could be incomplete
5. *label*: a label that marks the article as potentially unreliable
  - 1: unreliable

- 0: reliable

The dataset comprises the total 26,000 news articles. The training set consists of 10,387 fake news articles and 10,413 real news articles. The testing set consists of 2,339 fake news articles and 2,861 real news articles.

### ISOT Fake News Dataset

The ISOT Fake News Dataset contains variety of articles belongs to different topics and majority of these articles' emphasis on World and political news only. The dataset stored in two CSV files - "True.csv" consist of more than 21,417 real articles and "Fake.csv" contains more than 23,481 fake articles. Each article is consisting an information such as *article title*, *text*, *type*, and *publication date*. The data collected were cleaned and processed, however, the punctuations and mistakes that existed in the fake news were kept as it is. Figure 2 shows the first five rows of the unprocessed data in the training set that contains news articles indicated under "Text" field, moreover, 0 or 1 in "Value" field indicates fake and authenticate news, respectively. However, first column simply indicates the serial number of news article.

Figure 2. Sample news articles from ISOT Fake News Dataset (training set).

	Text	Value
0	The Republican candidates continue to sink to ...	0
1	MOSCOW (Reuters)- Donald Trump and Vladimir P...	1
2	ISLAMABAD (Reuters)- Pakistan s former prime ...	1
3	<a href="https://fedup.wpengine.com/wp-content/uploads/...">https://fedup.wpengine.com/wp-content/uploads/...</a>	0
4	NEW YORK/WASHINGTON (Reuters)- President-elec...	1

### Pre-processing

The news articles and headlines in the dataset consist of redundant information that do not play a vital role in efficient features' extraction. Each article in the dataset is then pre-processed and redundant information is dropped. So, we cleaned the textual part to ensure better generalization. The pre-processing phase involved removing the URLs and all irrelevant characters in the text (i.e., numbers and punctuation). All characters are converted into lowercase and the stop words are removed. We then performed stemming and lemmatization on all the words in the text to impose generality. We remove the words having a length of less than two. This list of tokens is converted back to the string.

We used the Keras Tokenizer class to automate the tokenization of our training data. First, we create the Tokenizer object, providing the maximum number of words to keep in the vocabulary. After tokenization, if tokens in vocabulary does not exists for encoding, then these previously-unseen words would simply be dropped from the vocabulary and mysteriously unaccounted. After creation of Tokenizer, we then fit it on the training data (later utilized to fit the testing data). Next, word index

is formed which is the by-product of the tokenization process, maps words in our vocabulary to their numeric representation i.e., a mapping which will be essential for encoding of sequences.

In pre-processing, all the inputs are padded to make the uniform lengths inputs. In general, pre-padding is better when multiple types of neural networks are combined to perform a unified task (Dwarampudi & Reddy, 2019). Now, to have a fixed sequence length for all the encoded sentences, we set the limit by first finding the longest encoded sequence and use that as maximum sequence length. Once, the length of the longest sequence is identified then we pad with zeroes to all other relatively smaller length encoded sequences. After padding, the sequences are converted from Python lists to NumPy arrays (distributed vectors) i.e. helpful to feed them into proposed Mc-DNN model.

## Classification Models

In DL, a CNNs or ConvNet is a class of deep neural networks, most commonly applied to analyzing visual imagery (Valueva *et al.*, 2020). They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics

RNN is a class of deep neural networks which is basically employed for the feature learning and classification from sequential or time series data. This allows it to exhibit temporal dynamic behaviour in a sequential input. Derived from feedforward neural networks, RNNs can use their internal state (memory) to process variable-length sequences of inputs (Dupond, 2019). This makes them applicable for the various sequential tasks such as unsegmented, connected handwriting recognition or speech recognition. In sequential or temporal input, data at a time-step has relevance over the data of the preceding time-steps. Prediction at any instant is not only determined by the instantaneous input but depends on past history also. In other words, another dimension i.e. temporal ordering is also taken care in all RNN model computations and its successors. This philosophy is the backbone of the RNN computation. RNN is represented by Eq. 1 (output of hidden) and Eq. 2 (output cells);

$$h_t = \tanh(W \times h_{t-1} + U \times x_t + b_s) \quad (1)$$

$$o_t = \sigma(V \times h_t + b_o) \quad (2)$$

LSTM is an architectural variants and immediate successor of RNN (Hochreiter & Schmidhuber, 1997) belongs to sequential deep learning architectures, designed to deal with vanishing gradient problem. Unlike, standard feedforward neural networks, LSTM has feedback connections. A LSTM cell is composed of three gates: an *input* gate, an *output* gate and a *forget* gate. The forget gate processes the input of current timestep and hidden output from the previous cell. Cell state gets manipulated due to the various gates operations in LSTM cell i.e. information is added, retained, or subsidized. Cell state gets modified by taking into account the input gate, forget gate and previous cell state. Output gate is responsible for the generation of the hidden state that shall be utilized in the next LSTM cell. The LSTM cell is represented by Eq. 3-8.

$$f_t = \sigma(W_f \times [x_t, h_{t-1}] + b_f) \quad (3)$$

$$i_t = \sigma(W_i \times [x_t, h_{t-1}] + b_i) \quad (4)$$



$$\tilde{C}_t = \tanh(W_c \times [x_t, h_{t-1}] + b_c) \quad (5)$$

$$C_t = C_{t-1} \times f_t + \tilde{C}_t \times i_t \quad (6)$$

$$o_t = \sigma(W_o \times [x_t, h_{t-1}] + b_o) \quad (7)$$

$$h_t = C_t \times \tanh(o_t) \quad (8)$$

Gated recurrent units (**GRUs**) are a gating mechanism in RNNs, introduced in 2014 by (Cho *et al.*, 2014), architecture for solving the same vanishing gradient problem, with lesser tensor operations as compared to LSTM (Britz, 2015). Similarly, a LSTM cells consistently outperform GRU cells as demonstrated by (Britz *et al.*, 2017). Update gate is responsible for the collective functioning of the *forget* and *input* gate of a LSTM cell whereas *reset* gate determines the amount of the past information to be kept or forgotten. The GRU cell is represented by Eq. 9-12.

$$r_t = \sigma(W_r \times [x_t, h_{t-1}] + b_r) \quad (9)$$

$$z_t = \sigma(W_z \times [x_t, h_{t-1}] + b_z) \quad (10)$$

$$\tilde{h}_t = \tanh(r_t \times [x_t, h_{t-1}] + b_h) \quad (11)$$

$$h_t = z_t \times \tilde{h}_t + (1 - z_t) \times h_{t-1} \quad (12)$$

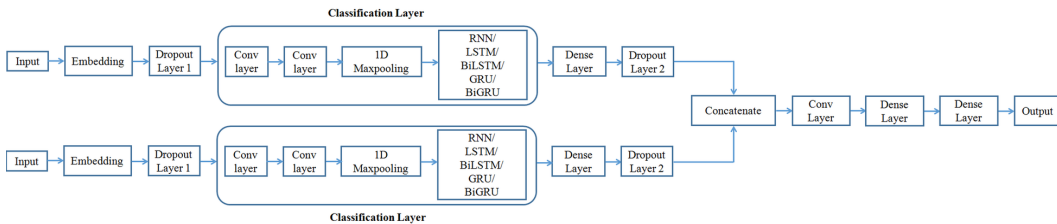
## Multi-channel Model

The most important aspect of the proposed model is its multi-channel concept. A sequential model would run in a straight way and analyzes an article from top to bottom as it passes through the layers progressively. We recommend to analyze the news headlines and the news articles both, to determine a connection between them as to find whether they play a vital role in predicting news as fake or real. In our model, there are two channels; first channel – we input the news headlines of the articles, and second channel – we input the news articles. The model trains along these two channels separately and the output of each channel is merged to predict the cumulative output as fake or real. The architecture of the proposed model ensemble deep model is presented in Figure 3.

Figure 3 shows the model based on the concept of multi-channel model where two neural network models run in parallel to each other on different data - first for news article and second for the news

headlines. News headlines and articles both are first represented by distributed dense vectors using “*Embedding*” blocks. Dropout is introduced on both the channels to achieve generalization and to reduce overfitting. The classification layer consists of convolutional layers followed by one of the sequence model i.e. RNN / LSTM / BiLSTM / GRU / BiGRU. The extracted features on both of these channels are fed to the dense or fully connected layers. Further, these extracted features are concatenated, and processed for the final classification.

Figure 3. Architecture of the proposed multi-channel model.



The model consists of an embedding layer as input, followed by the combination of CNNs and sequence models (i.e., RNN, LSTM, GRU, BiLSTM, BiGRU), and finally the prediction is performed by the output layer. All the layers of proposed architecture are discussed as follows:

**Input layer:** The first layer of our model is the embedding layer which takes input to the model. Embedding layer represents words and documents using a weighted dense vector representation i.e., it converts sentences into the distributed vectors. There are many pretrained models available for word embedding such as Glove (Pennington *et al.*, 2014) and Word2Vec (Mikolov *et al.*, 2013) but we decided to train our models with the traditional embedding layer which is purely created from the words present in the news articles. Glove and Word2Vec takes more time for training hence we have chosen traditional embedding due to large dataset. Our dictionary of words is very large owing to the huge datasets that we utilized for training proposed models.

**Dropout layer 1:** To improve the efficiency of our model and to reduce the chances of overfitting with smaller datasets, we introduced a dropout layer just after the embedding layer. The dropout of 20% is proposed along both the channels. Deep learning models tend to overfit with small datasets and produce less accurate results for the unseen data. This can be prevented by using dropout layers that randomly drop some neurons from the layer on which dropout is applied. These output layers known as nodes, creates a scenario where our models trains with different nodes in every epoch, reducing the chances of overfitting.

**Classification layer:** For the purpose of finding the best model for the task of detecting fake news, we experimented several combinations of CNN with RNNs models. The CNN takes a news article in the form of a distributed vector, assign weights and biases to each element of the vector. CNNs have the ability to learn the various filters of news articles and headlines and perform extremely well. Moreover, LSTM is proven to be highly effective in analyzing sequences that can help in investigating the truthfulness of news article. However, it sometimes fails to capture semantic of words in different contexts. BiLSTM can analyze the context in both forward and reverse direction and provide useful insights from it. Hence, we have combined CNN and BiLSTM which in turn produced benchmarking results. Subsequently, a one dimensional feature map is produced by the ensemble. Later, max-pooling layer performs a pooling operation and captures the dominating features in each output layer of the feature maps. The max-pooling layer reduces the dimensionality of the feature maps, hence reducing the number of parameters, resulting in reduction of the model complexity. The CNN and BiLSTM

generates precisely positioned feature maps. The max-pooling layers summarizes these extracted features effectively, which makes our model more robust.

**Dropout layer 2:** We used another dropout layer to remove any leftovers of overfitting and to regularize the pattern.

**Concatenation layer:** The concatenation layer adds the output of two multi-channel models running side-by-side. The output from the two channels is concatenated into a single vector and processed by Dense layers and an output layer. The result obtained here is directed to predict the output from the combined input data after processed from various layers.

**Dense layer:** The dense layer is a bunch of neural network layers that connects a node to all the nodes in the previous layer.

**Output layer:** The output layer is the final layer of the model that is a  $1 \times 1$  matrix with the value 0 or 1 representing fake or real news respectively.

Figure 4 shows the multi-channel model consist of two parallel layers of size  $1 \times 250$  which demonstrates the change in the tensor shape after applying “Layer (type)”. The two parallel models train simultaneously on the news articles and news headlines, respectively. The input layer is the embedding layer that takes input of maximum length of 250, after padding or truncating the data. The input dimension is the length of the vocabulary created and the output is of the shape of  $1 \times 250 \times 50$ . It is followed by a 20% dropout that do not affect the size of the vector. Next, the convolutional layer with kernel size 8, that gives tensors of dimension  $1 \times 243 \times 250$ , 250 filters are applied in these operations. Both, the convolutional layers have kernel size as 8 and valid padding. It is followed by a max-pooling layer (size and stride) that reduces the size of the tensor to  $1 \times 121 \times 250$  i.e., halves the input size. BiLSTM then reduces the dimension of the tensor and makes it to 1-D array of length 128. The dense layer adds the same number of neurons, resulting in an output of 1-D array of length 256. Concatenating the outputs of length  $1 \times 256$  from two parallel channel, we get the resultant output as a vector of length 512. The convolutional layer is applied to capture the combined features after concatenation. The first dense layer with “*relu*” activation reduces the output to only 10 in number and the final dense layer with “*sigmoid*” activation produces a  $1 \times 1$  vector with 1 or 0 values indicating the news is either true or false.

## RESULTS AND DISCUSSIONS

The results for each model is derived from testing data, which consist of several news articles and their corresponding headlines, which belongs to the class of fake or real from the aforementioned two datasets. It is found that the cloud computing (Bhushan & Gupta, 2019) is adopted to perform simulation and experimentation for various applications wherein it offers required privacy and security (Stergiou *et al.*, 2018). We have used Keras library for implementing the deep learning models. The models are trained on the Kaggle cloud server with 13 GB RAM using Tesla P100-PCI-E-16GB. We applied Adam optimizer and loss function as a Logistic loss or Binary cross entropy. The results on the FND and ISOT Fake News dataset using proposed model is discussed in the following sub-sections, respectively.

### Performance Evaluation of Mc-DNN on FND

We trained several models to achieve the highest accuracy in the case of FND (Fake News Dataset, 2020). Table 1 shows the testing performance measures obtained from different ensemble model i.e. CNN and RNN, CNN and GRU, CNN and LSTM, CNN and BiLSTM, and CNN and BiGRU under multi-channel environment for fake or real news detection. The highest accuracy of 94.68% is obtained for CNN and BiLSTM ensemble. The performance is evaluated based on accuracy, precision, recall, and F1-Score.

To evaluate the performance of our model, we compared the works of other researchers who developed the similar models on FND. Moreover, the comparison can be viewed in Table 2.

Figure 4. Multi-channel Ensemble Model (Specifically for CNN and BiLSTM)

Layer (type)	Input size	Output size
Embedding	$1 \times 250$	$1 \times 250 \times 50$
Embedding	$1 \times 250$	$1 \times 250 \times 50$
Dropout layer 1	$1 \times 250 \times 50$	$1 \times 250 \times 50$
Dropout layer 1	$1 \times 250 \times 50$	$1 \times 250 \times 50$
1D Cov	$1 \times 250 \times 50$	$1 \times 243 \times 250$
1D Cov	$1 \times 250 \times 50$	$1 \times 243 \times 250$
1D Cov	$1 \times 250 \times 50$	$1 \times 243 \times 250$
1D Cov	$1 \times 250 \times 50$	$1 \times 243 \times 250$
Maxpooling	$1 \times 243 \times 250$	$1 \times 121 \times 250$
Maxpooling	$1 \times 243 \times 250$	$1 \times 121 \times 250$
BiLSTM	$1 \times 121 \times 250$	$1 \times 128$
BiLSTM	$1 \times 121 \times 250$	$1 \times 128$
Dense layer	$1 \times 128$	$1 \times 256$
Dense layer	$1 \times 128$	$1 \times 256$
Dropout layer 2	$1 \times 256$	$1 \times 256$
Dropout layer 2	$1 \times 256$	$1 \times 256$
Concatenation	$1 \times 256$	$1 \times 512$
1D Cov	$1 \times 512$	$1 \times 512$
Dense layer	$1 \times 512$	$1 \times 10$
Dense layer	$1 \times 10$	$1 \times 1$
Output	$1 \times 1$	$1 \times 1$

### Performance Evaluation of Mc-DNN on ISOT Dataset

We evaluated the performance of MC-DNN models on ISOT Fake News dataset to detect whether the news is fake or real. Initially, the training accuracy is low and remains constant in the later epochs due to large sized dataset. The dropout layers in the model make sure that the model doesn't overfit. We calculate the training loss as an average of the losses over each batch during training data and the loss increases over the last batch of an epoch. Moreover, the testing loss is computed at the end of an epoch and hence the testing loss is more than the training loss.

Table 1. Performance Evaluation of various ensemble models on FND.

Model	Accuracy	Precision	Recall	F1-Score
CNN and RNN	89.06	89	89	89
CNN and GRU	91.63	92	92	92
CNN and LSTM	91.15	91	91	91
CNN and BiGRU	91.45	91	91	91
<b>CNN and BiLSTM</b>	<b>94.68</b>	<b>95</b>	<b>95</b>	<b>95</b>

Table 2. Comparison of Mc-DNN on FND.

Ref.	Accuracy	Precision	Recall	F1-Score
(Ghanem <i>et al.</i> , 2018)	48.80	-	-	59
(Singh <i>et al.</i> , 2017)	87.00	87	87	87
(Ahmed <i>et al.</i> , 2017)	89.00	90	-	-
(Ruchansky <i>et al.</i> , 2017)	89.20	-	-	89
(Ahmad <i>et al.</i> , 2020)	91.00	91	87	89
(Bhattacharjee <i>et al.</i> , 2017)	92.7	-	-	-
(Faustini & Covões, 2020)	94	-	-	94
(Wijeratne, 2021)	94	-	-	-
(Khan <i>et al.</i> , 2021)	87.00	87.00	87.00	87.00
(Khan <i>et al.</i> , 2021)	86.00	86.00	86.00	86.00
<b>CNN and BiLSTM</b>	<b>94.68</b>	<b>95</b>	<b>95</b>	<b>95</b>

To evaluate the performance of the models, we calculate the recall, precision, and F1-Score, this is calculated by comparing the testing set's outputs with ground truths. We found that the recall, precision, and F1-Score for the CNN and BiLSTM multi-channel model is 99%, 99%, and 99% respectively. We also performed the experimentations using BiGRU and BiLSTM in combination with CNN and achieved higher accuracy. Table 3 presents the results based on different models experimented on ISOT Fake News Dataset wherein we can observe the best accuracy achieved by CNN and BiLSTM is 99.23%. Table 4 shows the comparison of different models with the proposed model on ISOT Fake News Dataset, our model outperforms the existing models (Ahmed *et al.*, 2018; Wang, 2017; Ozbay & Alatas, 2020) and achieved an accuracy of 99.23%.

Table 3. Performance evaluation of various ensemble models on ISOT Fake News Dataset.

Model	Accuracy	Precision	Recall	F1-Score
CNN and RNN	91.42	92	91	91
CNN and GRU	98.96	99	99	99
CNN and LSTM	98.29	98	98	98
CNN and BiGRU	99.18	99	99	99
<b>CNN and BiLSTM</b>	<b>99.23</b>	<b>99</b>	<b>99</b>	<b>99</b>

Table 4. Comparison of Mc-DNN on ISOT Fake News Dataset.

Ref.	Accuracy	Precision	Recall	F1-Score
CNN (Wang, 2017)	87.00	84.00	90.00	87.00
BiLSTM (Wang, 2017)	86.00	92.00	78.00	84.00
(Ahmed <i>et al.</i> , 2018)	92.00	-	-	-
(Ozbay & Alatas, 2020)	96.80	96.30	97.30	96.80
(Nasir <i>et al.</i> , 2021)	99.00	99.00	99.00	99.00
(Hansrajh <i>et al.</i> , 2021)	98.48	98.4	98.4	98.4
<b>CNN and BiLSTM</b>	<b>99.23</b>	<b>99.00</b>	<b>99.00</b>	<b>99.00</b>

Table 5 demonstrates the improvements or variations in the results for both the datasets in fake news detection. As we can notice that single channel underperforms in comparison with multichannel architecture. We experimented with different sized “Dropout” module consistently over the entire architecture. Model fails to generalize in the absence of “Dropout” module. We observe a slight improvement in the accuracy while increasing the dropout rate from 10% to 20%. Moreover, we don’t observe any further improvements in increasing the dropout rate after 20%. Single channel operating with news articles performs slightly better than single channel operating with news headlines. The probable reason behind it is the comparatively more informative contents in the articles.

Table 5. Variations or Improvements in Accuracy on FND and ISOT Fake News Dataset.

Architecture	Components	FND (% Accuracy)	ISOT Fake News Dataset (% Accuracy)
Single Channel (Only News Headlines)	No Dropout	86.31	89.13
	10% Dropout	89.1	90.4
	20% Dropout	89.62	91.45
Single Channel (Only News Articles)	No Dropout	89.42	90.25
	10% Dropout	91.85	92.89
	20% Dropout	92.43	93.67
Multichannel (CNN + BiLSTM) with Input Padding	No Dropout	92.29	97.19
	10% Dropout	94.1	98.82
	20% Dropout	94.68	99.23

Table 6 shows the accuracy comparison of proposed models with existing models. Hence, we suggest the use of CNN and BiLSTM multi-channel model for fake news detection. On the downside, due to multi-channel and various layers added, the complexity of model increases and takes more time for training but the accuracy obtained is the highest.

The accuracy of ISOT Fake News Dataset (99.23%) is better than FND (94.68%). ISOT Fake News Dataset consists of twice the number of news articles and headlines as compared to FND. This could be the probable reason behind the benchmarking results in cases of ISOT Fake News Dataset. Moreover, the inherent characteristics of the datasets also play an important role in the model performance. The former dataset grasps the benefits of the proposed model and generalize well too.

The challenge faced while developing the model is lack of diverse set of datasets. Moreover, most of the datasets are revolve around 2016's presidential election. Thus, developing the generalized fake news dataset is the open research problem. Furthermore, the datasets don't usually consist of user information, which is necessary for understanding the user's sentiment while creating fake news. In addition, propagation details were also not available, as the time interval for which the news stays online plays an important role in determining the validity of the news article.

Though, we achieved better results in employing multichannel architecture for fake news detection instead of working either with news headlines or news' articles separately, the downside of the proposed approach is the overhead incurred in terms of model complexity. The number of parameters almost doubled in the introduced multichannel architecture but we also obtain significant gain in the performance, as illustrated in Table 5.

## CONCLUSION AND FUTURE SCOPE

Fake news is a festering wound and keeps growing exponentially. Social networks play a vital role in the propagation of fake news. Without the presence of real news with facts and figures, fake news has the capability of impersonating as a piece of real news and cause havoc. It's important to mark fake news at the earliest on social media sites. We presented a Mc-DNN model for the detection of fake news leverages the articles' headlines and articles' contents on altogether different channels. The Mc-DNN outperform most of the fact-checking websites and deep models created by the various researchers. The performance is analyzed for Mc-DNN using the combination of CNN and RNN, CNN and GRU, CNN and LSTM, CNN and BiGRU, and CNN and BiLSTM. We found that the CNN and BiLSTM Mc-DNN is best for the task of fake news detection and achieve the highest accuracy of 99.23% and 94.68% on ISOT fake news dataset and FND, respectively.

Table 6. Comparison with Fake News Detection Models on Different Datasets.

Ref.	Model	Dataset	Accuracy	Precision	Recall	F1-Score
(Thota <i>et al.</i> , 2018)	TF-IDF + Neural Net	FNC 1 (Fake News Challenge Stage 1, 2018)	94.31	-	-	-
(Bhatt <i>et al.</i> , 2017)	Neural + statistic + external features	FNC 1	83.08	-	-	-
(Borges <i>et al.</i> , 2019)	BiLSTM		82.23	-	-	-
(Riedel <i>et al.</i> , 2017)	TF + TF-IDF + MLP		81.72	-	-	-
(Mohtarami <i>et al.</i> , 2018)	Neural method + TF-IDF Neural method		88.57 87.27	81 78	- -	56 56
(Bhatt <i>et al.</i> , 2017)	Baseline - skip-thought embeddings Baseline - word2vec + hand-crafted features Neural baseline – BiLSTMs		76.18	-	-	-
			72.78	-	-	-
		63.11	-	-	-	
(Mahabub, 2020)	MLP X-Gradient Boosting LR	(Shu <i>et al.</i> , 2017)	93.83	85	85	93
			92.87	86	87	92
			98.21	85	90	93
(Khan <i>et al.</i> , 2021)	HAN Naïve Bayes	FND	87.00	87.00	87.00	87.00
			86.00	86.00	86.00	86.00
(Wijeratne, 2021)	XBG SVM		95.00	-	-	-
			94.00	-	-	-
(Kaliyar <i>et al.</i> , 2020)	FNDNet		98.36	99	96	98
(Ahmad <i>et al.</i> , 2020)	Linear SVM		98.00	98	98	98
(Ahmed <i>et al.</i> , 2018)	Linear SVM	ISOT	92.00	-	-	-
(Ozbay <i>et al.</i> , 2020)	Decision Tree		96.80	96.30	97.30	96.80
(Nasir <i>et al.</i> , 2021)	CNN + RNN		99.00	99.00	99.00	99.00
(Hansrajh <i>et al.</i> , 2021)	Ensemble		98.48	98.4	98.4	98.4
<b>Our</b>	<b>CNN and BiLSTM</b>	<b>FND</b>	<b>94.68</b>	<b>95</b>	<b>95</b>	<b>95</b>
	<b>CNN and GRU</b>	<b>ISOT</b>	<b>98.96</b>	<b>99</b>	<b>99</b>	<b>99</b>
	<b>CNN and LSTM</b>		<b>98.29</b>	<b>98</b>	<b>98</b>	<b>98</b>
	<b>CNN and BiLSTM</b>		<b>99.23</b>	<b>99</b>	<b>99</b>	<b>99</b>
<b>CNN and BiGRU</b>		<b>99.18</b>	<b>99</b>	<b>99</b>	<b>99</b>	



## REFERENCES

- Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. *Complexity*, 2020, 2020. doi:10.1155/2020/8885861
- Ahmed, H., Traore, I., & Saad, S. (2017, October). Detection of online fake news using n-gram analysis and machine learning techniques. In *Proceedings of International conference on intelligent, secure, and dependable systems in distributed and cloud environments* (pp. 127-138). Springer. doi:10.1007/978-3-319-69155-8\_9
- Ahmed, H., Traore, I., & Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Security and Privacy*, 1(1), e9. doi:10.1002/spy2.9
- Ajao, O., Bhowmik, D., & Zargari, S. (2018, July). Fake news identification on twitter with hybrid CNN and RNN models. In *Proceedings of the 9th international conference on social media and society* (pp. 226-230). doi:10.1145/3217804.3217917
- Al-Smadi, M., Qawasmeh, O., Al-Ayyoub, M., Jararweh, Y., & Gupta, B. (2018). Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels' reviews. *Journal of Computational Science*, 27, 386–393. doi:10.1016/j.jocs.2017.11.006
- AlZu'bi, S., Shehab, M., Al-Ayyoub, M., Jararweh, Y., & Gupta, B. (2020). Parallel implementation for 3d medical volume fuzzy segmentation. *Pattern Recognition Letters*, 130, 312–318. doi:10.1016/j.patrec.2018.07.026
- BBC News. (2020). <https://www.bbc.com/news/uk-36528256>
- Bhatt, G., Sharma, A., Sharma, S., Nagpal, A., Raman, B., & Mittal, A. (2017). On the benefit of combining neural, statistical and external features for fake news identification. arXiv preprint arXiv:1712.03935.
- Bhattacharjee, S. D., Talukder, A., & Balantrapu, B. V. (2017, December). Active learning based news veracity detection with feature weighting and deep-shallow fusion. In *Proceedings of IEEE International Conference on Big Data* (pp. 556-565). IEEE. doi:10.1109/BigData.2017.8257971
- Bhushan, K., & Gupta, B. B. (2019). Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment. *Journal of Ambient Intelligence and Humanized Computing*, 10(5), 1985–1997. doi:10.1007/s12652-018-0800-9
- BIC. (2020). <https://www.bic-rhr.com/projects/social-media-and-conflict-weighting-pros-and-cons>
- Borges, L., Martins, B., & Calado, P. (2019). Combining similarity features and deep representation learning for stance detection in the context of checking fake news. *Journal of Data and Information Quality*, 11(3), 1–26. doi:10.1145/3287763
- Britz, D. (2015). *Recurrent neural network tutorial, part 4 implementing a gru/lstm rnn with python and theano*. <http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4>
- Britz, D., Goldie, A., Luong, M. T., & Le, Q. (2017). Massive exploration of neural machine translation architectures. arXiv preprint arXiv:1703.03906. doi:10.18653/v1/D17-1151
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078. doi:10.3115/v1/D14-1179
- Dong, X., Victor, U., & Qian, L. (2020). *Two-path Deep Semi-supervised Learning for Timely Fake News Detection*. arXiv preprint arXiv:2002.00763.
- Dupond, S. (2019). A thorough review on the current advance of neural network structures. *Annual Reviews in Control*, 14, 200–230.
- Dwarampudi, M., & Reddy, N. V. (2019). *Effects of padding on LSTMs and CNNs*. arXiv preprint arXiv:1903.07288.
- Fake News Challenge Stage 1 (FNC 1): Stance Detection. (n.d.). <http://www.fakenewschallenge.org/>
- Fake News Dataset. (2020). <https://www.kaggle.com/c/fake-news/data>

- Faustini, P. H. A., & Covões, T. F. (2020). Fake news detection in multiple platforms and languages. *Expert Systems with Applications*, 158, 113503. doi:10.1016/j.eswa.2020.113503
- Fu, X., Liu, W., Xu, Y., & Cui, L. (2017). Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis. *Neurocomputing*, 241, 18–27. doi:10.1016/j.neucom.2017.01.079
- Gaurav, A., Gupta, B. B., Castiglione, A., Psannis, K., & Choi, C. (2020, December). A Novel Approach for Fake News Detection in Vehicular Ad-Hoc Network (VANET). In *International Conference on Computational Data and Social Networks* (pp. 386-397). Springer. doi:10.1007/978-3-030-66046-8\_32
- Ghanem, B., Rosso, P., & Rangel, F. (2018, November). Stance detection in fake news a combined feature representation. In *Proceedings of the first workshop on fact extraction and VERification (FEVER'18)* (pp. 66-71). doi:10.18653/v1/W18-5510
- GottfriedJ.ShearerE. (2016). <https://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>
- Hansrajh, A., Adeliyi, T. T., & Wing, J. (2021). Detection of Online Fake News Using Blending Ensemble Learning. *Scientific Programming*, 2021, 2021. doi:10.1155/2021/3434458
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735 PMID:9377276
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1), 215–243. doi:10.1113/jphysiol.1968.sp008455 PMID:4966457
- Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2020). FNDNet—a deep convolutional neural network for fake news detection. *Cognitive Systems Research*, 61, 32–44. doi:10.1016/j.cogsys.2019.12.005
- Khan, J. Y., Khondaker, M. T. I., Afroz, S., Uddin, G., & Iqbal, A. (2021). A benchmark study of machine learning models for online fake news detection. *Machine Learning with Applications*, 4, 100032. doi:10.1016/j.mlwa.2021.100032
- Mahabub, A. (2020). A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers. *SN Applied Sciences*, 2(4), 1–9. doi:10.1007/s42452-020-2326-y
- Mansouri, R., Naderan-Tahan, M., & Rashti, M. J. (2020, August). A Semi-supervised Learning Method for Fake News Detection in Social Media. In *Proceedings of 28th Iranian Conference on Electrical Engineering (ICEE'20)* (pp. 1-5). IEEE. doi:10.1109/ICEE50131.2020.9261053
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). *Distributed representations of words and phrases and their compositionality*. arXiv preprint arXiv:1310.4546.
- Mohtarami, M., Baly, R., Glass, J., Nakov, P., Márquez, L., & Moschitti, A. (2018). Automatic stance detection using end-to-end memory networks. arXiv preprint arXiv:1804.07581. doi:10.18653/v1/N18-1070
- Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, 1(1), 100007. doi:10.1016/j.jjimei.2020.100007
- Ozbay, F. A., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A*, 540, 123174. doi:10.1016/j.physa.2019.123174
- Pan, J. Z., Pavlova, S., Li, C., Li, N., Li, Y., & Liu, J. (2018, October). Content based fake news detection using knowledge graphs. In *Proceedings of International semantic web conference* (pp. 669-683). Springer. doi:10.1007/978-3-030-00671-6\_39
- Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on empirical methods in natural language processing (EMNLP'14)* (pp. 1532-1543). doi:10.3115/v1/D14-1162
- Pew Research Center. (2015). <https://www.journalism.org/2015/06/01/millennials-political-news>
- Reddy, H., Raj, N., Gala, M., & Basava, A. (2020). Text-mining-based fake news detection using ensemble methods. *International Journal of Automation and Computing*, 1-12.

- Review42. (2016). <https://review42.com/how-much-time-do-people-spend-on-social-media>
- Riedel, B., Augenstein, I., Spithourakis, G. P., & Riedel, S. (2017). *A simple but tough-to-beat baseline for the Fake News Challenge stance detection task*. arXiv preprint arXiv:1707.03264.
- Roy, A., Basak, K., Ekbal, A., & Bhattacharyya, P. (2018). *A deep ensemble framework for fake news detection and classification*. arXiv preprint arXiv:1811.04670.
- Ruchansky, N., Seo, S., & Liu, Y. (2017, November). CSI: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 797-806). doi:10.1145/3132847.3132877
- Saikh, T., De, A., Ekbal, A., & Bhattacharyya, P. (2020). *A Deep Learning Approach for Automatic Detection of Fake News*. arXiv preprint arXiv:2005.04938.
- Shu, K., Wang, S., & Liu, H. (2017). *Exploiting tri-relationship for fake news detection*. arXiv preprint arXiv:1712.07709.
- Singh, V., Dasgupta, R., Sonagra, D., Raman, K., & Ghosh, I. (2017). Automated fake news detection using linguistic analysis and machine learning. In *Proceedings of International conference on social computing, behavioral-cultural modeling, & prediction and behavior representation in modeling and simulation (SBP-BRiMS'17)* (pp. 1-3). Academic Press.
- Statista. (2019). <https://www.statista.com/statistics/262966/number-of-internet-users-in-selected-countries/>
- Statista. (2020). <https://www.statista.com/statistics/649234/fake-news-exposure-usa/>
- Stergiou, C., Psannis, K. E., Gupta, B. B., & Ishibashi, Y. (2018). Security, privacy & efficiency of sustainable cloud computing for big data & IoT. *Sustainable Computing: Informatics and Systems*, 19, 174–184. doi:10.1016/j.suscom.2018.06.003
- Thota, A., Tilak, P., Ahluwalia, S., & Lohia, N. (2018). Fake news detection: A deep learning approach. *SMU Data Science Review*, 1(3), 10.
- Valueva, M. V., Nagornov, N. N., Lyakhov, P. A., Valuev, G. V., & Chervyakov, N. I. (2020). Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. *Mathematics and Computers in Simulation*, 177, 232–243. doi:10.1016/j.matcom.2020.04.031
- Wang, W. Y. (2017). “Liar, liar pants on fire”: A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648. doi:10.18653/v1/P17-2067
- Wijeratne, Y. (2021). *How Much Bullshit Do We Need? Benchmarking Classical Machine Learning for Fake News Classification*.
- Wikipedia. (2020). *Fake News*. [https://en.wikipedia.org/wiki/Fake\\_news](https://en.wikipedia.org/wiki/Fake_news)
- Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake news early detection: A theory-driven model. *Digital Threats. Research and Practice*, 1(2), 1–25.
- Zhou, X., & Zafarani, R. (2018). *Fake news: A survey of research, detection methods, and opportunities*. arXiv preprint arXiv:1812.00315.

*Jitendra V. Tembhurne received the B. E. degree in Computer Technology from Kavikulguru Institute of Technology and Science (KITS), Ramtek, Nagpur University, Maharashtra, India, in 2003 and the M. E. degree in Computer Science and Engineering from MGM's College of Engineering, SRTMU Nanded, Maharashtra, India, in 2011 and the Ph.D. degree in Computer Science and Engineering from Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, India, in 2017. From 2005 to 2011, he was an Assistant Professor with Computer Technology Department, KITS Ramtek, Nagpur University, Maharashtra, India. From 2016 to 2018 he was an Assistant Professor with Computer Engineering Department, SVPCET, Nagpur, Maharashtra, India. Since 2018, he has been Assistant Professor with the Computer Science and Engineering Department, Indian Institute of Information Technology (IIIT), Nagpur, Maharashtra, India. He is the author of more than 20 articles published in international conferences and journals. His research interests include parallel computing on multi-core and many-core hardware, data science, deep learning, medical imaging, and cryptography and security. He is the member of IAENG Society.*

*Md. Moin Almin received his B. Tech. degree in Computer Science and Engineering from School of Engineering, Tezpur University, Tezpur, Assam, India in 2021. After his graduation, he worked at Zaloni Inc, Assam, India, A Big Data Company for 6 months. He is now working at Wizni, Inc., Haryana, India as a full-time Software Engineer. He is interested in deep learning, web development and integrated technology.*

*Tausif Diwan received his M. Tech. and Ph.D. degree from Department of Computer Science and Engineering, Visvesvaraya National Institute of Technology, Nagpur, India in 2011 and 2017 respectively. Since July 2019, he has been associated with the Department of Computer Science and Engineering, Indian Institute of Information Technology, Nagpur, India as an Assistant Professor. His research area includes parallel computing, Analysis of algorithm & machine learning. He has published several research papers in various international conferences and reputed journals.*