Journal Of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 7
July 2023

Automated Fake News Detection for societal benefit using Hybrid Deep Neural Network

Meer Tauseef Ali[1], Dr. Syed Asadullah Hussaini[2], Dr. S K Yadav[3]

[1]Research Scholar, Department of CSE, Shri JJT University, Rajasthan, India.
[2]Research Co-Supervisor, Department of CSE, Shri JJT University, Rajasthan, India.
[3]Research Supervisor, Department of CSE & Director R&D, Shri JJT University, Rajasthan, India.
meertauseefali@gmail.com[1], drsyedasadullahh@gmail.com[2], drskyadav@hotmail.com[3]

Abstract: In recent years, the emergence of Online Social Networks has led to a profusion of social news such as commercial advertisements, political news, information about celebrities, and many other types of information. Users of social media platforms like Facebook, Instagram, and Twitter have been influenced by false information. The proliferation of false news across several industries and government bodies is a direct result of the rise of social media and online discussion forums. As a result, people are less likely to put their faith in the media. There is a mountain of literature on the topic of Artificial Intelligence (AI) methods for spotting hoaxes. The aforementioned problem is solved by employing a hybrid Neural Network design that incorporates the strengths of both CNN and LSTM. Dimensionality reduction strategies were recommended for use in this study to help make feature vectors more manageable before being sent on to a classifier. To do this, we construct a multi-layered Convolutional Neural Network (CNN). We evaluate the suggested method by contrasting it to many standard templates. The proposed model was trained and tested using state-of-the-art benchmark datasets, with test-data accuracy of 98.36%. The results were verified using a battery of performance assessment metrics including false positives, true negatives, precision, recall, F1, accuracy, and more. These results reveal considerable gains in the field of fake news identification as compared to previous state-of-the-art results and validate the promise of our technique for identifying false news on social media. This study will help academics gain a deeper comprehension of how hybrid CNN-based deep models may be used to spot false news.

Keywords: Fake news detection, Deep learning, Rumor Detection, social media, Neural Network

1. Introduction

The proliferation of online technology has made it possible for anybody, everywhere to obtain any kind of information. While the Internet does make a tremendous quantity of information available, the reliability of that material can be affected by a number of reasons [1]. A vast volume of information is transmitted every day via internet and print media; yet, it is impossible to identify whether the material includes truth or is fraudulent. The material must be studied and analyzed thoroughly, which includes checking the facts by assessing the supporting sources, learning more about the source, establishing the legitimacy of the writers, and so on. False information is spread with the objective to damage or improve the public's perception of a certain group or individual, or for financial or political gain [2]. Such false information, or "fakeness," is harmful because it leads individuals in the wrong direction. Several of these fabricated tales, news articles, and modified photographs proliferated online, for instance, during the Indian election campaigns. We can no longer imagine our everyday lives, activities, or way of life apart from the ubiquitous presence of social media and the profound impact it has on all three [3]. The advent of instant messaging and blogging, among other social media, has revolutionized the global dissemination and consumption of knowledge. These days, the dissemination of news is triggered every second by people sharing and commenting on postings related to news stories. The distribution process is not usually complex, but it is a neuralgic aspect in how a piece of information spreads across a network [4]. The way that people see the news has shifted as a result of a new paradigm that has emerged recently. Fake news is a term used to describe information in the news media that is intentionally misleading or false. It's being thought about for potential political, private, or national benefits. False news detection is a challenging problem that has inspired a lot of study [5].

1.1. Impact and Challenges of Fake News

To determine the scale of this phenomenon, researchers [6] collected data from Twitter over a period of four weeks. Based on the data, it was determined that 16% of the accounts and over 19% of the tweets were created by social bots. The results of a survey asking American voters which news outlet they trusted the most during the 2016 election are shown in Fig. 1.

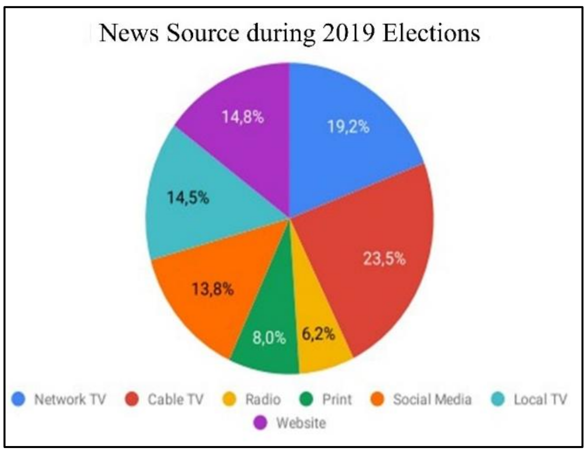


Figure 1. The most credible news outlets for the 2019 elections

The proliferation of false news is another risk associated with the growth of social media. According to forecasts, the number of social media users throughout the world will increase to over 2.77 billion in 2019, up from 2.46 billion in 2017 [7]. In 2017, 71% of all internet users were also active on at least one social network, and that share is only projected to grow. The exponential growth of the user base is seen in Fig 1.1. The exponential growth in population has an effect on the phenomena of Fake news by providing fuel for the unsustainable amount of disseminating incorrect information over the web and social media [8].

Automatically spotting newly-emerging, unchecked false news is a tricky task that requires familiarity with different viewpoints and "common sense." Fake news is also developing unchecked by evil actors, who want it to seem as actual news while having wrong or deceptive information that is hard to spot even for specialists. Misinformation spread over the internet has the potential to degrade society as a whole [9]. News broadcasts on television (T.V.) can influence people's thinking. When false information is presented on news broadcasts, viewers' capacity for empathy is diminished, and people's feelings are wounded. People will feel deceived and abandoned.

Meanwhile, they have false beliefs about how they should think and act. Without a solution to this issue, it will be difficult for society to go forward in a healthy and peaceful manner, and social unrest will grow [10]. The second issue is that broadcasters might lose money due to the legal repercussions of airing fake news that leads to infractions. Since Fake News violates the law of trustworthiness of news reporting, it can swiftly escalate into illegal acts, leading to litigation against the production team and the TV network. If a news editor publishes a hoax, he or she may be held criminally liable and ordered to pay restitution to the victims [11]. Everyone in the office would rather avoid seeing this happen. That's why it's so important to actively avoid spreading misinformation. Last but not least, false news will do serious damage to the reputation of the party or government, and phone news will hurt the TV station's reputation. All data must be representative of party and government agendas. In order to carry out ideological and political objectives for the entire population, government agencies must rely on news broadcasts. Therefore, the News is the embodiment of the voice of the nation. If we don't get a handle on this, not only will the credibility of government agencies be at jeopardy, but so will the prestige of television networks.

1.2. Fake news on social networks

Recently, social media have become one of the most important news sources for millions of people around the world because they are cheap, easy to use, and quickly spread. However, this comes at a price: readers run a high danger of falling for "fake news," which is intentionally disseminated with the intent to mislead. Automatic identification of bogus news imposes issues against the popular content-based analysis approaches. One reason for this is because current natural language processing algorithms [12, 13] still haven't caught up to human levels of 'common sense,' which is often necessary for making sense of the news. Recent empirical investigations have shown the propagation patterns that may be utilized to instantaneously identify fraudulent news on social media [14], showing that fake and legitimate news are spread differently. Figure 2 depicts possible solutions and open questions regarding the identification of disinformation spread via social media.

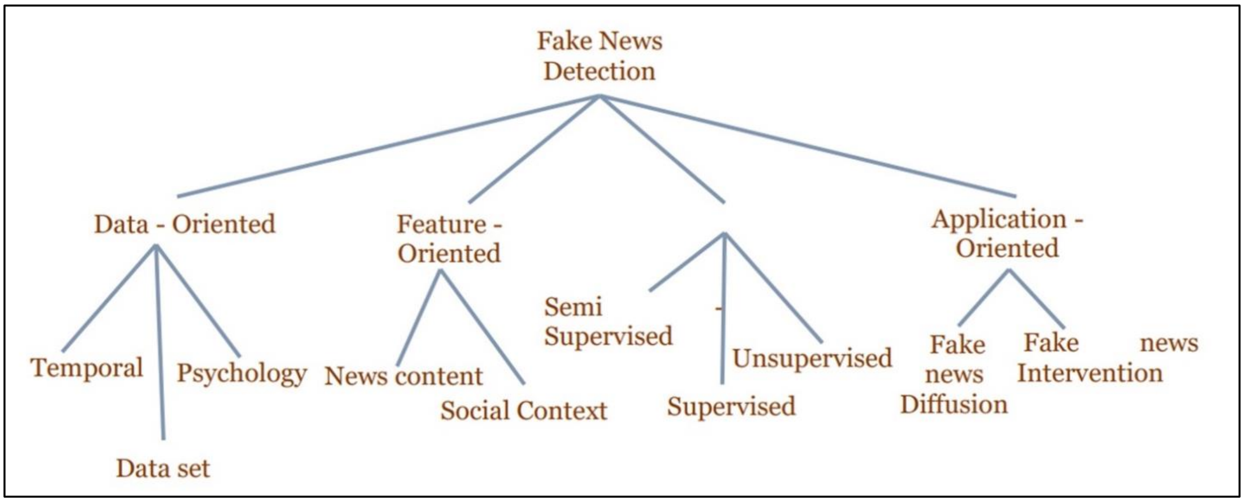


Figure 2. Prospects and unresolved questions for social media fake news detection

The inaugural false News Challenge (FNC-1) was established by [15] to promote the development of automated false news detection techniques utilizing AI technology and machine learning. Roughly fifty groups, from both business and academics, entered this competition. The goal of the FNC-1 test is to identify the article's position in relation to the headline provided. An article's position can fall into one of four categories. It may support, refute, or otherwise address the claims made in the headline. On their official website [16], you can find details on the FNC-1 task, its rules, the dataset, and the assessment measures.

Many natural language processing (NLP) tasks, including those that involve calculating semantic similarity between sentences [17, 18] and community-based question answering [19, 20], share similarities with the spread of fake news, and have been successfully implemented using deep learning models such as recurrent neural networks (RNN) and its variants [21, 22] and convolution neural networks (CNN) [23]. In [24], the semantic similarity of question pairs is calculated using Siamese MaLSTM. To evaluate the importance of each headline-body pair, we train a deep neural network to turn the text sequence into a fixed-length vector representation [25]. In this study, we present a technique for automatically assigning labels such as "agree," "disagree," "not relevant," and "discuss" to news stories. The amount of concordance between the headline and the allocated body is used to determine the categorization. Using keywords in article headlines, the suggested technique identifies relevant articles based on these observations. Some of the terms used in the headlines might help you find the most important parts of the text.

1.3. Motivation and Objective

Researchers in the field of artificial intelligence are increasingly interested in a number of developing subjects, one of which is the identification of fake news. Due to a lack of context-specific news data, the accuracy of fake news detection has not considerably improved despite extensive attention from the academic community. Deep learning is superior to the traditional feature-based model since it can discover the optimal feature set for a given problem or issue without any human intervention. Our goal is to use a Deep Convolutional Neural Network (FNDNet) to enhance the performance of existing false news detection systems.

2. Related Work

The identification of fake news is a topic that has aroused the interest of academics from all around the world. Numerous studies in the social sciences have examined the effects of fake news and people's responses to it. There are good and bad ways to use news on social media. News is a popular sort of content on social media platforms since they are widely used, accessible, and cost nothing to access. It also provides a venue for spreading deliberately false and otherwise low-quality news, which has the potential to harm both the general public and specific individuals. Therefore, the research of spotting fake news in social media gives unique characteristics that render conventional news media identification algorithms useless or inefficient.

In [26], the author suggested an adversarial cross-lingual learning CNN (Convolution Neural Networks) architecture based on user attention. The design relied on techniques for drawing the reader's focus to individual postings in order to identify the author's distinctive linguistic style. The attention-based CNN model was then fused with an adversarial cross-lingual learning framework for separating language-specific and independent features, leading to better user posts and resolving data scarcity problems by leveraging user characteristics as translation bridges. On datasets of English and Chinese microblog posts, their proposed methodology outperformed the vast majority of competing methods. Unknown ties between false news traits and news strips' authors, subjects, and events [27]. Their proposed false Detector will use AI to make credibility assessments of false news stories in real time. In order to learn about news strips, their authors, and subjects all at once utilizing many attributes from the text's explicit and hidden material, the proposed Fake Detector constructed deep diffusive networks [28]. The usefulness of the schema was proved in comprehensive experiments compared to many other algorithms on real-world false news datasets. In order to detect fake news in social networks, [29] offered a pattern-driven network approach, detailing the characteristics of false information, the individuals distributing it, and the relationships between them, as well as the occurrences of social psychology theories that provide an explanation for these tendencies. False information was detected using these patterns at the node, ego, triad, community, and network levels. In tests using real data, their strategy surpassed the majority of competing approaches. False news and misinformation, according to [30], affect public opinion and the political process. Therefore, it is essential to recognize misleading assertions in order to better appreciate their propagation and halt them. The ever-evolving nature of news strips presents serious difficulties for modern content-based methods of false article identification [31]. Therefore, trained classifiers acquired from reading published publications may become obsolete. Based on linguistic and web-markup features, this study proposed a TAG (Topic-AGnostic) classification strategy to deal with the problem of identifying fake news sites. Experimental results from many data sets were given, showing that the method improved false-news detection accuracy even when topics evolved over time. The authors of [32] presented a theory-driven framework for identifying fake News. This method examines news articles from several angles, including their vocabulary, grammar, semantics, and overall discourse. It was a symbol for cracking. All-Levels News Utilizing Reliable Forensic and Social Psychology Concepts. Then, under supervised ML frameworks, fraudulent reports are uncovered. This multidisciplinary study investigated the rise of fake news, improved the understandability of clickbait, false news, and deceptions/disinformation, and identified links between them all. On two real-world datasets, the strategy proposed in this paper outperformed most other strategies in benchmarks of limited data. In [33] advocated utilizing top-down and bottom-up rumor or rumor and rumor dispersions in a method called Bi-GCN (Bi-Directional Graph Convolution Networks) to analyze the quality of bogus news. Top-down directed graphs of rumor diffusion and adversarial graphs of rumor diffusion were employed in the study. Layers of GCNs also included data from the first posts, amplifying the impact of the rumor. Empirical data from the study showed that the approach was better to alternative methods.

In [34], the author proposes a method for detecting false news by examining supervised AI algorithms in social media profiles. Using freely available data, the author employed twenty-three different intelligent categorization methods. SentiWordNet was utilized by the author of [35] to take into account the text's cognitive clues in order to make opinion mining easier. In addition, it integrated objective elements like sentiment and credibility score from the GRNN (Bidirectional Gated Recurrent Neural Network) to give a platform for detecting bogus news. Additionally, in [36], the author proposes a unique multi-level voting ensemble model. The suggested system has been evaluated with twelve classifiers across three datasets. These machine learning classifiers are integrated using a false positive rate. Based on several performance measures, it has been shown that the Passive Aggressive, Logistic Regression, and Linear Support Vector Classifier (LinearSVC) all do best when utilizing the TF-IDF, CV, and HV feature extraction methods, in that order. Protecting social networks [37] and others [38,39] against various threats [4] is the focus of a wide range of alternative methods presented in the literature.

S. No	**Reference**	**Methods**	**Merits**	**Demerits**
1.	[3]	NLP tool	This tool was designed to aid users in recognizing and filtering possibly false news.	But it still has an issue with accuracy
2.	[1]	The mechanism for detecting deception	It looked into whether content-based signals and language clues are helpful for fraud detection.	It has an issue with a large dataset
3.	[12]	Deep bot detection model	This method provides high accuracy	This technique needs a long time to implement.
4.	[13]	Dataset analysis	It provides better sensitivity and specificity values	In a few cases, high dimensionality problem
5.	[24]	Recurrent and convolutional networks	On Twitter and Sina Weibo, it detected bogus news with accuracies of 85% and 92%.	However, it has a problem with an incomplete dataset
6.	[25]	CNN algorithm	It provides a higher precision metric	But it has an issue with error rates
7.	[16]	Fake Detector mechanism	It provides an automatic false news credibility inference model	Though, it has a problem with the accuracy value for the given dataset
8.	[7]	A pattern-driven network-based technique for detecting false News	Its goal was to investigate trends of fake news on social media platforms.	It has an issue with computational complexity
9.	[8]	Topic-agnostic approach	It achieved excellent accuracy in detecting fake news, even as themes shifted over time.	It includes limited availability of corpora
10	[9]	Theory-driven model	When there was minimal content, it allowed early detection of fake news.	It increases the computational time

A method for identifying false news that is based on an advanced learning model has been created by [22]. Multiple training models were used to analyze the first stages of news report preparation and editing. Then, an ensemble learning model comprised of four separate models was used for erroneous news identifications: the LSTM, the depth LSTM, the LIWC CNN, and the N-gram CNN. SAHS (Self-Adaptive Harmony Search) built an optimal weight of the ensemble learning model, which improved the accuracy of false news identifications. Cross-domain intractability was shown to be 72.3% accurate, while maximum accuracy was 99.4%.

These machine learning studies rely on attributes that were created by humans, though. These capabilities are restricted since they ignore the surrounding content. Furthermore, the majority of models cannot achieve satisfactory detection performance for the agree and disagree classes. We use principal component analysis (PCA) and Chi-square testing (Chi-square) in conjunction with CNN and LSTM layers to circumvent these restrictions. The suggested pipeline achieves a higher percentage of success (97.8%) than previously reported deep learning approaches.

3. Methodology

The primary innovation of this study is the proposal of hybrid deep learning models that combine two neural network layers (Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM)) for feature reduction. The suggested method outperforms conventional deep learning models in terms of prediction. Four data models are constructed to investigate the connection. In the first model, all the characteristics are used without preprocessing for categorization. After preprocessing, the second model makes use of the whole set of characteristics without any reductions. The third and fourth models are created with the use of principal component analysis (PCA) and Chi-square testing (Chi-square). This research delves deeper into the question of which models function best when combined with a hybrid CNN/LSTM model to process text input. The CNN-LSTM architecture receives the characteristics that have been picked by one of the aforementioned four models. In the embedding layer, each word in the input headlines and article bodies is transformed into a 100-element vector. Given that there are 5000 features, this layer will produce a 5000 × 100 matrix as its output. The weights for each word's vector are calculated using matrix multiplication and included in the output matrix. Contextual characteristics are extracted from these vectors by the CNN layer. In order to generate a single stance as the final result, the CNN layer's output is sent through an LSTM layer and subsequently to a fully connected dense layer. The proposed model is trained and evaluated using 32-sample subsets, as illustrated in fig.

3.

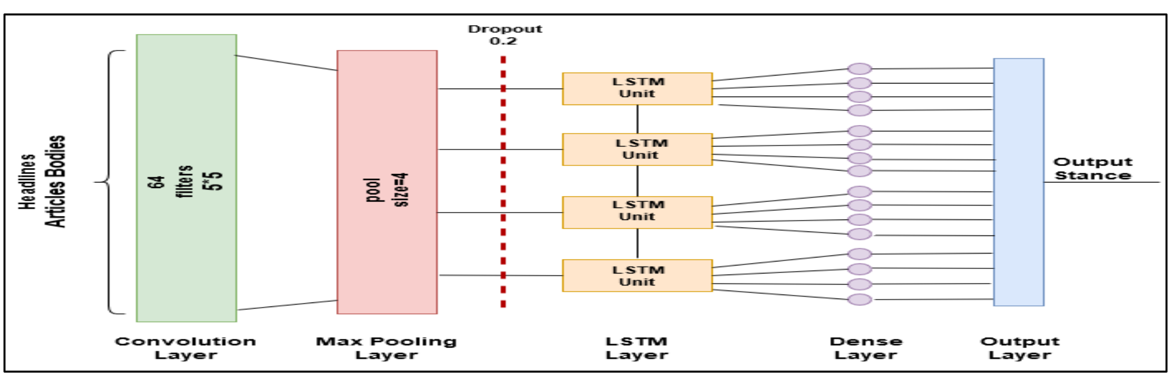


Figure 3. Proposed model architecture diagram.

3.1. Dataset description

The information comes from two sources so that we can tell the difference between authentic and fraudulent news. Kaggle.com is the source for both the fake and actual news dataset [22]. About 40,000 pieces, including both bogus and legitimate news, make up the dataset as a whole. Two distinct datasets, one containing fake news and one containing real news, both include around 20,000 items. But the authors of this piece [3] also call to another dataset: the glove Twitter data, which is also known to be pre-trained. Title, primary content, subject, and date are the four columns that make up the dataset.

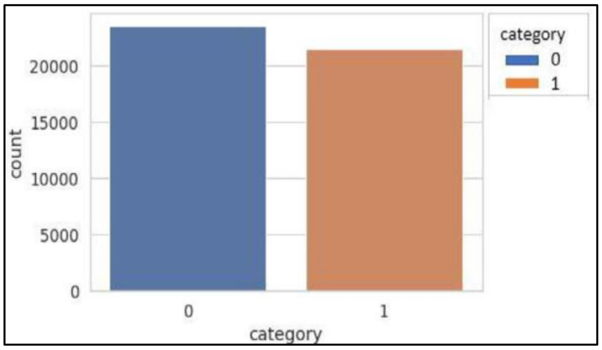


Figure 3. Articles in the news classified as either false or real

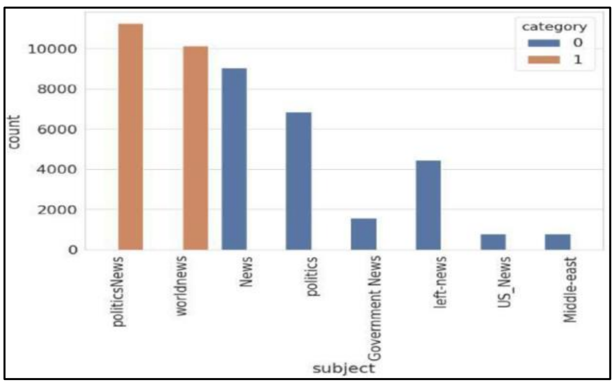


Figure 4. Current topics in both the fake news and real news industries

3.2. Data visualization and preprocessing

There are two types of categories in the dataset, with one labeled as true and the other as false. Displaying data in a visual framework, such maps or graphs, aids comprehension of what that data means in context. This makes it more intuitive for the human mind to analyze the data, which in turn facilitates the identification of trends, patterns, and outliers in huge datasets. The data set is divided between "fake news" and "real news" groups. Class '1' represents the first category of 'real news,' while class '0' represents the second category of 'false news.

Pre-processing data is an essential procedure that includes changing information before it is used.

Data cleansing and transformation are covered in the next section. As a first step in cleaning the data, Fig. 3 displays the total number of articles classified as bogus or authentic. It can be observed that the two datasets are very similar. The dataset has clearly stabilized. The false category of news is depicted in the figure by the blue '0' class, whereas the genuine category is depicted by the orange '1' class. Due to the fact that the subject lines of both types of articles are distinct, the prefixes, suffixes, and dates associated with each piece of news can be removed. The many types of data subjects are described in Fig. 4.

Figure 5. Word cloud visualization of the primary news dataset.

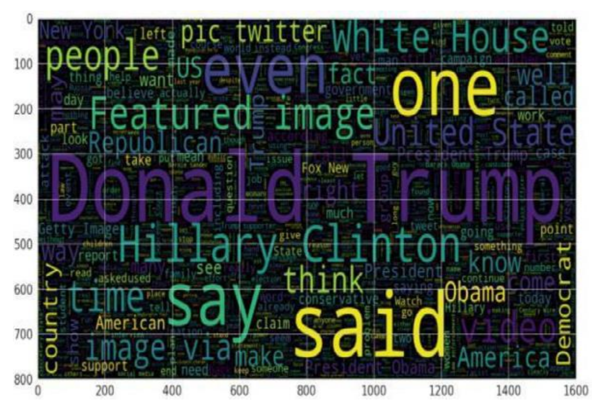
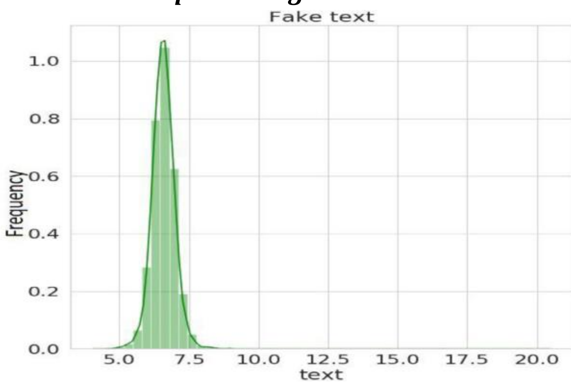


Figure 6. Representation of Word cloud for fake dataset

3.2.1. Data Preprocessing



Figures 5 and 6 display the terms from the authentic news collection and the disinformation set, respectively. A maximum of 2000 words are considered while making a word cloud for each topic. Word clouds are visual representations of large quantities of text highlighted and shown in various font sizes and alignments. White, House, United, State, Donald, Trump are all equally represented in the original news dataset, as seen by their relative size and volume in Fig. 5. While in Fig. 6, the bolder and louder terms are one, said, Donald, Trump, Hillary, Clinton which suggests that these phrases equally participate in the false news dataset. A term's size and thickness indicate how often it appears in a text and how much weight it carries. Two input categories, one containing false news and the other having actual news, are compared using their respective terms. The word cloud is constructed separately for each classification.

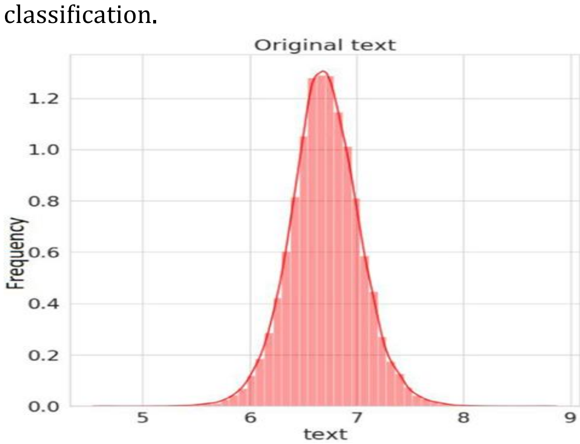


Figure 7. News word length categorization system.

Figure 8. Classification of Newspeak Phrase Length.

Data depiction of various outcomes along the X-axis is displayed in Figs. 7 and 8. The authentic or original news terms vary from 5 to 8 as indicated in 7. The frequency of sentences in original news with a score between 6 and 7 is highest, indicating that these scores represent the most common scores. In contrast, between 5 and 8 may be found in the dataset news for bogus terms.

A completely linked dense layer is used as the last layer in the proposed model, leading to a single output. After this layer, a softmax activation function is used. In order to do multi-class classification, softmax activation is utilized. Our dataset has four categories, thus we utilized softmax activation: agree, disagree, discuss, and unrelated. For this test, we employed Adam as the optimizer. For our tests, we employ a 32-person batch size and 50-epoch iterations.

4. Results and Discussion

This model can now distinguish between authentic and fraudulent news. The suggested model is tested for its efficacy using a variety of criteria, including accuracy, precision, recall, F1-score, and support. In addition, Table 2 displays the results of a comparison study. Model accuracy and publication year are included in this table's comparison analysis. When developing a model, testing and training are crucial steps. How far we train the model is crucial to its comprehensive foundation. Information on the number of epochs used for testing and training, together with the resulting accuracy, is displayed in Fig. 9. Figure shows that when the number of training epochs rises from 1 to 10, both the training and testing accuracy improves until the maximum number of epochs is considered. Comparative information about test and training losses over time is shown in Fig. 9. The generated graph shows that the training and testing loss are decreasing with time, indicating that the minimal loss has been achieved. The suggested model's confusion matrix is shown in Fig. 10. It demonstrates that the proposed model is quite accurate (97.80 percent).

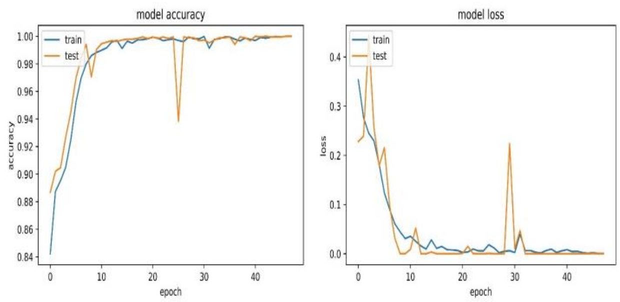


Figure 9. Training and Testing Accuracy and Loss of Proposed Model.

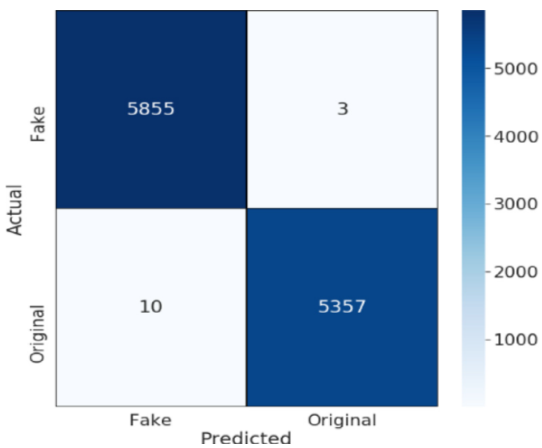


Figure 10. Confusion matrix for proposed model

4.1. Performance Evaluation

Our model is evaluated and compared using the following metrics: accuracy (A), precision (P), recall (R), and F1-score (F). Accuracy and memory are determined by solving equations 1 and 2. The F1-score, on the other hand, is the harmonic mean of the recall and accuracy scores, as shown in equation 1.

$$Pre = \frac{TP}{TP+FP} \quad (1)$$

The accuracy is determined by dividing the number of properly identified positive class instances by the total of the correctly and incorrectly classified positive class values. It provides information on how realistic the model is.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

The recall rate is determined by dividing the number of instances in which a positive class was assigned a value by the total number of instances in which a negative class was assigned a value. It provides information on the model's comprehensiveness.

$$F1 = 2 * \frac{Pre \cdot Recall}{Pre+Recall} \quad (3)$$

Model accuracy across classes is evaluated using the F1-score. In cases of data inequity, the F1-score measure is frequently employed. Since FNC-1's dataset is similarly extremely unbalanced, we utilize F1-score as evaluation metrics to demonstrate the completeness of the proposed model in terms of class-wise accuracy.

4.2. Comparison with deep learning models

The findings provided on the FNC- 1 challenge have employed the fine-tuning strategy, where all parameters are jointly fine-tuned and a basic classification layer is added to the pre-trained model (BERT stands for "Bidirectional Encoder Representations from Transformers" [67]). BERT makes predictions for each masked position individually. This means that it neglects relationships between anticipated masked locations during training. BERT has a pre-train fine-tune inconsistency because of the simultaneous decrease of some dependencies it learns. On the FNC-1 challenge, this model achieves 91.3% accuracy. When compared to our model and the F1-scores of the disagree, agree, and unrelated classes, BERT achieves a significantly lower score.

Table 2. CNN-LSTM model k-fold cross-validation with PCA.

K-folds	Accuracy	Precision	Recall	F-score
Fold-1	97.4%	97.4%	98.4%	97.9%
Fold-2	97.8%	96.2%	96.5%	96.3%
Fold-3	97.2%	96.3%	98.6%	97.4%
Fold-4	97.1%	96.4%	97.4%	96.9%
Fold-5	97.8%	97.9%	99.2%	98.5%
Fold-6	95.9%	96.3%	97.1%	96.7%
Fold-7	96.9%	96.8%	97.4%	97.1%
Fold-8	97.9%	97.5%	99.2%	98.3%
Fold-9	96.4%	96.9%	98.1%	97.5%
Fold-10	97.2%	97.3%	98.1%	97.7%
10-Fold Mean	**97.1%**	**96.9%**	**98.0%**	**97.4%**

In addition to avoiding independent predictions, XLNet incorporates a bidirectional context [68]. Instead of predicting tokens in a strict sequential fashion, "permutation language modeling" predicts them in a completely arbitrary order. XLNet, which is based on Transformer XL, beats BERT on 20 different tasks.

5. Conclusion and Future Scope

In contrast to prior research that focused solely on particular lines or phrases, this study developed a fake news stance detection algorithm based on both the headline and the body of the news. The proposed model combines convolutional neural networks (CNNs) and long short-term memories (LSTMs), with PCA and chi-square used to extract quality characteristics that are then fed into the CNN-LSTM model. We begin by feeding the full, unfiltered feature set into the neural network, both with and without any preprocessing. Applying dimensionality reduction methods and contrasting the outcomes follows. By excluding these elements from the feature vector, PCA improves the classifier's effectiveness in detecting bogus news by filtering out superfluous data. Up to 97.8% accuracy is achieved with this method, which is a significant improvement over earlier research. It is worth noting that dimensionality reduction techniques allow for fewer characteristics to be used without sacrificing classification accuracy. In the future, we plan to (a) test our model on more extensive data sets, (b) investigate if tree-based learning outperforms more straightforward methods, and (c) examine how fusing various textual elements might improve performance.

References

1. T. Mihaylov, G. Georgiev, and P. Nakov, ‘‘Finding opinion manipulation trolls in news community forums,’’ in Proc. 19th Conf. Comput. Natural Lang. Learn., Beijing, China, Jul. 2015, pp. 310–314. [Online]. Available: https://www.aclweb.org/anthology/K15-1032
2. T. Mihaylov, I. Koychev, G. Georgiev, and P. Nakov, ‘‘Exposing paid opinion manipulation trolls,’’ in Proc. Int. Conf. Recent Adv. Natural Lang. Process., Hissar, Bulgaria, Sep. 2015, pp. 443–450. [Online]. Available: https://www.aclweb.org/anthology/R15-1058
3. T. Mihaylov and P. Nakov, ‘‘Hunting for troll comments in news commu- nity forums,’’ in Proc. 54th Annu. Meeting Assoc. for Comput. Linguistics, vol. 2, 2016, pp. 399–405, doi: 10.18653/v1/P16-2065.
4. P. Bourgonje, J. Moreno Schneider, and G. Rehm, ‘‘From click- bait to fake news detection: An approach based on detecting the stance of headlines to articles,’’ in Proc. EMNLP Workshop: Natural Lang. Process. meets Journalism, 2017, pp. 84–89. [Online]. Available: https://www.aclweb.org/anthology/W17-4215
5. S. Vosoughi, D. Roy, and S. Aral, ‘‘The spread of true and false news online,’’ Science, vol. 359, no. 6380, pp. 1146–1151, Mar. 2018.
6. A. M. Michael Barthel and J. Holcomb. (2016). Many Americans Believe Fake News is Sowing Confusion. Accessed: Sep. 29, 2019. [Online]. Available: https://www.journalism.org/2016/12/15/many-americans- believe-fake-news-is-sowing-confusion/
7. A. K. Chaudhry, ‘‘Stance detection for the fake news challenge: identify- ing textual relationships with deep neural nets,’’ in Proc. Natural Lang. Process. Deep Learn., 2017, pp. 1–10.
8. S. Chopra, ‘‘Towards automatic identification of fake news: Headline- article stance detection with LSTM attention models,’’ Stanford Univ., Stanford, CA, USA, Tech. Rep., 2017.
9. G. Bhatt, A. Sharma, S. Sharma, A. Nagpal, B. Raman, and Mittal, ‘‘Combining neural, statistical and external features for fake news stance identification,’’ in Proc. Companion Web Conf. Web Conf. (WWW), Geneva, Switzerland, 2018, p. 1353, doi: 10.1145/3184558.3191577.
10. L. Konstantinovskiy, O. Price, M. Babakar, and A. Zubiaga, ‘‘Towards automated factchecking: Developing an annotation schema and bench- mark for consistent automated claim detection,’’ 2018, arXiv:1809.08193. [Online]. Available: http://arxiv.org/abs/1809.08193
11. D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M.

Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, ''The science of fake news,'' Science, vol. 359, no. 6380, pp. 1094–1096, 2018.

12. L. Borges, B. Martins, and P. Calado, ''Combining similarity features and deep representation learning for stance detection in the context of checking fake news,'' J. Data Inf. Qual., vol. 11, no. 3, pp. 1–26, Jul. 2019, doi: 10.1145/3287763.

13. M. A. Walker, P. Anand, R. Abbott, and R. Grant, ''Stance clas- sification using dialogic properties of persuasion,'' in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Tech- nol., Stroudsburg, PA, USA, 2012, pp. 592–596. [Online]. Available: http://dl.acm.org/citation.cfm?id=2382029.2382124

14. D. Sridhar, J. Foulds, B. Huang, L. Getoor, and M. Walker, ''Joint models of disagreement and stance in online debate,'' in Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics 7th Int. Joint Conf. Natural Lang. Pro- cess., vol. 1. Beijing, China, Jul. 2015, pp. 116–125. [Online]. Available: https://www.aclweb.org/anthology/P15-1012

15. S. Somasundaran and J. Wiebe, ''Recognizing stances in ideological on- line debates,'' in Proc. NAACL HLT Workshop Comput. Approaches Anal. Gener. Emotion Text, Los Angeles, CA, USA, Jun. 2010, pp. 116–124. [Online]. Available: https://www.aclweb.org/anthology/W10-0214

16. M. Lukasik, P. K. Srijith, D. Vu, K. Bontcheva, A. Zubiaga, and T. Cohn, ''Hawkes processes for continuous time sequence classification: an application to rumour stance classification in twitter,'' in Proc. 54th Annu. Meeting Assoc. Comput. Linguistics, vol. 2. Berlin, Germany, Aug. 2016, pp. 393–398. [Online]. Available: https://www.aclweb.org/anthology/P16-2064.

17. L. Derczynski, K. Bontcheva, M. Liakata, R. Procter, G. Wong Sak Hoi, and A. Zubiaga, ''SemEval-2017 Task 8: RumourEval: Determining rumour veracity and support for rumours,'' in Proc. 11th Int. Workshop Semantic Eval. (SemEval), Vancouver, BC, Canada, Aug. 2017, pp. 69–76. [Online]. Available: https://www.aclweb.org/anthology/S17-2006

18. C. Stab and I. Gurevych, ''Parsing argumentation structures in persuasive essays,'' Comput. Linguistics, vol. 43, no. 3, pp. 619–659, Sep. 2017. [Online]. Available: https://www.aclweb.org/anthology/J17-3005

19. D. P. Rao, (2017). Exploring How Artificial Intelligence Technologies Could be Leveraged to Combat Fake News. Accessed: Oct. 29, 2019. [Online]. Available: http://www.fakenewschallenge.org/

20. J. Chung, Ç. Gülçehre, K. Cho, and Y. Bengio, ''Empirical evaluation of gated recurrent neural networks on sequence modeling,'' CoRR, vol. abs/1412.3555, 2014. [Online]. Available: http://arxiv.org/abs/ 1412.3555

21. A. Graves and J. Schmidhuber, ''Framewise phoneme classification with bidirectional LSTM and other neural network architectures,'' Neural Netw., vol. 18, pp. 602–610, Jul./Aug. 2005.

22. P. Neculoiu, M. Versteegh, and M. Rotaru, ''Learning text similarity with siamese recurrent networks,'' in Proc. 1st Workshop Represent. Learn., Jan. 2016, pp. 148–157.

23. H. He, K. Gimpel, and J. Lin, ''Multi-perspective sentence similarity modeling with convolutional neural networks,'' in Proc. Conf. Empirical Methods Natural Lang. Process., 2015, pp. 1576–1586. [Online]. Avail- able: https://www.aclweb.org/anthology/D15-1181

24. R. Kiros, Y. Zhu, R. Salakhutdinov, R. S. Zemel, A. Torralba, R. Urtasun, and S. Fidler, ''Skip-thought vectors,'' in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 3294–3302.

25. K. S. Tai, R. Socher, and C. D. Manning, ''Improved semantic rep- resentations from tree-structured long short-term memory networks,'' Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics 7th Int. Joint Conf. Natural Lang. Process., vol. 1, 2015, pp. 1–11. [Online]. Available: http://dx.doi.org/10.3115/v1/p15-1150

26. L. Yang, Q. Ai, D. Spina, R.-C. Chen, L. Pang, W. B. Croft, J. Guo, and F. Scholer, ''Beyond factoid qa: Effective methods for non-factoid answer sentence retrieval,'' in ECIR, 2016. Y. Yang, W.-T. Yih, and C. Meek, ''WikiQA: A challenge dataset for open- domain question answering,'' in Proc. Conf.

Empirical Methods Natural Lang. Process., Lisbon, Portugal, Sep. 2015, pp. 2013–2018. [Online]. Available: https://www.aclweb.org/anthology/D15-1237

27. Z. Imtiaz, M. Umer, M. Ahmad, S. Ullah, G. S. Choi, and A. Mehmood, ''Duplicate questions pair detection using siamese MaLSTM,'' IEEE Access, vol. 8, pp. 21932–21942, 2020.

28. B. Ghanem, P. Rosso, and F. Rangel, ''Stance detection in fake news a combined feature representation,'' in Proc. 1st Workshop Fact Extraction Verification (FEVER), Brussels, Belgium, Nov. 2018, pp. 66–71. [Online]. Available: https://www.aclweb.org/anthology/W18-5510

29. Q. Zeng, ''Neural stance detectors for fake news challenge,'' Stanford Univ., Stanford, CA, USA, Tech. Rep., 2017.

30. S. R. Pfohl, ''Stance detection for the fake news challenge with atten- tion and conditional encoding,'' Stanford Univ., Stanford, CA, USA, Tech. Rep., 2017.

31. D. A., A. I., and S. S., ''A comparative study on using principle component analysis with different text classifiers,'' Int. J. Comput. Appl., vol. 180, no. 31, pp. 1–6, Apr. 2018, doi: 10.5120/ijca2018916800.

32. S. Karamizadeh, S. Abdullah, A. Manaf, M. Zamani, and A. Hooman, ''An overview of principal component analysis,'' J. Signal Inf. Process., vol. 4, no. 3B, p. 173, Aug. 2013.

33. M. Ahmad, A. M. Khan, J. A. Brown, S. Protasov, and A. M. Khattak, ''Gait fingerprinting-based user identification on smartphones,'' in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2016, pp. 3060–3067.

34. S. Deegalla and H. Bostrom, ''Reducing high-dimensional data by princi- pal component analysis vs. Random projection for nearest neighbor classi- fication,'' in Proc. 5th Int. Conf. Mach. Learn. Appl. (ICMLA), Dec. 2006, pp. 245–250.

35. M. Ahmad, D. I. U. Haq, Q. Mushtaq, and M. Sohaib, ''A new statistical approach for band clustering and band selection using K-means clustering,'' IACSIT Int. J. Eng. Technol., vol. 3, no. 6, pp. 606–614, Dec. 2011.

36. S. Mohammad, S. Kiritchenko, P. Sobhani, X. Zhu, and C. Cherry, ''SemEval-2016 task 6: Detecting stance in tweets,'' in Proc. 10th Int. Workshop Semantic Eval. (SemEval-2016), San Diego, CA, USA, Jun. 2016, pp. 31–41. [Online]. Available: https://www.aclweb.org/anthology/S16-1003

37. K. Popat, S. Mukherjee, J. Strötgen, and G. Weikum, ''Where the truth lies: Explaining the credibility of emerging claims on the Web and social media,'' in Proc. 26th Int. Conf. World Wide Web Companion, Apr. 2017, pp. 1003–1012.

38. C. Stab, T. Miller, and I. Gurevych, ''Cross-topic argument mining from heterogeneous sources using attention-based neural networks,'' 2018, arXiv:1802.05758. [Online]. Available: http://arxiv.org/abs/1802.05758

39. I. Augenstein, T. Rocktäschel, A. Vlachos, and K. Bontcheva, ''Stance detection with bidirectional conditional encoding,'' in Proc. 2016 Conf. Empirical Methods Natural Lang. Process., Austin, TX, USA, Nov. 2016, pp. 876–885.