Vol. 03, No. 02, pp. 81 –88 (2022)
ISSN: 2708-0757

JASTT

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

A New Approach to Detect Fake News Related to Covid-19 Pandemic Using Deep Neural Network

Awf Abdulrahman [1,*], Muhammet Baykara[2]

[1] *Department of public Health, Duhok Polytechnic University, Duhok, Iraq, awf.ramadhan@dpu.edu.krd*
[2] *Department of Software Engineering, Faculty of Technology, Firat University, Elazig, Turkey, mbaykara@firat.edu.tr*
**Corresponding Author awf.ramadhan@dpu.edu.krd*

Abstract

The fake news that accompanied the Covid-19 pandemic on social media platforms negatively affected people and led to a state of panic and fear of the unknown. This study aims to build a model for classifying textual news for four datasets related to COVID-19, binary classification (fake and real) with high performance. Two-hybrid deep learning models were built. The first model consists of three layers of a one-dimension convolutional neural network (1D-CNN) followed by two layers of long-short term memory neural network (LSTM). The second model consists of three layers of a 1D-CNN followed by two layers of bidirectional LSTM neural network (BiLSTM). Finally, the results obtained using hybrid models were compared with the results obtained by applying three machine learning classifiers (naïve Bayes, logistic regression, and k-nearest neighbor) on the same data sets. This study achieved promising results with an accuracy of (96.98%, 94.52%, 99.60%, and 99.90%) for the first model with all data sets and (97.15%, 95.32%, 99.40%, and 99.82%) for the second model with the same four data sets.

Keywords: *Fake news detection, BiLSTM, Deep learning COVID-19, LSTM, CNN*

Received: July 17th, 2022 / Accepted: December 05th, 2022 / Online: December 14th, 2022

I. INTRODUCTION

The outbreak of the Covid-19 epidemic from March 2019 until today, April 29, 2022, has infected about 510270667 people and death of more than 6 million people, according to World Health Organization statistics (WHO)[1]. This new strain was identified as a member of the Coronaries family. It was preceded by severe acute respiratory syndrome (SARS) in China in 2002 and Middle East Respiratory Syndrome (MERS) in the Kingdom of Saudi Arabia in 2013 [1]. WHO named this new virus 'COVID-19' in February 2020 and classified it as a global pandemic on March 11 2020 [1][2][3]. Coinciding with the outbreak of this pandemic, a massive number of fake news on many platforms, mainly social media sites such as Twitter, Facebook, Instagram, etc., due to its popularity, which allows the free dissemination and promotion of ideas, was published [4][5]. Social media has greatly contributed to spreading misinformation about COVID-19 through posts, likes, comments, and retweets [5].

Lots of fake news has caused panic among people. As governments and decision-making institutions are in a very difficult position in trying to contain the pandemic, manage the crisis, and control fake news by spreading awareness to get news from reliable sources [6]. Fake news about the pandemic was divided into several axes, including fake news about the conspiracy theory, fake news about 5G technologies that it claimed caused the spread of the pandemic, lack of medicines and medical supplies, and fake anti-vaccine news. All this news led to the reluctance of a large number of people from taking the Covid-19 vaccine. According to a field survey conducted in Germany and France, 10% of the participants indicated that they did not want to take the COVID-19 vaccine [7]. The foregoing shows the extent of the damage caused by the spread of fake news through various social media sites. Many researchers have made great efforts in analyzing digital content about the pandemic and categorizing it into fake news and real news using many applications. The use of machine learning and deep learning techniques became widespread in the early detection of fake news [8].

In this study, a hybrid deep learning model was built between the convolutional neural network (CNN) and the recurrent neural network (RNN) to classify news binary-class classification (fake and real classes) using four data sets. The built model was evaluated using parameters of global accuracy,

[1] https://covid19.who.int

doi: 10.38094/jastt302124

ipAcademia www.ipacademia.org

sensitivity, specificity, and area under the curve (AUC). The model achieved high performance with all data sets.

The rest of this paper is arranged as follows. In the second section, related work is discussed. The third section describes the databases that were used in this study. In the fourth section, the methods and materials used in the study are described. In the fifth section, the results are listed and discussed briefly. In the final section, the conclusion of this study is presented.

II. RELATED WORKS

In the study [8], four binary classification experiments were applied to detect fake news: logistic regression (LR), embedding with dense layer, embedding layer with LSTM layer, and bidirectional LSTM. In this study, the Webhose.io dataset was used. News unrelated to COVID-19 was removed from the dataset using three keywords in the filter: corona, coronavirus, and COVID-19. This study achieved 75.07% accuracy using the logistic regression model, 68.74% accuracy using the embedding model with dense layer, 68.7% accuracy using the embedding layer model with LSTM layer, and 70.64% accuracy using the bidirectional LSTM model. The study [9] aimed to analyze the problem of hostile and fake content in the Devanagari (Hindi) script and addressed the problem of fake news related to Covid-19 in English.

For the second aspect of the study, the term frequency-inverse document frequency (TF-IDF) technique was used to select useful words to be included in the bag of word (BoW) method. Also, an embed layer extracted from FastText and Word2Vec was used as inputs to the built model. This study achieved 93.45% fake news classification accuracy using the SVM classifier. The study presented [10] a new approach to classifying fake news based on reducing the number of corresponding features and increasing classification accuracy. In this study, the Koirala dataset was used. Three feature extraction algorithms were applied, namely particle swarm optimization (PSO), genetic algorithm (GA), and Salp swarm algorithm (SSA). The best classification accuracy achieved by this study was 75.4% using the k-nearest neighbour Binary-coded genetic algorithm (k-NN-BGA) method after reducing the number of features to 303. The study [11] provided solutions to the problem of fake news about COVID-19 and 5G technology by analyzing tweets related to conspiracy theories (COVID-19 and 5G) and classifying them into two and three classes. This study consists of two tasks: the first is to detect text-based news, and the second is to detect structure-based fake news.

This study suggested six solutions based on Bag of Words (BoW) and the emergence of Bidirectional Encoder Representation from Transformer (BERT) inclusion in the first task. In the second task, Graph Neural Networks (GNNs) were used. This study achieved F1 performance accuracy of 66% and 69.3% in binary classification using BoW and BERT, respectively. It also achieved an average ROC of 0.95% using Graph Neural Networks (GNNs) in the second task. The study [12], provided a solution to the problem of Constraint@AAAI2021 -COVID19 Fake News Detection in English. Eight pre-trained algorithms were used: BERT, Generative Pre-trained Transformer 2 (GPT-2), a generalized autoregressive pretraining method (XLNet), Robustly Optimized BERT pre-training Approach (RoBERTa), distillation BERT (DistilRoBERTa), A lite PERT (ALBERT), and Decoding-enhanced BERT (DeBERTa). Then it was followed by machine learning or deep learning classifiers: Support vector machine (SVM), Multi-Layer Perceptron (MLP), and RCNN. This study achieved the best performance accuracy of 97.9% using RoBERTa + MLP.

As can be understood from the studies in the literature, it has been seen that fake news in social media can be found with high accuracy with various deep learning algorithms. Similarly, machine learning methods often give high success rates. However, it is seen that the TF-IDF and BoW techniques are generally used in studies. Therefore, the repetition of studies with different methods based on fake textual news does not add scientific innovation to the literature. For this reason, a way must be found to detect fake news in different ways, such as photos, videos, and audio clips.

III. COVID-19 FAKE NEWS DATABASES

This study used four databases of COVID-19 fake news freely available to the public: (1) COVID19FN[2] [13]; (2) Covid-19 News Dataset Both Fake and Real[3] [14]; (3) The CoVID19-FNIR dataset[4] [15]; and (4) the Fake News database[5].

A. COVID19FN dataset

The first dataset contains four columns (title, text, country, and label) and more than 2800 news articles collected from Poynter and other sites. The COVID19FN also contains the source URL, publish date, and the country of the news.

B. COVID_19 Nwes Dataset Both Fake and Real

The second dataset contains 16,988 fake and real news. It was collected from two different sources, the first from CBC[6] and the second from different web portals[7].

C. The CoVID19-FNIR Dataset

The third dataset consists of 7,588 news related to the novel coronavirus (COVID-19) pandemic. Samples were collected from India, the USA, and Europe between February 2020 to June 2020. In addition, fake news was collected from Poynter and real news from Twitter.

D. Fake News Database

The fourth data set contains six columns (ID, title, text, subject, date, and label). More than 31,400 news articles have been collected from social media sites.

[2] https://data.mendeley.com/datasets/b96v5hmfv6/3
[3] https://zenodo.org/record/4722484#.YPG9NYJR3IV
[4] https://ieee-dataport.org/open-access/covid-19-fake-news-infodemic-research-dataset-covid19-fnir-dataset
[5] https://www.kaggle.com/c/classifying-the-fake-news/data?select=training.csv
[6] https://zenodo.org/record/4722470
[7] https://zenodo.org/record/4282522

Table I shows the details of all the data sets and the ratio of training and testing.

TABLE I. DATA SETS DETAILS AND THE RATIO OF THE TRAINING AND TESTING

Data set	***Symbol***	***Total rows***	***Training***	***Testing***	***Validation***
Data set 1	DB1	2821	2257	564	112
Data set 2	DB2	16990	13592	3398	680
Data set 3	DB3	7588	6070	1518	304
Data set 4	DB4	31430	25144	6286	1257

IV. METHODOLOGY

This study used two hybrid deep learning models (CNNs + RNNs) to detect fake COVID-19 news. In addition, the result obtained from both hybrid models were compared under the same conditions with three machine learning classifiers (NB, LR, and k-NN). Figure 1 shows this study's stages to reach the desired results. As shown in Figure 1, this study had two main steps, some of which included several sub-steps.

A. Splitting Datasets

In this step, we divided the data sets into training, testing, and validation sets. 80% for training and 20% for validation and testing. The proportion of the validation set is 20% of the testing set, as seen in Table I.

B. Vectorizing Methods

This part converts the text data to 0 and 1 vectors. In addition, this method converts the text file to new vectors [16]. Three vectorizing techniques were used for the machine learning part: TF-IDF, the N-Gram level vectorizer, the Count vectorizer (CV)), and the tokenizer with an embedding layer for deep learning.

1) TF-IDF Vectorizer

The term frequency-inverse document frequency (TF-IDF) method is considered one of the most popular text-to-vector conversion methods [17]. This method consists of two parts. In the first part, the term frequency (TF) is calculated as in Equation 1, and the inverse document frequency (IDF) is calculated in the second part, as in Equation 2. Then the two sections are calculated together as in Equation 3.

$$TF(t) = \text{No of } t \text{ in a doc.} / \text{Total term in the doc.} \quad (1)$$

$$IDF(t) = \text{Log (No. of Doc. / No. of doc. Containing } t). \quad (2)$$

$$TF\text{-}IDF = tf_d^i * Log(\frac{N}{df^i}) \quad (3)$$

Where [16] tf_d^i : represents how many times i appears in the document d. N : represents number of the total documents in the document set. df^i : these are the documents in which the term i is occurring.

2) N-Gram level Vectorizer

The N-gram method is a sub-method of the TF-IDF method. This method consists of a slice (N) of letters representing the degrees of TF-IDF. This method was first proposed to overcome the problem of selecting the correct features and their numerical value by using the TF-IDF classification associated with unigrams or bigrams [16][18].

3) Count Vectorizer

The CV method is a two-dimensional array. This matrix consists of rows and columns; each row contains a document from the data set, each column includes a term from the data set, and each cell has a number that represents the number of times the term is repeated in the document [16].

C. Classification Experiments

1) Machine Learning Classifiers

In this study, three machine learning classifiers were applied, which are the naive Bayes classifier (NB), the logistic regression (LR) classifier, and the k-nearest neighbour (k-NN) classifier.

An NB classifier is a supervised machine learning classifier that classifies an object into one of the predefined classes. NB classifier performs well if the amount of data is large [19]. Logistic regression (LR) is an algorithm based on a central mathematical concept, the logarithm-the natural log of the odds ratio. This algorithm is a machine learning algorithm that performs the process of predictive analysis of data by explaining the relationship between a binary dependent variable and one or more independent variables [20]. K-NN is a simple supervised machine learning algorithm aiming to solve classification problems. K-NN classifies the new data based on the similarity status of the new data to the closest available class. This algorithm classifies the training data into categories based on the output and K-value [21].

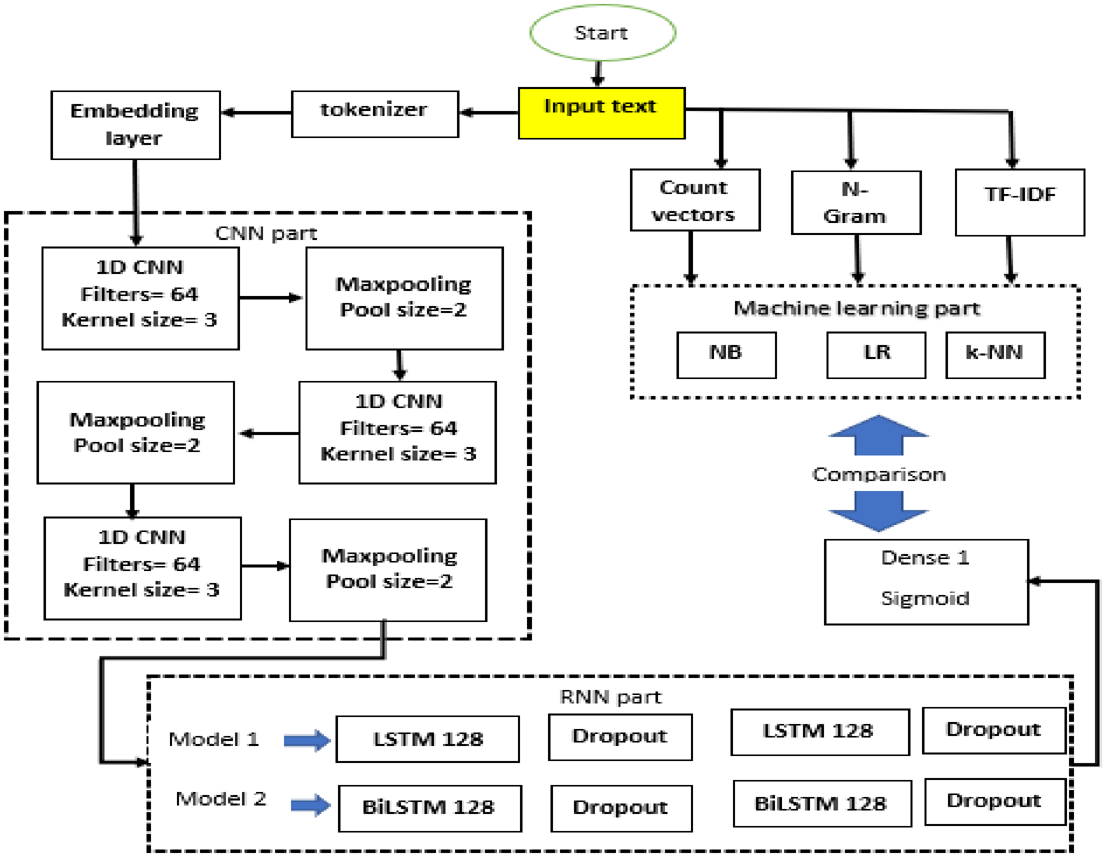

Fig. 1. The Block Diagram of this Study

2) Two Hybrid Models (Embedding with 1D CNN, LSTM and 1D CNN, BiLSTM Layers

In this study, two hybrid models are presented. The first model consists of an embedding layer of dimension 250 with pre-trained weights, followed by three 1D-CNN layers and two LSTM layers. The first 1D-CNN layer consists of a 64-size filter and a 3-kernel-size. The second 1D-CNN layer consists of a 128-size filter and a 3-kernel-size. And the last 1D-CNN layer consists of a 192-size filter and a 3-kernel-size, as shown in Figure 1 (CNN part). The LSTM layers consist of 128 units for each layer with a dropout layer with a 0.2 value, as shown in Figure 1 (RNN part). The second model consists of three layers 1D-CNN and two BiLSTM layers consisting of 128 units for each layer with a dropout layer with a 0.2 value. For the training, we used these parameters: Adam optimizer with binary cross-entropy loss, batch size of 32, random state of 42, the learning rate of 0.001, and the model is trained for 100 epochs. Table II shows the total layers and params using the first model.

3) LSTM Network

Long-Short term memory (LSTM) was first proposed to improve RNNs by Hochreiter and Schmidhuber (1997) [22]. An LSTM network is based on control gates that control the flowing data and prevent problems that unnecessary data may cause. Later, the LSTM network was updated by Gers et al. (2000) [23]. Forgetting gates have been added to forget useless memories from the memory cell. The number of gates in the LSTM network becomes three: input gate, forget gate, and output gate. Figure 2 shows the three gates of a single-cell structure of the LSTM network. Also, Figure 3 shows the recurrent process in the LSTM network, and The mathematical model for the single LSTM cell is shown in equations 4- 9 [24]. The LSTM network differs from the RNN in learning skills using a complex gate approach. The input gates learn the importance of the input information. The network gate learns the importance of the volume of relevant information and the forgotten gate. In 2012 [25], a study was presented in which LSTM was used for a modelling task using English and French. This study concluded that LSTM outperformed RNN by 8%. One of the advantages of LSTM networks is their ability to extract high-level textual information, and plays a pivotal role in natural language processing [25] [26].

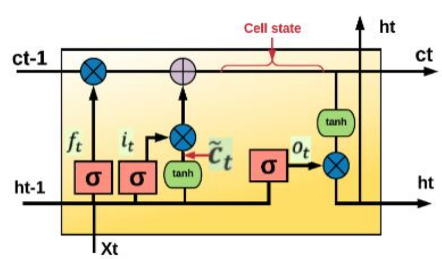

Fig. 2. Single-cell of LSTM network [27].

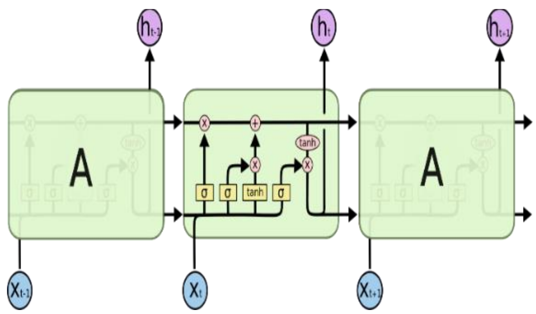

Fig. 3. The general architecture of the LSTM network [22].

TABLE II. THE TOTAL LAYERS AND PARAMS FOR THE FIRST MODEL (1D CNN + LSTM)

Layer	***Output shape***	***Params***
Embedding	(None, 1000, 250)	7163750
1D CNN	(None, 1000, 64)	48064
Max pooling	(None, 500, 64)	7588
1D CNN	(None, 500, 128)	24704
Max pooling	(None, 250, 128)	0
1D CNN	(None, 250, 192)	73920
Max pooling	(None, 125, 192)	0
LSTM	(None, 125, 128)	164352
Dropout	(None, 125, 128)	0
LSTM	(None, 128)	131584
Dropout	(None, 128)	0
Dense	(None, 1)	129
Total params	7606503	
Trainable params	7606503	
Non-trainable params	0	

$$f_t = \sigma(w_f \,.\, [h_{t-1}, x_t] + b_f) \tag{4}$$

$$i_t = \sigma(w_i \,.\, [h_{t-1}, x_t] + b_i) \tag{5}$$

$$\tilde{c}_t = tanh(w_c \,.\, [h_{t-1}, x_t] + b_c) \tag{6}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{c}_t \tag{7}$$

$$O_t = \sigma(w_o \,.\, [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = O_t * \tanh(C_t) \tag{9}$$

Where: W_f, W_i, W_c, W_o represents the LSTM weights, b_f, b_i, b_c, and b_o represent the used biases for each cell. And C_t represents the cell state.

4) BiLSTM Network

The bidirectional LSTM neural network (BiLSTM) was presented as a deep learning network for classification and prediction problems. Also, BiLSTM is an advanced LSTM network. The BiLSTM network combines a frontal hidden layer and a back hidden layer. The BiLSTM model encodes the information from back to front. The output of the forward layer (h_t^f) and backward layer (h_t^b) of the BiLSTM model are shown in equations 10 and 11 [28]. The BiLSTM network has access to both the previous and subsequent contexts. One of the advantages of the BiLSTM network is that it solves the sequential modelling problem better than LSTM. In this study, BiLSTM was applied to extract contextual information from features generated by the convolutional layer [25]. Figure 4 shows the general architecture of the BiLSTM network [25].

$$h_t = \alpha h_t^f + \beta h_t^b \tag{10}$$

$$y_t = \sigma(h_t) \tag{11}$$

where α and σ are the numerical factors respecting the equality $\alpha + \sigma = 1$ [28].

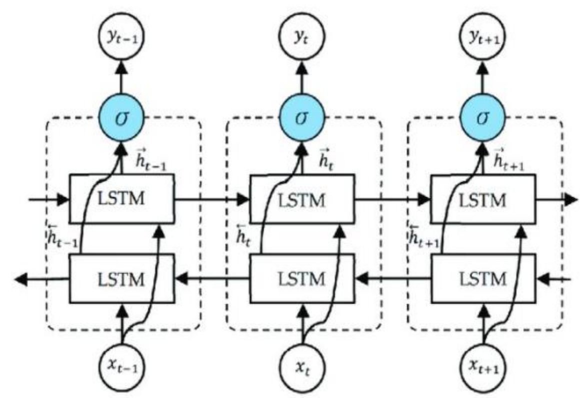

Fig. 4. The general architecture of the BiLSTM network.

D. *Evaluation Metrics*

Both proposed models were evaluated using several parameters: global accuracy, sensitivity, specificity, and ROC curve. A confusion matrix was used to check how well both models performed. Equations (12)– (14) give the global accuracy, sensitivity, and specificity performance measures, respectively [29].

$$acc = \frac{TP+TN}{TP+FP+FN+TN} \tag{12}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{13}$$

$$Specifcity = \frac{TN}{TN+FP} \tag{14}$$

Where TP: represents the true positive values, TN: represents the true negative values, FP: represents the false positive values, and FN: represents the false negative values.

V. RESULTS AND DISCUSSION

In this section, we will list the results we obtained using the proposed models and discuss them in detail. Then we will list the results of machine learning classifiers and compare them with the results of the proposed models. Finally, we will make a comparison with related works in literature.

A. Part one (proposed models)

Table III contains the global accuracy, sensitivity, specificity, and ROC curve results obtained after applying the first model (1D-CNN + LSTM). And Table IV has the global accuracy, sensitivity, specificity, and ROC curve results obtained after applying the second model (1D-CNN + BiLSTM). We notice from both tables that there is a relative superiority of the second model over the first model, as shown in Figure 5.

Also, we noticed that if we increased the number of samples of the dataset, the accuracy of both models increased. Whereas the number of samples in DB4 is 31430, it achieved a performance accuracy of 99.90% and 99.82% in both models, respectively. While the number of samples in DB1 is 2821, it achieved a performance accuracy of 96.98% and 97.15% in both models, respectively.

Figures 6-9 show the confusion matrix for some experiments applied to the proposed models. Figures 10-12 show the ROC curve for some experiments applied to the proposed models.

TABLE III. THE RESULTS OF THE FIRST MODEL (1D CNN + LSTM)

Data set	*Global accuracy*	*Sensitivity*	*Specificity*	*ROC*
DB1	96.98%	95.68%	98.06%	96.68%
DB2	94.52%	93.36%	95.42%	94.4%
DB3	99.60%	99.59%	99.61%	99.60%
DB4	99.90%	99.94%	99.86%	99.90%

TABLE IV. THE RESULTS OF THE SECOND MODEL (1D CNN + BILSTM)

Data set	*Global accuracy*	*Sensitivity*	*Specificity*	*ROC*
DB1	97.15%	95.29%	98.70%	97.00%
DB2	95.32%	92.90%	96.95%	94.95%
DB3	99.40%	99.06%	99.73%	99.40%
DB4	99.82%	99.81%	99.83%	99.82%

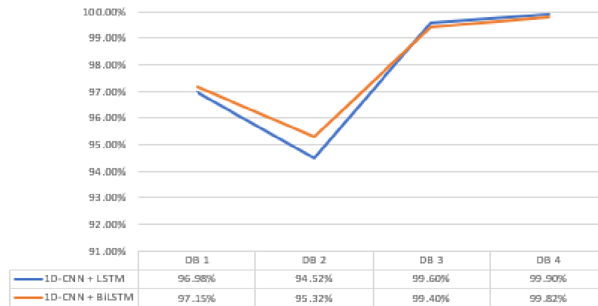

Fig. 5. The performance of hybrid models among all data sets using global accuracy.

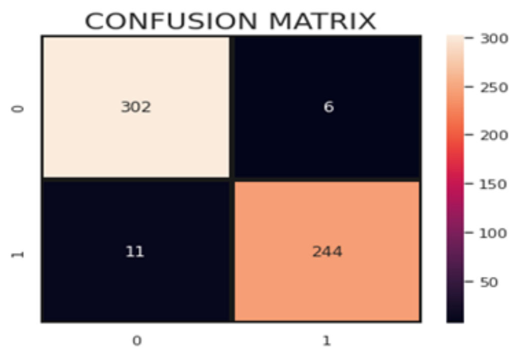

Fig. 6. The confusion matrix of DB1 using the CNN+LSTM model.

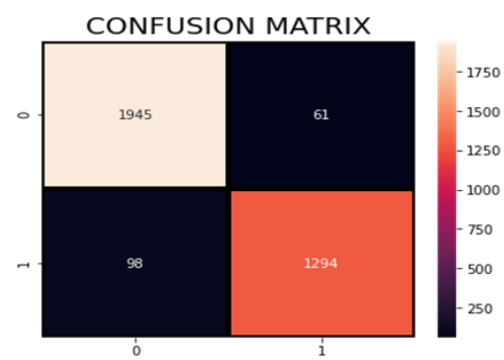

Fig. 7. The confusion matrix of DB2 using CNN +BiLSTM model.

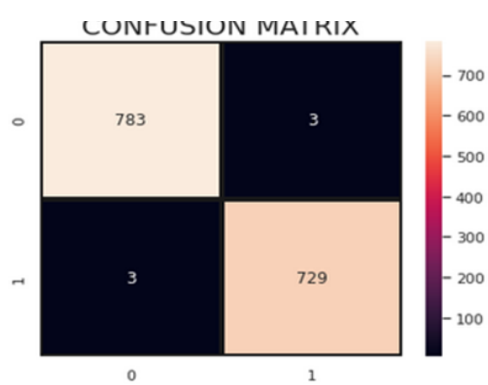

Fig. 8. The confusion matrix of DB3 using CNN +LSTM

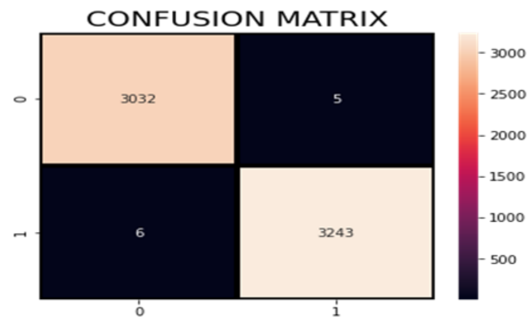

Fig. 9. The confusion matrix of DB4 using CNN +BiLSTM model.

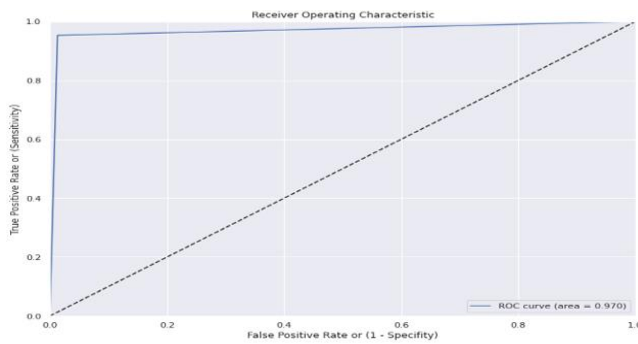

Fig. 10. The ROC curve of DB1 using CNN +BiLSTM model.

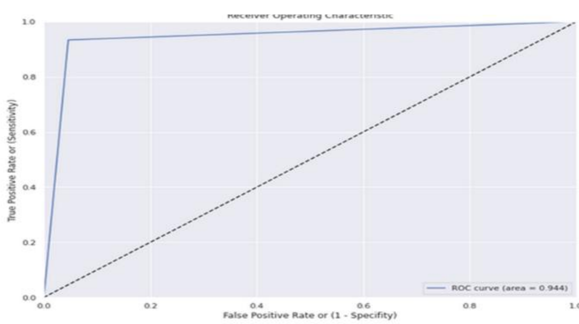

Fig. 11. The ROC curve of DB2 using CNN +LSTM model.

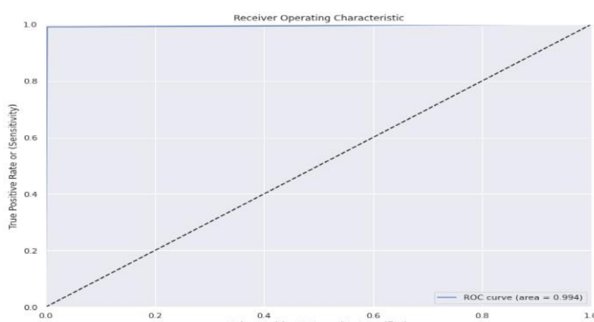

Fig. 12. The ROC curve of DB3 using CNN +BiLSTM model.

B. *Part Two (A comparison with the machine learning classifiers)*

Table V presents the results of the machine learning classifiers applied to the four datasets using three feature extraction methods.

TABLE V. THE RESULTS OF MACHINE LEARNING CLASSIFIERS.

Data set	***classifier***	***Count Vectors***	***TF-IDF Vectors***	***N-Gram Vectors***
DB1	NB	88.80%	91.11%	90.76%
	LR	93.67%	95.69%	94.80%
	KNN	91.65%	91.82%	49.91%
DB2	NB	92.93%	92.93%	89.64%
	LR	92.52%	93.02%	89.81%
	KNN	68.06%	93.29%	52.73%
DB3	NB	96.82%	96.31%	94.92%
	LR	98.34%	98.53%	98.20%
	KNN	69.30%	95.05%	60.54%
DB4	NB	94.11%	93.55%	95.24%
	LR	97.66%	97.34%	96.87%
	KNN	88.45%	68.91%	54.05%

Table V shows that the best performance was obtained using the LR classifier and TF-IDF Vectors method with an average accuracy of 96.15%. And the worst performance was obtained using the KNN classifier and N-Gram Vectors method with an average accuracy of 54.31%. Figure 13 shows the comparison process between the hybrid models and machine learning classifiers. Figure 13 shows that hybrid models are superior to machine learning models.

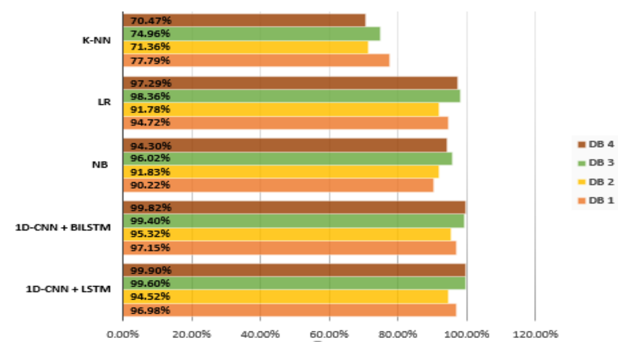

Fig. 13. The comparison process between the proposed models and machine learning classifiers.

C. *Part Three (a comparison with related works in the literature)*

Finally, Table VI compares this study with previous related work in the literature, including the methods and classifiers used in each study.

We can see from Table VI that deep learning techniques can detect fake news more accurately than machine learning algorithms.

TABLE VI. A COMPARISON OF THIS STUDY WITH PREVIOUS STUDIES.

Ref.	***Method***	***Classifier***	***Results***
[8]	Count Vectorizer, Glove vectors	LSTM, BiLSTM, LR	75.07 by LR
[9]	TF-IDF	SVM	93.45%
[10]	BGA, BSSA, BPSO	SVM, k-NN, RF,	75.43% using k-NN-BGA
[11]	BoW, BERT	Graph Neural Networks (GNNs)	ROC=95%

Ref.	***Method***	***Classifier***	***Results***
[12]	BERT, GPT-2, XLNet, RoBERTa, DistilRoBERTa, ALBERT, Bart, and DeBERTa	RCNN, SVM, MLP	97.9% using RoBERTa + MLP
This study	**TF-IDF, BoW, Count Vectors**	**NB, k-NN, LR, 1D-CNN+LSTM, BiLSTM**	**99.90% using 1D-CNN+LSTM**

VI. CONCLUSION

The COVID-19 pandemic is considered the most serious health crisis that has swept the world in the last decade. It paralyzed public life (cancelling gatherings, stopping factories, schools, etc.). The COVID-19 crisis coincided with a hazardous phenomenon of fake news on social media. In this paper, we address the problem of fake news by proposing two-hybrid deep learning models that combine a CNN and an RNN to detect fake news. Our study suggested the effective detection of fake news using two-hybrid deep learning models. The first model combines three CNN layers with two LSTM layers. We replaced the LSTM layers in the second model with two BiLSTM layers. An embedding layer precedes both models with 250 units. Experiments were conducted on four data sets related to fake news about COVID-19. Based on the obtained classification results, the proposed models achieve superior performance with an accuracy of no less than 94.5%.

For future work, we will apply the proposed models to other data sets in other areas. For example, the proposed models can be used to fake news data sets in the business and education sectors.

REFERENCES

[1] P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, "Detection of coronavirus disease (COVID-19) based on deep features and support vector machine," Int. J. Math. Eng. Manag. Sci., vol. 5, no. 4, pp. 643–651, 2020, doi: 10.33889/IJMEMS.2020.5.4.052.

[2] T. D. Pham, "Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning?," Heal. Inf. Sci. Syst., vol. 9, no. 1, 2021, doi: 10.1007/s13755-020-00135-3.

[3] F. Demir, "DeepCoroNet: A deep LSTM approach for automated detection of COVID-19 cases from chest X-ray images," Appl. Soft Comput., vol. 103, p. 107160, 2021.

[4] S. van Der Linden, J. Roozenbeek, and J. Compton, "Inoculating against fake news about COVID-19," Front. Psychol., vol. 11, p. 2928, 2020.

[5] O. D. Apuke and B. Omar, "Fake news and COVID-19: modelling the predictors of fake news sharing among social media users," Telemat. Informatics, vol. 56, p. 101475, 2021.

[6] A. Moscadelli et al., "Fake news and covid-19 in Italy: results of a quantitative observational study," Int. J. Environ. Res. Public Health, vol. 17, no. 16, p. 5850, 2020.

[7] D. Carrion-Alvarez and P. X. Tijerina-Salina, "Fake news in COVID-19: A perspective," Heal. Promot. Perspect., vol. 10, no. 4, p. 290, 2020.

[8] A. Koirala, "COVID-19 Fake News Classification with Deep Learning," Preprint, 2020.

[9] A. Gupta, R. Sukumaran, K. John, and S. Teki, "Hostility detection and covid-19 fake news detection in social media," arXiv Prepr. arXiv2101.05953, 2021.

[10] B. Al-Ahmad, A. Al-Zoubi, R. Abu Khurma, and I. Aljarah, "An Evolutionary Fake News Detection Method for COVID-19 Pandemic Information," Symmetry (Basel)., vol. 13, no. 6, p. 1091, 2021.

[11] A. Hamid et al., "Fake news detection in social media using graph neural networks and NLP Techniques: A COVID-19 use-case," arXiv Prepr. arXiv2012.07517, 2020.

[12] S. M. Shifath, M. F. Khan, and M. Islam, "A transformer based approach for fighting COVID-19 fake news," arXiv Prepr. arXiv2101.12027, 2021.

[13] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," Phys. Rev. E, vol. 71, no. 2, p. 21906, 2005.

[14] N. Boonnak, S. Kamonsantiroj, and L. Pipanmaekaporn, "Wavelet transform enhancement for drowsiness classification in EEG records using energy coefficient distribution and neural network," Int. J. Mach. Learn. Comput., vol. 5, no. 4, pp. 288–293, 2015.

[15] T. Maszczyk and W. Duch, "Comparison of Shannon, Renyi and Tsallis entropy used in decision trees," in International Conference on Artificial Intelligence and Soft Computing, 2008, pp. 643–651.

[16] A. Abdulrahman and M. Baykara, "Fake News Detection Using Machine Learning and Deep Learning Algorithms," in 2020 International Conference on Advanced Science and Engineering (ICOASE), 2020, pp. 18–23.

[17] S. Ghosh and M. S. Desarkar, "Class specific TF-IDF boosting for short-text classification: Application to short-texts generated during disasters," in Companion Proceedings of the The Web Conference 2018, 2018, pp. 1629–1637.

[18] A. Dey, M. Jenamani, and J. J. Thakkar, "Lexical TF-IDF: An n-gram feature space for cross-domain classification of sentiment reviews," in International Conference on Pattern Recognition and Machine Intelligence, 2017, pp. 380–386.

[19] D. M. Abdulqader, A. M. Abdulazeez, and D. Q. Zeebaree, "Machine Learning Supervised Algorithms of Gene Selection: A Review," Mach. Learn., vol. 62, no. 03, 2020.

[20] C.-Y. J. Peng, K. L. Lee, and G. M. Ingersoll, "An introduction to logistic regression analysis and reporting," J. Educ. Res., vol. 96, no. 1, pp. 3–14, 2002.

[21] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", 2003, pp. 986–996.

[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[23] L. Tai and M. Liu, "Deep-learning in mobile robotics-from perception to control systems: A survey on why and why not," arXiv Prepr. arXiv1612.07139, 2016.

[24] F. Rundo, "Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems," Appl. Sci., vol. 9, no. 20, p. 4460, 2019.

[25] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," Neurocomputing, vol. 337, pp. 325–338, 2019.

[26] A. Abdulrahman and M. Baykara, "A Comprehensive Review for Emotion Detection Based on EEG Signals: Challenges, Applications, and Open Issues.," Trait. du Signal, vol. 38, no. 4, 2021.

[27] R. A. Hamad, L. Yang, W. L. Woo, and B. Wei, "Joint Learning of Temporal Models to Handle Imbalanced Data for Human Activity Recognition," Appl. Sci., vol. 10, no. 15, p. 5293, 2020.

[28] M. Massaoudi, S. S. Refaat, H. Abu-Rub, I. Chihi, and F. S. Oueslati, "PLS-CNN-BiLSTM: An end-to-end algorithm-based Savitzky–Golay smoothing and evolution strategy for load forecasting," Energies, vol. 13, no. 20, p. 5464, 2020.

[29] A. A. Ramadhan and M. Baykara, "A Novel Approach to Detect COVID-19: Enhanced Deep Learning Models with Convolutional Neural Networks," Appl. Sci., vol. 12, no. 18, p. 9325, 2022.