MALATYA TURGUT OZAL UNIVERSITY JOURNAL of ENGINEERING and NATURAL SCIENCES Volume 5, Issue 2 (2024) 28-36

NATURENGS
E-ISSN 2717-8013
https://dergipark.org.tr/en/pub/naturengs

Hybrid CNN-LSTM Model for Fake News Detection

Anıl Utku

Munzur University, Department of Computer Engineering, Tunceli Türkiye – 62100

ARTICLE INFO

Received 22.10.2024
Accepted 27.11.2024

Doi: 10.46572/naturengs.1571897

ABSTRACT

In recent years, the way people access information has changed because of the increasingly digital world. Social media has begun to replace traditional news sources such as television and newspapers. Most people reach news about social, economic, and political developments worldwide through social media. Its fast, easy access and cost advantage have made social media widely used among users. In addition to these advantages, social media has become a suitable platform for disseminating fake news. Fake news can have hazardous consequences for individuals, societies, and governments. Therefore, detecting fake news on social media must be necessary. This research created a hybrid CNN-LSTM model for detecting fake news. The CNN component is responsible for analyzing subsequences, which serve as inputs to the LSTM, and extracting relevant features. While the CNN captures critical features from the input data, the LSTM is employed for the classification. The created model was tested with LR, RF, SVM, MLP, and LSTM. The experiments showed that the created model is more successful than the others, with 99.91% accuracy, 99.93% precision, and 99.89% recall. In addition, according to our research, more successful results were obtained in this study than in all studies in the literature using the ISOT dataset.

Keywords: Fake news detection; CNN; LSTM, Deep learning

1. Introduction

In traditional news sources such as newspapers and television, communication is one-way, and the news source is usually apparent. Therefore, it is easier to control them and prevent fake news. With the increase in the use of the Internet and social media, the sources of access to news have begun to vary [1]. Communication has become multi-faceted, and the readers' reactions to the news have also become fast and efficient [2]. Readers can interact with content differently, such as commenting, sharing, and liking/disliking [3]. Readers can also suggest or share news with other users. Since users also become news sources, checking whether all news content is true is impossible.

Since the rise of the Internet, social media has become a primary means for people worldwide to obtain information. Platforms like Twitter and Facebook have grown in popularity in recent years, gradually replacing traditional news outlets such as newspapers, TV, and radio [4]. The main reasons for the widespread use of these platforms are that they are low-cost, easy to access, and spread information quickly [5].

News and comments on social media significantly impact users [6]. The spread of low-quality or false information, called fake news, can negatively affect individuals and societies. Fake news can have dangerous consequences not only for individuals and societies but also for businesses and public administrations [7]. Therefore, fake news on social media must be detected and blocked.

The shift from traditional media to digital platforms has transformed how people access and interact with news. As social media has become more widespread, the ease of disseminating information has led to a flood of content that is difficult to verify. As a result, the challenge of detecting fake news has become a significant concern. This issue has been widely addressed in recent literature, and several studies have proposed advanced artificial intelligence models to address fake news detection.

Ozbay and Alataş [8] presented an application of Grey Wolf Optimization (GWO) and Salp Swarm Optimization (SSO) for Fake News detection. The study compared GWO, SSO, Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), Gradient Boosting, Ridor, J48 and SMO. The experiments showed that GWO outperformed the compared models with 0.875

* Corresponding author. e-mail address: anilutku@munzur.edu.tr
ORCID : 0000-0002-7240-8713

accuracy.

Jiang et al. [9] created a fake news detection system utilizing models such as Logistic Regression (LR), SVM, Convolutional Neural Networks (CNN), k-nearest Neighbors (k-NN), Long Short-Term Memory (LSTM), Decision Trees (DT), Gated Recurrent Units (GRU), and Random Forest (RF). Experiments show that the RF model is the most successful, with 99.94% test accuracy for the ISOT dataset. In contrast, the LR model is the most successful, with 96.05% accuracy for the KDnuggets dataset.

Abdul Nasir et al. [10] designed a hybrid fake news detection system that combines CNN and RNN models. The system was tested on the ISOT and FA-KES datasets. The performance of the created model was compared with several others, including RF, Stochastic Gradient Descent (SGD), LR, Multinomial Naive Bayes (MNB), k Nearest Neighbour (kNN), DT, Recurrent Neural Networks (RNN), AdaBoost (AB), and CNN. The results demonstrated that the model achieved 60% and 99% accuracy on the ISOT and FA-KES datasets.

Goldani et al. [11] have presented an applied analysis of CNN models using different embedding models and margin loss. They compared static and non-static word embedding approaches. The experimental studies using the ISOT and LIAR datasets showed an improvement of 7.9% for the ISOT dataset and 2.1% for the LIAR dataset.

Goldani et al. [12] have presented an application of capsule neural networks for fake news detection. In the study, different embedding methods were used for news of different lengths. Static embedding methods are used for short news, while non-static embedding methods are used for longer news. Experimental results showed that 7.8% performance improvement was achieved in the ISOT dataset and 1% in the LIAR dataset.

Alameri and Mohd [13] have detected fake news using LSTM, NB, a Neural Network with Keras and TensorFlow, and SVM. The experiments showed that LSTM outperforms other models, with an average accuracy of 94.21%.

Ozbay and Alatas [14] have created a Salp Swarm Optimization (SSO) method. They have applied standard SSO, Gray Wolf Optimization (GWO), and two adaptive SSO algorithms. The experiments showed that the created model outperforms the standard SSA and GWO.

Rajalaxmi et al. [15] created an optimized LSTM-based model for predicting fake news. The model's parameters were fine-tuned using hyper-parameter optimization techniques. Experiments indicated that the model achieved 99.65% and 45.23% accuracy on the ISOT and LIAR datasets.

Yıldırım [16] proposed the hybrid CNN-LSTM model for COVID-19 Fake News detection. The proposed model was compared with kNN, LR, NB, RF, Gradient Boosting, Discriminant Analysis and XGB. Experiments showed that the proposed model has 99.42% accuracy.

Yildiz et al. [17] proposed a hybrid CNN-LSTM model for the detection of chronic kidney disease. In the study, the proposed model was compared with kNN, SVM, LR, NB, RF, and AdaBoost. Experiments showed that the proposed model has 99.17% accuracy.

The key novelty of this study to the literature are as follows:

- Hybrid CNN-LSTM model was created for fake news detection.
- The created model was comprehensively compared with traditional classification models.
- Experimental results demonstrated that the created model achieved superior performance.
- It is seen that the proposed CNN-LSTM model produces more effective results thanks to its advantages in both feature extraction and processing sequential data.
- In the proposed model, CNN analyzes subsequences and extracts important features that can be effective in fake news detection, while LSTM processes these features extracted by CNN and performs classification.

2. Classification Models

LR is a statistical technique used to analyze datasets with one or more independent variables influencing an outcome [18]. It is primarily applied in binary classification, where the dependent variable can take only two distinct values. LR applies the logistic function to convert predicted values into probabilities confined within 0 to 1. This characteristic makes it particularly useful for binary classification tasks, where the goal is to determine the probability of one of two possible outcomes. While it is mainly used for binary classification, variations like multinomial logistic regression can handle multiple classes [19].

RF leverages the power of multiple decision trees to improve predictive performance. It is an ensemble method that combines the predictions of several individual trees, each built on a random sample of the training data [20]. Each tree is constructed using a different subset of the dataset, which helps to reduce the risk of overfitting that can occur with a single decision tree. When making predictions, RF aggregates the outputs of all the trees, typically through majority voting for classification tasks or averaging for regression tasks, leading to more accurate and stable results [21].

SVM scenarios by constructing a feature space, a finite-dimensional vector space where each dimension corresponds to a specific feature of an object. SVM aims to create a model that classifies new, unseen objects into distinct categories. It divides the feature space linearly into two categories, positioning an object above or below the hyperplane based on its features [22].

MLP is an advantageous model for tasks where data can be transformed through nonlinear relationships. MLP is a neural network model with a layered structure, and each node in the layers is fully connected to neurons in neighboring layers [23]. MLP processes the input data by applying biases, weights, and activation functions to learn complex patterns. MLP is used in classification, regression, and prediction problems in application areas such as medicine, environment, and finance [24].

LSTM is a variant of recurrent neural networks (RNNs) developed for processing sequential data. LSTM stands out in capturing long-term dependencies in data such as time series [25]. LSTM has memory cells that allow information in the network to be stored for long periods [26]. The input gate controls how much new information enters the cell. The forget gate determines which information is discarded. The output gate decides which information is passed to the next layer [27]. LSTM overcomes the limitations of RNNs, as they often handle long-term dependencies due to issues such as vanishing gradients. Their ability to dynamically remember and forget information allows them to be successful in tasks where temporal context is essential [28].

3. Deep Learning Based Fake News Detection

Social media has become an increasingly popular news source as the Internet has advanced. Individuals and organizations have extensively used online news websites and social media platforms to access information. Because it is faster, easier, and cheaper to access information, people widely use online news sources. Confirming this large amount of news data or determining whether there is valuable information to be extracted from the data has reached a level that cannot be overcome with classical methods or humans today. Consequently, employing artificial intelligence techniques for rapidly and accurately detecting fake news has emerged as a significant research field.

3.1. Dataset

In this study, the ISOT fake news dataset provided by Victoria University has been used to train and test the model [29]. The dataset consists of real news obtained through reuters.com and fake news tagged by PolitiFact. It contains content on various topics, mainly politics and world news. As shown in Fig. 1, 21417 real-labeled and 23481 fake-labeled news articles are in the dataset.

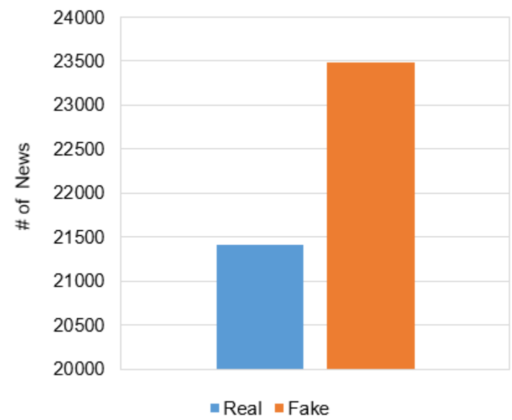

Figure 1. News articles in the dataset

Table 1 demonstrates the number of news in the dataset on the basis of the relevant topic. There are 44.898 news in total.

Table 1. Number of news articles in the dataset according to their content

Content	Number of news
World	10.145
Politics	11.272
Government	1.570
Middle east	778
US	783
Left	4.459
Politics	6.841
News	9.050
Total	44.898

Fig. 2 shows a sample from the dataset.

	title	text	subject	date
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

Figure 2. Sample data from the ISOT dataset

As seen in Fig. 2, the dataset consists of title, text, subject, and date columns. Fig. 3 shows the distribution of the frequency of characters that are used in the titles of real and fake news.

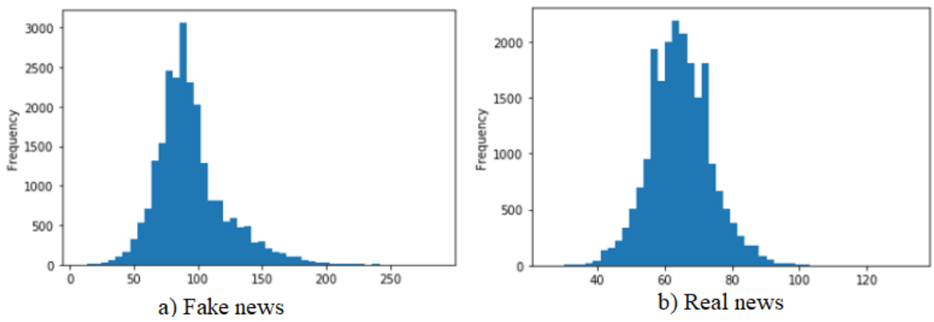

Figure 3. Distribution of character numbers used in fake news titles

As it can be seen in Fig. 3, shows that the average number of characters used in the titles of fake news is higher than real ones. Fig. 4 shows unique words distributions.

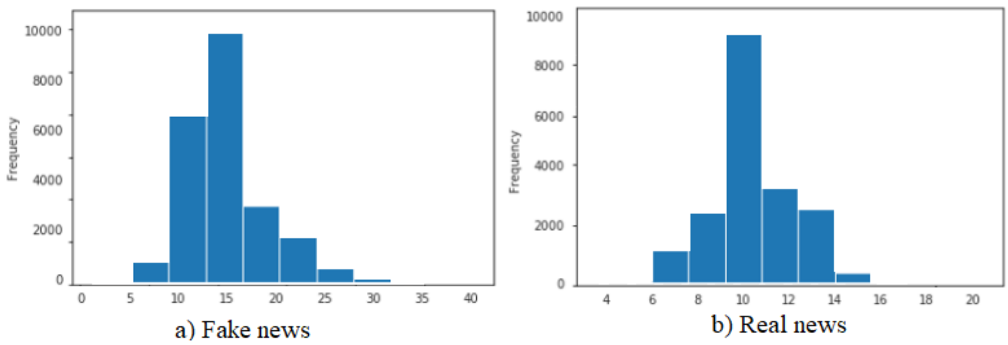

Figure 4. Distribution of unique words used in news titles

As it can be seen in Fig. 4, the fake news consists of more unique words than real ones.

3.2. Created model

The developed fake news detection system consists of data preparation and model creation phases. The data preparation process consists of data cleaning, augmentation, pre-processing, and loading embedding words.

After the meticulous data cleaning process, the data augmentation process is thoroughly implemented. This reassures the robustness of the model. During the data pre-processing phase, all characters were converted to lowercase, and all punctuation was eliminated. The NLTK library was then utilized to carry out tokenization, lemmatization, and the removal of stop words. Text data has been converted to vector format and Skip-Gram architecture has been embedded using Tensorflow's wiki-words-250. The dataset was organized by splitting 80% for training and 20% for testing. Within the training data, 90% was utilized for the training process, while 10% was designated for validation. The validation data was utilized to optimize the model parameters. GridSearchCV was employed to identify the best model parameters. In the created model, binary cross entropy was used as the loss function. The optimization function is Adam and the evaluation metric is accuracy. Fig. 5 shows the created model architecture.

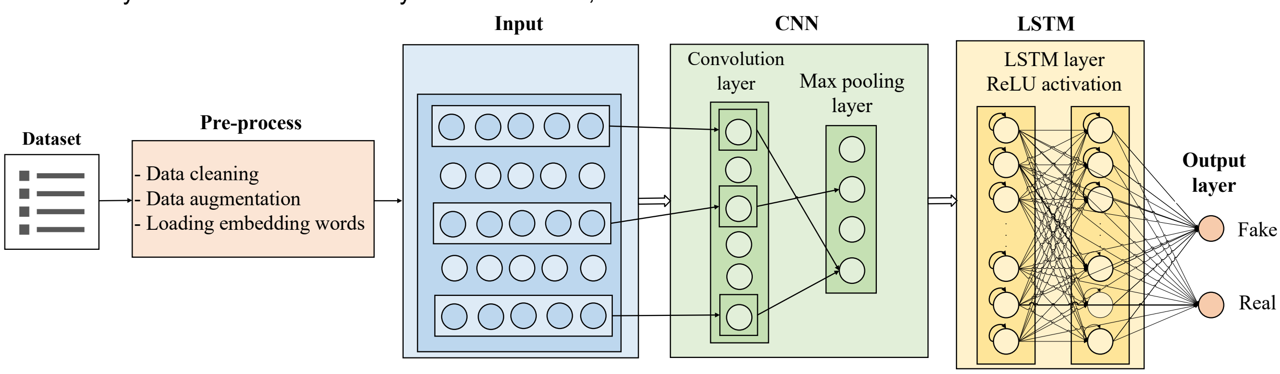

Figure. 5. The created model architecture

As seen in Fig. 5, the developed model was created using CNN and LSTM as a hybrid model. CNN was used to extract meaningful features in the text. LSTM was used to learn long-term dependencies in the data. Using 1D convolution, features carried by news texts in short contexts were extracted. LSTM takes feature maps extracted by CNN as input. LSTM was used to model long-term dependencies of word order and content structure in news texts. LSTM incorporates feedback loops from prior iterations, encoding contextual information related to a temporal sequence. For CNN, the number of filters is 64, kernel size is 3, pool size is 2, and activation function is ReLU. For LSTM, the number of neurons is 128, dropout rate is 0.2, optimizer is Adam, epoch number is 50, and batch size is 32. For a given input sequence the hidden states h_t and outputs y_t can be computed as follows:

$$h_t = H(W_{ih}F_t + W_{hh}h_{t-1} + b_h) \quad (1)$$

$$yt = W_{ho}h_t + b_o \quad (2)$$

where, W_{ih}, W_{hh}, W_{ho} are weight matrices between input, hidden and output layers. Basically, for x_t input at time *t*, long-term memory C_{t-1} and transaction memory h_{t-1} are transferred from the previous states to time *t*. The created model uses the following Eqns. for learning and prediction:

$$i_t = \sigma\left(w_{xi}\,\mathrm{x}_t + w_{hi}h_{t-1} + w_{ci}c_{t-1} + \mathrm{b}_i\right) \quad (3)$$

$$f_t = \sigma\left(w_{xf}\,\mathrm{x}_t + w_{hf}h_{t-1} + w_{cf}c_{t-1} + \mathrm{b}f\right) \quad (4)$$

$$c_t = \mathrm{f}_t \odot c_{t-1} + i_t \odot \tanh(w_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma\left(w_{xo}\,\mathrm{x}_t + w_{ho}h_{t-1} + w_{co}c_{t-1} + \mathrm{b}_o\right) \quad (6)$$

$$h_t = \mathrm{o}_t \odot h(c_t) \quad (7)$$

Here, i_t denotes the input gate, f_t represents the forget gate, o_t indicates the output gate, *c* is the cell activation vector, *b* signifies the bias vector, and *w* refers to the weight matrix. ʘ represents the element-wise product of two vectors. Additionally, x_t and h_t denote the input-output sequences, while g_t refers to the input string.

The created model integrates CNN layers with LSTM to extract features from the input data effectively. The process begins with the CNN, which identifies essential information and organizes it into multi-dimensional arrays through convolutional operations.

Once the CNN has processed the input, these multi-dimensional representations are then inputted into the LSTM for classification purposes. Within this architecture, the CNN focuses on feature extraction, while the LSTM handles the analysis and categorization of the features provided by the CNN. After the CNN processes the data, it forwards these subarray samples to the LSTM in a seamless data flow.

The architecture also includes a max-pooling layer that enhances the interpretation of the extracted features following the convolutional layer. This is succeeded by a dense layer that further refines the features obtained from the convolutional process. Since both the convolutional and pooling layers generate 3D data, a flattening layer converts the resulting feature maps into a one-dimensional vector. This vector then acts as the input for the LSTM, facilitating effective classification.

3.3. Performance Assessment Metrics

Classification models are designed to categorize binary or multiclass categorical values. Typically, metrics such as accuracy, recall, precision, and F1-score are utilized to assess the errors that arise and effectively determine the accuracy rates. These evaluation metrics are derived from the Confusion Matrix (CM) values. Fig. 6 displays the CM for binary classification.

		Actual values	
		Positive (1)	Negative (0)
Predicted values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 6. CM for binary classification

False Positive (FP) occurs when the model mistakenly labels a genuine news article as fake. True Positive (TP) refers to instances where the model correctly identifies fake news articles as fake. True Negative (TN) indicates the number of actual news articles the model correctly classifies as accurate. False Negative (FN) occurs when the model fails to identify a piece of fake news, labeling it as accurate instead.

Accuracy represents the overall correctness of the model in classifying news articles as fake or real, as shown in Eq. 8.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+FN+TN}} \quad (8)$$

Precision, recall, and F-score are key evaluation metrics used in classification tasks. Precision indicates how many of the news predicted as fake are actually fake. Precision is calculated using Eq. 9.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (9)$$

Recall indicates how many of the actual fake news were correctly identified by the model. Recall is computed using Eq. 10.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (10)$$

The F1-score is especially useful when there’s an imbalance between false positives and false negatives. This makes it a balanced metric for evaluating classification performance when precision and recall are equally important. The F1-score is computed as shown in Eq. 11.

$$F1-score = \frac{2.Recall.Precision}{Recall+Precision} \quad (11)$$

3.4. The Experimental Results

This paper presents a comparative analysis of the CNN-LSTM model, which hybridizes CNN and LSTM to detect fake news content, with LR, LSTM, RF, MLP, and SVM. GridSearchCV is used to optimize the parameters of the models. Using grid search, the most appropriate hyper-parameters with the highest accuracy value were selected for each model. Cross-validation is used to prevent over-fitting problems. A 10-fold cross-validation is used, and each model is tested on ten randomly generated datasets. The results of the experiments for each compared model are presented in Table 2.

Table 2. Experimental outcomes obtained for each model and run step according to the accuracy metric

CV	LR	RF	SVM	MLP	LSTM	CNN-LSTM
1	0.9860	0.9897	0.9958	0.9884	0.9962	0.9991
2	0.9858	0.9913	0.9956	0.9886	0.9974	0.9992
3	0.9881	0.9911	0.9958	0.9883	0.9979	0.9992
4	0.9866	0.9904	0.9960	0.9889	0.9969	0.9991
5	0.9855	0.9908	0.9957	0.9886	0.9969	0.9992
6	0.9864	0.9906	0.9958	0.9882	0.9969	0.9989
7	0.9866	0.9898	0.9961	0.9891	0.9968	0.9991
8	0.9870	0.9904	0.9958	0.9895	0.9966	0.9990
9	0.9875	0.9909	0.9955	0.9886	0.9969	0.9991
10	0.9874	0.9892	0.9959	0.9881	0.9965	0.9991
Average	0.9866	0.9904	0.9958	0.9886	0.9969	0.9991

The CM for LR can be seen in Fig. 7.

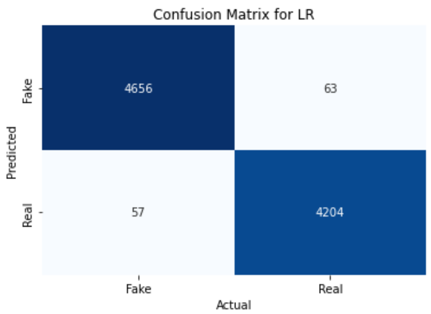

Figure 7. CM for LR

As seen in Fig. 7, LR successfully identified 4656 fake news samples and 4204 real news samples. In total, LR accurately classified 8860 news items, while 120 items were misclassified.

The CM for RF is presented in Fig. 8.

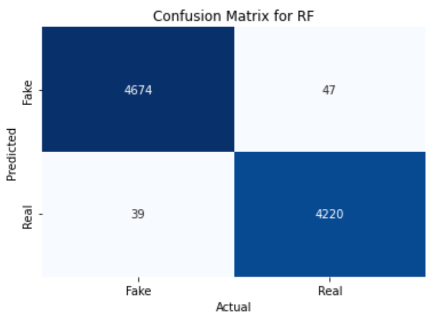

Figure 8. CM for RF

As seen in Fig. 8, RF accurately classified 4674 fake news samples and 4220 real news samples. Overall, RF successfully identified 8894 news items, with 86 instances misclassified.

The CM for SVM can be found in Fig. 9.

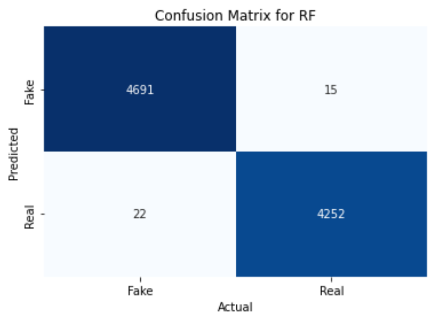

Figure 9. CM for SVM

As seen in Fig. 9, SVM accurately classified 4691 fake news samples and 4252 real news samples. In total, SVM correctly identified 8943 news items, with 37 cases misclassified.

The CM for MLP is presented in Fig. 10.

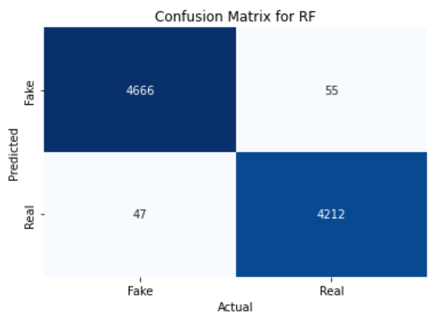

Figure 10. CM for MLP

As seen in Fig. 10, MLP model accurately classified 4666 fake news samples and 4212 real news samples. In total, the MLP successfully categorized 8878 news items, with 102 instances misclassified.

The CM for LSTM can be seen in Fig. 11.

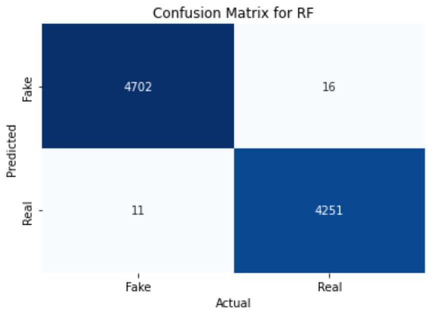

Figure 11. CM for LSTM

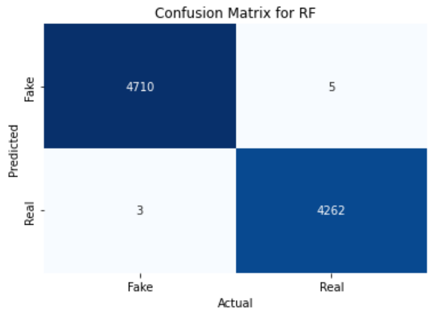

Figure 12. CM for CNN-LSTM

As seen in Fig. 11, the model accurately identified 4702 fake news samples and 4251 real news samples. The LSTM successfully classified a total of 8953 news items, with only 27 misclassified.

The CM for the CNN-LSTM is presented in Fig. 12.

As seen in Fig. 12, 4702 fake news samples and 4251 real news samples were accurately identified by the model. In total, the CNN-LSTM successfully classified 8953 news items, while only 27 were incorrectly classified.

The comparative experimental results are presented in Table 3 and Fig. 13.

Table 3. Comparative experimental results

Model	Accuracy	Precision	Recall	F1-score
LR	0.9866	0.9879	0.9866	0.9872
RF	0.9904	0.9889	0.9900	0.9899
SVM	0.9958	0.9953	0.9968	0.9960
MLP	0.9886	0.9900	0.9883	0.9891
LSTM	0.9969	0.9976	0.9966	0.9970
CNN-LSTM	0.9991	0.9993	0.9989	0.9990

According to Table 3 and Fig. 13, the CNN-LSTM outperforms the other models in comparison. The accuracy of the CNN-LSTM is 0.9991, with 0.9993 precision, 0.9989 recall, and 0.9990 F1-score.

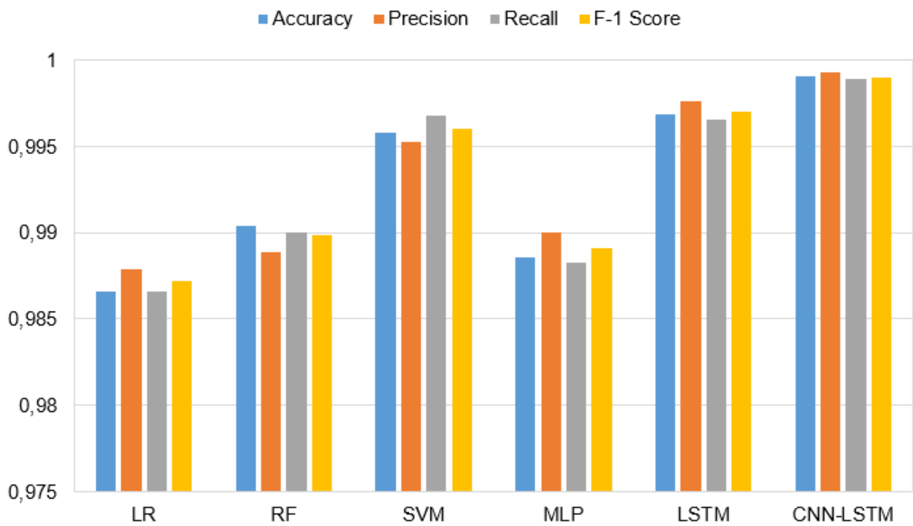

Figure 13. Comparative experimental results

Table 3 and Fig. 13 illustrate that the CNN-LSTM demonstrated superior classification performance in detecting fake news compared to other models. Following the CNN-LSTM, the LSTM, SVM, RF, MLP, and LR models achieved the next best results, respectively.

Fig. 14 displays the accuracy/loss graphs of the CNN-LSTM.

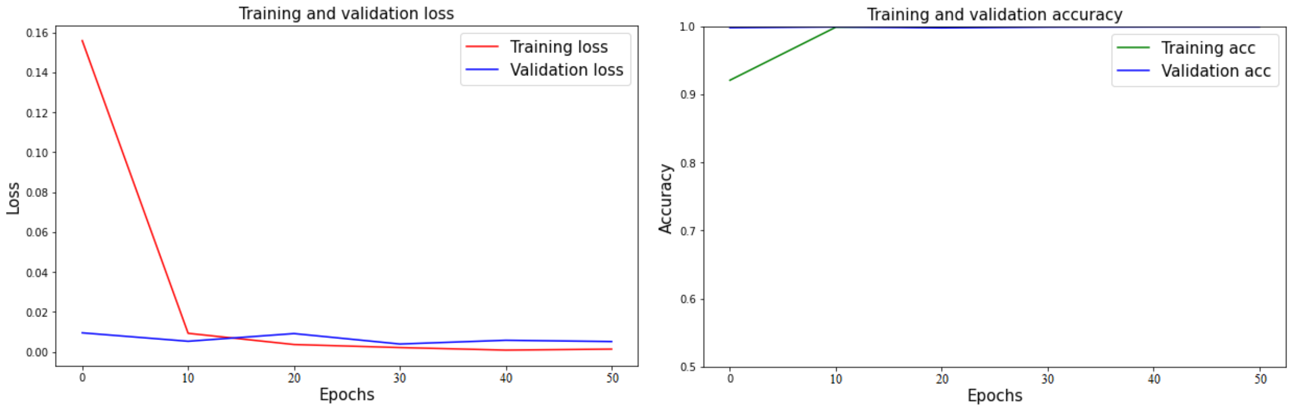

Figure 14. Training and validation loss/accuracy graphs

4. Conclusions

This research created a hybrid deep learning model combining CNN and LSTM architectures to predict fake news. The model's performance was evaluated against well-known machine learning and deep learning models, including LR, RF, SVM, MLP, and LSTM. Experimental evaluations were conducted using the ISOT dataset, focusing on accuracy, precision, recall, and F1-score metrics. The results unequivocally demonstrate the superiority of the proposed model, outperforming the others with an accuracy of 99.91%, precision of 99.93%, and recall of 99.89%.

The created model's superior performance can be attributed to CNN's effectiveness in feature extraction and LSTM's proficiency in processing sequential data. The findings indicate that SVM exhibits better classification performance compared to RF. It is suggested that RF may perform more effectively on structured data, such as audio, video, and text, which may explain MLP's relatively lower performance than RF. The categorical features present in the dataset utilized in this study contributed to RF's enhanced performance over MLP.

References

[1] **Safori, A. O.** (2018). Social Media's Impact on a Journalist's role. *Journal of science education*, 19(1), 148-62.

[2] **Flintham, M.,** Karner, C., Bachour, K., Creswick, H., Gupta, N., and Moran, S. (2018). Falling for fake news: investigating the consumption of news via social media. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1-10.

[3] **Dutceac Segesten, A.,** Bossetta, M., Holmberg, N., and Niehorster, D. (2022). The cueing power of comments on social media: how disagreement in Facebook comments affects user engagement with news. *Information, Communication & Society*, 25(8), 1115-1134.

[4] **van Erkel, P. F.,** and Van Aelst, P. (2021). Why don't we learn from social media? Studying effects of and mechanisms behind social media news use on general surveillance political knowledge. *Political Communication*, 38(4), 407-425.

[5] **Mallik, A.,** and Kumar, S. (2024). Word2Vec and LSTM based deep learning technique for context-free fake news detection. *Multimedia Tools and Applications*, 83(1), 919-940.

[6] **Colliander, J.** (2019). This is fake news: Investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media. *Computers in Human Behavior*, 97, 202-215.

[7] **Mutsvairo, B.,** and Bebawi, S. (2019). Journalism educators, regulatory realities, and pedagogical predicaments of the "fake news" era: A comparative perspective on the Middle East and Africa. *Journalism & Mass Communication Educator*, 74(2), 143-157.

[8] **Ozbay, F. A.,** and Alatas, B. (2019). A novel approach for detection of fake news on social media using metaheuristic optimization algorithms. *Elektronika ir Elektrotechnika*, 25(4), 62-67.

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

[10] **Nasir, J. A.,** Khan, O. S., and Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, 1(1), 100007.

[11] **Goldani, M. H.,** Safabakhsh, R., and Momtazi, S. (2021). Convolutional neural network with margin loss for fake news detection. *Information Processing & Management*, 58(1), 102418.

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

[13] **Alameri, S. A.,** and Mohd, M. (2021, January). Comparison of fake news detection using machine learning and deep learning techniques. In 2021 3rd International Cyber Resilience Conference (CRC), 1-6.

[14] **Ozbay, F. A.,** and Alatas, B. (2021). Adaptive salp swarm optimization algorithms with inertia weights for novel fake news detection model in online social media. *Multimedia Tools and Applications*, 80(26), 34333-34357.

[15] **Rajalaxmi, R. R.,** Narasimha Prasad, L. V., Janakiramaiah, B., Pavankumar, C. S., Neelima, N., and Sathishkumar, V. E. (2022). Optimizing Hyperparameters and Performance Analysis of LSTM Model in Detecting Fake News on Social media. *Transactions on Asian and Low-Resource Language Information Processing*.

[16] **Yildirim, M.** (2022). Detection of COVID-19 Fake News in Online Social Networks with the Developed CNN-LSTM Based Hybrid Model. *Review of Computer Engineering Studies*, 9(2).

[17] **Yildiz, E. N.,** Cengil, E., Yildirim, M., and Bingol, H. (2023). Diagnosis of chronic kidney disease based on CNN and LSTM. Acadlore transactions on ai and machine learning, 2(2), 66-74.

[18] **Liang, D.,** Tsai, C. F., and Wu, H. T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, 73, 289-297.

[19] **Farid, H.,** Izadi, Z., Ismail, I. A., and Alipour, F. (2015). Relationship between quality of work life and organizational commitment among lecturers in a Malaysian public research university. *The Social Science Journal*, 52(1), 54-61.

[20] **Utku, A.** (2024). Hindistan'daki Turistik Şehirlerin İklim Değişkenlerinin Tahminine Yönelik Hibrit ConvGRU Modeli. *Mühendislik Bilimleri ve Araştırmaları Dergisi*, 6(2), 165-176.

[21] **Yang, Y.,** and Chen, W. (2016). Taiga: performance optimization of the C4. 5 decision tree construction algorithm. *Tsinghua Science and Technology*, 21(4), 415-425.

[22] **Wang, H.,** and Shao, Y. (2024). Fast generalized ramp loss support vector machine for pattern classification. *Pattern Recognition*, 146, 109987.

[23] **Fu, S.,** and Avdelidis, N. P. (2024). Novel Prognostic Methodology of Bootstrap Forest and Hyperbolic Tangent Boosted Neural Network for Aircraft System. *Applied Sciences*, 14(12), 5057.

[24] **Muhammad Ehsan, R.,** Simon, S. P., and Venkateswaran, P. R. (2017). Day-ahead forecasting of solar photovoltaic output power using multilayer perceptron. *Neural Computing and Applications*, 28(12), 3981-3992.

[25] **Utku, A.** (2023). Deep learning based hybrid prediction model for predicting the spread of COVID-19 in the world's most populous countries. *Expert Systems with Applications*, 231, 120769.

[26] **Mirzaei, S.,** Kang, J. L., and Chu, K. Y. (2022). A comparative study on long short-term memory and gated recurrent unit neural networks in fault diagnosis for chemical processes using visualization. *Journal of the Taiwan Institute of Chemical Engineers*, 130, 104028.

[27] **Al-Selwi, S. M.,** Hassan, M. F., Abdulkadir, S. J., Muneer, A., Sumiea, E. H., Alqushaibi, A., and Ragab, M. G. (2024). RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. *Journal of King Saud University-Computer and Information Sciences*, 102068.

[28] **Kaya, Y.,** Yiner, Z., Kaya, M., and Kuncan, F. (2022). A new approach to COVID-19 detection from X-ray images using angle transformation with GoogleNet and LSTM. *Measurement Science and Technology*, 33(12), 124011.

[29] **Traore I.,** Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. Lecture Notes in Computer Science, 10618. Springer, 127- 138.