Journal of Universal Computer Science, vol. 29, no. 5 (2023), 491-507
submitted: 26/8/2022, accepted: 8/2/2023, appeared: 28/5/2023

Developed Models Based on Transfer Learning for Improving Fake News Predictions

Tahseen A. Wotaifi
(University of Babylon, Babil, Iraq
https://orcid.org/0000-0002-3833-3885, tahseen.ali@uobabylon.edu.iq)

Ban N. Dhannoon
(Al-Nahrain University, Baghdad, Iraq
https://orcid.org/0000-0002-8474-5565, ban.n.dhannoon@nahrainuniv.edu.iq)

Abstract: In conjunction with the global concern regarding the spread of fake news on social media, there is a large flow of research to address this phenomenon. The wide growth in social media and online forums has made it easy for legitimate news to merge with comprehensive misleading news, negatively affecting people's perceptions and misleading them. As such, this study aims to use deep learning, pre-trained models, and machine learning to predict Arabic and English fake news based on three public and available datasets: the Fake-or-Real dataset, the AraNews dataset, and the Sentimental LIAR dataset. Based on GloVe (Global Vectors) and FastText pre-trained models, A hybrid network has been proposed to improve the prediction of fake news. In this proposed network, CNN (Convolution Neural Network) was used to identify the most important features. In contrast, BiGRU (Bidirectional Gated Recurrent Unit) was used to measure the long-term dependency of sequences. Finally, multi-layer perceptron (MLP) is applied to classify the article news as fake or real. On the other hand, an Improved Random Forest Model is built based on the embedding values extracted from BERT (Bidirectional Encoder Representations from Transformers) pre-trained model and the relevant speaker-based features. These relevant features are identified by a fuzzy model based on feature selection methods. Accuracy was used as a measure of the quality of our proposed models, whereby the prediction accuracy reached 0.9935, 0.9473, and 0.7481 for the Fake-or-Real dataset, AraNews dataset, and Sentimental LAIR dataset respectively. The proposed models showed a significant improvement in the accuracy of predicting Arabic and English fake news compared to previous studies that used the same datasets.

Keywords: Fake News; Pre-trained Model; Hybrid Network; Improved Random Forest; Fuzzy Model
Categories: J, L
DOI: 10.3897/jucs.94081

1 Introduction

Fake news can be described as any material capable of making the reader (user) believe that the information is incorrect [Mishra et al., 2022]. Most information that tends to be misleading news is derived from articles and stories that are intentionally published to harm internet users. As a result, these rumours or misinformation are created and spread via social media or the internet, leading to other users believing these falsehoods and propagating them themselves [Aljwari et al., 2022]. Some studies have revealed that the spread of fake news is more than 100 times faster than the spread of real news. The reason is that the publisher chooses important and sensitive topics and releases them at the right time [Qayyum et al., 2019]. The spread of false information has harms and threats on a large scale, as this news gives incorrect impressions to individuals and society. In addition, the psychological effects caused by the spread of fake news in societies may threaten countries' security, civil wars, and many others [Fayaz et al., 2022].

Fake news detection has recently caught the attention of many researchers and academics [Mishra et al., 2022]. With the absence of censorship on the internet, detecting fake news is a considerable challenge. In addition, identifying fake news on social media and stopping its spread is a difficult task even for humans because this requires a complicated activity to collect evidence and sift the facts before making a decision [Elsaeed et al., 2021]. Even with some of the topics presented and available on the Internet, it is difficult for humans to detect fake news manually. One of the reasons behind the inability of experts and specialists to identify misinformation is the large volume and speed of information flow on social media platforms and the Internet. Therefore, recent research efforts have focused on the development in the field of Machine Learning (ML) and Natural Language Processing (NLP) to automate and classify news articles [Upadhayay and Behzadan, 2020].

Many current studies rely on the style in detecting fake news, that is, focusing on the style of writing text content as features in classifying news articles [Pan et al., 2018]. Although this approach is effective, it cannot determine what is fake in a targeted news article. On the other hand, the process of detecting fake news based on the content of the text (also called fact-checking) is more promising [Agarwal et al., 2019]. As such, this study aims to detect fake news based on content through NLP. This study introduces three publicly available datasets, two focusing on English fake news and the other on Arabic fake news. Pretrained models are included with both machine learning techniques and a hybrid deep neural network (see methodology section). Our main contributions are the following:

- Proposing a developed selection model to identify the most important speaker-based features using fuzzy logic.
- Modifying the behaviour of the random forest algorithm by controlling the selection of features to build trees instead of random selection.
- Building a hybrid neural network based on pre-trained models to improve prediction accuracy.

The remainder of the study is organized as follows: Section 2 presents related works that dealt with English and Arabic fake news. Technical background, relevant literature, and an overview of the datasets are explained in Section 3. Section 4 describes the methodology and building of machine learning and deep learning models. In Section 5, the results are presented and discussed. Finally, Section 6 shows the conclusions of this study.

2 Related Work

Many researchers have tried to solve the problem of fake news and limit its spread by building prediction models. Some studies have discussed the challenges of fake news detection from the perspective of NLP and data mining. This approach relies on content-based and user-based features and then adapts machine learning techniques. A

Recurrent Neural Network (RNN) is the first deep learning technique considered to solve this problem. This network is developed by deriving other improved versions such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU). Recently, researchers have contributed to a new development of NLP by building so-called pre-trained models or transfer learning (e.g., GloVe, BERT, FastText, and others) and adapting Convolutional Neural Networks (CNN) to deal with texts. On the other hand, other studies have proven that the complexity of the model is not always the best solution. Therefore, simple models may outperform complex models if appropriate parameters and necessary data are selected. For all of the above, this study presents some of the works that strived to reduce the problem of spreading fake news in our study. Some related studies were conducted on the datasets used in our study, and the results found that our proposed models outperformed these studies.

In [Mishra et al., 2022], the authors used three datasets, including the LIAR dataset. This dataset contains speaker-based features and sentence in each record, so several machine learning and deep learning techniques were used to predict fake news. The best accuracy achieved for the models built on the LIAR dataset is (0.60) and reached (0.93) for the other datasets used in the study.

As for the work by [Aljwari et al., 2022] the authors applied machine learning techniques based on the AraNews dataset to predict Arabic fake news. In this study, the Term Frequency-Inverse Document Frequency was used to represent words into features (word vectors). Among the techniques used, the best accuracy obtained was 0.866 through the Random Forest Model.

In [Fayaz et al., 2022], the authors utilized the ISOT dataset that contains 23 features to detect misinformation. Four feature selection techniques were used to identify the relevant features. Through these techniques, 14 features were identified out of all the features. Finally, the prediction process was performed using machine learning techniques, and the random forest model achieved the best accuracy.

As for the study by [Upadhayay and Behzadan, 2020], the authors built the models based on the Sentimental LIAR dataset. A convolution Neural Network (CNN) based on BERT base pre-trained model has been proposed to detect fake news. Although the study presented a hybrid and valuable method, the results were not up to the level of ambition.

3 Theoretical Background

3.1 Deep Learning

Deep learning is the most influential branch of artificial intelligence as it processes unstructured data [Alshdaifat et al., 2022]. Based on artificial neural networks, the philosophy of deep learning techniques is to distribute tasks between layers, where each layer is responsible for a specific function. There are many techniques involved in deep learning; the most important are CNN and RNN [Al-Hmouz, 2020].

3.1.1 Bidirectional Gated Recurrent Unit

GRU is a development of an RNN and is similar to LSTM in terms of the forget gate, but the number of parameters is less [Becerra-Rico et al., 2020]. The architecture of this network is very suitable for text processing and deals with small datasets. Also, the problem of vanishing gradient related to the RNN was solved through the GRU network. Figure 1 represents the GRU framework, whereas Equations 1, 2, 3, and 4 show the output of the hidden and output neurons in this network [Demirel et al., 2021].

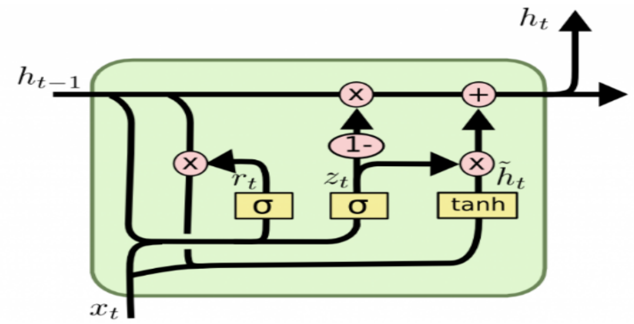

Figure 1: BiGRU Framework

$$Update\ Gate: z_t = \sigma(W_z . x_t + U_z . h_{t-1}) \quad (1)$$
$$Reset\ Gate: r_t = \sigma(W_r . x_t + U_r . h_{t-1}) \quad (2)$$
$$Hidden\ State: \hat{h}_t = tanh(W x_t + r_t . U h_{t-1}) \quad (3)$$
$$The\ output: h_t = (1 - z_{t)}. h_{t-1} + z_t . \hat{h}_t \quad (4)$$

Where: x is the input vector. W and U are weight vectors.

3.1.2 Convolution Neural Network

CNN is a deep neural network commonly used in image processing and computer vision. Due to the characteristics of this network, researchers have recently been interested in adapting it for NLP and text classification [Hussien and Dhannoon, 2020]. The importance of the CNN in NLP is to extract important features from the text through the convolution layer and the max-pooling layer. Briefly, text classification through this network is achieved in four steps: 1) converting words into vectors through embedding techniques; 2) identifying kernel and obtaining the local features by convolution layer; 3) selecting the most important features through max-pooling layer; 4) applying dense layer (fully connecting layer) to classify the texts or news articles. Figure 2 shows the general idea of CNN with texts [Zhang et al., 2021].

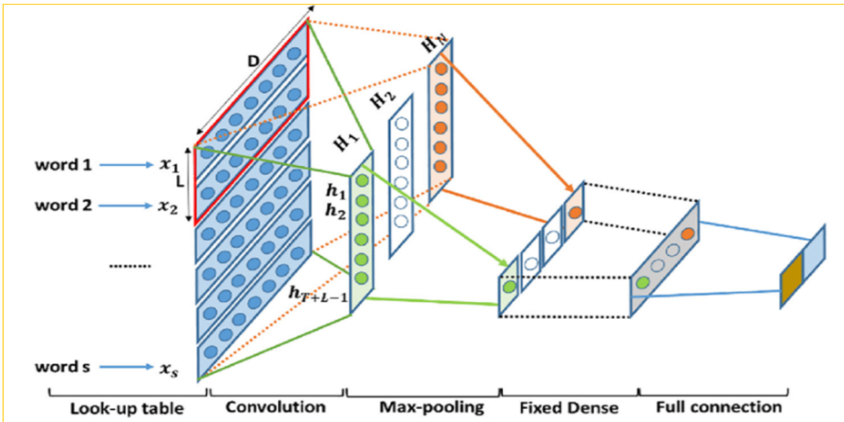

Figure 2: Text-CNN Framework

3.1.3. Pre-trained Models

Pre-trained models were proposed for word representation. These models are trained on huge datasets to learn a certain task, saved, and then utilized for solving other similar tasks [Becerra-Rico et al., 2020]. Pre-trained models are a recent trend in NLP and are characterized by their efficiency for two reasons: (1) they solved the problem of a small dataset; (2) through these models, no need to train the model from scratch. GloVe, BERT, and FastText are three common pre-trained models for processing Arabic and English datasets. These models assign a vector or feature to a word based on the semantic meaning of that word [Hussien and Dhannoon, 2020].

3.2 Machine Learning

Machine learning (ML) is an important area of artificial intelligence concerned with building and understanding models that learn from data [Wotaifi and Dhannoon, 2022]. The concept of ML involves many techniques that build the model on a sample of data called a training set [Ahmed et al., 2018]. Then the quality of the model or algorithm is measured by predicting the second part of the data called the testing set [Hamzah and Dhannoon, 2021].

3.2.1 Logistic Regression

Logistic Regression (LR) focuses on the relationship between the categorical variable (the dependent variable or the so-called target class) and a set of independent variables (features) [Boateng and Abaye, 2019]. In contrast to linear regression, which is used with continuous values, logistic regression can be used with discrete values as it treats data with a different distribution and shape. Figure 3 shows a simple example of a data distribution that can be processed by logistic regression using the sigmoid function, whereas the dependent variable (target class) is described by Equation 5 [Tomar et al., 2018].

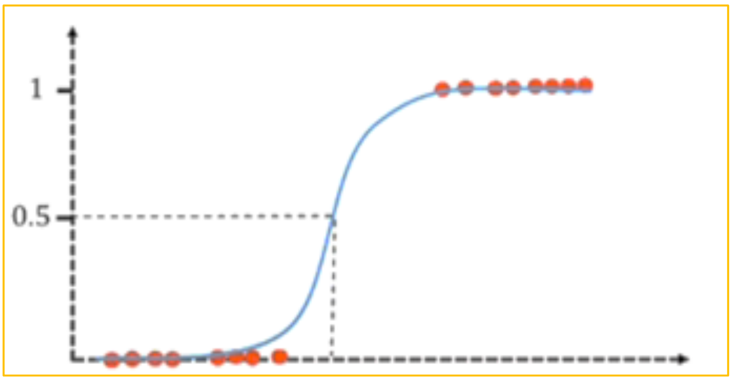

Figure 3: LR using the Sigmoid Function

$$Y = \frac{1}{1 + e^{-\theta^T x}} \qquad (5)$$

3.2.1 Random Forest

Random Forest (RF) is an ensemble technique that makes the final decision based on a set of decision tree models. Each decision tree in the random forest is constructed by randomly selecting a subset of the features in the original dataset [Wotaifi and Al-Shamery, 2019]. After that, each tree's outputs are considered a vote, and thus the final prediction is based on the majority voting. In RF, the number of features considered in each subset is according to Equation 6 [Wotaifi and Al-Shamery, 2018].

$$F = log_2 f + 1 \qquad (6)$$

Where: F is the number of selected features and f is the number of all features in the dataset.

3.3 Fuzzy Logic

In classical (traditional) logic, element x either belongs or does not belong to set A. Therefore, the membership function of this element can be expressed through the following formula [Toffano and Dubois, 2020].

$$\mu_A = \begin{cases} 1 \ if \ x \in A \\ 0 \ if \ x \notin A \end{cases}$$

On the other hand, in fuzzy logic, a value or weight between (0 - 1) is assigned to each element through Membership Functions (MF). These values represent the belonging degree of this element in the fuzzy set [Jain et al., 2020].

In fuzzification, the Triangular MF is one of the essential MFs that convert clear values into fuzzy values. This function is described by Equation 7 [Bennouna and Tkiouat, 2018].

$$MF = \frac{X_i - a}{\mathrm{b} - a} \tag{7}$$

Where: X is the real value within the universe of discourse. a and b are the coordinates of the triangular heads.

In defuzzification, the Centre of Gravity (COG) is the function applied to each fuzzy value (resulting from the MF) to obtain a final and clear value. The COG is represented by Equation 8 [Wotaifi and Al-Shamery, 2020].

$$\mu o = \frac{\sum_{i=1}^{n} \mu(Xi) * Xi}{\sum_{i=1}^{n} \mu(Xi)} \tag{8}$$

Where: $\mu(Xi)$ is an MF of input X

3.4 Datasets

The lack of tight censorship on social media and biased news agencies (biased media) has led to the discredit of many articles by recipients or people. To limit this problem, a set of models has been proposed to predict Arabic and English fake news based on three data sets. One of these datasets is an Arabic dataset while the other two are English datasets.

The AraNews dataset is an Arabic dataset available to the research community. It includes many topics, such as politics, economics, sports, and others. This dataset was collected from news agencies and newspapers in 15 Arab countries as well as the United Kingdom and the United States. Figure 4 shows more details about this dataset [Aljwari et al., 2022].

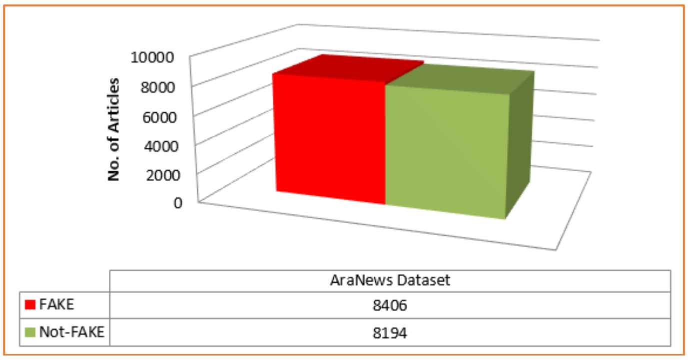

Figure 4: Number of News Articles in AraNews Dataset

Sentimental LIAR is a modified version of the LIAR dataset containing multi-class. In addition to the sentence, this dataset contains a set of speaker-based features. As in [6] this dataset is converted to a binary class by changing false, barely true, pants-fire, and half-true labels to Fake, and the remaining labels to Not-Fake [Upadhayay and Behzadan, 2020]. For more details see Figure 5.

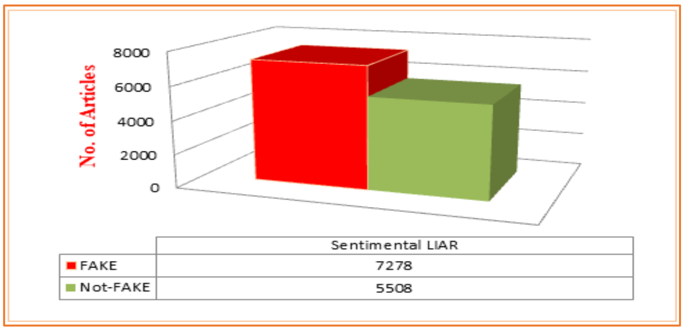

Figure 5: Number of News Articles in Sentimental LIAR Dataset

The Fake-or-Real dataset was collected from real-world sources. The real news in this dataset was collected from an authoritative news site called Reuters.com. As for the fake articles, it was collected from a fake dataset on kaggle.com. For each article, there is some information available such as the article title, article type, article date, and class label [Ahmed et al., 2018]. The number of real and fake articles and other details are shown in Figure 6.

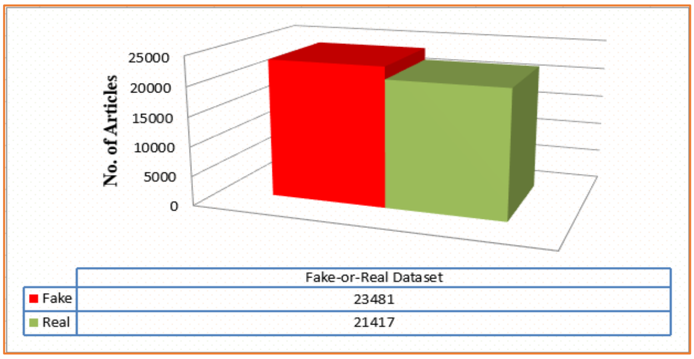

Figure 6: Number of News Articles in Fake-or-Real Dataset

4 Methodology

Our proposed methodology referred to in Figure 7 aims to detect fake news in both Arabic and English. Two improved models have been proposed based on machine learning and deep learning. Concerning machine learning, the random forest model is improved by controlling the selection of the essential features. A fuzzy model based on feature selection methods has identified these essential features. As for deep learning, a hybrid architecture is built based on CNN and BiGRU.

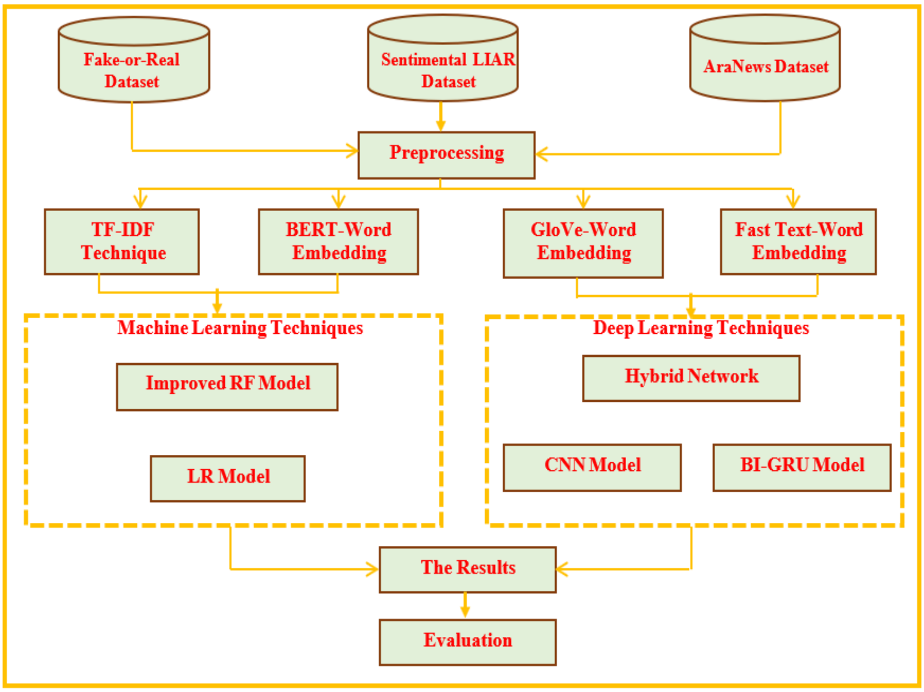

Figure 7: The Proposed Methodology

As shown in Figure 7, our methodology consists of four stages: pre-processing, feature extraction, building prediction models, and finally evaluation. The three datasets have been cleaned in the preprocessing stage. Each news article has been pre-processed in five steps: deleting symbols, deleting punctuation, deleting consecutive spaces, deleting stop words, and stemming. Then, tokenization is conducted by splitting any news article into words (tokens) based on spaces between them. The next stage is the process of representing words into vectors or features. Three pre-trained models are used: GloVe, FastText, and BERT. Through these models, it was allocated embedding values to each word in datasets based on the meaning.

After representing the datasets in a modellable format, prediction models have been used to predict fake news articles. Two improved models are proposed to improve the prediction of this news (more details are explained in the next sub-sections). Finally, the quality of these models was measured and the results of our proposed models are compared with other prediction models during the evaluation stage. Important evaluation metrics such as accuracy, precision, recall, and f-score were used to measure the quality and performance of the prediction models. Equations 9, 10, 11, and 12 show general formulas for these evaluation metrics [Hamzah and Dhannoon, 2021].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F - Scoree = \frac{2 * TP}{2 * TP + FP + FN} \tag{12}$$

Where: TP is a true positive, TN is a true negative, FP is a false positive, and FN is a false negative.

4.1 Pre-trained Model-Based Hybrid Network

To improve fake news detection, the properties of CNN and BiGRU are mixed into one effective Hybrid Network. This network was used to predict Arab fake news by training it on the AraNews dataset. It has also been used to predict English fake news by training it on the Fake-or-Real dataset. Figure 8 shows the main architecture of the Hybrid Network.

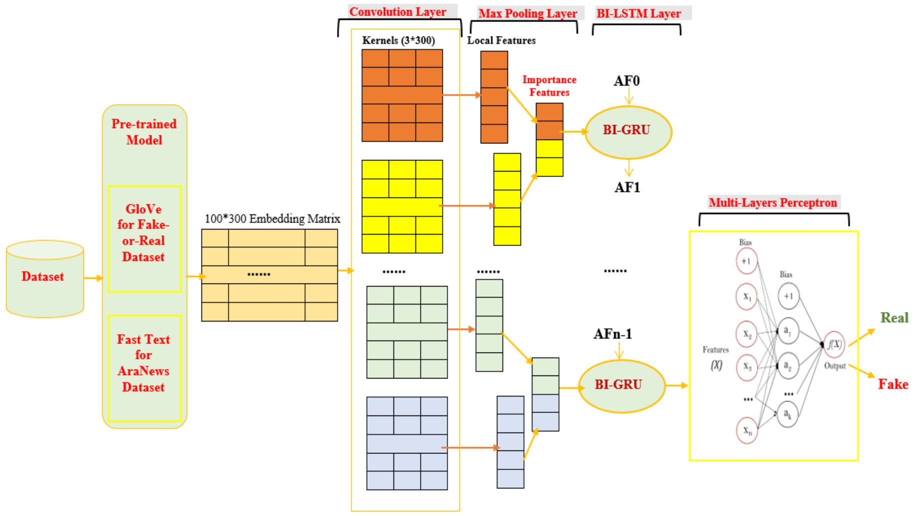

Figure 8: Hybrid Network

Because the pre-trained models are neural networks that were trained on millions of words and usually led to better results, three of these models have been used. The FastText model is used to represent Arabic words into vectors or features, whereas GloVe is applied for English word representation. The output of these models is a large embedding matrix; each row represents a vector for each word in the dataset. This matrix is added as an embedding layer (non-trainable) in the proposed network structure. On the output of this layer, the CNN is applied to determine the relevant features. Two layers of this network are used: a convolution layer is added to identify the local features, whereas the max-pooling layer is added to select the relevant features. To measure the long-term dependency of the sequences, a BiGRU network has been applied. This network is fed with important features extracted from the previous layer (max-pooling layer). Finally, MLP is used to classify news articles as fake or real. In MLP, three layers are added: a fully connected layer, the dropout layer, and the classification layer (output layer). In the last layer, the sigmoid activation function was used.

4.2 Pre-trained Model-Based Improved RF

In this study, the random forest model is improved by focusing on the relevant features in the Sentimental LIAR dataset. This dataset contains the sentence in addition to 23 speaker-based features in each object, so the process of building the improved model is achieved in stages as shown in Figure 9.

Regarding texts (sentences), unlike most studies that use traditional methods such as TF-IDF with machine learning techniques, this study uses BERT pre-trained model for word representation. The benefit of using BERT is that the model assigns a vector to the whole sequence or sentence in the dataset. In other words, this study uses BERT to convert every text or sentence in the dataset into a vector containing embedding values for the entire sentence, not for each word in the sentence. Thus, 12786 vectors are extracted, which is equal to the number of sentences in the Sentimental LIAR dataset.

As for the speaker-based features, a developed selection method based on fuzzy logic was proposed. This Fuzzy Model (FM) identifies the relevant features based on feature selection methods. Initially, speaker-based features have been evaluated in three methods: Mean Decrease Of Impurity (MDI), Drop-Feature Selection (DFS), and Information Gain (IG). Based on a random forest model, the MDI method is proposed for evaluating features based on the contribution of each feature in impurity reduction when it is selected as a splitting node in the tree. Then the average contributions of each feature to all trees in the forest are taken as the weight or rank of that feature.

DFS is also proposed to evaluate features based on their contributions to increase accuracy. Initially, accuracy is calculated by all features. Then we recalculated the accuracy after dropping the first feature. The difference between accuracy with all features and accuracy after dropping the first feature is the weight or rank of that feature. The same procedure is performed for all other features. Finally, the speaker-based features have been evaluated a third time using one of the filtering methods, which is IG.

The FM takes the results of the above three methods and allocates one weight for each feature. In the fuzzification, Triangular MF is applied to calculate the degree of membership for each weight obtained from the three selection methods above. Equation 9 is used to calculate fuzzy values using the Triangular MF where a is the minimum weight obtained through the three feature selection methods and equal to 0.06, and b is the maximum weight and equal to 0.87. The process for calculating membership values is shown in Figure 10.

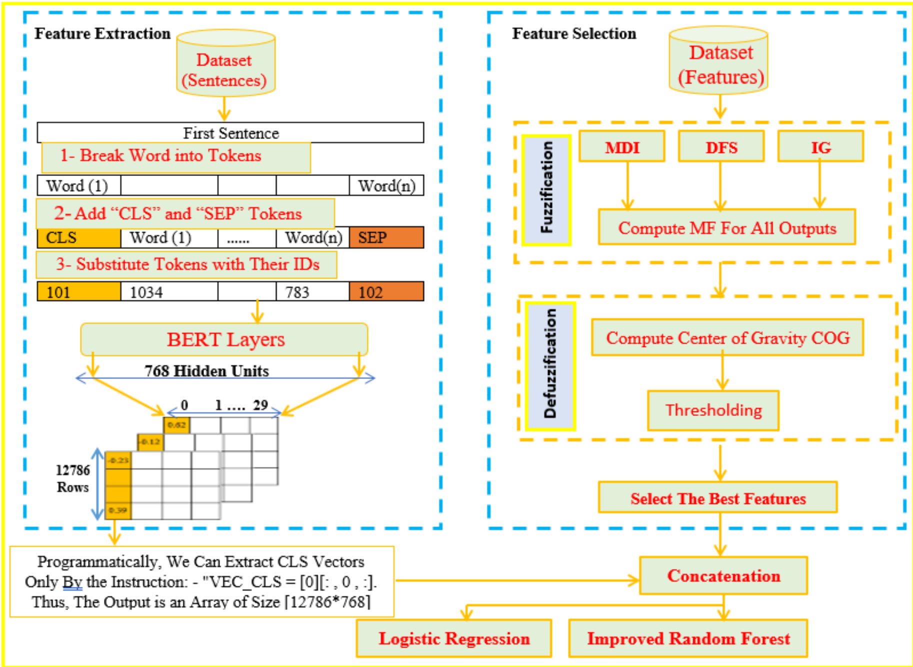

Figure 9: Improved RF Based on BERT and Fuzzy Model

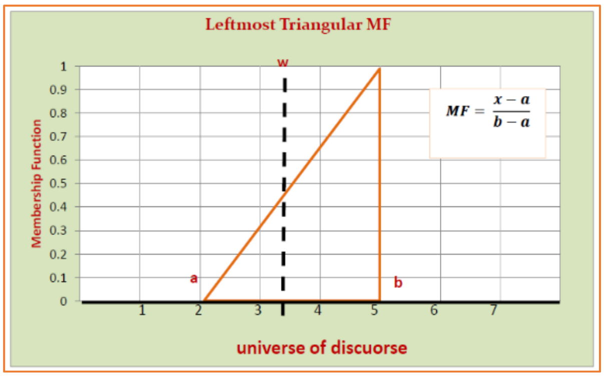

Figure 10: Triangular MF

In defuzzification, the COG is applied to obtain one weight for any feature. This is achieved according to Equation 10, which considers three membership values and three weights obtained from the feature selection methods. Finally, the most important five relevant features are selected according to a predefined threshold.

To improve prediction accuracy, a random forest classifier is modified by controlling the selection of features in each tree in the forest. This improved model is built on features extracted from the BERT pre-trained model and relevant speaker-based features (the five features) selected from the fuzzy model. Instead of the random approach used by the random forest, the improved model controls the selection process by selecting the five speaker-based features in all trees of the forest. The improved random forest model is set by selecting 100 trees, with each tree having 12 features. In all these trees, this model selects the five features that have been identified by the fuzzy model and the other seven features are randomly chosen from the features extracted by the BERT model.

5 Results and Discussion

To measure the quality of our proposed models, experiments are conducted on the AraNews dataset, Fake-or-Real dataset, and Sentimental LIAR dataset. After preprocessing these datasets and cleaning them from noise, the features were extracted through the pre-trained models. With the FastText model that is applied to the AraNews dataset, the embedding dimension is 300, whereas the maximum length of the article is set to 100. Likewise, both the embedding dimension and maximum length for GloVe are set to 300 and 100 respectively. On the other hand, the BERT model assigns to each sequence (sentence), a vector of length 768. Because the Sentimental LIAR dataset contains short sentences, the maximum length is set to be 30. Finally, the datasets are split so that the training set is set to be 0.8 while the test set is 0.2.

The Hybrid Network architecture is built by Keras. The proposed hybrid network has been used to predict Arabic fake news and English fake news. Through the GloVe pre-trained model (English language) and the FastText pre-trained model (Arabic language), an embedding matrix of size 100 * 300 has been obtained. Then seven layers have been added: Convolution layer, Max pooling layer, BiGRU layer, fully connected layer, Drop-out layer, and Classification layer (Output layer).

The role of CNN lies in identifying the relevant features through the convolution layer and the Max pooling layer. In the convolution layer, Rectified Linear Unit Activation Function (ReLU activation function) and 32 kernels of size 3 to each one is applied. To measure the long-term dependency of sequences (sentence), the BiGRU layer is added with 256 units. After that, three layers are used to classify the news articles. The first layer represents a fully connected layer (Dense layer) consisting of 128 nodes with a ReLU activation function. To reduce the overfitting problem and optimize the network, a dropout layer with a probability of 0.3 is added. The last layer, the classification layer (output layer), is used to classify the news article as fake or real. In this layer, the sigmoid activation function is used. With a batch size equal to 64 and 15 epochs, Table 1, Figure 11, Table 2, and Figure 12 show the results of our proposed model based on the AraNews dataset and Fake-or-Real dataset respectively. These results are compared with other studies that used the same datasets.

Models	Accuracy	Precision	Recall	F-Score
Hybrid Network (The Proposed Model)	**0.9473**	**0.9301**	**0.9537**	**0.9418**
CNN	0.8721	0.8439	0.9001	0.8711
BiGRU	0.9083	0.9033	0.9288	0.9150
Previous Study (Aljwari et al., 2022)	0.866	----	----	----

Table 1: The results of the proposed network and previous studies based on AraNews dataset

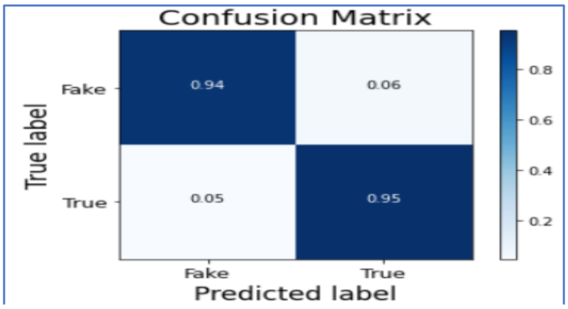

Figure 11: Confusion Matrix (AraNews Dataset)

Models	Accuracy	Precision	Recall	F-Score
Hybrid Network (The Proposed Model)	**0.9935**	**0.9947**	**0.9902**	**0.9924**
CNN	0.9886	0.9793	0.9862	0.9827
BiGRU	0.9830	0.9781	0.9886	0.9780
Previous Study (Ahmed et al., 2018)	0.92	----	----	----

Table 2: The results of the proposed network and previous studies based on Fake-or-Real dataset

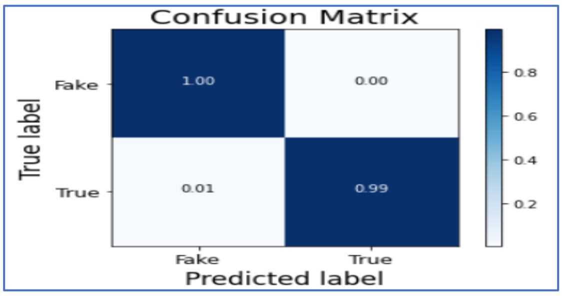

Figure 12: Confusion Matrix (Fake-or-Real Dataset)

The improved random forest model was built on a Sentimental LIAR dataset. Each sentence was analysed and transformed into features through the BERT pre-trained model. This model assigns a vector of length 768 to each sentence through the CLS token added to the sentence (The output of this model is a matrix of size (‘12786*786’). On the other hand, a FM is proposed to select the best speaker-based features. The top five features whose weights exceed 0.35 have been identified as the best features. Table 3 presents some details of these features.

Feature	Description	Rank
sentiment_score	Percentage or score of sentiment for the speaker or user.	0.504
sentiment_anger	The score of the anger of the speaker	0.501
sentiment_fear	The score of fear of the speaker	0.499
speaker_job	The job of the speaker	0.384
Subject	The topic being addressed by the speaker	0.377

Table 3: The Best Five Selected Features

A random forest classifier selects samples of features randomly to build decision trees. The number of features in each sample is determined according to Equation 6. The random approach of this classifier is improved by controlling the selection of features in each sample. In the improved model, 12 features have been identified in each sample (tree). Then the five best user features (Table 3) are chosen in all samples and the other seven are randomly selected from the features extracted from the BERT model. In the evaluation stage, it was found that our proposed model outperformed the studies conducted on the same dataset. Table 4 shows details of the results and comparisons with previous studies.

Models	Accuracy	Precision	Recall	F-Score
Improved RF	**0.7481**	**0.7005**	**0.7319**	**0.7146**
LR	0.6863	0.6779	0.6931	0.6854
Previous Study (Upadhayay and Behzadan, 2020)	0.70	----	----	----

Table 4: The results of the proposed model and comparisons with previous studies

6 Conclusion

Recently, it has been noticed that people are interested in following up on events and activities through social media. On the other hand, due to the massive growth of content on the internet, it becomes difficult to detect what is real and what is fake. Therefore, this problem has received great attention from researchers by including modern methods for fake news detection. As such, this study aims to include machine learning and deep learning for classifying fake news in Arabic and English. This study uses three pre-trained models for word representation: Glove, FastText, and BERT. Then the prediction accuracy is improved by developing a hybrid deep neural network. This network relies on CNN to identify relevant features and BiGRU to follow the sequence of text. On the other hand, the behaviour of the random forest classifier is modified in selecting the features used to construct the decision trees. The improved random forest model focuses on the relevant speaker-based features identified through a fuzzy model.

The proposed models are trained on three datasets: Fake-or-Real dataset, AraNews dataset, and the Sentimental LIAR dataset. The results have shown that the accuracy of our proposed models in detecting fake news is better compared to previous studies conducted on the same datasets.

In future work, the attention model can be applied to improve prediction accuracy by identifying the more informative words. It is also possible to include the bidirectional long short-term memory with the architecture of this study because this may achieve better results with both short articles and long articles.

References

[Agarwal et al., 2019] Agarwal, V., Sultana, H. P., Malhotra, S., Sarkar, A.: "Analysis of classifiers for fake news detection"; (2019), Procedia Computer Science, 165, 377-383.

[Ahmed et al., 2018] Ahmed, H., Traore, I. & Saad, S.: "Detecting opinion spams and fake news using text classification"; (2018), Security and Privacy, 1, 1-15.

[Al-Hmouz, 2020] Al-Hmouz, R.: "Deep learning autoencoder approach: Automatic recognition of artistic Arabic calligraphy types"; (2020), Kuwait Journal of Science, 47.

[Aljwari et al., 2022] Aljwari, F., Alkaberi, W., Alshutayri, A., Aldhahri, E., Aljojo, N., Abouola, O.: "Multi-scale Machine Learning Prediction of the Spread of Arabic Online Fake News"; (2022), Postmodern Openings, 13, 01-14.

[Alshdaifat et al., 2022] Alshdaifat, N. F. F., Osman, M. A., Talib, A. Z.: "An improved multi-object instance segmentation based on deep learning"; (2022). Kuwait Journal Of Science, 49.

[Becerra-Rico et al., 2020] Becerra-Rico, J., Aceves-Fernández, M. A., Esquivel-Escalante, K., Pedraza-Ortega, J. C.: "Airborne particle pollution predictive model using Gated Recurrent Unit (GRU) deep neural networks. Earth Science Informatics";(2020), 13, 821-834.

[Bennouna and Tkiouat, 2018] Bennouna, G., Tkiouat, M.: "Fuzzy logic approach applied to credit scoring for microfinance in Morocco"; (2018), Procedia Computer Science, 127, 274-283.

[Boateng and Abaye, 2019] Boateng, E. Y. & Abaye, D. A.: "A review of the logistic regression model with emphasis on medical research"; (2019), Journal of data analysis and information processing, 7, 190-207.

[Demirel et al., 2021] Demirel, U., Handan, Ç. & Ramazan, Ü.: "Predicting stock prices using machine learning methods and deep learning algorithms"; (2021), The sample of the Istanbul Stock Exchange. Gazi University Journal of Science, 34, 63-82.

[Elsaeed et al., 2021] Elsaeed, E., Ouda, O., Elmogy, M. M., Atwan, A., El-Daydamony, E.: "Detecting Fake News in Social Media Using Voting Classifier"; (2021), IEEE Access, 9, 161909-161925.

[Fayaz et al., 202] Fayaz, M., Khan, A., Bilal, M., Khan, S. U.: "Machine learning for fake news classification with optimal feature selection"; (2022). Soft Computing, 1-9.

[Jain et al., 2020] Jain, M., Madeira, A., Martins, M. A.: "A fuzzy modal logic for fuzzy transition systems"; (2020), Electronic Notes in Theoretical Computer Science, 348, 85-103.

[Mishra et al., 2022] Mishra, S., Shukla, P. & Agarwal, R.: "Analyzing machine learning enabled fake news detection techniques for diversified datasets"; (2022), Wireless Communications and Mobile Computing, 2022, 1-18.

[Hamzah and Dhannoon, 2021] Hamzah, N. A., Dhannoon, B. N.: "The Detection of Sexual Harassment and Chat Predators Using Artificial Neural Network" (2021); Karbala International Journal of Modern Science, 7, 301-312

[Hussien and Dhannoon, 2020] Hussien, Z. K., Dhannoon, B. N.: Anomaly Detection Approach Based on Deep Neural Network and Dropout"; (2020), Baghdad Science Journal, 17, 701-709.

[Pan et al., 2018] Pan, J. Z., Pavlova, S., Li, C., Li, N., Li, Y., Liu, J.: "Content based fake news detection using knowledge graphs"; (2018), International semantic web conference, Springer, 669-683.

[Qayyum et al., 2019] Qayyum, A., Qadir, J., Janjua, M. U., Sher, F.: "Using blockchain to rein in the new post-truth world and check the spread of fake news"; (2019), IT Professional, 21, 16-24.

[Toffano and Dubois, 2020] Toffano, Z., Dubois, F.: "Adapting logic to physics: the quantum-like eigenlogic program"; (2020), Entropy, 22, 139.

[Tomar et al., 2018] Tomar, A. S., Singh, M., Sharma, G., Arya, K.: "Traffic management using logistic regression with fuzzy logic"; (2018), Procedia computer science, 132, 451-460.

[Upadhayay and Behzadan, 2020] Upadhayay, B., Behzadan, V.: "Sentimental LIAR: Extended Corpus and Deep Learning Models for Fake Claim Classification"; (2020), IEEE International Conference on Intelligence and Security Informatics, 1-6.

[Wotaifi and Dhannoon, 2022] Wotaifi, T. A., Dhannoon, B. N." "Improving Prediction of Arabic Fake News Using Fuzzy Logic and Modified Random Forest Model,"; (2022), Karbala International Journal of Modern Science, 8(3), 477-485.

[Wotaifi and Al-Shamery, 2018] Wotaifi, T. A., Al-Shamery, E. S.: "Fuzzy-Filter Feature Selection for Envisioning the Earnings of Higher Education Graduates"; (2018), Compusoft, 7, 2969-2975.

[Wotaifi and Al-Shamery, 2019] Wotaifi, T. A., Al-Shamery, E. S.: "Mining of completion rate of higher education based on fuzzy feature selection model and machine learning techniques"; (2019), International Journal of Recent Technology and Engineering, 8, 393-400.

[Wotaifi and Al-Shamery, 2020] Wotaifi, T. A., Al-Shamery, E. S.: "Modified Random Forest based Graduates Earning of Higher Education Mining"; (2020), Int. J. Comput. Inf. Syst. Ind. Manag. Appl., 12, 56-65.

[Zhang et al., 2019] Zhang, A., Lipton, Z. C., LI, M., SMOLA, A. J.: "Dive into deep learning", arXiv preprint, (2021).