International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 1, February 2025, pp. 836~845
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i1.pp836-845

Arabic fake news detection using hybrid contextual features

Hussain Mohammed Turki[1], Essam Al Daoud[1], Ghassan Samara[1], Raed Alazaidah[1], Mais Haj Qasem[1], Mohammad Aljaidi[1], Suhaila Abuowaida[2], Nawaf Alshdaifat[3]
[1]Department of Computer Science, Faculty of Information Technology, Zarqa University, Zarqa, Jordan
[2]Department of Computer Science, Faculty of Information Technology, Al al-Bayt University, Mafraq, Jordan
[3]Faculty of Information Technology, Applied Science Private University, Amman, Jordan

Article Info

Article history:

Received Mar 24, 2024
Revised Aug 16, 2024
Accepted Oct 1, 2024

Keywords:

AraBERT
AraNews
Fake news
Tokenization
WaraBERT

ABSTRACT

Technology has advanced and social media users have grown dramatically in the last decade. Because social media makes information easily accessible, some people or organizations distribute false news for political or commercial gain. This news may influence elections and attitudes. Even though English fake news is widely detected and limited, Arabic fake news is hard to recognize owing to a lack of study and data collection. Wara Arabic bidirectional encoder representations from transformers (WaraBERT), a hybrid feature extraction approach, combines word level tokenization with two Arabic bidirectional encoder representations from transformers (AraBERT) variants to provide varied features. The study also discusses eliminating stopwords, punctuations, and tanween markings from Arabic data. This study employed two datasets. The first, Arabic fake news dataset (AFND), has 606,912 records. Second dataset Arabic news (AraNews) has 123,219 entries. WaraBERT-V1 obtained 93.83% AFND accuracy using the bidirectional long short-term memory (BiLSTM) model. However, the WaraBERT-V2 technique obtained 81.25%, improving detection accuracy above previous researchers for the AraNews dataset. These findings show that WaraBERT outperforms word list techniques (WLT), term frequency-inverse document frequency (TF-IDF), and AraBERT on both datasets.

Corresponding Author:

Suhaila Abuowaida
Department of Computer Science, Faculty of Information Technology, Al al-Bayt University
Mafraq 25113, Jordan
Email: sabuoweuda@zu.edu.jo

1. INTRODUCTION

With the wide spread of the internet in the world, many platforms of social media like Instagram, Facebook, Twitter, and many others, have become popular among people and they are used daily as they facilitate getting information and rapidly spreading the news. The number of social media users around the world is increasing dramatically, the number of users in 2017 was about 2.73 billion people, while in 2019, the number of users increased to 3.51 billion, and in 2022 they became 4.59 billion people, which constitutes about 58% of the world's population [1]. With the increase in the number of users, the number of people who depend on social media as a source from which they take news increases. Recent studies have been conducted in America about the number of social media users showing that 71% of these people use the Facebook platform, and about 52% of them rely on Facebook as a source of news, which means that 36% of Americans get their news from Facebook [2].

All of these factors made fake news a part of our online daily routines. As fake news had an enormous harmful impact, it gained popularity among politicians, journalists, researchers, and the public at large. It is prepared or disseminated to deceive the public and harm the reputation of an entity, person, or agency, for political or financial gain. Fake news may affect people's opinions negatively by targeting a specific institution or personality [3]–[5]. In 2016, Mark Zuckerberg confessed that fake news on Facebook posts and stories supported by Russians, with political goals presented to over 126 million people in America [6]. Deep learning (DL) can play an important role in reducing the spread of fake news by providing more sophisticated techniques to identify inaccurate or misleading information [7].

As DL models are capable of automatically learning and extracting high-level features from raw text data, therefore using DL can be better than traditional machine learning (ML) methods for Arabic fake news detection. The main goal of this paper is to distinguish fake news from real news, in addition to building a system to detect this news to limit its spread on social media, These goals can be summarized as follows: create a new hybrid feature extraction method by concatenating different feature extraction methods (such as word-level tokenization, term frequency-inverse document frequency (TF-IDF), and Arabic bidirectional encoder representations from transformers (AraBERT)) to enhance the detection accuracy, comparing AraBERT versions (V1 and V2) in the field of feature extraction (tokenization) and determining the best one, and finally indicate the effect of applying data cleaning (such as removing stop words, removing punctuation, and normalizing text) on the dataset.

2. RELATED WORKS

Khalil *et al.* provided a large, labeled, and diverse Arabic fake news dataset (AFND) that is collected from public Arabic news websites. AFND consists of 606,912 public news [8]. Saadany *et al.* created a dataset consisting of 3,185 articles from two Arabic satirical news websites [9]. They proposed a model that achieved high accuracy, up to 98.6%, in identifying satirical fake news in Arabic. On the other hand, an innovative method for automatically creating Arabic-manipulated news stories was proposed by Nagoudi *et al.* [10]. They gathered and published a new part-of-speech-tagged (POS-tagged) Arabic news dataset known as Arabic news (AraNews). Alkhair *et al.* developed an Arabic corpus for fake news analysis, concentrating on the most rumored concepts. As ML classifiers, they used multinomial naïve Bayes (MNB), decision trees (DT), and support vector machines (SVM) [11]–[13]. Conversely, Azad concentrated on creating Arabic fake news datasets of superior quality and creating an accurate classification for Arabic fake news. It provides insights and recommendations for researchers in the field of fake news recognition in Arabic [14]–[16]. Nassif used a range of contextualized Arabic embedding algorithms to create multiple transformer models to detect fake news in the Arabic language. Eight transformer models were used in the study [17]. whereas ML approaches were used to train the model by Himdi *et al.* [18].

Awajan highlighted the critical issue of fake news on social media, especially on Twitter, where false information is increasingly being spread. Using the Twitter application programming interface (API), the researcher gathered a dataset comprising 206,080 tweets [19]. Sorour and Abdelkader used a hybrid method that integrates traditional neural networks (NNs) with long short-term memory methods (LSTM) methods. Shishah created and implemented jointBERT, an innovative technique for detecting fake news in Arabic datasets [20]. Extensive experiments were carried out employing actual Arabic fake news datasets such as coronavirus disease 2019 (COVID-19) fakes, Satirical, AraNews, and others [21]. Awajan *et al.* noted that using ensemble models achieved better accuracy than using a single machine learning classifier. An online repository was included to provide continuous updates and resources related to fake news, including educational programs, publications, new methods, datasets, and other relevant resources [19]. Rahab *et al.* addressed the challenge of fake news detection and elimination, particularly in the context of user-generated content on social media platforms [22]. Alotaibi identified Arabic fake news tweets related to the COVID-19 pandemic and classified them into six categories entertainment, health, politics, religious, social, and sports [23]. A disinformation detection framework specifically designed for combating the spread of fake information related to the COVID-19 pandemic on Arabic social media platforms was proposed by Elaziz *et al.* [24].

The researcher employed a combination of multi-task learning (MTL), a pre-trained transformer-based model AraBERT, and an optimization algorithm fire hawk optimizer in their framework. Ameur and Aliane addressed the issue of an “infodemic” where false and misleading information about COVID-19 has emerged and complicated response efforts. The researchers highlight the role of social networking sites in contributing to the spread of rumors, conspiracy theories, hate speech, xenophobia, racism, and prejudice [25]. Four types of features were used: count vector, word-level TF-IDF, n-gram-level TF-IDF, and character-level TF-IDF. Six classifiers were employed (naïve Bayes, logistic regression, support vector machine, multilayer perceptron, random forest bagging, and extreme gradient boosting) [26]–[28].

3. DATA COLLECTION

In this paper, two datasets were used for the experiments: the first dataset was collected by Khalil *et al.* [8], a large and diverse AFND from public Arabic news websites. It contains about 606,912 public news articles gathered over 6 months from 134 public news websites in 19 different Arab countries. This dataset was used to compare the proposed feature extraction method with other typical feature extraction methods. The second dataset, AraNews, a significant, multi-topic, and multi-country Arabic news dataset, was created by [10]. It consists of 123,219 records. This dataset was used to compare the results of the proposed model with other researchers' results.

Data cleaning is an essential process when working on Arabic textual data, this process includes several steps:

- Removing punctuations: like commas, question marks, periods, and many others which usually do not have an important semantic value.
- Removing stop words: stop words are popular words that are used usually in the language but have little or no value, such as prepositions, and conjunctions.
- Removing tanween marks.

4. METHOD

Despite the enormous development in the field of artificial intelligence and the availability of many DL algorithms [29]–[32], natural language processing (NLP) suffers from the lack of available feature extraction techniques for Arabic text. Therefore, it was necessary to develop a hybrid feature extraction method that combines more than one feature extraction method as shown in Figure 1. This new model will help improve the work of these algorithms.

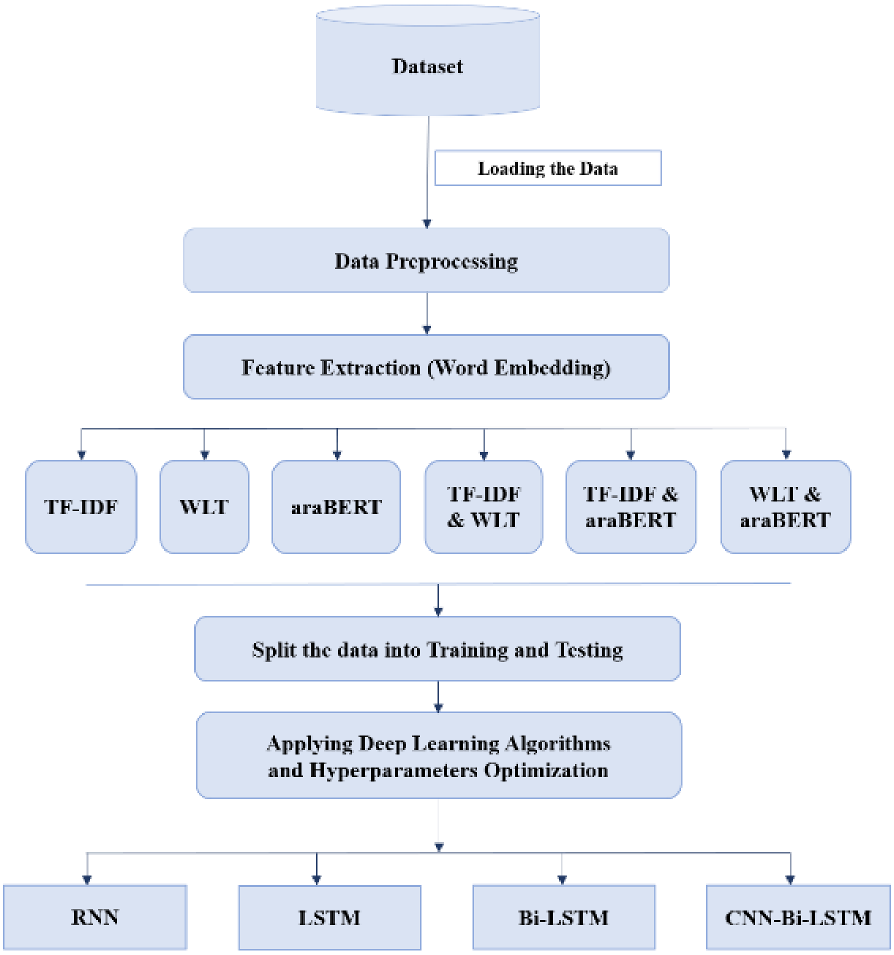

Figure 1. General proposed framework

When dealing with textual data, one of the most important points is extracting features from texts after being converted into numerical data. Therefore, the results are affected by the tokenization methods. Table 1 shows four models word list techniques (WLT), TF-IDF, araBERT-V1, and araBERT-V2 and its tokenization using a sample sentence. In this research proposes new hybrid models that concatenate between these models as following:

- WaraBERT: this model combines WLT and araBERT models. The vocabulary size of araBERT is static 64,000, while the vocabulary size of WLT is a hyperparameter. The maximum length for both models must be determined before calling the models for tokenization. After tokenization and indexing the articles using each model alone, the output of those models would be concatenated. The concatenated output will be used for training and testing. And so as not to get duplicate values, the value of the araBERT vocabulary size 64,000 must be added for each value in the WLT. Thus, the range of araBERT from 0 to 639,99, while the WLT values start from 64,000. Notable that WaraBERT has two versions WaraBERT-V1, and WaraBERT-V2 depending on the chosen version of the araBERT tokenizer.
- AraTFIDF: it refers to the combination of araBERT with the TFIDF model. It has the same architecture as WaraBERT. In addition, it also has two versions AraTFIDF-V1, and AraTFIDF-V2 depending on the chosen version of the araBERT tokenizer.
- TokenTFIDF: the part Token refers to the word tokenization. This model combines WLT with TF-IDF. It also has the same procedure as the previously proposed models.

Table 1. AFND versions

	Removing stopwords	Removing punctuations	Removing tanween marks
AFND-V1	No	No	No
AFND-V2	Yes	Yes	No
AFND-V3	Yes	Yes	Yes

5. EXPERIMENTAL RESULTS AND DISCUSSION

To determine the most appropriate DL model for Arabic fake news detection, the most known and powerful four DL models were used to perform the training process to determine the most appropriate one: recurrent neural network (RNN), long short-term memory (LSTM), bidirectional long short-term memory (BiLSTM) and convolutional neural network-bidirectional long short-term memory (CNN-BiLSTM). As the AFND is a novel and huge dataset and no experiments have been conducted on it by other researchers, this paper uses 30,000 samples of the dataset and applies the experiments on three versions of the dataset see Table 1.

- AFND-V1: This version contains the original data, without applying data cleaning (removing punctuations, stopwords, and tanween marks).
- AFND-V2: This version includes the data after removing punctuations and stopwords.
- AFND-V3: In this version, all data cleaning processes had been conducted on the original dataset. This means that punctuation, stopwords, and tanween marks are removed.

For all three versions of AFND, eight feature extraction methods, and four DL models were used in the experiments. There are three main goals behind these experiments: the first goal is to understand the effect of removing punctuations, stopwords, and tanween marks on the detection process. The second goal is to determine the best feature extraction method, in addition to checking the proposed methods. The last goal is to select the most appropriate DL model. Notable that the training was repeated five times to achieve accurate results. For all DL models, the vocabulary size for WLT, araBERT, and TF-IDF is 176,000, 64,000, and 1,000 consecutively. They have the same length (number of attributes) which is 160. All models are evaluated based on the accuracy metric. The average test accuracy is calculated using K-fold, where K=5.

The first architecture is RNN, where the hyperparameters that are selected to train the model are shown in Table 2. The vocabulary size of the feature extraction method represents the input dimension, and the output dimension for all methods is 150 neurons. The Simple RNN layer is the second layer with 146 neurons with a recurrent dropout rate of 0.12. The next hidden layers are five dense layers with 212 neurons and a rectified linear unit (ReLU) activation function for each layer. The output layer with one neuron uses the sigmoid activation function 0 or 1. Table 3 describes RNN model hyperparameters and Table 3 shows the results using different features.

Comparing the results, WaraBERT-V1 achieves higher accuracy on AFND-V1 84.83%, AFND-V2 83.35%, and AFND-V3 75.83%. WaraBERT-V2 shows similar trends to WaraBERT-V1 but generally has lower accuracies across all AFND versions. Table 4 presents LSTM hyperparameters that are selected to train the model. The first layer is the embedding layer and the vocabulary size of the feature extraction method represents the input dimension, and the output dimension for all methods is 344 neurons. The LSTM layer contains 144 neurons and a recurrent dropout rate 0.2. The next hidden layers are four dense layers containing 77 neurons with a dropout rate 0.17 and a ReLU activation function for each layer. Table 5 shows the results of LSTM using different features.

Table 6 Shows Bi-LSTM hyperparameters that are selected to train this model, where 120 neurons and a recurrent dropout rate 0.12, four dense layers of 77 neurons with a dropout rate 0.12 and a ReLU activation function for each layer. Table 7 shows the results of Bi-LSTM using different features.

WaraBERT-V1 achieves higher accuracy on AFND-V1, AFND-V2, and AFND-V3 92.67%, 92.37%, and 89.91% consecutively. While WaraBERT-V2 has similar patterns to WaraBERT-V1, but often performs less accurately via all other AFND versions. The hyperparameters of the fourth architecture CNN-BiLSTM are shown in Table 8. The second layer is the convolutional layer with 10 filters and the length of filters is 20, the Bi-LSTM layer contains 120 neurons and the dropout rate is 0.22, the dense layers contain 128 neurons with a dropout rate 0.23 and a ReLU activation function for each layer. Table 9 shows CNN-BiLSTM results using different features.

Table 2. RNN training hyperparameters

RNN neurons	Other hidden layers neurons					Batch size	Total epochs
	L1	L2	L3	L4	L5		
146	212	212	212	212	212	337	6

Table 3. RNN results using different features

Model	Vocabulary size	Max length	Trainable parameters	Test accuracy		
				AFND - V1	AFND - V2	AFND - V3
WLT	176,000	160	26,655,000	75.79	68.86	64.47
araBERT-V1	64,000	160	9,855,000	76.10	74.02	59.44
araBERT-V2	64,000	160	9,855,000	81,48	74.19	60.05
WaraBERT-V1	176,000 + 64,000	160 + 160	36,255,000	84.83	83.35	75.83
WaraBERT-V2	176,000 + 64,000	160 + 160	36,255,000	83.40	79.53	73.11
TokenTFIDF	176,000 + 1,000	160 + 160	26,805,000	51.24	50.01	49.13
AraTFIDF -V1	64,000 + 1,000	160 + 160	10,005,000	76.23	70.88	72.66
AraTFIDF -V2	64,000 + 1,000	160 + 160	10,005,000	75.58	77.14	71.14
WLT	176,000	320	26,655,000	59.71	54.19	52.97
araBERT-V1	64,000	320	9,855,000	61.02	79.17	73.30
araBERT-V2	64,000	320	9,855,000	66.85	71.23	55.62

Table 4. LSTM training hyperparameters

Embedding vector len.	LSTM neurons	Other hidden layers neurons				Batch size	Total epochs
		L1	L2	L3	L4		
344	144	77	77	77	77	180	4

Table 5. LSTM results using different features

Model	Vocabulary size	Max length	Trainable parameters	Test accuracy		
				AFND - V1	AFND - V2	AFND - V3
WLT	176,000	160	60,854,000	85.02	84.30	70.13
araBERT-V1	64,000	160	22,326,000	87.12	87.09	83.36
araBERT-V2	64,000	160	22,326,000	86.27	86.99	83.35
WaraBERT-V1	176,000 + 64,000	160 + 160	82,870,000	89.30	88.33	85.31
WaraBERT-V2	176,000 + 64,000	160 + 160	82,870,000	88.52	87.95	84.36
TokenTFIDF	176,000 + 1,000	160 + 160	61,158,000	60.08	60.33	54.17
AraTFIDF -V1	64,000 + 1,000	160 + 160	22,630,000	86.55	86.81	73.94
AraTFIDF -V2	64,000 + 1,000	160 + 160	22,630,000	85.09	82.21	71.18
WLT	176,000	320	60,854,000	59.18	56.62	54.12
araBERT-V1	64,000	320	22,326,000	87.55	86.54	83.77
araBERT-V2	64,000	320	22,326,000	84.68	84.02	82.50

Table 6. Bi-LSTM training hyperparameters

Embedding vector len.	BiLSTM neurons	Other hidden layers neurons				Batch size	Total epochs
		L1	L2	L3	L4		
400	120	128	128	128	128	144	4

Table 7. Bi-LSTM results using different features

Model	Vocabulary size	Max length	Trainable parameters	Test accuracy		
				AFND - V1	AFND - V2	AFND - V3
WLT	176,000	160	70,980,000	91.09	91.02	88.76
araBERT-V1	64,000	160	26,180,000	86.51	86.69	83.3
araBERT-V2	64,000	160	26,180,000	85.39	84.48	81.45
WaraBERT-V1	176,000 + 64,000	160 + 160	96,580,000	92.67	92.37	89.91
WaraBERT-V2	176,000 + 64,000	160 + 160	96,580,000	92.14	92.09	89.42
TokenTFIDF	176,000 + 1,000	160 + 160	71,380,000	91.08	90.93	88.17
AraTFIDF -V1	64,000 + 1,000	160 + 160	26,580,000	87.27	86.43	84.22
AraTFIDF -V2	64,000 + 1,000	160 + 160	26,580,000	84.21	83.18	82.37
WLT	176,000	320	70,980,000	90.69	89.47	88.77
araBERT-V1	64,000	320	26,180,000	84.52	86.75	83.72
araBERT-V2	64,000	320	26,180,000	84.30	85.52	82.65

Table 8. CNN-BiLSTM training hyperparameters

Embedding vector len.	No. of filters	Filter length	BiLSTM neurons	Other hidden layers neurons				Batch size	Total epochs
				L1	L2	L3	L4		
400	10	20	120	128	128	128	128	144	4

Table 9. CNN-BiLSTM results using different features

Model	Vocabulary size	Max length	Trainable parameters	Test accuracy		
				AFND - V1	AFND - V2	AFND - V3
WLT	176,000	160	38,970,000	91.02	90.23	88.71
araBERT-V1	64,000	160	14,330,000	85.12	84.74	84.09
araBERT-V2	64,000	160	14,330,000	85.37	84.87	81.26
WaraBERT-V1	176,000 + 64,000	160 + 160	53,050,000	91.96	91.64	89.46
WaraBERT-V2	176,000 + 64,000	160 + 160	53,050,000	91.74	91.24	89.03
TokenTFIDF	176,000 + 1,000	160 + 160	39,190,000	91.55	91.22	88.74
AraTFIDF -V1	64,000 + 1,000	160 + 160	14,550,000	87.89	88.43	84.52
AraTFIDF -V2	64,000 + 1,000	160 + 160	14,550,000	87.83	87.63	82.66
WLT	176,000	320	38,970,000	90.78	90.54	88.22
araBERT-V1	64,000	320	14,330,000	87.24	87.46	84.41
araBERT-V2	64,000	320	14,330,000	86.1	86.43	82.83

Although WaraBERT-V2 and TokenTFIDF achieved high accuracy compared with other methods, WaraBERT-V1 achieved higher accuracy on AFND-V1, AFND-V2, and AFND-V3 91.96%, 91.64%, and 89.46% consecutively as shown in Table 9. To analyzing the results of all DL models according to the AFND version and the effect of removing stopwords and tanween marks, as well as comparing the performance of tokenization methods, we can make the following observations:

– Comparing the proposed hybrid models with other typical models: WaraBERT-V1 consistently outperforms both WLT and araBERT methods across all AFND.
– AFND-V1, which represents the original data without removing stopwords and tanween marks, achieves the highest accuracy in most cases.
– AFND-V2, in which the stopwords were removed, shows a slight decrease in accuracy compared to AFND-V1.
– AFND-V3, in which both stopwords and tanween marks were removed, exhibits the lowest accuracies, indicating the potential loss of important semantic information.
– Removal of stopwords and tanween marks generally leads to a decrease in accuracy across all models, highlighting the importance of preserving stopwords and tanween marks.

6. HYPERPARAMETERS FINE TUNING

Depending on the results of the previous experiments, using WaraBERT tokenization with BiLSTM for the training process achieves the best accuracy. Therefore, a new model was built to enhance the accuracy and avoid overfitting. Concerning the AFND, only punctuation marks had been removed for the dataset. The dataset was split into 75% for training and 25% for testing. Table 10 shows the new hyperparameters. To avoid overfitting, we increased the complexity of the model (increasing the neurons of the layers) and increased the dropout rate for each layer. The vocabulary size used is 214,000 (150,000 for WLT and 64,000 for araBERT). The batch size of 280 is selected. According to the experiments, the batch size of 280 is the best since the batch size of 144 increases the number of iterations for every epoch, slowing down the training process, while the batch size of 444 needs more memory. A learning rate of 0.0001 was used for training, which needs a lower learning rate for fine-tuning. Proven by the experiments, early stopping at the patience of 1 is beneficial for avoiding overfitting and providing a chance for the model to enhance.

The same procedures used in the AFND were implemented on the AraNEWS dataset, WaraBERT-V2 was used for the feature extraction, the BiLSTM model was used for the training process, and only punctuation marks were removed from the data. The dataset was split into 80% for training and 20% for testing. Selected hyperparameters are shown in Table 11. The vocabulary size used is 224,000 (160,000 for WLT and 64,000 for araBERT-V2). A batch size of 70 is selected. A learning rate of 0.0001 was used for training, which needs a lower learning rate for fine-tuning. As proven by the experiments, early stopping at the patience of 1 is beneficial for avoiding overfitting. Comparing the results of the proposed model with those achieved by other researchers, WaraBERT-V2 enhanced the test accuracy by 1.25%, it achieved 81.25%.

Table 10. Bi-LSTM Tuned hyperparameters for AFND dataset

Layer		Parameters		
Embedding layer		Input dimension 214,000	Embedding vector length 768	
BiLSTM layer		Neurons 512	Dropout 0.6	Activation function ReLU
Other hidden layers	Layer 1	Neurons 512	Dropout 0.6	Activation function ReLU
	Layer 2	Neurons 512	Dropout 0.6	Activation function ReLU
	Layer 3	Neurons 256	Dropout 0.6	Activation function ReLU
	Layer 4	Neurons 256	Dropout 0.5	Activation function ReLU
Output layer		Neurons 1	Activation function Sigmoid	
Trainable parameters			170,583,000	
Test loss			0.16	
Test accuracy			93.83	

Table 11. Bi-LSTM tuned hyperparameters for AraNews dataset

Layer		Parameters		
Embedding layer		Input dimension 224,000	Embedding vector length 222	
BiLSTM layer		Neurons 128	Dropout 0.12	Activation function ReLU
Other hidden layers	Layer 1	Neurons 128	Dropout 0.17	Activation function ReLU
	Layer 2	Neurons 128	Dropout 0.17	Activation function ReLU
	Layer 3	Neurons 128	Dropout 0.17	Activation function ReLU
	Layer 4	Neurons 128	Dropout 0.17	Activation function ReLU
	Layer 5	Neurons 128	Dropout 0.17	Activation function ReLU
Output layer		Neurons 1	Activation function Sigmoid	
Trainable parameters			50,186,000	
Test accuracy			81.25	

7. CONCLUSION

In summary, this paper has three primary objectives. Firstly, it aims to develop a new hybrid feature extraction technique that surpasses existing methods. Secondly, it seeks to investigate the effects of eliminating punctuation, stopwords, and tanween marks on the detection process. Lastly, it endeavors to identify the most suitable deep learning model RNN, LSTM, BiLSTM, or CNN-BiLSTM for enhancing Arabic fake news detection. To achieve these objectives, the study utilized two datasets. The Arabic fake news dataset was employed to compare the proposed feature extraction method with conventional ones and to assess the impact of data cleaning on training. The AraNews dataset was used to compare the model's performance with that of previous research. The investigation into the removal of stopwords and tanween marks utilized three versions of AFND: AFND-V1 original data, AFND-V2 stopwords removed, and AFND-V3 both stopwords and tanween marks removed. AFND-V1 consistently yielded the highest accuracy, while AFND-V3 exhibited the lowest, indicating potential loss of crucial semantic information. Across all models, removing stopwords and tanween marks generally resulted in decreased accuracy. Several feature extraction methods were proposed and evaluated across the three AFND versions using four DL models. Among these methods, WaraBERT-V1 consistently outperformed others, including WLT and araBERT, across all versions of AFND. WaraBERT-V2 displayed similar patterns to WaraBERT-V1 but tended to achieve lower accuracy. The results underscored the importance of retaining stopwords and tanween marks for accurate fake news detection.

ACKNOWLEDGEMENTS

This research is funded by the deanship of scientific research in Zarqa University, Jordan.

REFERENCES

[1] S. J. Dixon, “Number of social media users worldwide from 2017 to 2028,” *Statista*, 2022. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ (accessed Nov. 21, 2022).
[2] E. Shearer and E. Grieco, “Americans are wary of the role social media sites play in delivering the news,” *Pew Research Center*, vol. 2, pp. 1–23, 2019.
[3] M. Alzyoud *et al.*, “Diagnosing diabetes mellitus using machine learning techniques,” *International Journal of Data and Network Science*, vol. 8, no. 1, pp. 179–188, 2024, doi: 10.5267/j.ijdns.2023.10.006.
[4] M. Haj Qasem, M. Aljaidi, G. Samara, R. Alazaidah, A. Alsarhan, and M. Alshammari, “An intelligent decision support system based on multi-agent systems for business classification problem,” *Sustainability*, vol. 15, no. 14, pp. 1–14, Jul. 2023, doi: 10.3390/su151410977.
[5] R. Alazaidah, G. Samara, S. Almatarneh, M. Hassan, M. Aljaidi, and H. Mansur, “Multi-label classification based on associations,” *Applied Sciences*, vol. 13, no. 8, pp. 1–16, Apr. 2023, doi: 10.3390/app13085081.
[6] A. Denton, “Fake news: the legality of the Russian 2016 Facebook influence campaign,” *BU Int’l LJ*, vol. 37, pp. 183–210, 2019.
[7] D. M. J. Lazer *et al.*, “The science of fake news,” *Science*, vol. 359, no. 6380, pp. 1094–1096, Mar. 2018, doi: 10.1126/science.aao2998.
[8] A. Khalil, M. Jarrah, M. Aldwairi, and M. Jaradat, “AFND: Arabic fake news dataset for the detection and classification of articles credibility,” *Data in Brief*, vol. 42, pp. 1–7, Jun. 2022, doi: 10.1016/j.dib.2022.108141.
[9] H. Saadany, E. Mohamed, and C. Orasan, “Fake or real? a study of Arabic satirical fake news,” *arXiv preprint arXiv:2011.00452*, pp. 1–11, 2020.
[10] E. M. B. Nagoudi, A. Elmadany, M. Abdul-Mageed, T. Alhindi, and H. Cavusoglu, “Machine generation and detection of Arabic manipulated and fake news,” *arXiv preprint arXiv:2011.03092*, pp. 1–15, 2020.
[11] M. Alkhair, K. Meftouh, K. Smaïli, and N. Othman, “An Arabic corpus of fake news: collection, analysis and classification,” in *Arabic Language Processing: From Theory to Practice: 7th International Conference, ICALP 2019, Nancy, France, October 16–17, 2019, Proceedings 7*, 2019, pp. 292–302, doi: 10.1007/978-3-030-32959-4_21.
[12] H. Gharaibeh *et al.*, “Arabic sentiment analysis of Monkeypox using deep neural network and optimized hyperparameters of machine learning algorithms,” *Social Network Analysis and Mining*, vol. 14, no. 1, Jan. 2024, doi: 10.1007/s13278-023-01188-4.
[13] R. Alazaidah, G. Samara, M. Aljaidi, M. Haj Qasem, A. Alsarhan, and M. Alshammari, “Potential of machine learning for predicting sleep disorders: a comprehensive analysis of regression and classification models,” *Diagnostics*, vol. 14, no. 1, pp. 1–19, Dec. 2023, doi: 10.3390/diagnostics14010027.
[14] R. Azad, “A novel taxonomy for Arabic fake news datasets,” *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 159–166, Jul. 2023, doi: 10.12785/ijcds/140115.
[15] A. K. J. Saudagar and H. V. Mohammed, “Arabic character extraction and recognition using traversing approach,” *The International Arab Journal of Information Technology*, vol. 15, no. 3, pp. 360–368, 2018.
[16] F. Hanandeh, M. Akour, E. Al Daoud, R. Alshorman, and I. Alsmadi, “KP-trie algorithm for update and search operations,” *International Arab Journal of Information Technology*, vol. 13, no. 6, pp. 1–10, 2016.
[17] A. B. Nassif, A. Elnagar, O. Elgendy, and Y. Afadar, “Arabic fake news detection based on deep contextualized embedding models,” *Neural Computing and Applications*, vol. 34, no. 18, pp. 16019–16032, Sep. 2022, doi: 10.1007/s00521-022-07206-4.
[18] H. Himdi, G. Weir, F. Assiri, and H. Al-Barhamtoshy, “Arabic fake news detection based on textual analysis,” *Arabian Journal for Science and Engineering*, vol. 47, no. 8, pp. 10453–10469, Aug. 2022, doi: 10.1007/s13369-021-06449-y.
[19] A. Awajan, M. Alazab, R. A. Khurma, R. Alsaadeh, and M. Wedyan, “Fake news detection and prevention using artificial intelligence techniques: A review of a decade of research,” *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 14, pp. 326–337, 2022.
[20] S. E. Sorour and H. E. Abdelkader, “AFND: Arabic fake news detection with an ensemble deep CNN-LSTM model,” *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 14, pp. 5072–5086, 2022.
[21] W. Shishah, “JointBert for detecting Arabic fake news,” *IEEE Access*, vol. 10, pp. 71951–71960, 2022, doi: 10.1109/ACCESS.2022.3185083.
[22] H. Rahab, A. Zitouni, and M. Djoudi, “Arabic fake news and spam handling: Methods, resources and opportunities,” in *2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP)*, Nov. 2021, pp. 1–7, doi: 10.1109/AI-CSP52968.2021.9671174.
[23] F. L. Alotaibi and M. M. Alhammad, “Using a rule-based model to detect Arabic fake news propagation during Covid-19,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 112–119, 2022, doi: 10.14569/IJACSA.2022.0130114.
[24] M. Abd Elaziz, A. Dahou, D. A. Orabi, S. Alshathri, E. M. Soliman, and A. A. Ewees, “A hybrid multitask learning framework with a fire hawk optimizer for Arabic fake news detection,” *Mathematics*, vol. 11, no. 2, pp. 1–15, Jan. 2023, doi: 10.3390/math11020258.
[25] M. S. Hadj Ameur and H. Aliane, “AraCOVID19-MFH: Arabic Covid-19 multi-label fake news & hate speech detection dataset,” *Procedia Computer Science*, vol. 189, pp. 232–241, 2021, doi: 10.1016/j.procs.2021.05.086.

[26] A. R. Mahlous and A. Al-Laith, "Fake news detection in Arabic tweets during the COVID-19 pandemic," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 778–788, 2021, doi: 10.14569/IJACSA.2021.0120691.
[27] V. Thanasekaran and A. Selvaraj, "Low dimensional multi class steganalysis of spatial LSB based stego images using textural features," *The International Arab Journal of Information Technology*, vol. 21, no. 2, pp. 233–242, 2024, doi: 10.34028/iajit/21/2/6.
[28] M. Alqarni, "Embedding search for Quranic texts based on large language models," *The International Arab Journal of Information Technology*, vol. 21, no. 2, pp. 243–256, 2024, doi: 10.34028/iajit/21/2/7.
[29] H. Abu Owida *et al.*, "The performance of artificial intelligence in prostate magnetic resonance imaging screening," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 2, pp. 2234-2241, Apr. 2024, doi: 10.11591/ijece.v14i2.pp2234-2241.
[30] A. Al Sharah, H. Abu Owida, F. Alnaimat, and S. Abuowaida, "Application of machine learning in chemical engineering: outlook and perspectives," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 1, pp. 619-630, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp619-630.
[31] O. Tarawneh *et al.*, "The effect of pre-processing on a convolutional neural network model for dorsal hand vein recognition," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 3, pp. 1284–1289, 2024, doi: 10.14569/IJACSA.2024.01503126.
[32] S. Abuowaida *et al.*, "Proposed enhanced feature extraction for multi-food detection method," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 24, pp. 8140–8146, 2023.

BIOGRAPHIES OF AUTHORS

Hussain Mohammed Turki earned his bachelor's degree in computer science from the University of Anbar in Iraq and his MSc in computer science from Zarqa University in Jordan. His specializations are artificial intelligence, deep learning, and machine learning. He can be contacted via email at 20219109@zu.edu.jo.

Essam Al Daoud earned his Ph.D. in computer science from Universiti Putra Malaysia. As a professor at Zarqa University, he excels in research, curriculum development, and teaching, with expertise in data science, cryptography, and bioinformatics. He has also taught at institutions such as Multimedia University in Malaysia. Driven by a passion for academic excellence, he continues to contribute significantly to his field through research, teaching, and leadership roles. He can be contacted via email at essamdz@zu.edu.jo.

Ghassan Samara holds BSc. and MSc. in computer science and Ph.D. in computer networks. He obtained his Ph.D. from Universiti Sains Malaysia (USM) in 2012. His specializations are the internet of things, cryptography, vehicular ad hoc networks, wireless sensor networks, internet access protocols, accident prevention, alarm systems, cloud computing, computational complexity, computer crime, computer network security, data integrity, data privacy, delays, digital forensics, directed graphs, energy conservation, energy management systems, fog computing, graph theory, home automation, intelligent robots, and intelligent transportation systems. Currently, Dr. Samara is an associate professor and vice dean of the College of Information Technology at Zarqa University, Jordan. He can be contacted via email at gsamara@zu.edu.jo.

Raed Alazaidah received the B.Sc. degrees in computer information system, M.Sc. degrees in computer science from Al al-Bayt University, Jordan, in 2012 and 2015, respectively, and the Ph.D. degree in computer science from Universiti Sains Malaysia, Malaysia, in 2023. Dr. Abuowaida is currently working as an assistant professor in the Computer Science Department at Zarqa University, Jordan. Her research interests include deep learning, depth estimation, point cloud processing, and computer vision. He can be contacted via email at razaidah@zu.edu.jo.

Mais Haj Qasem received his bachelor's degree in computer science from the University of Jordan, Jordan, in 2011; the Master's degree in computer science from the University of Jordan in 2016; and his Ph.D. degree in computer science and artificial intelligence from the University of Jordan in 2020. Currently, he works as an assistant professor in the Data Science and Artificial Intelligence Department at Zarqa University, Jordan. His research interests include machine learning and multi-agent systems. She can be contacted via email at m.hajqasem@zu.edu.jo.

Mohammad Aljaidi received his B.Sc. (with distinction) in computer science from Zarqa University, Jordan, in 2014; the M.Sc. (with distinction) in computer science from Zarqa University, Jordan, in 2017; and the Ph.D. degree in computer science and artificial intelligence from the University of Northumbria at Newcastle, UK, in 2022. Dr. Aljaidi is currently working as an assistant professor with the Computer Science Department, Zarqa University, Jordan. His research interests include but are not limited to, EVs charging management and development, wireless sensor networks (WSNs), AI, cybersecurity, optimization, reinforcement learning. He can be contacted at email: razaidah@zu.edu.jo.

Suhaila Abuowaida received the B.Sc. degrees in computer information system, M.Sc. degrees in computer science from Al al-Bayt University, Jordan, in 2012 and 2015, respectively, and the Ph.D. degree in computer science from Universiti Sains Malaysia, Malaysia, in 2023. Dr. Abuowaida is currently working as an assistant professor in the Computer Science Department at Zarqa University, Jordan. Her research interests include deep learning, depth estimation, point cloud processing, and computer vision. She can be contacted via email at sabuoweuda@zu.edu.jo.

Nawaf Alshdaifat received the B.Sc degrees in computer science from AL al-Bayt University and M.Sc. degrees in computer science from the University of Jordan, Jordan, in 2002 and 2011, respectively, and the Ph.D. degree in computer science from Universiti Sains Malaysia, Malaysia, in 2023. His research interests include deep learning and machine learning. He can be contacted via email at n_alshdaifat@asu.edu.jo.