OKU Fen Bilimleri Enstitüsü Dergisi
7(3): 973-993, 2024

OKU Journal of The Institute of Science and Technology, 7(3): 973-993, 2024

Osmaniye Korkut Ata Üniversitesi
Fen Bilimleri Enstitüsü
Dergisi

Osmaniye Korkut Ata University
Journal of The Institute of Science
and Technology

# Fake News Detection and Classification with Recurrent Neural Network Based Deep Learning Approaches

**Halit ÇETİNER**[1*]

[1]Isparta University of Applied Sciences, Vocational School of Technical School, Computer Technology, Isparta

[1]https://orcid.org/0000-0001-7794-2555
*Corresponding author: halitcetiner@isparta.edu.tr

**Research Article**

***Article History:***
Received: 05.11.2022
Accepted:06.04.2023
Published online: 25.06.2024

***Keywords:***
Convolutional neural network
Natural language processing
Fake news classification
BiLSTM
GRU

**ABSTRACT**

Events happening in the world are transmitted to the end user through the news channel. The information transmitted from the news is generally considered to be accurate. However, there may be errors or lies in the information that circulates on the news channels. At the same time, this news has an impact on serious environments, such as the economy. In social networks where data sharing is increasing, news data is piling up uncontrollably. In these data piles, there is real information as well as different information that is not real commercial, political, or sales-orientated. False information and data expand faster as a result of sharing false information by users. This news directly affects users, causing erroneous transactions, misinformation, or financial loss. For the stated reasons, automatic fake news classification systems are proposed in this article by combining natural language processing with Recurrent Neural Network (RNN) based deep learning methods. The proposed systems were tested on a dataset containing 23.481 fake news and 21.417 real news using general performance metrics. As a result of the test processes, the proposed BiLSTM method provided 99.72% accuracy, while the proposed GRU method accessed 97.50% accuracy.

# Tekrarlayan Sinir Ağı Tabanlı Derin Öğrenme Yaklaşımları ile Sahte Haber Tespiti ve Sınıflandırması

**Araştırma Makalesi**

***Makale Tarihçesi:***
Geliş tarihi: 05.11.2022
Kabul tarihi: 06.04.2023
Online Yayınlama: 25.06.2024

***Anahtar Kelimeler:***
Convolutional neural network
Natural language processing
Sahte haber sınıflandırma
BiLSTM
GRU

**ÖZ**

Dünyada olup biten olaylar son kullanıcıya haber mecrası aracılığıyla aktarılmaktadır. Haberlerden aktarılan bilgilerin genellikle doğru olduğu düşünülmektedir. Ancak haber kanallarında dolaşan bilgilerde hata ya da yalan olabilmektedir. Aynı zamanda bu haberlerin ekonomi gibi ciddi ortamlarda etkisi de bulunmaktadır. Veri paylaşımının artış gösterdiği sosyal ağlarda haber verileri kontrolsüz bir şekilde yığılmaktadır. Bu veri yığınları içerisinde gerçek bilgiler olduğu gibi gerçek olmayan ticari, siyasi ya da satış hedefli farklı bilgilerde bulunmaktadır. Gerçek olmayan bilgiler, kullanıcılar tarafından paylaşılması sonucunda sahte bilgi ve veriler daha hızlı bir şekilde genişlemektedir. Bu tür haberler kullanıcıları doğrudan etkileyerek hatalı işlem yapmaya, yanlış bilgi sahibi olmaya veya maddi bir kayba neden olmaktadır. Belirtilen sebeplerden dolayı bu makalede doğal dil işleme Tekrarlayan Sinir Ağı (TSA) tabanlı derin öğrenme yöntemleri ile birleştirerek otomatik sahte haber sınıflandırma sistemleri önerilmiştir. Önerilen sistemler genel performans metrikleri kullanılarak 23,481 adet sahte haber, 21,417 adet gerçek haber içeren bir veri setinde test edilmiştir. Yapılan test işlemleri sonucunda önerilen BiLSTM yöntemi %99,72 doğruluk oranı sağlarken, önerilen GRU

yöntemi %97,50 doğruluk oranına ulaşmıştır.

**To Cite:** Çetiner H. Fake News Detection and Classification with Recurrent Neural Network Based Deep Learning Approaches. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2024; 7(3): 973-993.

## 1. Introduction

People use online social networking sites to maintain past or present friendships, meet new people, and build professional relationships. Online social networking sites such as WhatsApp, Twitter, Facebook and Instagram allow users to easily share information over the internet. For the main reasons, such as sharing information on these social networks is easy, fast and cheap, they are becoming more and more popular every day. At the same time, these social networks are the most used basic communication channels (Sahoo and Gupta, 2021). The popularity of online social networks in recent years has led to a rapid proliferation of messages with different social content, such as political news and product sales promotions.

Users who benefit from sharing and news sites, especially social networking sites, are affected by fake news. Some users increase the reputation of fake news by sharing fake news. However, this fake news also misleads other users as it is questionable information. Deliberate spreading of fake news to make people believe. It is difficult to detect false or fake news from deliberately disseminated fake news content. Given the speed with which fake news spreads through sharing, there is a need for approaches to automatically detect fake news.

Due to the increase in the use of the Internet, fake political talk, satire, fake rumours, and misleading content are growing to target online media platforms. It is stated that the political attitudes of citizens watching fake news have changed (Balmas, 2014). Furthermore, it is stated that trust in politicians has been shaken by people exposed to real and fake news. It is reported that the effect of perceiving fake news as real in the formation of this situation is great. It is stated that accounts were created to spread fake news on social networking sites in the political elections of different countries (Sahoo and Gupta, 2021).

Information is also spread on social networking platforms to create mistrust by confusing users. Detecting fake news on social networks is a tedious and laborious task. In such an environment, fake news can spread very quickly. This situation can affect millions of users in the virtual world and change the balance in the real world. Sometimes fake news can contain real evidence, although with false meanings (Feng et al., 2012). So much so, in fact, that since 2016, the creation and dissemination of fake news, with the support of academics and researchers, has been aimed at influencing real life (Horne and Adali, 2017). It is understood that this desired goal has been achieved with the intensive use of the FaceCheck.org and PolitiFact.com websites by users to determine whether the news is fake or false (Sahoo and Gupta, 2021). Fake news detection websites, such as FaceCheck.org and PolitiFact.com, play an important role in the detection of false and fake news on the Internet. However, these systems operate manually. In other words, the news that the system wants to detect if it is a lie is manually searched and checked. For this to work properly, continuously, and quickly, a sufficient number of experts are needed to detect false and fake news. In general, many profiles in social networking systems spread fake news via social networking platforms by preparing different posts (Sahoo and Gupta, 2019). In this direction, hybrid approaches that use machine learning techniques to detect malicious profiles on the Twitter social networking platform appear to be suggested (Sahoo and Gupta, 2019).

To detect whether fake news is real news or not, it is necessary to obtain the correct features. The fact that a site may have thousands of readers is no guarantee that the site will not have false or fake news stories. In addition, there are many types of fake news such as false statements, false advertisements, satirical news, and conspiracy theories. This type of news can affect people's lives in all ways. At the same time, it has the feature of directly affecting the opinions, interests, and decisions of the public. Among all these effects and difficulties, the detection of fake news on social networking platforms becomes difficult. In general, the detection of malicious content in the form of fake news, which is formed as a result of spreading misleading information by users who use more than one social network to gain a certain financial gain, is becoming increasingly important day by day. Rumors that occur without knowing the effect of fake news can create confusion (Barthel et al., 2016). Due to misleading information, division and racism, activities are seen to increase among people due to increasing polarization (Bakshy et al., 2015). Fake news about an honest, respectable, nationalistic, and successful businessman or politician can suddenly be a blow to that politician's credibility (Sunstein, 2009).

The main contributions of the study on fake news detection, which was developed motivated by these reasons, to the literature are presented below.

- Two different deep learning models, based on Recurrent Neural Networks (RNN), are proposed to make it possible to perform automated fake news detection
- Accuracy rates of 99,72% and 97,50% were achieved using the pretrained Keras embedding layer with the proposed Bidirectional Long Short-Term Memory (BiLSTM) and Gated recurrent units (GRU) models, respectively.
- Although the BiLSTM method holds information in memory for a long time in forward and backward hidden situations, it has been found that its success depends on the dataset.
- The GRU method gave successful results in terms of F1 score, precision, recall and accuracy, similar to the BiLSTM method.

The article consists of 3 sections, not including the introduction. The second section provides information on a recent dataset of real and fake news on which experimental studies have been conducted. In the third section, deep learning models are presented that allow for the automatic classification of a dataset consisting of fake news and real news. In the final section, the results of the study are evaluated and concluded.

## 2. Literature Review

This section examines the automatic fake news detection systems that have been proposed due to the rapid spread of false information and fake news on social networks. Different research methods, including deep neural networks and text-based representations (Karimi et al., 2018), semi-supervised (Guacho et al., 2018), and supervised or unsupervised (Hosseinimotlagh and Papalexakis, 2018) detection, are found to detect and classify false and fake news with the stated effects.

Yang et al. determined that there are those who aim to create commercial and political perceptions by publishing fake news from official or personal social media accounts (Yang et al., 2018). Yang et al. proposed a model that classifies real-world fake news with a Convolutional Neural Network (CNN) based model (Yang et al., 2018). They proposed a model that enables the classification of rapidly spreading fake news with deep learning based on Long Short Term Memory (LSTM) and RNN (Wu and Liu, 2018). Sabeeh et al. performed fake news classification with a CNN-based model (Sabeeh et al., 2020). User comments are seen as the most important signal representing user intentions. In general, fake comments have been determined to be written by illegitimate users (Sabeeh et al., 2020). Sabeeh et al. do opinion mining on user comments to perform reliability analysis of Twitter metadata. Their proposed method uses the information source to extract attributes about a particular event. They also used the SentiWordNet method to assess cognitive clues. They proposed an interpretation system based on objective factors such as sensitivity and reliability score using Bidirectional Gated Recurrent Neural Network. Karimi et al. proposed a system to detect fake news of various degrees to prevent fake news spread through the media from misleading readers (Karimi et al., 2018). They conducted a comprehensive study for the detection of fake news in multi-source data. Hosseinimotlagh and Papalexakis extract the full potential of news content, capturing hidden relationships and contextual relationships between terms (Hosseinimotlagh and Papalexakis, 2018). They create a tensor model from the captured information. They can group field news on real data with an average of 80% accuracy.

Wang et al. report that fake news content is a nightmare for the public and states, as reading news on social media becomes more common today (Wang et al., 2018). They declare that one of the most important issues in social media is how to determine whether a new event is fake or not. For the stated reason, they proposed a neural network from which event-independent features can be derived. This approach consists of three main components. The first component extracts textual and visual features from the texts. The second component learns the discrimination for the detection of fake news with the information from the feature extractor. The third component removes event-specific features and retains the remaining features. Experimental studies were conducted on data collected from the Weibo and Twitter social sharing systems. Reis et al. conducted studies on the detection of news in the field using classical machine learning-based methods. They achieved an 81% accuracy rate with the XGBoost machine learning method (Reis et al., 2019). On the other hand, Özbay and Alatas detected fake news with a two-stage method (Ozbay and Alatas, 2020). The performance of the implemented system has been tested on three real datasets. Their proposed method consists of two stages. In the first phase, unstructured raw datasets undergo some preprocessing. In the second stage, an experimental study was carried out with 23 artificial intelligence algorithms. Kaur et al. proposed an ensemble algorithm fed by twelve classifiers to classify the fake information obtained from three different feature extraction methods (Kaur et al., 2020). Their proposed algorithm outperformed passive aggressive, logistic regression, and linear svc models. When their proposed algorithm is examined, it is seen that they create a dataset by feeding from new trends, Kaggle and Reuters sources. Preprocesses were applied to the created dataset. In the preprocesses, it is seen that the repeating words and stop words are removed. In another preprocessing step, empty unnecessary values are shown to be removed. After the preprocessing step, the dataset is seen to be used as 67% training data and 33% test data. No cross-validation processes were used in any way. Afterwards, the feature extraction stage was started in order to extract meaningful features from the textual data. At this stage, the feature is extracted with the methods of term frequency, count vectorizer and hashing vectorizer. The classification algorithms are fed with the features obtained from the specified algorithms. Twelve different classifiers such as Adaboost and Support Vector Classifier are used. In the classification phase, it is seen that the classification is done according to the best results from the classifiers with multilevel voting. Through the model formed by combining these processes, real and fake news detection is carried out. Tacchini et al. applied logistic regression to determine whether Facebook posts were fake (Tacchini et al., 2017). They tried to find fake news by analysing the linguistic clues of texts with machine learning (Conroy et al., 2015).

Nasir et al. report that fake news is becoming more common as people do less research and filtering than before (Nasir et al., 2021). They proposed a hybrid deep learning model based on CNN and RNN to prevent the rapid spread of fake news. Classifiers mean that input texts in vector or numeric format should be fed. Every word found in fake news texts is represented by vectors. These word vectors are defined as word embedding (Nasir et al., 2021). The features obtained from the one-dimensional convolution layers are given as input to the RNN-based LSTM algorithms. With the proposed method, they aimed to capture local and sequential features in the input data. Nasir et al. separated the dataset they used in the experimental study as 80% training and 20% testing.

## 3. Material and Methods

In this section of the study, the deep learning methods proposed to perform an automatic classification of fake news created with political and commercial concerns are explained. At this point, real-world problems prompt researchers to turn to academic research to offer solutions. Political and commercial concerns and events related to fake news have allowed this study to be conducted.

*3.1. Material*

The fake news content used in the article is taken from Kaggle.com (Ahmed et al., 2018a). Actual news was obtained from the Reuters.com news site. The content of each of the news content is longer than 200 characters. The news content consists of English text content. Articles with real news contain more nouns and adjectives. On the other hand, fake news contains more adverbs and verbs. Figure 1 shows the numerical distribution of the topics in the dataset. How many contents from each topic is presented numerically in graphics. The number of unique words after preprocessing is 108.705. The maximum number of words in a preprocessed content is 4.405.

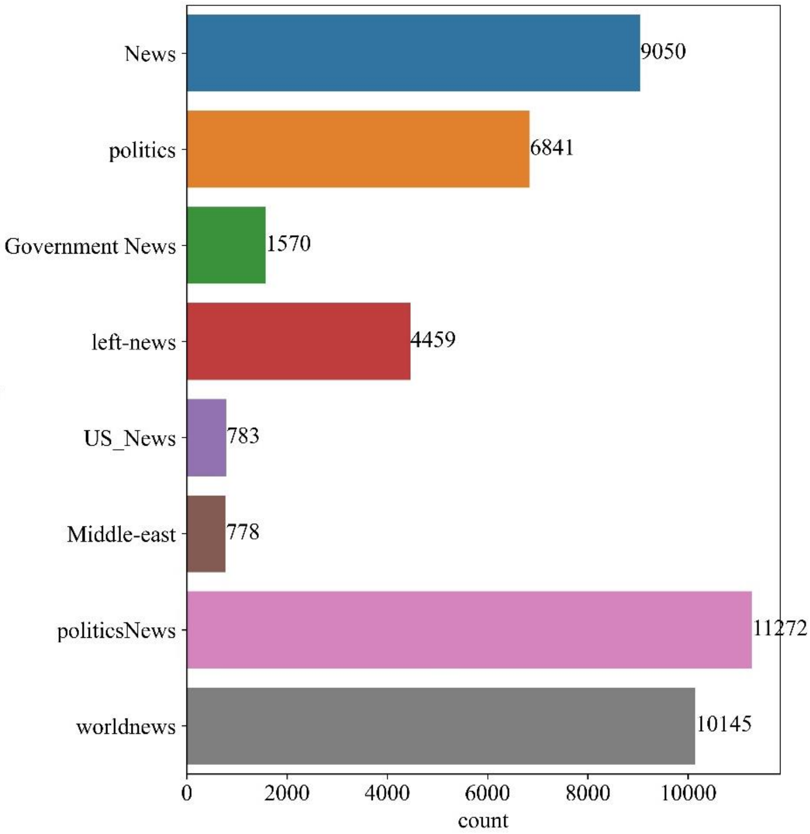


**Figure 1.** Numerical distribution of topics in the dataset

When Figure 1 is examined in general, it can be said that political news is dominant. The dataset, which was created by bringing together news content on different topics, from US news to world news, was used in experimental studies.

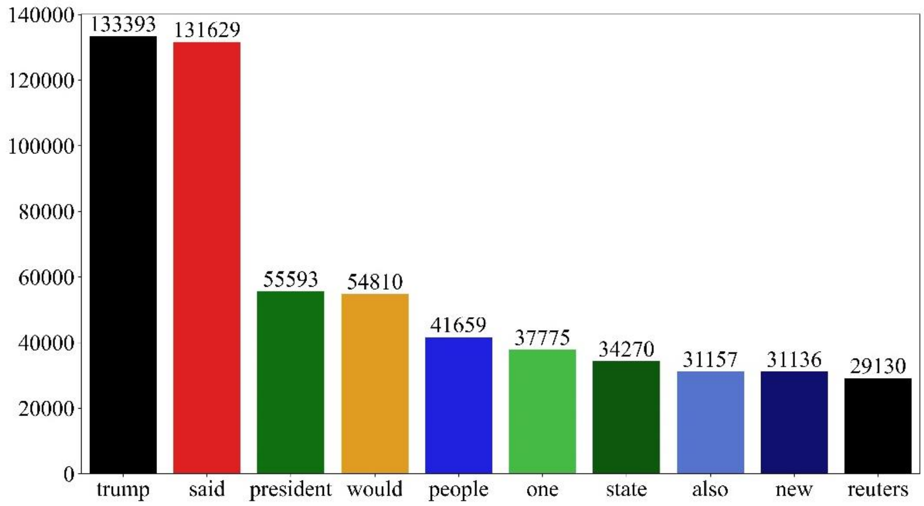


**Figure 2.** Words with high word frequency in the dataset

In Figure 2, high-frequency words that are frequently used in the preferred dataset for the experimental study are shown. Frequency values of different words are shown, from 133.393 trump words to 29.130 Reuters words. A publicly available dataset was chosen to carry out a study on the automatic detection of fake news (Ahmed et al., 2017; Ahmed et al., 2018). In the dataset used by the article study, there are 23.481 fake news and 21.417 real news.

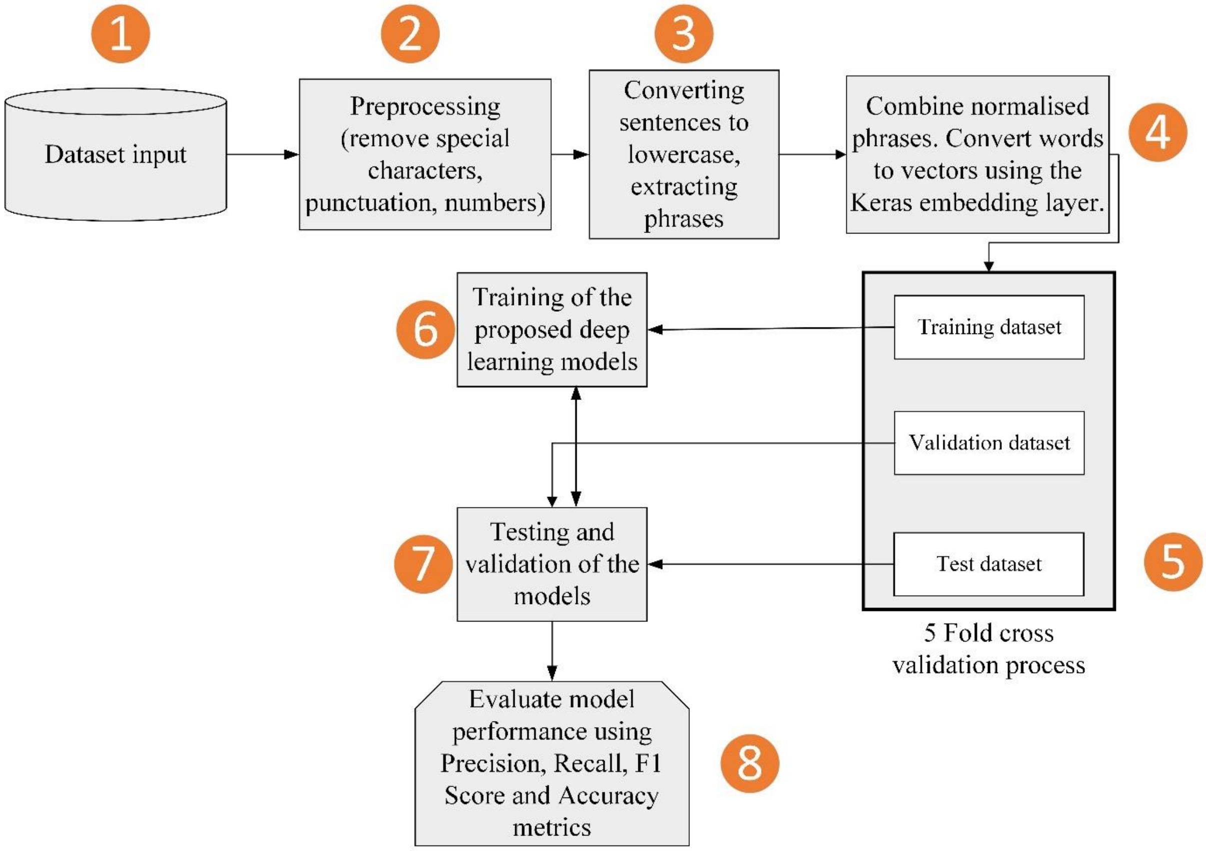


**Figure 3.** Flow diagram of proposed models

Figure 3 shows the flow diagrams of the proposed models. In this flow diagram, it is desired to perceive whether a news is real or fake news. When the specified fake and real news output is specified as the dependent variable, all remaining inputs and parameters can be defined as independent variables. The flowchart of proposed models consists of eight steps. The first step represents the dataset used in the article. In the second and third steps, the methods expressed in the preprocessing title are defined. In the fourth step, vector transformation with Keras word embedding is expressed. The fifth step shows the partitioning of the data with the 5 fold cross validation technique. In the sixth and seventh steps, the training and testing processes of each proposed model are defined. In the eighth step, the performance graphs of the proposed models are measured.

### *3.2. Preprocessing*

The dataset is divided into training and test data according to the 5 fold cross validation technique. Later preprocessing methods brought the news in all sections to the same standard. Preprocessing removes the proposed systems from words that have no effect on feature extraction. For this purpose, all texts in the dataset were converted to lowercase letters, and structures such as 'won't' were converted to 'will not' structures. Extra spaces and unwanted words have been removed, especially in all site extensions such as URL. The removal of the stop words that do not make any sense on their own contributed to the normalization of the dataset. The texts have been made ready for digitization by removing the characters that make it difficult to understand the words before and after, such as special symbols and numeric characters.

Almuzaini et al. seem to attempt to find out how text classification is affected by word embedding and stemming strategies (Almuzaini and Azmi, 2020). Both are very important topics in natural language processing. Although the Keras embedding method is used in this article, no experimental study has been conducted on stemming strategies. In further studies, the effect of stemming strategies and word embedding techniques such as Global Vectors for Word Representation (GloVe), Word2Vec, Bert on classification performance can be investigated in detail on different datasets similar to the dataset used in this article.

### *3.3. Keras word embedding*

Word embedding is a method that evaluates the semantic and syntax meanings of a vocabulary. Word embedding, in other words, is a natural language processing learning technique that is formed as a result of matching the phrases in the dataset with real numbers (Gulli and Pal, 2017). In any natural language processing task, text needs to be represented by vectors. In this way, the way to process digitized texts is opened with standard machine learning algorithms and deep learning models.

In this study, Keras library, which is a deep learning library, was used. Keras is an open source library written in Python programming language. With the help of the tokenizer, text_to_sequences, and pad_sequences functions of the Keras library, the texts are separated. None of the ngram techniques were used. The Keras embedding layer transforms the text inputs defined in the inputs of deep learning models into vectors. It is also a flexible layer trained to embed words that match random weight values. Embedding is a learning paradigm that paves the way for the structure developed within the scope of this article to be used in other models (Li, 2018). In the Keras embedding layers used in the first layers of the proposed models, 10.000 word inputs expressing the maximum number of words were used. The value 32 is given as the output.

### *3.4. GRU*

In this article, RNN structures that can control the flow of information without a memory unit in the automatic analysis of fake news content are called GRUs. In this structure, all confidential situations can be used without information control. Although GRU is based on the RNN structure, it is represented with fewer parameters. As a result, it has less processing load. At the same time, it can be processed faster as a result of processing with fewer parameters. Although GRU is stated to have less success in retrospective transactions, it is preferred due to rapid training modelling (Wang et al., 2022). GRU architecture consists of two different gates, the current and reset gates in the GRU structure (Hu et al., 2021). Between recurrent networks, the gate that decides to merge the current network with the memory structure of the previous network is called the reset gate. The other gate determines how long the information in the memory will be retained. If it is necessary to define the two different doors briefly mentioned, the relevant doors are defined by the formula (Equations 1-4).

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{1}$$

$$Update\ gate\ (z_t) = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{2}$$

$$\tilde{h}_t = \tanh\big(W_h x_t + r_t \odot (U_h h_{t-1})\big) + b_h \tag{3}$$

$$Reset\ gate\ (r_t) = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{4}$$

$W$ represents the weight at time $t$. $h_{t-1}$ represents the values of the hidden layer at time $t-1$. $U$ denotes cell units and $\sigma$ denotes sigmoid activation function (Hu et al., 2021). $z_t$ decides how much new information to add to keep the old information. While $x_t$ shows the inputs to the model, $r_t$ controls the effect of the parameters at time $t-1$ on the current situation.

### *3.5. BiLSTM*

CNN methods are successful in extracting local meanings from spatial data while having difficulty in learning sequential data (Liu and Guo, 2019). RNNs, on the other hand, are more successful with sequentially arrayed data. When fake news classification is considered as text classification, RNNs can also be used easily. In recurrent neural networks, the processing time increases as the depth of the input increases. At the same time, these methods have loss functions with variable sensitivities. The presence of different gradient values in model layers based on recurrent neural networks causes variation in the loss in each layer (Aggarwal, 2018). Problems occur as a result of disappearance and gradient bursts, which are often encountered in recurrent neural (Liu and Guo, 2019). This problem arises during the backpropagation of RNN-based methods. Although such problems are frequently encountered, it gives good results for short-term transactions (Pang et al., 2020).

Although RNN-based methods can remember in short-term transactions, they suffer memory loss in long-term transactions. As a result of research to eliminate memory loss, LSTM (Hochreiter and Schmidhuber, 1997) architecture has been developed to provide long-term recall in memory. The LSTM architecture effectively solves the gradient burst and vanishing gradient problems, which are fundamental in RNN architectures. Effective at extracting meaningful information, including multilabel or large text. LSTM can capture short-term dependencies as well as long-term dependencies.

BiLSTM (Zhang et al., 2019) on the other hand, is an advanced structure that allows the LSTM method to navigate between bidirectional series. It also represents an interconnected structure between forwarding and backward hidden states. The structures of BiLSTM methods combining advanced and hidden layers in text series are encountered (Chen et al., 2017; Niu et al., 2017; Nowak et al., 2017). LSTM solves the disappearing gradient problem by replacing hidden units with memory blocks. Memory blocks store information using the memory cell. Due to this storage feature, LSTM is better than the classical RNN structure in finding and using long-term information series. Memory units in LSTM structures are the units that decide when the network learns and forgets new information.

$$f_t = \sigma\left(W_f X_t + U_f h_{t-1} + b_f\right) \tag{5}$$

$$i_t = \sigma(W_i X_t + U_i h_{t-1} + b_t) \tag{6}$$

$$o_t = \sigma(W_o X_t + U_o h_{t-1} + b_o) \tag{7}$$

$$g_t = tanh\left(W_g x_t + R_g h_{t-1} + b_g\right) \tag{8}$$

$h_t$ shows the hidden states. $h_{t-1}$ shows the hidden state at time $t-1$, that is, the hidden state before the current state. $x_t$ represents the inputs to the units. Structures $f_t$, $i_t$, and $o_t$ at the $t$ show gates. $g_t$ is the state layer at the $t$. $W$, $R$, and $b$ represent weight, recurrent weight, and bias values (Equations 5-8).

## 4.Experimental Results and Discussion

In this article, a classification study was carried out on the dataset with 23.481 fake news and 21.417 real news. For this purpose, preprocessing steps are applied to the raw data and the data is cleaned. These preprocessing steps consist of removing prepositions, such as numeric or special characters. Apart from these, all characters that do not make sense on their own, such as quotation marks, comment lines, and symbols, are cleaned from the raw data. After these processes, the preprocessed data was obtained. Subsequently, the preprocessed data are divided into training, testing, and validation data. The area up to this step is common to both proposed models. However, draught model drawings were made to explain the next steps for both models separately.

The proposed models were developed using the tensorflow and Keras libraries in a python 3.8 environment on a computer with a NVIDIA GeForce RTX 3060 graphics card. The performances of the proposed models are calculated using accuracy, recall, precision and F1 score metrics (Equations 9-12). In the hyperparameter adjustment of the models, 128 batch size values were used with the Adam optimization method.

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

$$Recall = \frac{TP}{TP+FN} \tag{10}$$

$$Precision = \frac{TP}{TP+FP} \tag{11}$$

$$F1 = 2x\frac{PrecisionxRecall}{Precision+Recall} \tag{12}$$

The flow diagram of the proposed BiLSTM model is shown in Figure 4. The proposed BiLSTM model consists of 10 layers. The activation functions and the number of neurones used in the layer are explained in 11 steps. The tenth and eleventh steps represent the classification layer. Fake news classification is considered a binary classification (Kishwar and Zafar, 2023). Kishwar and Zafar classified the new dataset they created for the country of Pakistan using the 1D CNN model and the sigmoid activation function (Kishwar and Zafar, 2023). Sigmoid activation is preferred because there is no multiple classification. Both proposed models have an epoch number of 10. The batch size is set to 32. The results of the performance of the experimental study were effective in determining these values. In Figure 4, the data are given as input in the first step. In the second step, the input words are heavily digitized. In the third step, a BiLSTM layer with 256 neurones that can navigate between the previous and next hidden states has been added. In the fourth step, another 128-neurone BiLSTM layer was added, similar to the BiLSTM layer in the third step. In the fifth step, to prevent over-learning of the model, a 0.1 neuron dropout layer was added. In the sixth step, a batch normalization layer was added to normalize the mini steps between the layers.

In the seventh step, a dense layer with 128 neurons was added and the linking process was performed with the previous connections. In the eighth step, a dropout layer was added, which performs a 0,2% drop in neurones, similar to the layer in the fifth step. In the ninth step, a fully connected layer with 64 neurons with ReLU activation function, distant and near sentences from previous layers are connected to a small region. In the tenth step, the classification values are obtained with a classification layer with a sigmoid activation function. In the eleventh step, the estimated class value is determined based on the values obtained.

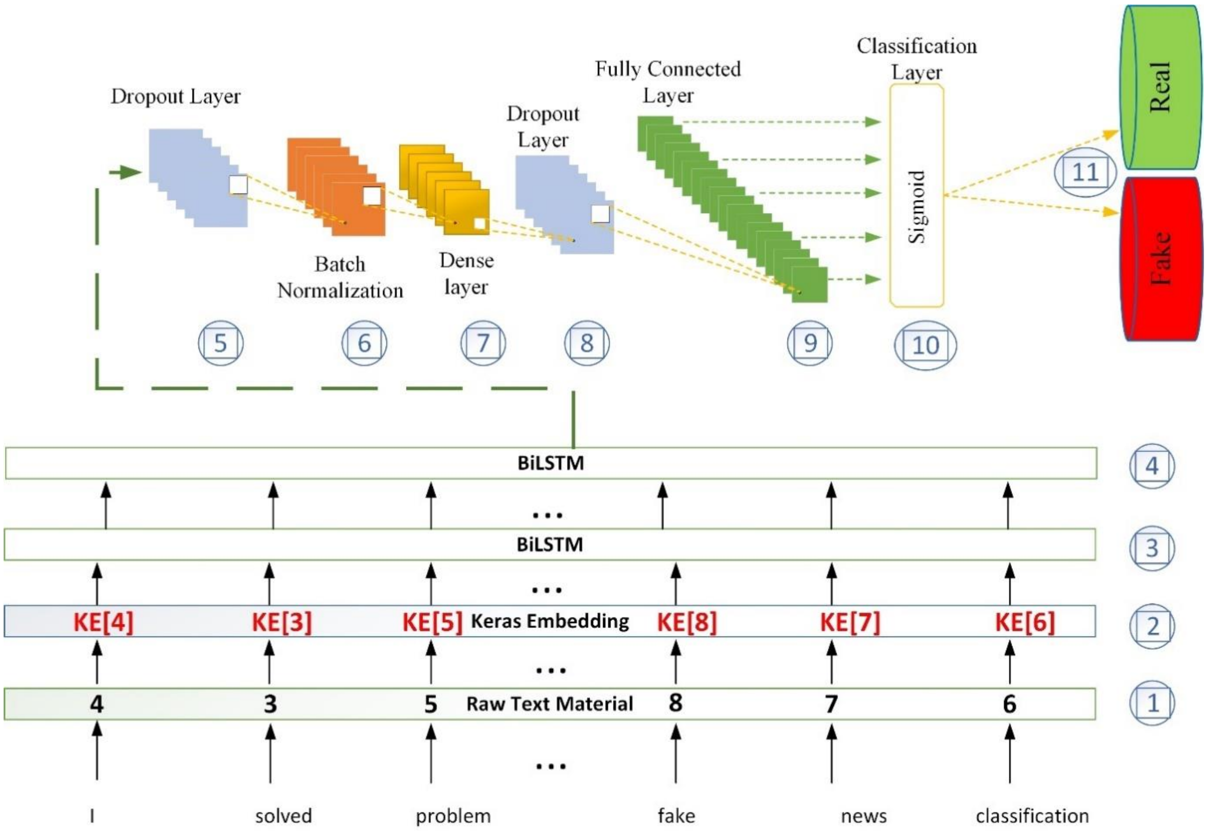


**Figure 4.** Proposed BiLSTM model

The performance results of the Proposed BiLSTM model according to the 5 fold cross validation technique are given in Table 1. Graphic drawings of the BiLSTM model were made by choosing the model with the lowest accuracy value according to the given performance results. At the same time, detailed analysis of F1 score, recall, precision and accuracy performance measurements were made according to the preferred 5 fold cross validation model.

**Table 1.** Proposed BiLSTM model 5 fold cross validation results

| Fold | Accuracy | Loss |
|---|---|---|
| Fold 1 | 0.9917 | 0.0377 |
| Fold 2 | 0.9982 | 0.0078 |
| Fold 3 | 0.9991 | 0.0041 |
| Fold 4 | 0.9997 | 0.0007 |
| Fold 5 | 0.1000 | 1.3283 |
| Average | 0.9977 | 0.0101 |

The experimental results of the BiLSTM model are presented in Figure 5. When the presented are examined, the training accuracy rate is close to the test accuracy rate. The fact that both curves are close to each other clearly shows that the proposed model is a fundamental study that can contribute to the literature. Detailed versions of the performance results shown in Figure 5 are presented in Table 2. When the presented data are examined, it is observed that the precision, recall, F1 score, and accuracy data are compatible with each other. While the training accuracy rate was 0,9989, the test accuracy rate reached 0,9972.

**Table 2.** Performance results of the proposed BiLSTM model according to Fold 1

| Model | Precision | Recall | F1 score | Accuracy |
|---|---|---|---|---|
| BiLSTM model (training) | 0.9988 | 0.9988 | 0.9988 | 0.9989 |
| BiLSTM model (test) | 1.00 | 1.0 | 0.9977 | 0.9972 |

In Figure 5, it is seen that the starting point for the training and testing processes is 0.87 and above. This is possible by making the words weighted with the Keras embedding layer. In Figure 5, while the epoch number of the model is shown on the x axis, the accuracy or loss information is plotted on the y axis. In other words, Adam (Kingma and Ba, 2014) with the y-axis ratio of 0.01 shows the loss and accuracy of the information calculated using the optimization method.

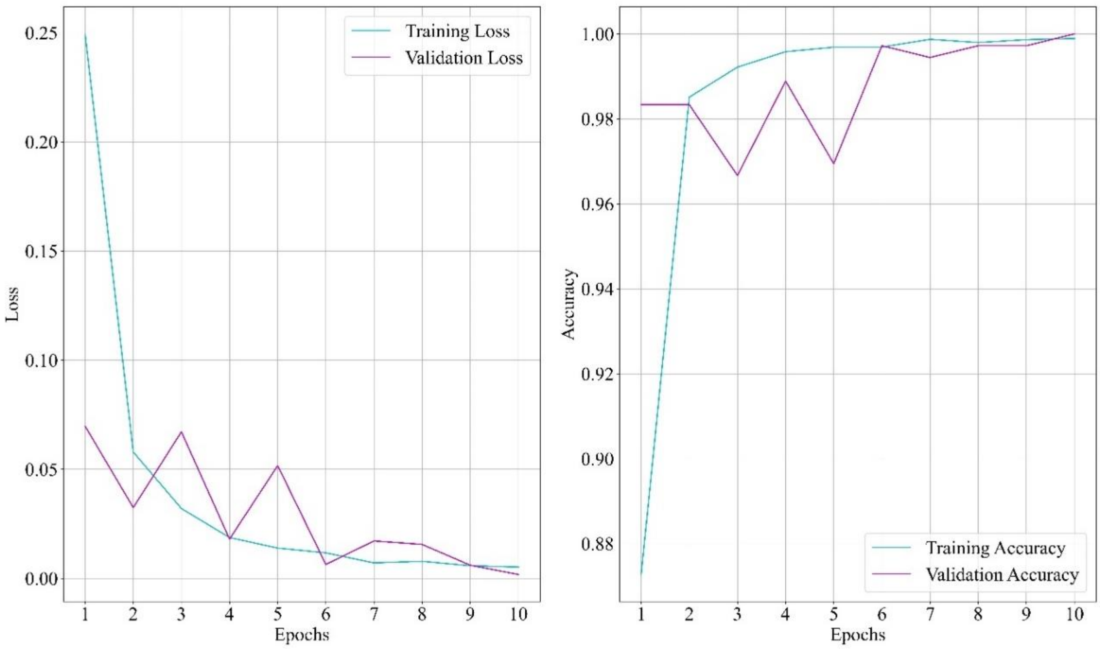


**Figure 5.** Proposed BiLSTM model performance results

The flow diagram of the proposed GRU model is shown in Figure 6. The proposed GRU model consists of 10 layers. The activation functions and the number of neurones used in the layer are explained in 11 steps. The tenth and eleventh steps represent the classification layer. Sigmoid activation is preferred because there is no multiple classification. In Figure 6, the data are given as input in the first step. In the second step, the input words are predominantly digitised.

In the third step, a GRU layer with 256 neurons that can navigate between the previous and next hidden states has been added. In the fourth step indicated in Figure 6, another layer of GRU with 128 neurons was added, similar to the layer of GRU in the third step. In the fifth step, to prevent overlearning of the model, a 0.1 neuronal dropout layer was added. In the sixth step, a batch normalization layer was added to normalize the mini steps between the layers.

In the seventh step in Figure 6, a dense layer with 128 neurones was added and the connection process was performed with the previous connections. In the eighth step, a dropout layer was added, which causes a 0.2% drop in the neurone, similar to the layer in the fifth step.

In the ninth step, the far and near sentences from the previous layer are connected to a small region on the fully connection layer with 64 neurons with a ReLU activation function. In the tenth step, the classification values are obtained with a classification layer with a sigmoid activation function.

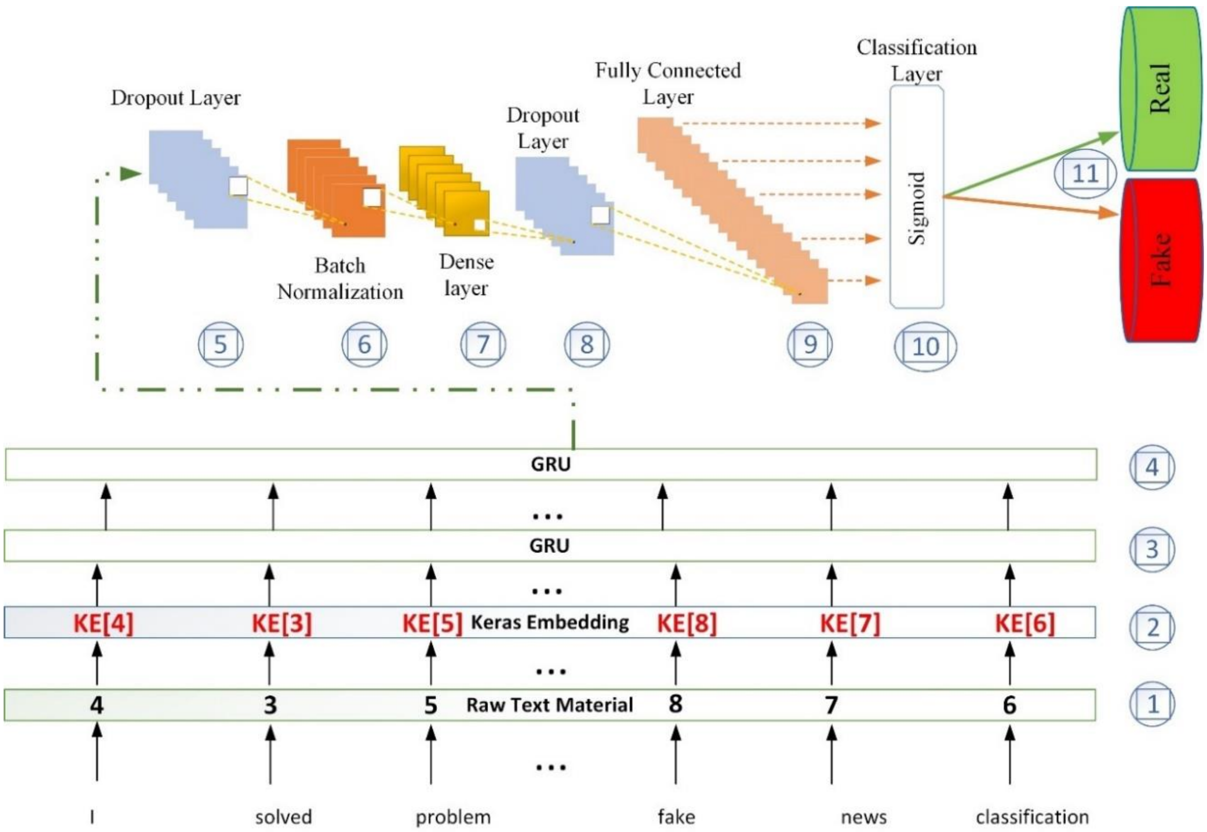


**Figure 6.** Proposed GRU model

In the eleventh step, the estimated class value was determined based on these values obtained. When both proposed model structures are examined, all layers, except the layers specific to BiLSTM and GRU neurones, are kept the same in the proposed models to make a correct performance comparison.

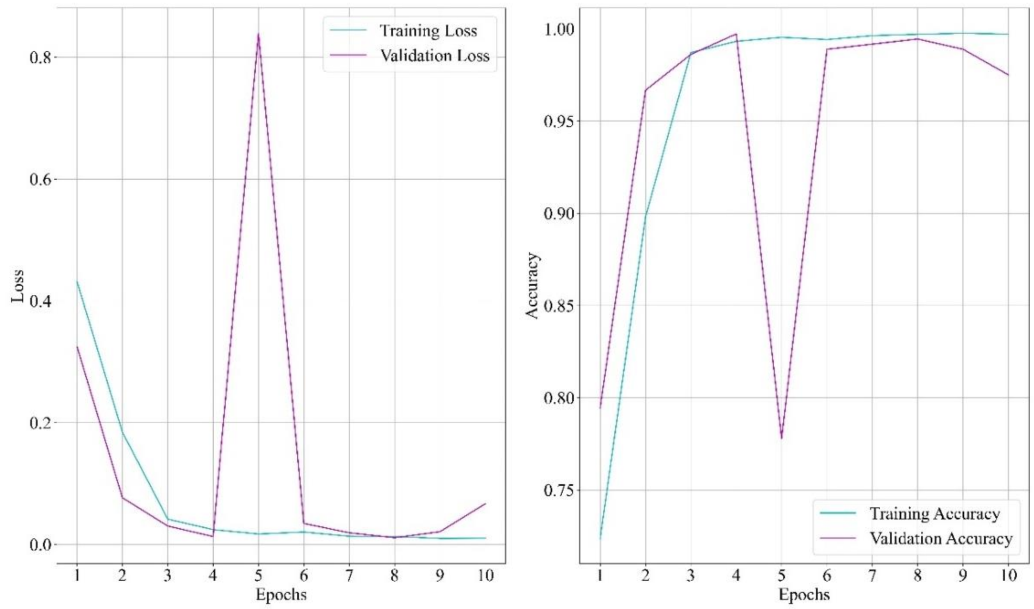


**Figure 7.** Proposed GRU model performance results

The performance results obtained according to the 5 fold cross validation technique of the proposed GRU model are given in Table 3. On the basis of the performance results given, the model with the lowest accuracy value was preferred, and the graphic drawings of the GRU model were made. At the same time, detailed analysis of F1 score, recall, precision and accuracy performance measurements were made according to the preferred 5 fold cross validation model.

**Table 3.** Proposed GRU model fold results

| Fold | Accuracy | Loss |
|---|---|---|
| Fold 1 | 0.9829 | 0.0752 |
| Fold 2 | 0.9976 | 0.0102 |
| Fold 3 | 0.9994 | 0.0020 |
| Fold 4 | 0.9995 | 0.0024 |
| Fold 5 | 0.9996 | 0.0024 |
| Average | 0.9958 | 0.0184 |

The experimental results of the GRU model, are presented in Figure 7. When the data presented are examined, the training accuracy rate is very close to the test accuracy rate. The fact that the training and test curves are close to each other clearly shows that the proposed model is a fundamental study that can contribute to the literature. Detailed versions of the performance results shown in Figure 7 are presented in Table 4. When the data presented are examined, it is observed that the precision, recall, F1 score, and accuracy data are compatible with each other. Although the training accuracy rate was 0.9971, the test accuracy rate reached 0.9750.

**Table 4.** Performance results of the proposed GRU model according to Fold 1

| Model | Precision | Recall | F1 score | Accuracy |
|---|---|---|---|---|
| GRU model (training) | 0.9969 | 0.9970 | 0.9967 | 0.9971 |
| GRU model (test) | 0.9940 | 0.9555 | 0.9740 | 0.9750 |

Although the BiLSTM model is stated to give better results than RNN-based methods in text analysis (Liu and Guo, 2019), this study has shown that this situation varies depending on the dataset and the sequence of the data.

**Table 5.** Comparison results with similar studies in the literature

| Model | Precision | Recall | F1 score | Accuracy |
|---|---|---|---|---|
| (Bali et al., 2019) | - | - | 0.91 | 0.88 |
| (Kishwar and Zafar, 2023)'s CNN | 0.93 | 0.93 | 0.93 | 0.93 |
| (Kishwar and Zafar, 2023)'s LSTM | 0.86 | 0.85 | 0.86 | 0.94 |
| (Ngada and Haskins, 2020)'s DT | 0.95 | 0.95 | 0.95 | 0.94 |
| (Ngada and Haskins, 2020)'s KNN | 0.96 | 0.86 | 0.91 | 0.91 |
| Proposed GRU model | 0.9940 | 0.9555 | 0.9740 | 0.9750 |
| Proposed BiLSTM model | 1.00 | 1.0 | 0.9977 | 0.9972 |

The comparison table made with the articles using the same or similar dataset used in this article is given in Table 5. In the study of (Kishwar and Zafar, 2023), a new dataset was created by collecting fake news content of only Pakistan from different datasets, including the dataset used in this article. Instead of creating a new deep learning model, a dataset is created by collecting data from different data sources. Bali et al. (2019) obtained a better fake news classification result with the Gradient Boosting algorithm than other machine learning algorithms in their study. Ngada and Haskins (2020) used the same dataset used in this article. The classification results obtained with the Decision Tree (DT) and K Nearest Neighbors (KNN) algorithms, which are machine learning algorithms, are presented in Table 5. The differences of the methods proposed within the scope of this article are presented below:

- Instead of dividing the training and test data into fixed values such as 80% and 20% as in the study by (2020), the 5 fold cross validation process was applied. Due to this, a different result is prevented every time the code is run.

- As in Ngada and Haskins (2020) study, new deep learning models have been developed that detect the most distinctive features instead of increasing the cost of the system by using weight values such as pretrained FakeNewsCorpus.
- The datasets used in the Bali et al. (2019) and (Kishwar and Zafar, 2023) studies are different from the dataset we used in this study. With these studies, the common features are ready to extract features from the data by applying preprocessing on each dataset.

When the given results are examined carefully, it is seen that the proposed model has similar competence to the studies in the literature. Experimental studies have been conducted with CNN-based approaches to classify fake news datasets with the same dataset used in this article (Çetiner, 2022). However, the inadequacy of the research and the fact that CNN methods do not have as much success in natural language processing as RNN-based methods necessitated looking at the study from a different perspective. At the same time, the proposed models are divided into training and testing according to cross validation methods. According to similar studies, its success is no longer relative. In In addition to the items mentioned, the proposed system, Newsbag Jindal et al. (2020) to analyze and classify fake news datasets.

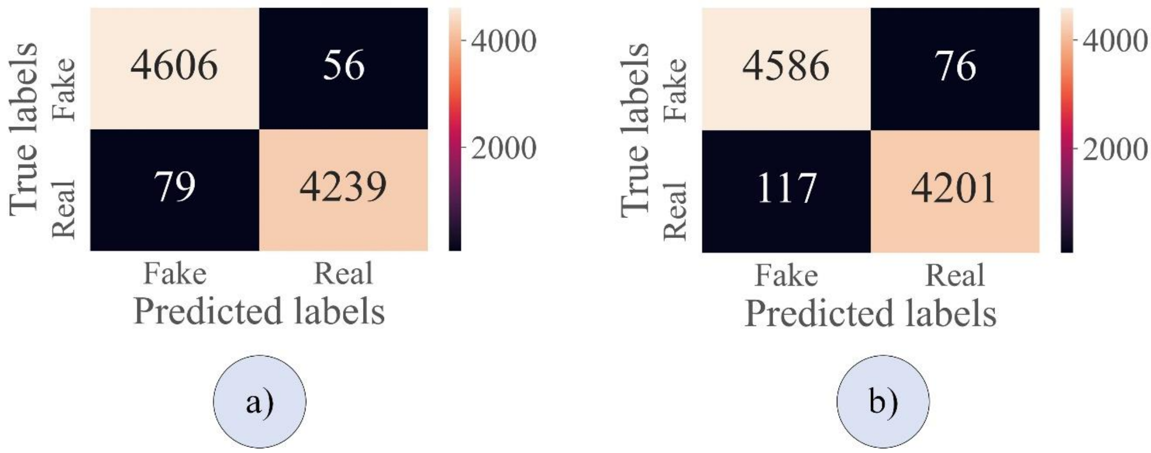


**Figure 8.** a) confusion matrix of proposed BiLSTM model, b) confusion matrix of proposed GRU model

The results of the confusion matrix of the proposed models that confirm the results given in Table 5 are presented in Figure 8. The proposed BiLSTM model gave more successful results in terms of the confusion matrix than the proposed GRU model.

## 5. Conclusions

Academic research was conducted to prevent the increase in false and fake news due to the prevalence and speed of social sharing systems. As a result of academic research, a study was carried out for automatic classification using a public dataset using deep learning models. To automatically classify fake news, experimental studies were carried out with the proposed BiLSTM and GRU model in a dataset open to researchers with 23.481 fake news and 21.417 real news. The performance values obtained are satisfactory. The fact that the BiLSTM method keeps the text context in memory for a long time in connection with the backward and forward hidden states makes this method even better. The GRU model provided as much accuracy as the BiLSTM method. The sequence and features of the data are believed to be effective in achieving this. Although the BiLSTM model has been reported to be more successful than the GRU method in the literature, it was found to be more effective in terms of training time and accuracy rate in this study. While the recommended training time for the GRU model is 8.56 minutes, the recommended training time for the BiLSTM model is 25.56 minutes. In this article, the effect of deep learning in terms of speed and performance in the automatic classification of fake news has been observed. It is seen that with the (Ngada and Haskins, 2020)'s DT decision support systems available in the literature, it has achieved a stable result in terms of precision, recall, F1 score, and accuracy values. On the other hand, in the proposed GRU model, the results are not as close to each other as in (Ngada and Haskins, 2020)'s DT study. The disadvantage of the proposed GRU model is that the difference between precision, recall, F1 score and accuracy parameters will be even less, which will increase the flexibility of the model. However, in general, promising results were obtained using the content presented in Table 5 and the word vectors that represent the news. Better results can be obtained by developing a hybrid embedding structure that will be obtained by selecting the best features of more than one embedding layer in terms of meaning and content, instead of Keras-based features representing each news content.

**Conflict of Interest Statement**

The article author declares that there is no conflict of interest.

**Contribution Rate Statement Summary of Researchers**

The author declares that he has contributed 100% to the article.

## References

Aggarwal CC. Neural networks and deep learning. Springer 2018; 10: 973–978.

Ahmed H., Traore I., Saad S. Detecting opinion spams and fake news using text classification. Security and Privacy 2018; 1(1): 15.

Ahmed H., Traore I., Saad S. Detection of online fake news using n-gram analysis and machine learning techniques. International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments 2017; 10618: 127-138.

Almuzaini HA., Azmi AM. Impact of stemming and word embedding on deep learning-based arabic text categorization. IEEE Access 2020; 8: 127913–127928.

Bakshy E., Messing S., Adamic LA. Exposure to ideologically diverse news and opinion on facebook. American Association for the Advancement of Science 2015; 348(6239): 1130–1132.

Bali APS., Fernandes M., Choubey S., Goel M. Comparative performance of machine learning algorithms for fake news detection. ICACDS 2019; 420–430.

Balmas M. When fake news becomes real: Combined exposure to multiple news sources and political attitudes of inefficacy, alienation, and cynicism. Communication Research Sage Publications 2014; 41(3): 430–454.

Barthel M., Mitchell A., Holcomb J. Many Americans believe fake news is sowing confusion. Pew Research Center 2016.

Çetiner H. Sahte haber verilerinin konvolüsyonel sinir ağı ile sınıflandırılması. 8th International Mardin Artuklu Scientific Researches Conference 2022.

Chen T., Xu R., He Y., Wang X. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. Expert Systems with Applications 2017; 72: 221–230.

Conroy NK., Rubin VL., Chen Y. Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology 2015; 52(1): 1–4.

Feng S., Banerjee R., Choi Y. Syntactic stylometry for deception detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics 2012; 171–175.

Guacho GB., Abdali S., Shah N., Papalexakis EE. Semi-supervised content-based detection of misinformation via tensor embeddings. 2018 IEEE/ACM International Conference on Advances In Social Networks Analysis and Mining (ASONAM) 2018; 322–325.

Gulli A., Pal S. Deep learning with Keras. Packt Publishing Ltd 2017.

Hochreiter S., Schmidhuber J. Long short-term memory. Neural Computation 1997; 9(8): 1735–1780.

Horne B., Adali S. This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. Proceedings of the International AAAI Conference on Web and Social Media 2017; 759–766.

Hosseinimotlagh S., Papalexakis EE. Unsupervised content-based identification of fake news articles with tensor decomposition ensembles. Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2) 2018.

Hu L., Wang C., Ye Z., Wang S. Estimating gaseous pollutants from bus emissions: A hybrid model based on GRU and XGBoost. Science of the Total Environment 2021; 783: 146870.

Jindal S., Sood R., Singh R., Vatsa M., Chakraborty T. Newsbag: A multimodal benchmark dataset for fake news detection. CEUR Workshop Proc. 2020; 138–145.

Karimi H., Roy P., Saba-Sadiya S., Tang J. Multi-source multi-class fake news detection. Proceedings of the 27th International Conference on Computational Linguistics 2018; 1546–1557.

Kaur S., Kumar P., Kumaraguru P. Automating fake news detection system using multi-level voting model. Soft Computing 2020; 24(12): 9049–9069.

Kingma D., Ba J. Adam: A method for stochastic optimization. International Conference on Learning Representations 2014.

Kishwar A., Zafar A. Fake news detection on Pakistani news using machine learning and deep learning. Expert Systems with Applications 2023; 211: 118558.

Li S. Application of recurrent neural networks in toxic comment classification. UCLA 2018.

Liu G., Guo J. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing 2019; 337: 325–338.

Nasir JA., Khan OS., Varlamis I. Fake news detection: A hybrid CNN-RNN based deep learning approach. International Journal of Information Management Data Insights 2021; 1(1): 100007.

Ngada O., Haskins B. Fake news detection using content-based features and machine learning. IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE) 2020; 1–6.

Niu X., Hou Y., Wang P. Bi-directional LSTM with quantum attention mechanism for sentence modeling. International Conference on Neural Information Processing 2017; 178–188.

Nowak J., Taspinar A., Scherer R. LSTM recurrent neural networks for short text and sentiment classification. International Conference on Artificial Intelligence and Soft Computing 2017; 553–562.

Ozbay FA., Alatas B. Fake news detection within online social media using supervised artificial intelligence algorithms. Physica A: Statistical Mechanics and its Applications 2020; 540: 123174.

Pang Z., Niu F., O'Neill Z. Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. Renewable Energy 2020; 156: 279–289.

Reis JCS., Correia A., Murai F., Veloso A., Benevenuto F. Supervised learning for fake news detection. IEEE Intelligent Systems 2019; 34(2): 76–81.

Sabeeh V., Zohdy M., Mollah A., Al Bashaireh R. Fake news detection on social media using deep learning and semantic knowledge sources. International Journal of Computer Science and Information Security (IJCSIS) 2020; 18(2).

Sahoo SR., Gupta BB. Multiple features based approach for automatic fake news detection on social networks using deep learning. Applied Soft Computing 2021; 100: 106983.

Sahoo SR., Gupta BB. Hybrid approach for detection of malicious profiles in twitter. Computers & Electrical Engineering 2019; 76: 65–81.

Sunstein C. On rumors how falsehoods spread, why we believe them, what can be done. Princeton University Press Publisher, 2009.

Tacchini E., Ballarin G., Della Vedova ML., Moret S., de Alfaro L. Some like it hoax: automated fake news detection in social networks. arXiv preprint arXiv:1704.07506 2017.

Wang J., Zhang Y., Yu LC., Zhang X. Contextual sentiment embeddings via bi-directional GRU language model. Knowledge-Based Systems 2022; 235: 107663.

Wang Y., Ma F., Jin Z., Yuan Y., Xun G., Jha K., Su L., Gao J. Eann: Event adversarial neural networks for multi-modal fake news detection. Proceedings of the 24th ACM Sigkdd International Conference on Knowledge Discovery & Data Mining 2018; 849–857.

Wu L., Liu H. Tracing fake-news footprints: characterizing social media messages by how they propagate. Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining 2018; 637–645.

Yang Y., Zheng L., Zhang J., Cui Q., Li Z., Yu PS. TI-CNN: Convolutional neural networks for fake news detection. arXiv preprint arXiv:1806.00749 2018.

Zhang Y., Zhang Z., Miao D., Wang J. Three-way enhanced convolutional neural networks for sentence-level sentiment classification. Information Sciences 2019; 477: 55–64.