Journal of Artificial Intelligence and Data Science (JAIDA)
e-ISSN: 2791-8335
https://dergipark.org.tr/pub/jaida

Vol. 4, No. 1, pp. 22-32, 2024
Research Article
Received: 17/04/2024 Accepted: 03/06/2024

Classification of Fake News Using Machine Learning and Deep Learning

Muhammed Baki ÇAKI [1], Muhammet Sinan BAŞARSLAN [2,*]

Abstract

The rapid advancement of technology has led to an increase in the spread of fake news, which has a detrimental effect on people in various fields, particularly in their daily lives. The negative impacts of fake news can be mitigated through the use of artificial intelligence. The development of AI technologies has made the detection of fake news a prominent area of research within natural language processing. This study explores style-based fake news detection using machine learning and deep learning methods. The texts were processed using natural language processing techniques and investigated with different models on the open-source ISOT dataset. The models utilised text processing, text representations (TF-IDF, word2Vec), and different machine learning (ML) methods (K-Nearest Neighbor, Naïve Bayes, Logistic Regression) as well as Long Short-Term Memory (LSTM). The performance of the models was evaluated using accuracy (Acc), precision (P), recall (R), and F1-score. Among the tested models, the LSTM model demonstrated the highest performance, with an accuracy of 99.2%. The development of state-of-the-art methods for text representation and classification, including preprocessing in text classification, and the application of these methods in practical settings can significantly reduce the prevalence of fake news.

Keywords: Deep learning, Fake news detection, Machine learning, Style based detection.

1. Introduction

Artificial Intelligence (AI) is a field that is divided into many sub-headings with its potential and is the subject of many researches in order to produce better solutions to our problems. Natural Language Processing (NLP), Machine Learning (ML) and Deep Learning (DL) are the main sub-topics of AI. Fake news is one of the problems we want to solve. The rapid spread of false content produced for various reasons causes social and economic damage to individuals, organizations and societies. This problem is growing with the increasing speed of communication. Misinformation and disinformation have negative effects on society. Therefore, new and effective methods are needed to detect and prevent fake news.

The main purpose of our study is to contribute to existing studies to find solutions to this problem with AI. In order to classify and distinguish between fake and real news, linguistic features of news texts are processed and analyzed with NLP. Then ML and DL models are built. After the models are trained, prediction is made for the given news text to be real or fake. In this study, various models are built using different NLP techniques and ML algorithms and the results obtained are analyzed. The results of the study show that NLP and ML models have a significant potential in fake news detection.

In the second part of the study, similar studies in the literature are presented. The third section discusses the dataset, preprocessing, vectorization, ML, DL and performance criteria. The fourth section describes the experimental setup. Section five presents the experimental results. The discussion and conclusion in sections six and seven provide an overall assessment and future works.

2. Related Works

Similar studies on fake news detection in the studied ISOT dataset will be described in this section. In their study, the researchers created models with Logistic Regression (LR), Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest (RF) and deep neural network. They achieved 91% Acc with neural [1] . After GloVe, the best performance with 92% Acc was obtained by using Linear Support Vector Machine (LSVM) as a classifier [2]. After vectorization with Term Frequency - Inverse Document Frequency (TF-IDF) on ISOT dataset, classifier

*Corresponding author

Muhammed Baki ÇAKI; Istanbul Medeniyet University, Faculty of Engineering and Architecture, Computer Engineering Department, Türkiye; e-mail: muhammedbakicaki@gmail.com; 0009-0005-2651-4047

Muhammet Sinan BAŞARSLAN; Istanbul Medeniyet University, Faculty of Engineering and Architecture, Computer Engineering Department, Türkiye; e-mail: muhammet.basarslan@medeniyet.edu.tr; 0000-0002-7996-9169

models were created for fake news detection with various ML algorithms. Among these models, the best result was obtained with Decision Tree (DT) with 96.8% Acc [3]. They created models with ML methods such as SVM, LSVM, K-Nearest Neighbor (KNN), DT on the data of fake news collected by ISOT and themselves. LSVM gave the highest Acc result with 92% [4]. After vectorization with the Word2Vec method called Maithi-Net, they obtained 97.28% Acc result in fake news detection with this method [5]. They obtained 74% Acc with NB after CBOW [6]. Word2Vec obtained 82.67% Acc with conditional random fields (CRF) classifier after CBOW [7]. They obtained 99% Acc in a study on fake news classification with ensemble learning after TF-IDF [8]. If we look at other fake news studies other than this dataset; They obtained 99.10% Acc in fake news classification model by hybridizing Recurrent Neural Network (RNN) and LSTM on Liar dataset [9]. In their study on the detection of false news in the pandemic, they obtained 96.19% Acc and 95% F1 with the Convolutional Neural Network (CNN) they proposed by optimizing hyperparameters after embedding methods such as GloVe [10].

It is seen that ML methods are frequently used after TF-IDF, Word2Vec text representation methods on ISOT dataset [2] and similar content data related to fake news. After the text representation methods TF-IDF and Word2Vec, which are frequently studied in the literature, ML (KNN, NB, LR), and DL (LSTM) models were created for the classification of fake news.

The contribution of this study to the detection of fake news on ISOT, an open source shared and balanced dataset, is listed below:

- It is investigated which of the popularly preferred text representation methods such as TF-IDF and Word2Vec has more impact on the performance of the models.
- The holdout discrimination results of the models built with classical ML (KNN, SVM, NB, LR) and DL (LSTM) are investigated.

3. Materials and Methods

In this section, dataset, preprocessing, vectorization, ML, DL, and performance criteria will be explained. In this study, a text analysis-based approach is adopted for automatic fake news detection. In this approach, the characteristics of the texts are extracted and processed by NLP methods, then modelled and predicted by ML methods. The field of study is in the combination of ML, DL, and NLP fields in Figure 1.

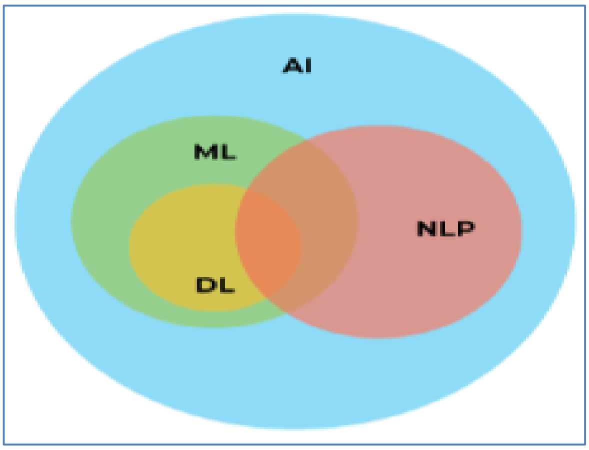


Figure 1. *Fields of AI*

The experimental steps carried out in the study are given in Figure 2.

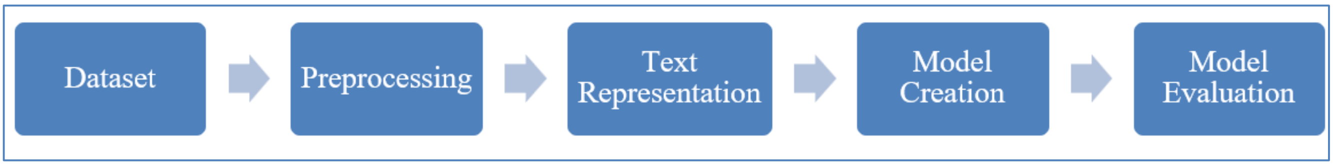


Figure 2. *Study Pipeline*

As seen in Figure 2, the dataset was preprocessed and then subjected to text representation (TF-IDF, Word2Vec). Then the model was created with LSTM, KNN, NB, LR with 75%-25% training-test separation. F1, P, R, Acc were used to evaluate the models.

3.1. Dataset

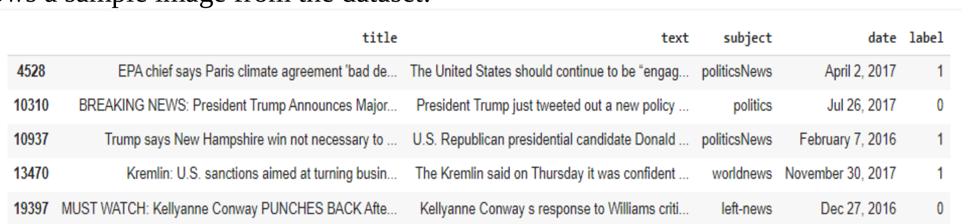
In the study, ISOT Fake News Dataset [2], which consists of fake and real news data created by the researchers with news collected from the internet between 2016-2017, was used.

The distribution of the dataset, which consists of 44,898 news in total, is given in Table 1. According to the researchers, news with true content was collected from the reuters.com website, while news with false content was collected from various websites marked as unsafe by Polifact [2]. Table 1 provides information about the content of the dataset.

Table 1. *Dataset*

News	Number of articles	Subjects	
		Type	**Articles size**
Real	21417	Government-News	1570
		Politics-News	11272
		Type	**Articles size**
Fake	23481	US News	783
		Left-news	4459
		Politics	6841
		News	9050

Figure 3 shows a sample image from the dataset.



	title	text	subject	date	label
4528	EPA chief says Paris climate agreement 'bad de...	The United States should continue to be "engag...	politicsNews	April 2, 2017	1
10310	BREAKING NEWS: President Trump Announces Major...	President Trump just tweeted out a new policy ...	politics	Jul 26, 2017	0
10937	Trump says New Hampshire win not necessary to ...	U.S. Republican presidential candidate Donald ...	politicsNews	February 7, 2016	1
13470	Kremlin: U.S. sanctions aimed at turning busin...	The Kremlin said on Thursday it was confident ...	worldnews	November 30, 2017	1
19397	MUST WATCH: Kellyanne Conway PUNCHES BACK Afte...	Kellyanne Conway s response to Williams criti...	left-news	Dec 27, 2016	0

Figure 3. *Summary Visualization of the Dataset*

Figure 3 shows the title of the news item, the text of the news item, the date of publication of the news item and the class label of whether it is fake or real. Figure 4 shows the graph of class distribution.

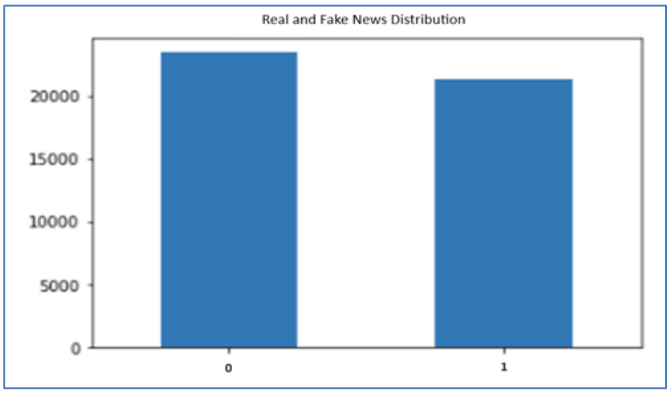


Figure 4. *News Class Distribution*

In Figure 4, according to the class distribution of the dataset; we see that the skewness coefficient is calculated as -0.08 and the kurtosis coefficient is calculated as -1.18. Since the skewness is very close to zero, we can accept the distribution as symmetric. In the light of this result, the dataset is balanced.

3.2. Text Representation

In this section you will find information about text representation.

3.2.1. Term Frequency Inverse Document Frequency

The method called Term Frequency - Inverse Document Frequency is based on the principle of extracting the attributes of the text by weighting each word in the text according to its importance. In this method, the importance of words is determined by analyzing how many times they occur in the examined text and how many times they occur in other texts. The TF-IDF representing the term in sentence t, document d is given in equation (1) [8].

$$TF(t,d) = \frac{Number\ of\ times\ term\ t\ appears\ in\ document\ d}{Total\ number\ of\ terms\ in\ document\ d} \tag{1}$$

D is the collection of all documents (corpus), the addition of 1 to the denominator is to prevent the term from dividing by zero if it is not found in any document IDF is given in equation (2) [8].

$$IDF(t,d) = \frac{Total\ number\ of\ documents\ in\ the\ corpus\ N}{Number\ of\ documents\ containing\ term\ t{+}1} \tag{2}$$

The TF-IDF in the document is given in Equation (3) [9].

$$TF - IDF(t,d,D) = TF(t,d) \times IDF(t,D) \tag{3}$$

According to this equation, frequent occurrence of a word in the relevant document increases its importance. If it is a common word in other documents, it decreases its importance. In this way, stopwords in documents also become unimportant.

3.2.2. Word2Vec

Word2Vec is a method of converting words into vectors of real numbers using artificial neural networks. Words with close meaning are also numerically close in vector representation. In this way, the semantic proximity and context information of the words are kept [11].

3.3. Machine Learning

In this section, ML methods are described.

3.3.1. K-Nearest Neighbors

KNN algorithm is a lazy learning algorithm used in classification and regression problems in ML. In the space where the data points are represented, prediction is made based on the distance of the relevant point to other points. For the classification task, the distances to the k nearest points are calculated. It is predicted as belonging to the class with the least total distance. The reason why it is categorized as a lazy learning algorithm is that there is no learning phase before the data to be predicted arrives. It takes two basic parameters, 'number of neighbours' and 'distance metric' [12].

The number of neighbours is the value 'k', which is also in the name of the algorithm. Distance is calculated with the k nearest neighbours. The distance metric is the algorithm to be used to measure the distance. The most commonly used equation for distance calculation is given in (Euclidean distance) equation (4).

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{4}$$

3.3.2. Naïve Bayes

NB Classifier is a probability-based prediction method that uses a simplified version of Bayes' Theorem in probability. It is often used in classification tasks. Bayes' Theorem allows the calculation of the probability of event A occurring when event B occurs; when the probabilities of event A occurring, event B occurring, event B occurring when event A occurs are known. Its formula is given in equation (5) [13]:

$$P(B) = \frac{P(A) * P(B|A)}{P(B)} \tag{5}$$

In the NB Classifier, the denominator part of the Bayes equation is ignored since the aim is to find the class with high probability instead of finding the exact value. For a two-class classification task, the probabilities of the data belonging to classes X and Y are calculated with the help of the equation. Whichever class the probability of belonging to is calculated to be higher, is predicted to belong to that class.

3.3.3. Logistic Regression

LR is an algorithm frequently used in classification problems in ML. It is based on probability-based class prediction by fitting the data to the logistic function. It is more suitable for binary classification task. Multi-class classification can also be performed. Figure 5 shows a visual of LR [14].

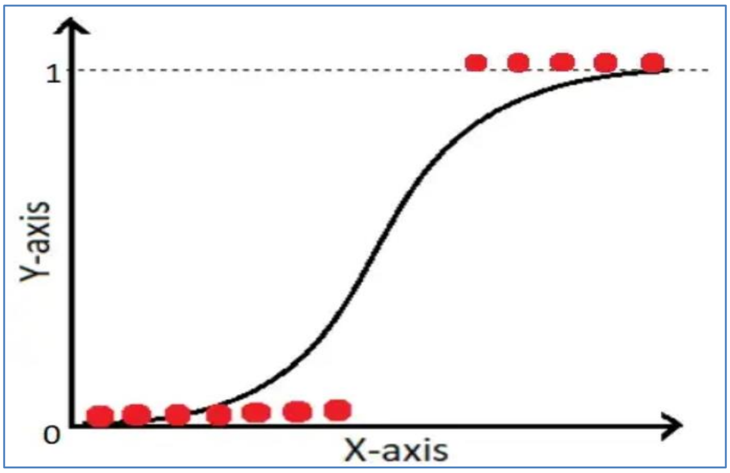


Figure 5. *LR [14].*

3.4. Deep Learning

In this section, LSTM, one of the DL methods used in this study, will be explained.

3.4.1. Long Short-Term Memory

Long short-term memory is a DL architecture that is an advanced version of the RNN model. In RNN, as each output affects the next input, a memory structure is formed. This memory is short term. In long inputs, the effect of past data on the new input decreases rapidly and disappears (gradient vanishing problem). In LSTM architecture, input, output and forget gates are used in addition to RNN. In this way, by creating a short-term and long-term memory structure at the same time, context information can be preserved in long inputs such as paragraphs. LSTM architecture is frequently used in the development of sequence-based prediction systems such as anomaly detection and time series [15]. Figure 6 shows the LSTM architecture.

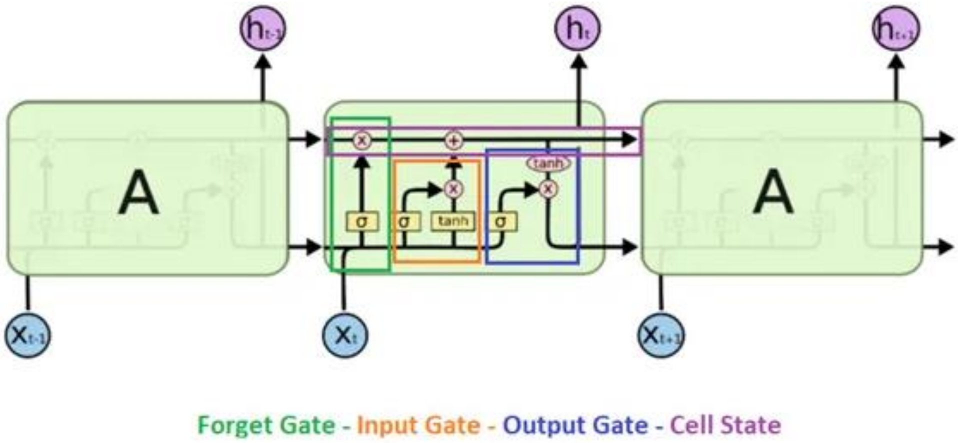


Figure 6. *LSTM Architecture* [15].

3.5. Performance Criteria

In this section, the metrics used in the evaluation of the performance criteria used for model evaluation and the confusion matrix that is used to obtain this metrics are explained.

3.5.1. Confusion Matrix

The confusion matrix is constructed by comparing the predicted and actual values of the test data. In each cell, the total number of samples belonging to that cell is recorded. In this way, the test result of the model can be analysed on a single table [16].

- True Positive (TP) if the true value of the data is positive and predicted as positive,
- False Negative (FN) when the true value is positive and estimated as negative,
- False Positive (FP) when the true value is negative and estimated as positive,
- A negative true value and a negative predicted value constitute True Negative (TN) cases.

Figure 7 shows a visualization of the confusion matrix.

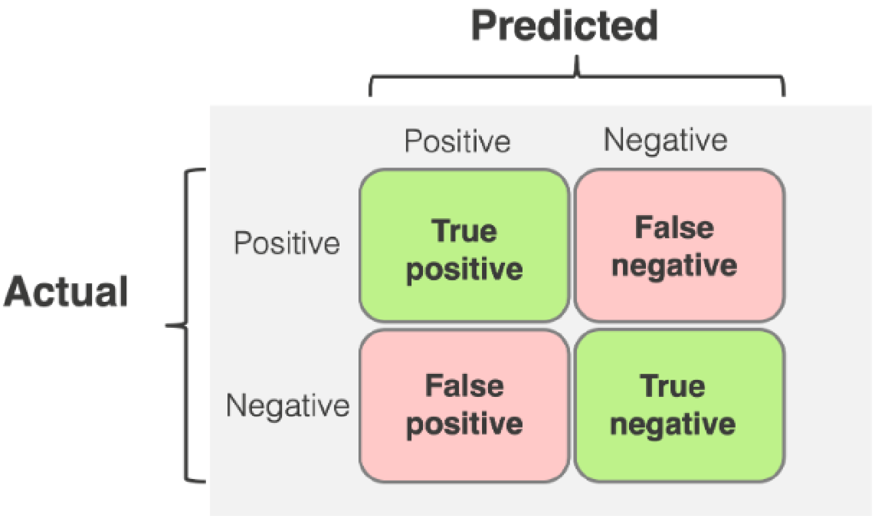


Figure 7. *Confusion Matrix* [17]

Various metrics are calculated to measure how well the created models make predictions. The predictions made by the model are processed to the relevant part in the confusion matrix. Metrics are calculated using the relevant fields in the confusion matrix. The four most important metrics used for evaluating classification tasks are explained

Accuracy:

It is the ratio of the model's correct predictions to all predictions [18-19]. This ratio is given in Equation (6).

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision:

It is the metric that measures the success ratio of the model when it predicts the outcome as positive [19]. This metric is given in Equation (7).

$$P = \frac{TP}{TP+FP} \quad (7)$$

Recall:

It is the metric that measures the extent to which the model can accurately detect situations that are actually positive [20]. This metric is given in Equation (8).

$$R = \frac{TP}{TP+FN} \quad (8)$$

F1:

In cases such as measuring high Acc and low P values in an unbalanced dataset, the Acc value may be misleading about the success of the model. F1 metric is obtained by the harmonic mean of P and R values and indicates the stability of the prediction [21]. This metric is given in Equation (9) [18].

$$F1 = 2 * \frac{P*R}{P+R} \quad (9)$$

4. Experimental setup

The work was done in Google Colab [22], which allows the Python programming language [23] to run on a Jupyter notebook. In this chapter, preprocessing of texts, TF-IDF and Word2Vec followed by modelling with ML and DL will be explained.

4.1. Preprocessing

The operations performed within the scope of the study will be explained in this section. Attributes other than text and label have been removed. Label attribute is a binary data type containing 0 for false news and 1 for true news. The preprocessing processes are listed below:

1. The news source at the beginning of real news has been removed.
2. Letters and characters other than the '@' sign were removed from the news texts.
3. News texts were converted to lower case.
4. Stopwords in the news texts were removed.
5. The words in the news texts were stemmed.

4.2. Text Representation

Since the models to be created cannot process text data, the text must be matrixised and given as input to the model. For this reason; TF-IDF with its implementation in sci-kit learn library, Word2Vec text representation methods were used with its implementation in the gensim library. The Word2Vec parameters used in the study are presented in Table 2.

Table 2. *TF-IDF and Word2Vec parameters*

	Parameters	Value
Word2Vec	Vector size	200
	Window	5
	Min count	5
	Methods (sg)	CBOW (0)
	Workers	4

4.3. Model Creation

In the study, a total of seven models were created with TF-IDF and Word2Vec text representation method, KNN, NB, LR, and LSTM. The data to be used to train and test these models are divided into 75% training and 25% test data. In the study, sklearn library was used to create models with KNN, LR, and NB, which are traditional ML methods, and keras library was used for LSTM, which is a DL model.

4.3.1. Machine learning model

In the study, for the creation of ML models, Multinomial NB Classifier and LR Classifier implementations in the sci-kit learn library were used to create models with default parameters. For the KNN model, the implementation of the same library was used and the grid search algorithm was used for hyperparameter optimization.

The TF-IDF vectorization method was tested with the grid search algorithm for the number of neighbours (k) parameter of the generated KNN model for values (1-10) and the optimal value was observed to be 1. It was observed that the optimum k value for the model created with Word2Vec vectorization method was 5. The distance metric parameter was chosen as "euclidean".

4.3.2. Deep learning model

In the study, an LSTM model was created using the LSTM module in the Keras library. The model was optimized by creating and comparing models with different hyperparameters. Used and preferred hyperparameters are shown in table 3.

Table 3. *TF-IDF and Word2Vec parameters*

Output Dimension	Neuron	Dropout	Epochs	Batch size	Loss Function	Opitimizer	Activation Function
100, 200	10-**100**	**0.2**, 0,3	5, 8, **10**	32, **64**	binary crossentropy	adam	sigmoid

**Optimum values in bold*

Confusion matrix and score metrics obtained from confusion matrix were used to measure the model performance. The loss functions and Acc values of the training and validation data were monitored to observe that overfitting/underfitting situations do not occur in the training phase.

4. Experimental Results

Dataset is splitted as 75% and 25% for training and testing respectively. The results of models created after text representation with TF-IDF and Word2Vec are given in Table 4 and Table 5.

Table 4. *Results with TF-IDF*

	KNN	NB	LR
Acc	79.3	92.3	**97.4**
P	72.8	91.4	**97.5**
R	95.2	93.8	**97.5**
F1	82.5	92.6	**97.5**

As seen in Table 4, the model created with LR is ahead of the other models in Acc, P, R, and F1 in 75%-25% hold-out separation after TF-IDF.

Table 5. *Results with Word2Vec*

	KNN	NB	LR	LSTM
Acc	94.3	89.2	97.1	**99.2**
P	96.0	88.2	97.5	**98.8**
R	92.8	91.3	96.8	**99.6**
F1	94.4	89.7	97.2	**99.2**

As seen in Table 5, the model created with LSTM is ahead of the other models in A, P, R, and F in 75%-25% hold-out discrimination after Word2Vec. When Table 4 and Table 5 are evaluated together, the results obtained with Word2Vec in KNN are ahead. However, in LR and NB, the difference between TF-IDF and Word2Vec is not much compared to KNN.

LSTM was the model that gave the best results among all models. The graphs of train-validation Acc and loss values of the LSTM neural network are given in Figure 8.

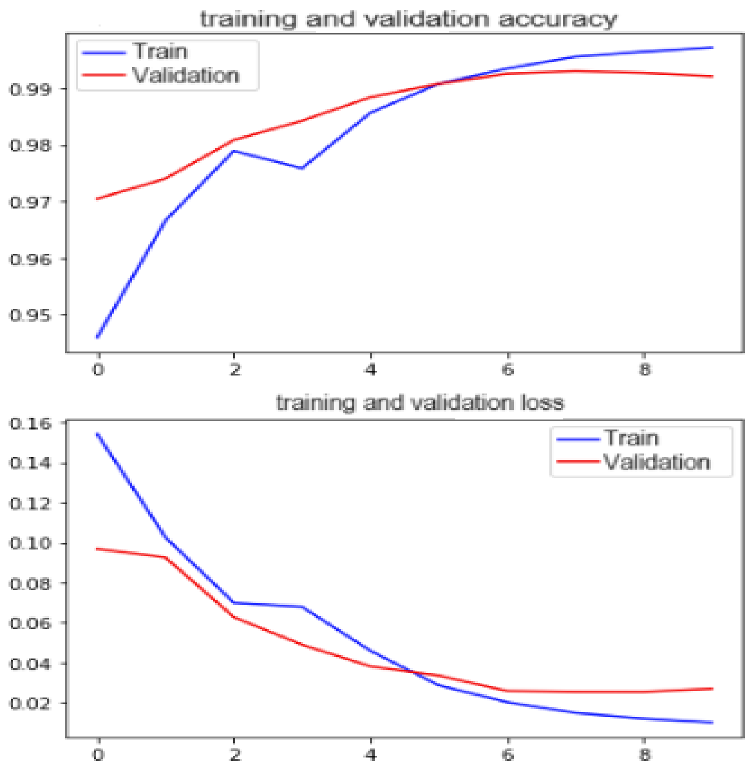


Figure 8. *Confusion Matrix*

As predicted, the success rate of the LSTM model exceeded the classical ML algorithms. However, it was observed that the training time of the model was considerably high compared to others.

5. Conclusion and Discussion

In the study where various traditional ML, DL models and text representation methods were compared for the Fake News Detection task, it was observed that the best result was obtained with the DL Model, the performance of the KNN model was more affected by the vectorization method, and NB and LR models obtained close and good results in both vectorization methods. In particular, the LR Model was found to be the best model in terms of efficiency for this binary text classification study. Similar studies on the same dataset are given in Table 6. Acc results were used because the dataset is balanced.

Table 6. *Previous studies on the ISOT dataset*

Authors	Text Representation	Model	Acc (%)
[2]	GloVe	LSVM	92.0
[3]	TF-IDF	DT	96.8
[4]	BOW	LSVM	92.0
[5]	Word2Vec	CNN	97.28
[6]	BOW	NB	74.0
[7]	Word2Vec	CRF	82.6
[8]	TF-IDF	Ensemble learning (DT)	99.0
This study	Wod2Vec	LSTM	**99.2**

In the models created for the classification of fake news with ML (KNN, NB, LR) and DL (LSTM) after TF-IDF and Word2Vec on the fake news dataset (ISOT), the best result was obtained with Word2Vec LSTM with 99.2% ACC. The model created in the study is a model that competes with the literature as seen in Table 6. The models created with Word2Vec are more successful than the models created with TF-IDF in most cases. This situation will be investigated in future studies by working on different datasets.

In the future, the process can be repeated with different data sets and the hyperparameters of the LSTM model can be further optimized. In addition, newer and successful state-of-the-art models such as BERT, RoBERTa can be used. As these models have been pre-trained with very large datasets, they have brought great advances in the fields of NLP and ML. For this reason, it is predicted that the success of the study will increase.

Declaration of Interest

The authors declare that there is no conflict of interest.

Author Contributions

Muhammed Baki ÇAKI; data analysis, experiments and evaluations, manuscript draft preparation. Muhammet Sinan BAŞARSLAN; defining the methodology, evaluations of the results, and original draft, supervision.

References

[1] C. K. Hiramath and G. C. Deshpande, "Fake News Detection Using Deep Learning Techniques," in 2019 1st International Conference on Advances in Information Technology (ICAIT), IEEE, Jul. 2019, pp. 411–415. doi: 10.1109/ICAIT47043.2019.8987258.

[2] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," 2017, pp. 127–138. doi: 10.1007/978-3-319-69155-8_9.

[3] F. A. Ozbay and B. Alatas, "Fake news detection within online social media using supervised artificial intelligence algorithms," Physica A: Statistical Mechanics and its Applications, vol. 540, p. 123174, Feb. 2020, doi: 10.1016/j.physa.2019.123174.

[4] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning--based Text Classification," ACM Comput Surv, vol. 54, no. 3, pp. 1–40, Apr. 2022, doi: 10.1145/3439726.

[5] D. Muduli, S. K. Sharma, D. Kumar, A. Singh, and S. K. Srivastav, "Maithi-Net: A Customized Convolution Approach for Fake News Detection using Maithili Language," in 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3), IEEE, Jun. 2023, pp. 1–6. doi: 10.1109/IC2E357697.2023.10262664.

[6] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), IEEE, May 2017, pp. 900–903. doi: 10.1109/UKRCON.2017.8100379.

[7] A. Priyadarshi and S. K. Saha, “Towards the first Maithili part of speech tagger: Resource creation and system development,” Comput Speech Lang, vol. 62, p. 101054, Jul. 2020, doi: 10.1016/j.csl.2019.101054.

[8] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, “Fake News Detection Using Machine Learning Ensemble Methods,” Complexity, vol. 2020, pp. 1–11, Oct. 2020, doi: 10.1155/2020/8885861.

[9] A. K. Shalini, S. Saxena, and B. S. Kumar, “Automatic detection of fake news using recurrent neural network—Long short-term memory,” Journal of Autonomous Intelligence, vol. 7, no. 3, Dec. 2023, doi: 10.32629/jai.v7i3.798.

[10] M. Akhter et al., “COVID-19 Fake News Detection using Deep Learning Model,” Annals of Data Science, Jan. 2024, doi: 10.1007/s40745-023-00507-y.

[11] J. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, “Distributed representations of words and phrases and their compositionality,” in Advances in Neural Information Processing Systems, 2013. [Online]. Available: https://proceedings.neurips.cc/paper

[12] R. Ahmed, M. Bibi, and S. Syed, “Improving Heart Disease Prediction Accuracy Using a Hybrid Machine Learning Approach: A Comparative study of SVM and KNN Algorithms,” International Journal of Computations, Information and Manufacturing (IJCIM), vol. 3, no. 1, pp. 49–54, Jun. 2023, doi: 10.54489/ijcim.v3i1.223.

[13] T. Öztürk, Z. Turgut, G. Akgün, and C. Köse, “Machine learning-based intrusion detection for SCADA systems in healthcare,” Network Modeling Analysis in Health Informatics and Bioinformatics, vol. 11, no. 1, p. 47, Dec. 2022, doi: 10.1007/s13721-022-00390-2.

[14] H. Canlı and S. Toklu, “Design and Implementation of a Prediction Approach Using Big Data and Deep Learning Techniques for Parking Occupancy,” Arab J Sci Eng, vol. 47, no. 2, pp. 1955–1970, Feb. 2022, doi: 10.1007/s13369-021-06125-1.

[15] R. Vankdothu, M. A. Hameed, and H. Fatima, “A Brain Tumor Identification and Classification Using Deep Learning based on CNN-LSTM Method,” Computers and Electrical Engineering, vol. 101, p. 107960, Jul. 2022, doi: 10.1016/j.compeleceng.2022.107960.

[16] H. Canli and S. Toklu, “Deep Learning-Based Mobile Application Design for Smart Parking,” IEEE Access, vol. 9, pp. 61171–61183, 2021, doi: 10.1109/ACCESS.2021.3074887.

[17] M. Z. Khaliki and M. S. Başarslan, “Brain tumor detection from images and comparison with transfer learning methods and 3-layer CNN,” Sci Rep, vol. 14, no. 1, p. 2664, Feb. 2024, doi: 10.1038/s41598-024-52823-9.

[18] S. N. Başa and M. S. Basarslan, “Sentiment Analysis Using Machine Learning Techniques on IMDB Dataset,” in 2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), IEEE, Oct. 2023, pp. 1–5. doi: 10.1109/ISMSIT58785.2023.10304923.

[19] F. Kayaalp, M. S. Basarslan, and K. Polat, “TSCBAS: A Novel Correlation Based Attribute Selection Method and Application on Telecommunications Churn Analysis,” in 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), IEEE, Sep. 2018, pp. 1–5. doi: 10.1109/IDAP.2018.8620935.

[20] Öztürk, T., Turgut, Z., Akgün, G. et al. Machine learning-based intrusion detection for SCADA systems in healthcare. Netw Model Anal Health Inform Bioinforma 11, 47 (2022). https://doi.org/10.1007/s13721-022-00390-2

[21] Ardaç, H.A., Erdoğmuş, P. Question answering system with text mining and deep networks. Evolving Systems (2024). https://doi.org/10.1007/s12530-024-09592-7

[22] Google LLC, “Colab.” https://colab.research.google.com/. Accessed 1 Feb 2023

[23] Python, “Python.” https://www.python.org/downloads/. Accessed 1 Feb 2023