Detection of Turkish Fake News in Twitter with Machine Learning Algorithms

Suleyman Gokhan Taskin^{1,2} · Ecir Ugur Kucuksille² · Kamil Topal³

Received: 30 March 2021 / Accepted: 12 September 2021 / Published online: 1 October 2021
© King Fahd University of Petroleum & Minerals 2021

Abstract

Social media has affected people's information sources. Since most of the news on social media is not verified by a central authority, it may contain fake news for various reasons such as advertising and propaganda. Considering an average of 500 million tweets were posted daily on Twitter alone in the year of 2020, it is possible to control each share only with smart systems. In this study, we use Natural Language Processing methods to detect fake news for Turkish-language posts on certain topics on Twitter. Furthermore, we examine the follow/follower relations of the users who shared fake-real news on the same subjects through social network analysis methods and visualization tools. Various supervised and unsupervised learning algorithms have been tested with different parameters. The most successful F1 score of fake news detection was obtained with the support vector machines algorithm with 0.9. People who share fake/true news can help in the separation of subgroups in the social network created by people and their followers. The results show that fake news propagation networks may show different characteristics in their own subject based on the follow/follower network.

Keywords Fake news detection · Machine learning · Natural language processing · Social network analysis

1 Introduction

In traditional news sharing, the number of news sources is restricted because communication is one-way. Therefore, since it is easier to control these sources and to check the content of the news in advance, exposure of the society to fake news can be prevented. Upon transition to the digital era, people's sources of access to news have begun to vary a lot. The communication has ceased to be one-way, and the responses of news receivers to the news have also been digitized. These reactions can be in different styles such as sharing, liking / disliking, and commenting on. Because people become the source of news by sharing it, it is not possible for supervisory institutions to check the correctness of the news with real persons (supervisors). Considering the rapid spread of

fake news sharing, from the first 2 hours to the 20th hour, instant detection of fake news by automatic systems are of great importance [1].

Even though the concept of fake news is a very old problem, its popularity has increased with the Brexit vote made in 2016 so as to vote on the UK's decision to quit the European Union, and with the US presidential elections. Its advance has increased. The biggest reason for fake news to increase its influence in recent years is the increase in the use of social media platforms where all users can be in the role of sharing the news. The World Digital, published in 2021, reported that 4.66 million of the 7.83 billion world population use the internet and 4.2 billion of them use the social media. It also stated that people living in Turkey use the social media for approximately 3 hours a day. It is also stated in the same report that 55% of the users use the Internet to get up-to-date information about news and events [2]. The heavily used social media platforms have gradually turned into news sources. It was also stated in the same report that 55% of users use the Internet to obtain up-to-date information regarding the news and events [2]. The social media platforms, which are used extensively, have gradually turned into news sources.

Every user can be a source of news by sharing news on the social media. Users follow other users with whom they

✉ Suleyman Gokhan Taskin
taskinster@gmail.com

¹ Department of Software Engineering, Bandirma Onyedi Eylul University, 10200 Bandirma, Balikesir, Turkey

² Department of Computer Engineering, Suleyman Demirel University, 32260 Isparta, Turkey

³ Department of Computer Engineering, Balikesir University, 10145 Balikesir, Turkey

agree. Since users trust the shares coming from the accounts they follow, they tend to share these shares without verifying their correctness. Furthermore, because they follow similar groups or people, and since they frequently encounter with the shares coming from these people, they perceive these news as if they were true news over time, even if they are fake news, and they share this fake news with people who are in their own social networks. In addition, the high tendency of users to share news that will prove the accuracy of their beliefs and opinions gained from their previous experiences increases the sharing of fake news and the news, drawing attention, can change the ideas of people having different opinions.

The fact that the social media increases the rate of sharing fake news in a remarkable ratio has increased the studies for struggling against the fake news. In order to verify the news shared by the users on the social networks, verification platforms via which experts detect the fake news on Turkish news have been created. As examples for these platforms, the platforms such as Teyit.org [3], Dogrulukpayi.com [4] can be counted. In this study, teyit.org platform, which started its broadcasting life in 2016, was used so as to confirm the fake news. Teyit.org is a member of the International Fact-Checking Network (IFCN), which brings together verification platforms around the world, and was established so as to increase the auditing and accountability of these platforms. Teyit.org not only provides the detection of the fake news, but also shares several publications that make people aware so that they can detect the fake news.

On such platforms where the fake news is detected by experts, the speed of informing users about the fake news is slower when it is compared to automatic fake news detection systems. Considering the large amount of sharing of the fake news in a 2-hour time period, the detection of the fake news to be made by automatic detection systems is of a great importance, instead of these platforms where experts detect the fake news. There are a lot of studies in English regarding automatic detection of the fake news. The studies for struggling with the fake news in Turkish language is insufficient when compared to the those for struggling with the fake news in English language, which increases the exposure of people living in Turkey to the fake news. In the reports Newman et al. released, it was stated that the rate of people who are exposed to the fake news among countries is the highest in Turkey, with a rate of 49% [5]. In the same report, 87% of the people in Turkey stated that they follow the news from social media and “online” sources. Again, according to their research in 38 countries in Newman 2019’s reports, on average 55% of people living in these countries are concerned about their ability to distinguish between real and fake news on the internet. This rate is 63% for those people who live in Turkey [6]. This situation causes this news to be spread very

quickly in Turkey due to the fact that the fake news is shared by people who cannot distinguish the fake news.

In most social media platforms, detection of fake news is carried out by experts. In social media platforms with high share traffic, thus, it is not possible to detect fake news in a short time with this working logic in the social media in which heavy sharing traffic exists. Therefore, this causes fake news to be shared by many people in a short time. For the reason, semi-automatic and automatic fake news detection systems can detect fake news in a shorter time than the systems where experts work. In order to detect fake news in a short time, it is essential that automatic detection systems be developed. Studies regarding detection of fake news have been carried out with machine learning methods in various languages. No automatic fake news detection study was encountered in Turkish language until the year of 2020.

In this study, in the tweet messages in Turkish language shared on Twitter platform [7], fake news detection was made with natural language processing methods and, with social network analysis methods, and the interaction between those spreading fake-real news and their followers were studied. In August, 2019, tweet messages about 3 subjects, the users who shared these tweets and the followers of these users were collected. In order to tag the collected tweets, a user web interface was created and the tweets were tagged by us. The words in the tagged tweets were transformed into vectors by using word representation methods and given as an introduction to supervised and unsupervised machine learning algorithms. Moreover, users who spread fake and real news and their followers were analyzed by using social network analysis methods.

The following difficulties were encountered in the study;

- Collecting data in the Twitter API free version is quite slow. For this reason, only the followers of users who shared fake and real news were collected. Level 2 user interaction was not used in this study, as it would take a long time to collect the followers of these collected followers.
- Turkish is an agglutinative language, it has a very different structure from English. Suffixes of Turkish words should also be analyzed. There are very few open source tools on the analysis of Turkish word suffixes.
- There are few pre-trained word embeddings in Turkish. There are no word embeddings trained with short texts like social media texts. So, word embeddings with a limited number of words were used.
- Sentence analysis tools in Turkish are not yet effective enough.

With this study, the following contributions to the literature were made;

- It has been shown that it can be determined by machine learning methods whether the Turkish news spread on social media is fake or real.
- It is one of the first studies in the detection of fake news in Turkish.
- It has been determined that the users who spread fake news in Turkish are in a certain group by examining the follower/follower analysis with the social network analysis method.

In the following section, we analysed academic studies on fake news detection and social media analysis. In the next section, we have explained the methods that we used in our study. In sect. 4, we analyzed the outputs of our fake news detection study and the social network analysis methods. And in the final section, we discussed the results of our fake news detection study and provided information about future studies.

2 Literature Review

When the literature is reviewed, studies that detect fake news in social media and make social media network analysis about fake news using NLP and ML algorithms are taken into consideration. Using these two criteria, the following studies were found.

The studies regarding the detection of fake news started with the studies of Zhao and Jiang in 2011 and these studies increased with the conspiracy theories in the US Presidential Elections in 2016 [8]. In the automatic fake news detection studies, news articles in English language, to a large extent, and supervised machine learning algorithms have been used.

In the studies regarding the fake news detection made with non-ANN-based (non-Artificial Neural Network based) supervised learning algorithms, algorithms of Naive Bayes (NB), Linear Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Stochastic Gradient Descent (SGD) and HBLC (Harmonic Boolean Label Crowdsourcing) were used. The studies made using these algorithms, the data sets they used and their study results are shown in Table 1. In the fake news detection problem, it can be seen that the SVM algorithm from the non-ANN-based supervised learning network gives more successful results than other algorithms do.

In the non-ANN-based supervised learning algorithms, as well as the fake news detection studies made in English language, there are also studies conducted in other languages such as Turkish, Indonesian [9] and Chinese [10]. Only studies made, by Mertoglu et al, in long texts were found in the Turkish language [11].

In the fake news detection studies, conducted with artificial neural network-based supervised learning algorithms,

there are some studies such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), Gated Recurrent Unit (GRU), Bidirectional Gated Recurrent Unit (BiGRU) and Convolutional Neural Network (CNN) algorithms and those studies such as three level hierarchical attendance network (3HAN) and Self Multi-Head Attention-based CNN (SMHA-CNN) that use their own methods. The studies made using these algorithms, the data sets they use and the results of the study are shown in Table 2. Artificial neural network-based algorithms yielded more successful results than non-ANN-based supervised learning algorithms. In the fake news detection program, the GRU, LSTM, BiGRU and BiLSTM algorithms, which have memory units, among the artificial neural network-based algorithms, were more successful than the CNN and RNN algorithms.

In addition to fake news detection studies with supervised learning algorithms, there are also studies conducted with unsupervised learning algorithms. These studies revealed results with different data sets based on Bayesian networks [12–14].

Another aspect of this study is that to understand the structure of fake news by examining the relationship of social media users among themselves. Social network analysis is a field that analyzes relationships between individuals and is used in various fields [15–17]. In this study, PageRank and HITS social network analysis methods are used. And various studies such as PageRank algorithm [18], website search algorithm [19–21], search in social networks [22], finding influential people in social networks [23–26] are also used. Similarly, the HITS (Hyperlink- Induced Topic Search) algorithm [27], the website search algorithm [28] were used in various studies for search social networks [29,30] and to find influential people on social networks [31,32].

Considering the fake news detection studies conducted through social network analysis, various studies were encountered. In one of these studies, Mocanu et al. examined the behavior of 2.3 million Facebook users in Italian choices and showed that users interact more with fake news that was deliberately shared [33]. They have demonstrated what they have entered. In another study, Kwon et al. used the linguistic and temporal properties of the tweet messages they gathered on Twitter and the graph properties of the users and suggested a model for the problem of fake news detection [34]. They classified the model, they proposed, into binary categories as rumor and non-rumor by using Decision Tree (DT), Random Forest (RF) and SVM supervised learning algorithms. The most successful result was obtained by the RF algorithm and 0.89 F1-score value. Nguyen et al. proposed a method, in their studies, so as to detect the network spreading the fake news with social network analysis by using greedy algorithms [35]. They tried the method, they proposed, in 3 different real net-

Table 1 Fake news detection studies made by using non-ANN-based supervised learning algorithms.

Refs.	Dataset	ML	Success ML	Performance measure	Best result
[10]	Mobile phone reviews from mobile01.com	LR, SVM	SVM	F1-score	0.61
[74]	Hotel reviews, restaurant review, gay marriage, and gun control, fake and real news articles from kaggle.com	SGD, SVM, KNN, LR, DT	SVM	Accuracy	0.9
[75]	3 large Facebook pages each from the right and from the left and Facebook pages of Politico, CNN and ABC News	NB	NB	Accuracy	0.75
[76]	The authors present a list of Facebook pages divided into two categories: scientific news sources and conspiracy news sources.	LR, HBLC	HBLC	Accuracy	0.99
[11]	Articles on sport, politics, rumor, health and other were collected with web crawler.	SVM, NB	SVM	F1-score	0.79
[77]	2 satirical news sites (The Onion and The Beaverton) and 2 legitimate news sources (The Toronto Star and The New York Times): varying across 4 domains (civics, science, business, and “soft” news)	SVM	SVM	F1-score	0.87
[78]	Collecting legitimate news from mainstream news websites such as CNN, FoxNews, Bloomberg, and CNET and collecting fake news using crowdsourcing.	SVM	SVM	F1-score	0.73
[9]	News articles from Google with web crawler	NB	NB	Accuracy	0.79

Table 2 Fake news detection studies made by using ANN-based supervised learning algorithms.

Refs.	Dataset	ML	Success ML	Performance measure	Best result
[79]	Kaggle open source dataset of fake news article and signalmedia open source dataset of not fake news article.	LR, RNN, GRU, LSTM, BiLSTM, CNN	GRU	F1-score	0.84
[80]	Using a set of articles flagged as false by Snopes, and a set of real articles from news organizations such as NDTV, CNN etc..	KNN, SVM, LSTM	LSTM	F1-score	0.90
[81]	FNC-1 open source dataset of articles	LSTM, GRU	GRU	FNC-score	69.08
[82]	The form of (headline, body) pairs from leading news organizations such as NDTV, CNN etc..	RNN, LSTM, BiLSTM, GRU, BiGRU	BiLSTM	Accuracy	0.84
[83]	FNC-1 open source dataset of articles.	MLP	MLP	FNC-score	83.08
[84]	Tweets on Twitter, discussion topics on Weibo and users.	CSI, DT, SVM, LSTM, GRU	CSI	Accuracy	0.95
[85]	19 fake news article websites (20,372 article) labeled by polifact, 9 real news article websites (20,932 article) listed by forbes.	3HAN, GRU	3HAN	Accuracy	0.97
[86]	Tweets from 174 suspicious propaganda accounts identified by PropOrNot and manually constructed a list of 252 trusted news accounts by writers.	LR, RNN, CNN	RNN, CNN	F1-score	0.92
[87]	LIAR open source dataset of articles.	LR, SVM, BiLSTM, CNN	CNN	Accuracy	0.27
[88]	LIAR open source dataset of articles.	LR, SVM, RNN, GRU, LSTM, Bi-LSTM, CNN	CNN	Accuracy	0.27
[89]	Open source dataset from fakenews.mit.edu.	GRU, LSTM, BiLSTM, SMHA-CNN	SMHA-CNN	F1-score	0.96

works, including Facebook, and stated that the method, they proposed, was the best.

When this study is compared with the studies about natural language processing, it was observed that it was the first study made for fake news detection in the Turkish language in short texts. In addition, in the fake news detection problem, it is thought that it was the first study to use the social network analysis methods in Turkish language.

3 Material-Method

3.1 Dataset

In this study, the two most shared topics on Twitter among the current fake news on many topics on Teyit.org platform on August 2019 were determined. In addition, the transfer of the football player named Radamel Falcao to Galatasaray Football Club, which was discussed on Twitter platform in August 2019, was used in the study. All shares made on this issue were collected via Twitter Scraper. Whether the posts shared were fake or real news was tagged with the tagging platform developed by the article authors. Finally, the followers of the owners of these posts shared were aggregated through the Twitter Search API. The topics in question are listed as follows:

Topic-1: The first topic is the claim that Bayburt Airport was built by the T.R. Ministry of Transport and Infrastructure, with 2 million passengers guaranteed annually. For the real tweet message on this subject, the tweet message of “Baybut will have an airport with capacity of 2 million passengers per year Ahxnbabxbabd” can be given as an example. And as an example of the fake tweet message, “An airport with 2 million passengers guaranteed for Bayburt. Let a confident economics professor tell us about it” tweet messages can be given. When the fake and real tweet message examples were analyzed, it can be seen that the words “guaranteed” and “capacity” indicate whether the tweet message was real or fake. The tweet messages on this subject were parsed by using the search phrase of “Bayburt airport”, and include tweet messages posted between the dates of July 2009 and September 2019.

Topic-2: The second topic was the claim that Galatasaray football club had transferred a football player named Radamel Falcao. As an example of the real tweet message on this topic, tweet messages of “We are obsessed with Falcao. It is true, he should come anymore if he is likely to come, Galatasaray” or

“We waited tonight, too, but you did not come @FALCAO ultrAslan Galatasaray” can be given. And as examples of the fake tweet message, the tweet messages of “@FALCAO Good luck to our Galatasaray community. or ”And Galatasaray informed the Stock Exchange that they agreed with Falcao!” can be given. Unlike the first subject, in tweet messages on this subject, the fake tweets messages and real tweet messages could be made up of completely different words. The tweet messages on this subject were parsed using the search phrase of “alcao Galatasaray” and included tweet messages from the date of April 2011 to the date of August 2019. Galatasaray football team could not transfer the player named Radamel Falcao on these dates, but this transfer realized on the date of 02nd October, 2019.

Topic-3: The third topic was the claim that the Kuleli Military High School was sold by T.R. Ministry of Culture and Tourism. As an example of real tweet messages on this topic, the tweet messages of “Culture and Tourism Minister Mehmet Nuri Ersoy stated that the news on the social media and press about the sale of Kuleli Military High School is completely unfounded.” can be given. And as an example of fake tweet messages, the tweet message of “Kuleli Military High School has been sold to the Arabs” can be given. When the tweet messages on this subject were analyzed, the fake tweet messages contain the word “sold”, while real tweet messages were made up of different words. The tweet messages on this subject were parsed using the search phrase of “Kuleli Military High School” and included the tweet messages posted between the dates of May 2009 and August 2019.

In this study, four different data sets consisting of tweet texts for use in natural language processing methods have been studied. These are respectively; the data set containing all tweet messages, the data set containing Topic-1 tweet messages, the data set containing Topic-2 tweet messages and the data set containing Topic-3 tweet messages. From the total tweet messages of 76738 tweet messages, retweets, spam and unnecessary messages were eliminated, and 18021 tweet messages remained. These tweet messages were confirmed by teyit.org platform and tagged one by one with the user interface created for tagging. Personal comments and irrelevant messages made on these news have been eliminated. 846 (by 783 different users) of these tagged tweet messages were tagged as fake news and 441 (by 420 different users) of them were tagged as real news. Three social network data sets have been created, one for each topic. In Table 3, the num-

Table 3 Numerical information about dataset

	F	F'	FU	F'U	FF	F'F
Topic-1	230	145	210	140	2.001.523	393.967
Topic-2	222	155	209	152	849.073	1.066.007
Topic-3	394	141	364	128	2.246.878	2.865.459
Total	846	441	783	420	5.097.474	4.325.433

bers tweets and social network of the data sets consisting of Topic-1, Topic-2 and Topic-3 tweet messages are shown. In the table, F indicates the number of fake news tweets, F' indicates the number of correct tweet news, FU indicates the number of users who spread the fake news, F'U indicates the number of users who spread the correct news, FT indicates the number of followers of the users spreading fake news, and F'. T indicates the number of followers of the users spreading the correct news. While the users (FU and F'U) in this table represent the nodes of the social network graph, the follower interactions (FT and F'T) represent the edges of the social network graph.

After the tagging process in the data set, the similarities of the tweet messages were measured through the Levenshtein distance algorithm. Tweet messages with a similarity of less than 0.5 were selected so as to use in machine learning algorithms. In order to receive equal numbers of both fake and real tweets in each subject, the minimum number of tweets remaining after similarity measurement was taken into account. The minimum number of tweets, with 90 tweets, was the number of real news tweet messages of Topic-3 tweet messages. For the reason, 90 real news tweets and 90 fake news tweet messages were selected on every topic. A total of 180 tweet messages were used on every topic.

3.2 Feature Extraction

In NLP (Natural Language Processing) studies conducted with machine learning, it is essential that textual expressions be converted into vector expressions in order to perform mathematical operations. In this study, Term frequency-Inverse document frequency (TF-IDF) and Word2vec word representation methods are used to represent words as numerical vectors.

3.2.1 TF-IDF

TF-IDF is a statistical method that finds the value of how important a word is to make predictions in the data, which will be given as input to machine learning. In this method, term frequency value is calculated firstly. In the term frequency, it is calculated how many times each word is used in each document. Equation 1 is used to calculate the TF value.

$f(t, d)$ = the frequency of the t term in d document

$$TF(t, d) = \frac{f(t, d)}{n_d} \quad (1)$$

Here, n_d represents the size of the d document sequence (the number of tweet messages in this study). After the calculation of the TF score, the reverse document frequency is calculated. In reverse document frequency, it is checked how many times a word is used in a document. Equation 2 is used to calculate the reverse document frequency.

$$IDF(t) = \log\left(\frac{1 + D}{D_t}\right) \quad (2)$$

In the equation, D represents the number of documents (in this study, the number of tweet messages), and D_t represents the number of documents in which the term t is used. The log function which is found in the equation is the dampening function. The TF-IDF score of each t term in d document is calculated with Eq. 3;

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (3)$$

Before converting the text tweet messages into vectors with TF-IDF, the following preliminary processing works were carried out on the text, respectively. All letters are converted into lower cases. Numeric expressions, punctuation marks, and non-textual expressions were removed from tweet messages. Then, using the Turkish stop words library in the NLTK (Natural Language Toolkit) library [36], these stop words were removed from the tweet messages. Later, the texts were converted into TF-IDF vectors by using the TfidfVectorizer method of the Sklearn (Scikit-Learn) library [37].

3.2.2 Word2vec

Before converting text tweet messages to vectors through word-2vec, as it was made before converting to TF-IDF vectors, preliminary processing works were performed in the text. In order to increase the success of the algorithms, the previously trained word2vec model was used. In the platform, formed by the University of Oslo Language Technology Group, there included pre-trained word representation models that have been uploaded by many researchers [38]. In this study, by using CoNLL (Computational Natural Language Learning) data set consisting of 3633786 words in Turkish language, and texts collected by Wikipedia and a general browser, and with the skip-gram method, the trained word2vec model was used. Then, a matrix, with the size of 50x100, was formed for each tweet message with word2vec vectors corresponding to each word of the tweet messages converted into word vector. Tweet messages that are shorter than 50 words were filled with zero vectors in the size of

1x100 to complement them to 50x100. This matrix, with the size of 50x100, is given as input to deep learning algorithms. Supervised learning algorithms and unsupervised learning algorithms without ANN basis accept vectors as input. Therefore, it is necessary to represent the matrix, consisting of 50x100 word2vec vectors, with a vector. In many essays, the average vector of the word2vec matrix is taken for this process [39–43]. There are also studies representing the vectors with a single vector by adding them one under the other. In a study, the use of TF-IDF and Word2vec vectors together was given as an input to the algorithms [44]. In this study, the addition of vectors one under the other and average vector methods are tried. It was seen that the average vector method was found to be quite successful compared to the adding one under the other method. Therefore, in this study, the average vector is given as input to the supervised learning and unsupervised learning algorithms without ANN basis.

The vectors, which are obtained in Word2Vec methods, can consist of positive and negative numbers. NMF (Non-Negative Matrix Factorization) method only accepts non-negative numbers as inputs. For the reason, the smallest negative numbers in the vectors created by Word2Vec were found and the absolute values of these numbers were added to the all numbers in the vector. In this way, it is ensured that the word vector consists of non-negative numbers.

3.3 Unsupervised Machine Learning

In unsupervised learning algorithms, untagged data is given to the algorithm as input. Moreover, it is specified how many clusters the algorithm will allocate this data. By learning the relationships and structures of the data by the algorithm, the data is divided into, the most meaningful way, clusters in the number given to the algorithm [45].

In this study, K-means, NMF and Linear Discriminant Analysis (LDA) unsupervised learning algorithms are used. It is seen that the K-means algorithm is also used in text clustering problems in the literature [46,47].

The K-means algorithm, which is the most widely used unsupervised learning algorithm, takes the number of clusters from the outside as a hyper parameter. By assigning a random center point for each of the given k pieces of clusters, the elements closest to the center point are included in that cluster. In this way, the algorithm tries to create k number of clusters. In this study, KMeans class of Sklearn [37] library is used. The argument of the “n clusters” hyperparameter, which indicates the number of clusters, is set to 2 in way that they represent fake and non-fake tags. In this study, both “k-means ++” and random center point selection methods are used as an argument in the “init” hyperparameter, which determines the initial center points.

NMF was proposed as an alternative to the k-means clustering algorithm and principal components analysis methods

[48]. The purpose in this method is to approach to a given nonnegative matrix in terms of the type of two nonnegative matrices multiplication. In this study, the NMF class of the Sklearn library is used. The argument of the “n components” hyperparameter, which indicates the number of clusters, is set as 2 in a way that they represent fake news and real news tags. The algorithm is trained by using the TF-IDF and Word2vec word representation methods. Furthermore, when NMF class of the Sklearn library is used, the selection of the starting matrix can be determined by the hyperparameter “init” and the selection of the solver can be determined by the “solver” hyperparameter. In this study, as an argument to the “init” hyperparameter, random, NNDSVD (Nonnegative Double Singular Value Decomposition), NNSVDA, filling the zero values with the matrix averages, and NNSVDAR filling the zero values with random small value. And as argument to “solver”, which is the hyper-parameter of the selection of solver, Coordinate Descent-CD and Multiplicative Update-MU are given.

LDA specifies a hyperplane to cluster the data based on their similarities [49]. In this study, the “Linear Discriminant Analysis” class of the Sklearn library is used. As an argument to the n components hyperparameter, which indicates the number of clusters, it is set as 2 in a way that they represent fake and non-fake tags. Singular Value Decomposition (SVD) and Least Squares Solution (LSQR) algorithms are given as an argument to the “solver” hyper-parameter used to determine the solver algorithm. The algorithm is trained with the TF-IDF and word2vec word representation methods.

3.4 Supervised Machine Learning

Machine learning is a science that focuses on the design and development processes of algorithms that enable the classification, grouping or prediction of a large number of data collected from sensors and detectors or accumulated in the database. It is aimed, with machine learning, to give computers the ability to recognize complex patterns and to make decisions. It has close relation to such areas as statistics, probability, data mining, pattern recognition, and artificial intelligence [50]. In supervised learning algorithms, data and tags of these data are also given to the machine learning algorithm as input. By learning a relational network among the data given to the algorithm by the algorithm, it is predicted which output the new data will have [45].

In this study; KNN, SVM and RF non-ANN-based supervised learning algorithms are used. In addition, 6 different supervised deep learning algorithms such as RNN, GRU, LSTM and the bidirectional algorithms of these algorithms (BiRNN, BiGRU and BiL-STM) are used.

3.4.1 Supervised Learning Algorithms which are not ANN based

The KNN algorithm takes one k hyper parameter from outside. This k parameter specifies the nearest k neighbors to a data to be classified. The distance is calculated with the nearest k elements to the data to be classified. These distances can be calculated with different methods such as Euclidean, Manhattan, Minkowski, Hamming and Cosine distance methods. In this study, the “KNeighborsClassifier” class of Sklearn library has been used. In this method, k hyper parameter is given as $k = 2$. Manhattan, Euclidean and minkowski distance functions are given as an argument to the hyperparameter “ p ” which determines the distance calculation method. If the hyperparameter p of the KNeighborsClassifier method is given 1 as an argument, Manhattan function can be used, if 2 is given Euclidean function can be used and if 3 or more is given then minkowski functions can be used.

SVM aims to reduce the generalized upper bound error rather than the local training error. This is one of the main advantages of the SVM model compared to the traditional machine learning methods [51]. The SVM algorithm determines the hyper planes, which are the most optimal so as to classifying the vectors contained in the training data into different classes. Training examples as points in space, and then divides the space using hyperplanes to give the best separation among the classes [52]. Kernel functions are used so as to determine nonlinear planes [53]. Mostly used kernel functions are Linear, Polynomial, Sigmoid and Radial Based C Kernel Functions [54]. In this study, the “SVC” class of Sklearn [37] library is used. Kernel function required to determine the nonlinear monitoring is determined with the “kernel” hyperparameter. As an argument for this hyperparameter, linear, poly (2nd, 3rd, 4th, 5th and 6th order polynomial), RBF and sigmoid values are given.

In the RF algorithm, which is a decision tree algorithm, the algorithm creates different decision trees by selecting different subsets. Each occurring decision tree makes an estimate. Afterwards, in the classification, the value with the highest vote for these estimates is chosen. On account of the creation of multiple subsets, the problem of excessive learning of decision trees is reduced in this algorithm [55]. In this study, the “Random Forest Classifier” class of the Sklearn library is used. The Values of 50, 100, 500 and 1000 are given as an argument to the “ n estimators” hyperparameter, which is the hyperparameter that indicates the number of trees in the forest. The “gini” and “entropy” values are given as an argument to the “criterion” hyperparameter, which defines division quality criterion.

3.4.2 Deep Learning Algorithms

As people read a text, they extract meanings from the text by keeping the previous words and sentences in their minds. They realize the learning event by examining the formerly learned information with the newly learned information. In the neural networks, which imitates human nature, the previous data are not taken into account while making extractions from traditional neural networks. However, in RNNs, extraction is made by considering the previous data. Because of its this feature, RNN is often used in the NLP field. The RNN model is created by repetition of more than one neural network. The output in each step is transmitted to the next neural network.

With RNN, data in the first parts of the sentence or the text in long sentences or texts (words or phrases) cannot be carried over until the last parts. Therefore, the important data found in the beginning parts in long sentences and texts can be forgotten at the end of the sentences or text. For example, in the sentence of “Clouds are in the sky.”, when the word of “clouds” given as input to the RNN model, the phrase of “is in the sky” can be guessed. Yet, in the sentence of “I grew up in Turkey and I can speak Turkish fluently.”, when the word of Turkish is tried to be estimated, it is necessary that the word of “in Turkey” at the beginning of the sentence be remembered. Therefore, the memory of RNN models is short and remains insufficient. As a solution for this problem, LSTM [56] and GRU [57] models are recommended. The LSTM method can also be used in current problems other than NLP [58–60].

In the “double-oriented” RNN architecture recommended, while one layer progresses with feed-forward and another layer, by starting from the end, progresses with feed-back [61]. Both the feed-forward and feed-back layers update the output value in the t time.

Python programming language is used to create and model deep learning algorithms used in this study. In order enable the models operate, Keras Sequence Model application programming interface of Tensorflow library is used [62]. Each of the supervised deep learning algorithms were operated with 3 different models and 2 different learning rates (LR) and the results were tested. The model, having entrance layer, hidden layer consisting of 50 units, hidden layer consisting of 10 units, exit layer with 1 unit, is called Model-1, the model, having entrance layer, hidden layer consisting of 50 units, hidden layer consisting of 25 units, hidden layer consisting of 10 units, exit layer with 1 unit, is called Model-2, and finally, the model, having entrance layer, hidden layer consisting of 100 units, hidden layer consisting of 50 units, hidden layer consisting of 25 units, hidden layer consisting of 10 units, exit layer with 1 unit, is called Model-3 (Fig. 1). These 3 models were tested with learning coefficients of 0.001 and 0.0001. As the activation function, tanh activation function is used in hidden layers and sigmoid activation function in

Table 4 Confusion matrix

	F'	R'	
F	TP	FN	P
R	FP	TN	N
	P'	N'	

exit layer. The Adam optimization algorithm was used as the optimization algorithm. And for the loss function, the binary cross entropy loss function is used.

3.5 Performance Metrics

One of the methods which are used to measure the success of machine learning algorithms is the use of the confusion matrix. There are various studies in the literature using confusion matrix performance metrics [63–66]. In the Table 4, structure of the confusion matrix is shown. In this matrix, it can be seen that how many correct predictions are made and how many incorrect predictions are made. Considering this study in the chart, F indicates the fake news, R indicates the real news, F' indicates the fake news predictions, and R' indicates the real news predictions. P indicates the total number of the fake news, N indicates the total number of the real news, P' indicates the total number of predicted fake news, and N' indicates the total number of predicted real news. TP represents the number of correctly estimated fake news. On the other hand, FP represents the number of that data that are predicted as fake news, but are supposed to be the real news. TN represents the number of real news predicted as the real ones. Finally, FN represents the number of data that are estimated as being real news, but are actually the fake news.

By using the numerical data about the correct and incorrect predictions in the confusion matrix, the accuracy of the model, measuring criterion can be calculated. Some of these criteria are shown in the Eqs. refreq:precision, 5, 6 and 7. The precision value indicates how much of the data predicted by the algorithm is correctly predicted in the first cluster. Considering the problem of detecting fake and real news, it gives the ratio of how many fake news in the data set were selected as fake news. The higher the precision value is, the greater the number of correctly tagged elements in the first set is. The equation of precision metric value is given in Eq. 4. If it is desired to predict exactly the elements correctly in the first cluster in the solution of a problem, it is necessary that the Precision value be high. For example, in the detection of an epidemic disease such as Covid-19, it is important that the Precision value be high when it is desired to, at a high rate, predict that someone is sick.

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{P'} \quad (4)$$

The recall value, on the other hand, gives the ratio of how much of the data tagged by the algorithm in the first cluster is correctly tagged. Again, when the problem of fake news detection is considered, it gives the ratio of the data, marked by the algorithm as fake news, to all fake news. In Eq. 5, the equation for the recall metric value is given. In the solution of a problem, the accuracy of the predicted data is important in both clusters, it is important to keep the precision and recall values together high. For example, in the fake news detection problem, it is essential that both precision and recall metric values be high.

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (5)$$

The harmonic mean of the Precision and Recall values gives the F1-metric value.

$$F1\text{-score} = 2x \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

The accuracy value is the ratio of predictions to be made correctly, as a result of the algorithm, to all predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{P + N} \quad (7)$$

Another metric used in measuring the success of machine learning algorithms is purity. Purity is the ratio of the biggest number, as numerical, of elements in each cluster to the total number of data [67]. With the purity value, it is possible to learn which class of data will be used more in the clusters of unsupervised learning algorithms. Through Eq. 8, the purity value can be calculated.

$$Purity = \frac{\sum_{i=1}^K \frac{|c_i^d|}{|c_i|}}{K} \quad (8)$$

K refers to the number of clusters, and c_i^d refers to the number of data belonging to the dominant class in the c_i cluster. In this study, the F1-score value is taken into account for the evaluation of machine learning algorithms. Because the parameters of the machine learning algorithms were chosen randomly in the exercise, instead of considering a single outcome of each model, each model was operated 100 times, and the average of these 100 different F-1 score values was taken into account. The steps of the method of fake news detection through machine learning algorithms are given in Fig. 2.

3.6 Social Network Analysis

Social network analysis is a method that examine, with the graph theorem, the users of social networks such as Face-

Fig. 1 Deep Learning Models for Fake News Detection.

Fig. 2 Steps of machine learning algorithms in fake news detection problem

book and Twitter and the interaction of these users with each other. In social networks, users are represented by nodes (vertex / nodes) and their relations with each other are represented by edges. The case of friendship with a person on the Facebook platform is double-oriented. So, when a person is added as a friend, both are considered to be friends of one another. Therefore, the social network graph can be used as an omnidirectional graph for friendship status on the Facebook platform. The following up system on the Twitter platform can be one-way. In other words, while a person is following someone, the person he follows may not follow him. Therefore, the social network graph on the Twitter platform must be a directed graph [68]. PageRank is a link ranking algorithm that is also used in social network analysis. It has been found in order to sort the closest pages for the search to be made in a search engine. In the PageRank algorithm, if too many nodes point to a node, then the PageRank score of that node will be higher. At the same time, if a few people are following a node but the PageRank is pointing to a high value, then again, the PageRank value of that node will be high. Calculation of the PageRank value of each node is shown in Eq. 9.

$$PR(u) = c \sum_{v \in B_u} \frac{R'(v)}{N_v} \quad (9)$$

In Eq. 9, u refers to a website, B_u , refers to the cluster of sites hosting the link of the u website, $PR(u)$ and $R'(v)$ refer to the rank value of websites u and v , respectively, N_v , refers to the number of links on the v page and c refers to the normalization factor.

In the HITS algorithm, the recommended sites are also brought, apart from the websites that are directly related to the search result. Web sites directly related to the search result are called hubs, and the recommended web sites are called authorities [69]. While automobile sites may represent hubs in a car-related search, auto-mobile-related forum sites may represent authorities. The hubs that show more good authorities than the other hub nodes are called “good hubs”. Authorities indicated by more good hubs than any other authority nodes are called “good authority”. In the HITS algorithm, while calculating the authority and hub scores, they are initially taken as 1. Then, the authority score of each node is updated using Eq. 11 and the hub score of each node is updated by using Eq. 9.

$$auth(n) = \sum_{q \in N_{to}} hub(q) \quad (10)$$

In Eq. 10, n represents the node, authority score of which is to be calculated, and N_{to} represents nodes that indicate n node. In other words, the authority score of the node is the

total of the hub scores of all the nodes that point to that node.

$$hub(n) = \sum_{q \in N_{from}} auth(q) \quad (11)$$

In Eq. 11, n represents the node, hub score of which is to be calculated, and N_{from} represents the nodes to which the n node points. In other words, the hub score of the node is the total of the authority scores of all nodes to which that node points.

In this study, whether people share fake news or not with the follower interaction of users on the Twitter social network are examined. In our problem, answers are sought for the questions “to what extent do users who spread fake news follow other users who spread fake news?”, “To what extent do users who spread fake news follow users who do not spread fake news?”. From the data set obtained, followers of the users who spread fake news and of the users who do not spread fake news were collected from the twitter platform. In this way, a graph has been created from the follow-followers relation. For the use and visualization of social network analysis methods, Gephi application, which is a graph visualization and analysis software for social network analysis, was used [70]. A social network analysis was performed using PageRank and HITS algorithms from Gephi application.

4 Research Findings

In order to reduce the prejudice of the algorithms, each algorithm was studied 100 times. In each step, the tweet messages on the topics were randomly mixed and it was ensured that all messages could be found in both the training set and the test set. Of the data set mixed in each step, 70% was used as the training set to train the model and remaining 30% was used as the test set to measure the success of these algorithms. Due to the small number of data in the dataset and to ensure that more data is available in the test dataset, 70% training and 30% test dataset were used.

In the tables shown in this section, the dark colored values indicate the biggest line-based F1-score value.

4.1 Unsupervised Learning Algorithms

Unsupervised learning algorithms, unlike supervised learning algorithms, do not need tagged data. For the reason, it does not require a very costly process such as data tagging. But in this study, solving relationships between data, as well as supervised learning algorithms, were not successful enough. While the TF-IDF word representation method was more successful in supervised learning algorithms, word2vec algorithms, which reveal the relation between data more in

unsupervised learning algorithms, yielded better results. In the study, K-means, LDA and NMF algorithms from unsupervised learning algorithms were operated with all parameters described in sect. 4.3, and their results were examined. The results of the K-means and NMF algorithms did not yield a sufficient success. LDA algorithm obtained F1-score average between 0.70 and 0.76 when word2vec word representation method of all data sets was used.

4.2 Supervised Learning Algorithms, which are not ANN based

The KNN algorithm was tested with different distance functions and different numbers of neighbors. In Table 5, F1-metric averages of the results obtained with different neighbor numbers and different distance algorithms by using the TF-IDF word representation method of the KNN algo-

Table 5 F1-score values obtained by using the TF-IDF word representation method, with different parameters, of the KNN algorithm.

		k=2	k=3	k=4	k=5	k=6	k=7	k=8
All	Manhattan	0.75	0.74	0.75	0.74	0.75	0.75	0.76
	Euclidean	0.77	0.80	0.80	0.81	0.82	0.82	0.83
	Minkowski	0.74	0.78	0.78	0.77	0.77	0.76	0.76
Topic-1	Manhattan	0.69	0.73	0.74	0.73	0.73	0.71	0.69
	Euclidean	0.66	0.73	0.72	0.77	0.77	0.78	0.78
	Minkowski	0.65	0.72	0.73	0.74	0.74	0.74	0.73
Topic-2	Manhattan	0.71	0.73	0.74	0.75	0.75	0.75	0.76
	Euclidean	0.75	0.78	0.78	0.79	0.79	0.80	0.81
	Minkowski	0.75	0.76	0.77	0.74	0.75	0.73	0.74
Topic-3	Manhattan	0.72	0.73	0.73	0.73	0.73	0.72	0.74
	Euclidean	0.78	0.8	0.81	0.82	0.82	0.82	0.83
	Minkowski	0.71	0.73	0.74	0.69	0.69	0.62	0.65

Table 6 F1-score values obtained by using Word2vec word representation method, with different parameters, of the KNN algorithm

		k=2	k=3	k=4	k=5	k=6	k=7	k=8
All	Manhattan	0.76	0.78	0.79	0.78	0.79	0.78	0.78
	Euclidean	0.76	0.79	0.79	0.78	0.79	0.78	0.8
	Minkowski	0.76	0.77	0.79	0.78	0.79	0.78	0.79
Topic-1	Manhattan	0.77	0.79	0.77	0.78	0.78	0.80	0.77
	Euclidean	0.77	0.78	0.78	0.77	0.77	0.78	0.78
	Minkowski	0.78	0.78	0.77	0.77	0.78	0.77	0.79
Topic-2	Manhattan	0.74	0.75	0.75	0.74	0.74	0.74	0.74
	Euclidean	0.75	0.75	0.75	0.74	0.75	0.74	0.75
	Minkowski	0.75	0.75	0.74	0.76	0.74	0.74	0.75
Topic-3	Manhattan	0.82	0.82	0.81	0.81	0.82	0.8	0.81
	Euclidean	0.81	0.81	0.77	0.82	0.82	0.82	0.81
	Minkowski	0.82	0.82	0.81	0.82	0.82	0.81	0.80

Table 7 F1-score values of the RF algorithm obtained with different parameters.

		Entropy-50	Entropy-100	Entropy-500	Entropy-1000	Gini-50	Gini-100	Gini-500	Gini-1000
All	TF-IDF	0.84	0.84	0.85	0.85	0.84	0.85	0.85	0.85
	W2v	0.79	0.8	0.8	0.81	0.78	0.79	0.8	0.81
Topic-1	TF-IDF	0.89	0.89	0.90	0.9	0.88	0.89	0.9	0.89
	W2v	0.8	0.81	0.83	0.83	0.8	0.82	0.83	0.84
Topic-2	TF-IDF	0.78	0.78	0.77	0.78	0.77	0.77	0.79	0.78
	W2v	0.76	0.77	0.77	0.77	0.75	0.76	0.76	0.77
Topic-3	TF-IDF	0.8	0.81	0.82	0.82	0.8	0.8	0.81	0.81
	W2v	0.83	0.83	0.84	0.84	0.83	0.83	0.84	0.84

Table 8 F1-score values of the SVM algorithm obtained with different parameters.

		Linear	Poly-2	Poly-3	Poly-4	Poly-5	Poly-6	RBF	Sigmoid
All	TF-IDF	0.86	0.87	0.83	0.75	0.63	0.54	0.87	0.86
	W2v	0.79	0.81	0.82	0.82	0.82	0.81	0.81	0.7
Topic-1	TF-IDF	0.89	0.83	0.75	0.61	0.42	0.44	0.87	0.89
	W2v	0.84	0.82	0.83	0.83	0.82	0.82	0.83	0.8
Topic-2	TF-IDF	0.82	0.82	0.79	0.7	0.61	0.55	0.83	0.8
	W2v	0.78	0.77	0.77	0.77	0.76	0.74	0.77	0.75
Topic-3	TF-IDF	0.86	0.84	0.78	0.66	0.54	0.44	0.86	0.86
	W2v	0.87	0.86	0.86	0.85	0.85	0.84	0.85	0.8

algorithm are given. When the Euclidean distance algorithm is used in all data sets, and with 8 neighbor parameters, the most successful results were obtained.

In Table 6, F1-metric averages of the results obtained, with different neighbor numbers and different distance algorithms, by using the Word2vec word representation method of the KNN algorithm are given. Unlike the KNN algorithm used with the TF-IDF word representation method, it was observed that the most successful results could be obtained with different parameters on each subject.

The RF algorithm was tested with measurement algorithms of different division quality and tree numbers. In Table 7, F1-metric averages of the results obtained by using 4 different data sets with entropy and gini algorithms of the RF algorithm and 50, 100, 500 and 1000 network numbers are given. When the results are examined, successful results, with at least 500 trees count, are obtained.

SVM algorithm was tested with the linear, radial-based function, sigmoid and polynomial algorithms. In Table 8, the F1-score averages of the results of SVM algorithm, obtained using linear, radial-based function, sigmoid and polynomial algorithms, TF-IDF and Word2vec word representation methods and 4 different data sets, are given. When the results are examined, it can be seen that more successful results can be obtained with the Linear kernel function.

4.3 Deep Learning Algorithms

In this study, 6 different supervised deep learning algorithms, including RNN, GRU, LSTM, and the bidirectional' algorithms of these algorithms, BiRNN, BiGRU and BiLSTM, were used.

The results of one-way supervised deep learning algorithms obtained with a data set consisting of tweet messages of all subjects are shown in Table 9. GRU and LSTM algorithms were more successful than the RNN algorithm. And in learning rate, the algorithms, having a higher learning rate, were more successful.

The results of the bidirectional supervised deep learning algorithms obtained with the data set consisting of tweet mes-

Table 9 F1-score values of the Deep Learning algorithms obtained with different parameters.

		LR=0.001			LR=0.0001		
		M-1	M-2	M-3	M-1	M-2	M-3
All	RNN	0.59	0.58	0.56	0.58	0.57	0.57
	GRU	0.79	0.79	0.79	0.68	0.7	0.72
	LSTM	0.8	0.8	0.8	0.76	0.77	0.76
Topic-1	RNN	0.57	0.57	0.57	0.57	0.56	0.56
	GRU	0.83	0.83	0.81	0.42	0.44	0.39
	LSTM	0.82	0.83	0.81	0.7	0.78	0.79
Topic-2	RNN	0.57	0.57	0.56	0.56	0.56	0.57
	GRU	0.81	0.83	0.81	0.42	0.42	0.38
	LSTM	0.8	0.8	0.82	0.73	0.76	0.80
Topic-3	RNN	0.57	0.57	0.56	0.57	0.55	0.55
	GRU	0.79	0.82	0.81	0.4	0.41	0.41
	LSTM	0.81	0.81	0.81	0.73	0.76	0.8

sages of all subjects are shown in Table 10. Models with a lower learning rate were more successful. And it was seen that Model-1 gave more successful results than other models.

The most successful F1-score values of non-ANN-based supervised learning algorithms and deep learning algorithms used with TF-IDF and Word2vec word representation method for all data sets are shown in Fig. 3. When the results are examined, it can be observed that the TF-IDF word representation method was more successful. In the Topic-2 dataset, it was observed that the deep learning algorithms increased the success of the word2vec word representation method.

4.4 Social Network Analysis

For the social network analysis, the users who spread fake and real tweet messages on every subject, and a maximum of 1000 followers of them, were saved from the database as a csv extension file so as to be analyzed in the Gephi application. Later, these user nodes and the edges indicating the follower relationship were given to the Gephi application. After determining the edges and nodes for each three topics,

Fig. 3 Error bars graph indicating F1-metric averages and standard deviations of non-ANN-based supervised learning algorithms, deep learning algorithms, and unsupervised learning algorithms in all data sets.

Table 10 F1-score values of the Bi-directional Deep Learning algorithms obtained with different parameters

		LR=0.001			LR=0.0001		
		M-1	M-2	M-3	M-1	M-2	M-3
All	Bi-RNN	0.66	0.65	0.64	0.63	0.62	0.63
	Bi-GRU	0.78	0.78	0.78	0.77	0.78	0.79
	Bi-LSTM	0.8	0.79	0.79	0.79	0.79	0.78
Topic-1	Bi-RNN	0.61	0.61	0.62	0.58	0.57	0.57
	Bi-GRU	0.82	0.81	0.79	0.68	0.7	0.75
	Bi-LSTM	0.82	0.81	0.81	0.81	0.82	0.83
Topic-2	Bi-RNN	0.63	0.62	0.61	0.58	0.58	0.58
	Bi-GRU	0.79	0.79	0.79	0.67	0.7	0.75
	Bi-LSTM	0.81	0.8	0.82	0.82	0.82	0.83
Topic-3	Bi-RNN	0.62	0.62	0.6	0.58	0.56	0.58
	Bi-GRU	0.8	0.8	0.79	0.67	0.69	0.74
	Bi-LSTM	0.82	0.82	0.81	0.81	0.82	0.82

the PageRank and HITS scores of each topic were calculated in the Gephi application.

4.4.1 HITS Algorithm

In Topic-1 tweet messages, 1 of 50 people having the highest hub score spread fake news, while the remaining 49 other users did not share any news about this topic. All of these 50 users followed mostly the users who spread fake news. In Topic-2 tweet messages, 4 of the 50 people having the highest hub score spread real news, while the remaining 46 other users did not share any news about this topic. Of these 50 users, 43 followed users who spread more fake news, and 7 followed users who spread more real news. In Topic-3 tweet messages, 1 of 50 people having the highest hub score spread real news, while the remaining 49 other users did not share any news about this topic. Of these 50 users, 44 followed users who spread more fake news, and 6 followed users who

spread more real news. Considering the hub score, it can be seen that a vast majority of 50 people having the highest hub points follow users who spread fake news more than users who spread real news. The users having high authorization points are those users in the social network followed by people having high hub points. When this algorithm is operated in the data set of users who spread fake and real news, the users having the highest authorization score were composed of users who spread fake and real news. After calculating the link sorting algorithms on the Gephi application, a ranking was made from the one with the highest authority score to the one with the lowest score. As a result of the ranking, for each topic, after 200 users, the authority score was reset to zero. These users were not taken into account because they were ineffective users in terms of authorization scores. The users, whose authority scores, are not zero were analyzed.

From the users who spread fake and real news in Topic-1 tweet messages and their followers, 254 people having the highest authorization scores were taken. It was observed that the score of the people after these 254 people was zero. 127 of these users having the highest authorization scores consisted of users sharing fake news and 127 consisted of users sharing real news. One hundred twenty-seven users spreading fake news follow 302 accounts spreading fake news and follow 52 accounts spreading real news. One hundred twenty-seven users spreading real news follow 177 users spreading fake news and follow 47 users spreading real news. Two hundred fifty-four people having higher authorization scores follow users spreading fake news more, regardless of whether they were spreading fake news or real news (Fig. 4).

In Fig. 4, each of the red nodes indicates users spreading fake news about the Topic-1, and having the highest authorization scores, and each of the green nodes indicates users spreading real news about the Topic-1, and having the highest authorization scores. As it can be seen from the figure, users having high authorization scores followed people who spread fake news.

Fig. 4 **a** Topic-1 users, having high authorization scores. **b** Topic-1 users, having high authorization scores, and who shared real news. **c** Topic-1 users, having high authorization scores, and who shared fake news

From users who spread fake and real news in Topic-2 tweet messages and their followers, 247 people, having the highest PageRank scores, were taken. It was observed that the score of the people after these 247 people was zero. Of these 247 users, having the highest PageRank scores, 120 consisted of user spreading fake news and 127 consisted of users spreading real news. 120 users spreading fake news follow 304 accounts spreading fake news, and follow 319 accounts spreading real news. And 127 users spreading real news follow 423 users spreading fake news, and follow 395 users spreading real news. Considering the Topic-2 tweet messages, it was observed that 247 people, having high PageRank scores follow almost equal numbers of users spreading fake and real news (Fig. 5).

In Fig. 5, each of the red nodes indicates users spreading fake news about the Topic-2, and having the highest authorization scores, and each of the green nodes indicates users spreading real news about the Topic-2, and having the highest authorization scores. As it can be seen from the figure, users having high authorization scores follow the people who spread fake and real news at an equal rate.

From users who spread fake and real news in Topic-3 tweet messages and their followers, 246 people, having the highest PageRank scores, were taken. It was observed that the score of the people after these 246 people was zero. Of these 246 users, having the highest PageRank scores, 137 consisted of user spreading fake news and 109 consisted of

users spreading real news. 137 users spreading fake news follow 768 accounts spreading fake news, and follow 167 accounts spreading real news. And 109 users spreading real news follow 184 users spreading fake news, and follow 315 users spreading real news. It was observed that 137 users, spreading fake news, from 246 users, having high PageRank scores follow again users spreading fake news, and 109 users spreading real news follow again users spreading real news (Fig. 6).

In Fig. 6, each of the red nodes indicates users spreading fake news about the Topic-3, and having the highest authorization scores, and each of the green nodes indicates users spreading real news about the Topic-3, and having the highest authorization scores. As it can be seen from the figure, users spreading fake news among those having high authorization scores follow the people who spread fake news, and users spreading real news follow the people who spread real news.

4.4.2 Pagerank Algorithm

After the link ranking algorithms were made to be calculated on the Gephi application, a ranking was made from the one, having the highest PageRank score to the one, having the lowest PageRank. As a result of the ranking, the PageRank score started to repeat itself, for each topic, after a certain user. Users who were up to the repetition part of the PageRank

Fig. 5 **a** Topic-2 users, having high authorization scores. **b** Topic-2 users, having high authorization scores, and who shared real news. **c** Topic-2 users, having high authorization scores, and who shared fake news.

Fig. 6 **a** Topic-3 users, having high authorization scores. **b** Topic-3 users, having high authorization scores, and who shared real news. **c** Topic-3 users, having high authorization scores, and who shared fake news.

score were taken and analyzed. It showed exactly the same results as the HITS algorithm in all data sets.

5 Discussion and Conclusion

When the studies made are considered, it can be seen that the exposure rate of the people living in Turkey to fake news is very high, and that the rate of those who distinguish fake news is low. In addition, when it is considered that fake news is shared very much in the first 2-hour period, the use of automatic detection systems in the detection of fake news is of great importance.

In this study, fake news detection was carried out by using non-ANN-based supervised learning, deep learning and unsupervised learning algorithms. Non-ANN-based supervised learning algorithms and unsupervised learning algorithms were tested with TF-IDF and word2vec word representation methods. In addition, social network analysis methods were used so as to examine the interactions of fake and real news users. The follower relationships of all users were collected. Then, effective people were found and visualized with HITS algorithm and PageRank algorithm.

In machine learning algorithms, TF-IDF word representation method yielded more successful results than word2vec word representation method. The reason of this is that the data, used for education, of the pre-trained word2vec dataset consists of regular articles. Yet, tweet messages consist of messages shared by people independent of the official spelling rules. For example, there are words and word groups that are frequently used in street language such as “Come now Falcao”, “I hope he will come”, in Topic-2 data set and “making present” in Topic-3 data set. There is no a pre-trained word2vec model that has been trained with enough tweet messages data. Therefore, pre-trained models that were trained with the tweet messages data set were more unsuccessful compared to the TF-IDF word representation method. Furthermore, the data are given as vectors for non-ANN-based supervised and unsupervised learning algorithms. Hence, taking the average of all word vectors found from the tweet message with word2vec word representation affected again the success of the word representation.

Considering the F1-metric values of unsupervised learning algorithms, word2vec word representation method was more successful than TF-IDF word representation method in all data sets. Considering the cost of tagged data, it can be seen that the results of unsupervised learning algorithms, which do not need tagged data, are successful. It can also be thought that unsupervised learning algorithms can be used when necessary. Among the non-ANN-based supervised learning algorithms, the SVM algorithm was more successful than other algorithms in training with both topic-based data sets and data sets containing key topics or yielded a similar result

to the algorithm which was successful. The algorithm that distinguished the differences between fake and real news texts and gave the best results was the SVM algorithm.

GRU, LSTM, BiGRU and BiLSTM algorithms were successful in all data sets because they determine the relationships between words in deep learning algorithms by being affected, with memory elements in their structures, from past words. RNN and BiRNN, on the other hand, could not establish a meaning relationship between words because they are simpler. The results of the algorithms LSTM, GRU, BiLSTM and BiGRU, which gave the best results, were very close to each other. The RNN and BiRNN algorithms, on the other hand, yielded unsuccessful results because they did not consider the previous and next words in the sentence on account of their structure. As a result, it was seen that supervised learning algorithms were more successful in detecting fake news with tagged data than unsupervised learning algorithms. has been seen. The reason of this is that it is because the Unsupervised Learning Algorithms do not need a tag so as to cluster the data. Unsupervised learning algorithms cluster the data according to the similarities between the data. In Supervised Learning algorithms, since tagged data was given to the model, they were more successful in solving similarities between the same tags than unsupervised learning algorithms.

In the social network analysis, the center, authority and PageRank scores of the users were calculated by using HITS and PageRank algorithms. It was observed that users having high hub and PageRank scores consisted of the same users. The people, who could potentially spread fake news, or be exposed to fake news, were determined by analyzing the users, who spread fake and real news, and follower relationships of these users.

When the studies in the literature were examined, it was seen that there were many studies in the English language. Yet, no detailed study was observed in the Turkish language. It was observed that it was the first study conducted the Turkish language on social media. Moreover, the data set is a specific one which is collected for this study. In this study conducted, the detection of fake news was made by an automatic system and it enabled the detection of fake news in very short time. In this way, it is possible to detect and prevent the fake news, before it is shared too much, in a short time and without too many users being exposed to this fake news.

In subsequent studies to be conducted, different results can be obtained by examining the followers of the users who spread fake or real news, and of their followers. Moreover, follower analysis can be given as input to machine learning algorithms. Algorithm success can be increased by using a word2vec model which is trained with sufficient tweet messages instead of pre-trained word2vec models trained with regular texts. In addition, for synonyms, which we could not use since there is not a pre-trained model, with a sufficient

number of words, in the Turkish language, algorithm success can be increased by using newer models such as ELMo [71], BERT [72] and GPT-3 [73] which creates multiple different vectors.

References

- Del Vicario, M.; Bessi, A.; Zollo, F.; Petroni, F.; Scala, A.; Caldarelli, G.; Stanley, H.E.; Quattrociochi, W.: The spreading of misinformation online. *Proceed. Natl. Acad. Sci.* **113**(3), 554–559 (2016)
- Simon Kemp. Digital 2021: Global Overview Report, 2021.
- Teyit.org. teyit.org, 2016.
- Dogrulukpayi.com. Dogrulukpayi.com, 2014.
- Newman, N.; Fletcher, R.; Kalogeropoulos, A.; Nielsen, R.: Digital News Report 2018. Technical report, Reuters Institute for the Study of Journalism (2018)
- Newman, R.; Fletcher, N.; Kalogeropoulos, R.; Nielsen, A.: Digital News Report 2019. Technical report, Reuters Institute for the Study of Journalism (2019)
- Twitter. Twitter Inc., 2006.
- Zhao, W. X.; Jiang, J.; Weng, J.; He, J.; Lim, E. P.; Yan, H. and Li, X.: Comparing Twitter and Traditional Media Using Topic Models. In *ECIR 2011: Advances in Information Retrieval*, pp. 338–349 (2011)
- Pratiwi, I. Y. R.; Asmara, R. A.; Rahutomo, F.: Study of hoax news detection using naïve bayes classifier in Indonesian language. In *2017 11th International Conference on Information & Communication Technology and System (ICTS)*, pp 73–78. IEEE(2017)
- Chen, Y.-R.; Chen, H.-H.: Opinion Spam Detection in Web Forum: A Real Case Study. In *Proceedings of the 24th International Conference on World Wide Web - WWW '15*, pages 173–183, New York, New York, USA, 2015. ACM Press.
- Mertoğlu, U.; Sever, H.; Genc, B.: Savunmada Yenilikçi Bir Dijital Donusum Alani. In: *Savtek 2018*, 9, pp. 771–778. Savunma Teknolojileri Kongresi. METU, Ankara (2018)
- Zhao, B.; Rubinstein, B.I.; Gemmell, J.; Han, J.: A Bayesian approach to discovering truth from conflicting sources for data integration. *Proceed. VLDB Endowment* **5**(6), 550–561 (2012)
- Li, Q.; Li, Y.; Gao, J.; Zhao, B.; Fan, W.; Han, J.: Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation. In *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, pages 1187–1198, New York, NY, USA, 2014. ACM.
- Yang, S.; Shu, K.; Wang, S.; Renjie, G.; Fan, W.; Liu, H.: Unsupervised Fake News Detection on Social Media: A Generative Approach. *Proceed. AAAI Conf. Artif. Intell.* **33**, 5644–5651 (2019)
- Anderson, J.G.: Evaluation in health informatics: social network analysis. *Comput. Biol. Med.* **32**(3), 179–193 (2002)
- Otte, E.; Rousseau, R.: Social network analysis: a powerful strategy, also for the information sciences. *J. Inf. Sci.* **28**(6), 441–453 (2002)
- Borgatti, S.P.; Mehra, A.; Brass, D.J.; Labianca, G.: Network analysis in the social sciences. *Science* **323**(5916), 892–895 (2009)
- Brin, S.; Page, L.: The anatomy of a large-scale hypertextual web search engine. *Comput. Netw.* **30**, 107–117 (1998)
- Haveliwala, T.H.: Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search. *IEEE Trans. Knowl. Data Eng.* **15**(4), 784–796 (2003)
- Xing, W.; Ghorbani, A.: Weighted PageRank algorithm. In *Proceedings. Second Annual Conference on Communication Networks and Services Research*, 2004. pages 305–314. IEEE, 2004.
- Cailan, Z.; Kai, C.; Shasha, L.: Improved PageRank algorithm based on feedback of user clicks. In *2011 International Conference on Computer Science and Service System (CSSS)*, pages 3949–3952. IEEE, 2011.
- Kwak, H.; Lee, C.; Park, H.; Moon, S.: What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web - WWW '10*, page 591, New York, New York, USA, 2010. ACM Press.
- Weng, J.; Lim, E.-P.; Jiang, J.; He, Q.: Twitterrank: Finding Topic-Sensitive Influential Twitterers. In *Proceedings of the third ACM international conference on Web search and data mining - WSDM '10*, page 261, New York, New York, USA, 2010. ACM Press.
- Gupta, P.; Goel, A.; Lin, J.; Sharma, A.; Wang, D.; Zadeh, R.: WTF: the who to follow service at Twitter. In *Proceedings of the 22nd international conference on World Wide Web - WWW '13*, pages 505–514, New York, New York, USA, 2013. ACM Press.
- Ngaffo, A. N.; El Ayeb, W.; Choukair, Z.: Mining User Opinion Influences on Twitter Social Network: Find that Friend who Leads your Opinion Using Bayesian Method and a New Emotional PageRank Algorithm. In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pages 680–685. IEEE, 2019.
- Moreno, F.; González, A.; Valencia, A.: NewFriends: an algorithm for computing the minimum number of friends required by a user to get the highest PageRank in a social network. *Int. J. Comput. Math.* **91**(2), 278–290 (2014)
- Kleinberg, J.M.: Authoritative sources in a hyperlinked environment. *J. ACM (JACM)* **46**(5), 604–632 (1999)
- Li, L.; Shang, Y.; Zhang, W.: Improvement of HITS-based algorithms on web documents. In *Proceedings of the eleventh international conference on World Wide Web - WWW '02*, page 527, New York, New York, USA, 2002. ACM Press.
- Yang, M.-C.; Lee, J.-T.; Lee, S.-W.; Rim, H.-C.: Finding Interesting Posts in Twitter Based on Retweet Graph Analysis. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '12*, page 1073, New York, New York, USA, 2012. ACM Press.
- Abilhoa, W.D.; de Castro, L.N.: A keyword extraction method from twitter messages represented as graphs. *Appl. Math. Comput.* **240**, 308–325 (2014)
- Yang, C.; Harkreader, R.; Zhang, J.; Shin, S.; Gu, G.: Analyzing spammers' social networks for fun and profit. In *Proceedings of the 21st international conference on World Wide Web - WWW '12*, pages 71–80, New York, New York, USA, 2012. ACM Press.
- Yang, M.-C.; Rim, H.-C.: Identifying interesting Twitter contents using topical analysis. *Expert Syst. Appl.* **41**(9), 4330–4336 (2014)
- Mocanu, D.; Rossi, L.; Zhang, Q.; Karsai, M.; Quattrociochi, W.: Collective attention in the age of (mis)information. *Comput. Hum. Behav.* **51**, 1198–1204 (2015)
- Kwon, S.; Cha, M.; Jung, K.; Chen, W.; Wang, Y.: Prominent Features of Rumor Propagation in Online Social Media. In *2013 IEEE 13th International Conference on Data Mining*, pages 1103–1108. IEEE, 2013.
- Nguyen, N. P.; Yan, G.; Thai, M. T.; Eidenbenz, S.: Containment of Misinformation Spread in Online Social Networks. In *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*, pages 213–222, New York, New York, USA, 2012. ACM Press.
- Bird, S.; Klein, E.; Loper, E.: Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc., 2009.
- Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É.: Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2012)

38. Language Technology Group at the University of Oslo. NLPL word embeddings repository, 2018.
39. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E.: Hierarchical Attention Networks for Document Classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 1480–1489, Stroudsburg, PA, USA, 2016. Association for Computational Linguistics.
40. Sahin, G.: Turkish document classification based on Word2Vec and SVM classifier. In 2017 25th Signal Processing and Communications Applications Conference (SIU), pages 1–4. IEEE (2017)
41. Elsaadawy, A.; Torki, M.; Ei-Makky, N.: A Text Classifier Using Weighted Average Word Embedding. In 2018 International Japan-Africa Conference on Electronics, Communications and Computations (JAC-ECC), pages 151–154. IEEE, 2018.
42. Bilgin, M.: Kelime Vektörü Yöntemlerinin Model Oluşturma Sürelerinin Karşılaştırılması. *Bilişim Teknolojileri Dergisi*, pages 141–146, 2019.
43. Karcioğlu, A. A.; Aydin, T.: Sentiment Analysis of Turkish and English Twitter Feeds Using Word2Vec Model. In 2019 27th Signal Processing and Communications Applications Conference (SIU), pages 1–4. IEEE (2019)
44. Lilleberg, J.; Zhu, Y.; Zhang, Y.: Support vector machines and Word2vec for text classification with semantic features. In 2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), pages 136–140. IEEE, 2015.
45. Goodfellow, I.; Bengio, Y.; Courville, A.: Deep learning. The MIT Press, Cambridge, MA (2017)
46. Abualigah, L.M.; Khader, A.T.; Hanandeh, E.S.: A combination of objective functions and hybrid Krill herd algorithm for text document clustering analysis. *Eng. Appl. Artif. Intell.* **73**, 111–125 (2018)
47. Abualigah, L.M.; Khader, A.T.; Hanandeh, E.S.: A new feature selection method to improve the document clustering using particle swarm optimization algorithm. *J. Comput. Sci.* **25**, 456–466 (2018)
48. Lee, D.D.; Sebastian Seung, H.: Learning the parts of objects by non-negative matrix factorization. *Nature* **410**(6755), 788–791 (1999)
49. Fisher, R.A.: The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**(2), 179–188 (1936)
50. Alpaydin, E.: Machine Learning: The New AI. The MIT Press, Cambridge, MA (2016)
51. Faizollahzadeh Ardabili, S.; Najafi, B.; Shamshirband, S.; Minaei Bidgoli, B.; Deo, R.C.; Chau, K.W.: Computational intelligence approach for modeling hydrogen production: a review. *Eng. Appl. Comput. Fluid Mech.* **12**(1), 438–458 (2018)
52. Taylor, V. F.; Spolaor, R.; Conti, M.; Martinovic, I.: AppScanner: Automatic Fingerprinting of Smartphone Apps from Encrypted Network Traffic. In 2016 IEEE European Symposium on Security and Privacy (EuroSP), pages 439–454. IEEE (2016)
53. Vapnik, V.: The Nature of Statistical Learning Theory. Springer (1995)
54. Basakin, E.E.; Ekmekcioglu, O.; Ozger, M.: Drought Analysis with Machine Learning Methods. *Pamukkale Univ. J. Eng. Sci.* **25**(8), 985–991 (2019)
55. Leo B.: Random Forests. In Machine Learning, chapter 45, pages 5–32. Springer, 2001.
56. Christopher O.: Understanding LSTM Networks, 2015.
57. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y.: Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1724–1734, Stroudsburg, PA, USA, 2014. Association for Computational Linguistics.
58. Shamshirband, S.; Rabczuk, T.; Chau, K.-W.: A Survey of Deep Learning Techniques: Application in Wind and Solar Energy Resources. *IEEE Access* **7**, 164650–164666 (2019)
59. Aceto, G.; Ciunzio, D.; Montieri, A.; Pescapé, A.: Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges. *IEEE Trans. Netw. Serv. Manage.* **16**(2), 445–458 (2019)
60. Fan, Y.; Kang, X.; Hui, W.; Zheng, Y.; Tao, B.: Spatiotemporal modeling for nonlinear distributed thermal processes based on KL decomposition MLP and LSTM network. *IEEE Access* **8**, 25111–25121 (2020)
61. Schuster, M.; Paliwal, K.K.: Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **45**(11), 2673–2681 (1997)
62. TensorFlow. The Sequential model- Keras, 2019.
63. Montieri, A.; Ciunzio, D.; Bovenzi, G.; Persico, V.; Pescapé, A.: A dive into the dark web: Hierarchical traffic classification of anonymity tools. *IEEE Trans. Netw. Sci. Eng.* **7**(3), 1043–1054 (2020)
64. Banan, A.; Nasiri, A.; Taheri-Garavand, A.: Deep learning-based appearance features extraction for automated carp species identification. *Aquacult. Eng.* **89**, 102053 (2020)
65. Wu, C.L.; Chau, K.W.: Prediction of rainfall time series using modular soft computing methods. *Eng. Appl. Artif. Intell.* **26**(3), 997–1007 (2013)
66. Taormina, R.; Chau, K.-W.: ANN-based interval forecasting of streamflow discharges using the LUBE method and MOFIPS. *Eng. Appl. Artif. Intell.* **45**, 429–440 (2015)
67. Ali, S. and Karacan, H.: Akan Veri Kümeleme Teknikleri Üzerine Bir Derleme. *European Journal of Science and Technology*, pages 17–30, 2018.
68. Galan-Garcia, P.; Puerta, J.G.D.L.; Gomez, C.L.; Santos, I.; Bringas, P.G.: Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying. *Logic J. IGPL* **24**(1), 42–53 (2016)
69. Kadry, S. and Al-Taie, M. Z.: Başlık: Social Network Analysis : An Introduction with an Extensive Implementation to a Large-scale Online Network Using Pajek. eBook Collection (EBSCOhost), 2014.
70. Gephi. Gephi-open source graph visualization software, 2020.
71. Peters, M.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L.: Deep Contextualized Word Representations. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pages 2227–2237, Stroudsburg, PA, USA, 2018. Association for Computational Linguistics.
72. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North, pages 4171–4186, Stroudsburg, PA, USA, 2019. Association for Computational Linguistics.
73. Brown, T. B.; Benjamin, M.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Ariel H.-V.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D. M.; Wu, J.; Winter, C.; Hesse, C.; Chen, M.; Sigler, E.; Litwin, M.; Gray, S.; Chess, B.; Clark, J.; Berner, C.; McCandlish, S.; Radford, A.; Sutskever, I. and Amodei, D.: Language Models are Few-Shot Learners, 2020.
74. Hadeer, A.: Detecting Opinion Spam and Fake News Using N-gram Analysis and Semantic Similarity. Msc. thesis, University of Ahram Canadian, 2017.
75. Granik, M.; Mesyura, V.: Fake news detection using naive Bayes classifier. In 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pages 900–903. IEEE, (2017)
76. Eugenio, T.; Gabriele, B.; Della Vedova, M. L.; Moret, S. and de Alfaro, L.: Some Like it Hoax: Automated Fake News Detec-

- tion in Social Networks. In *Proceedings of the Second Workshop on Data Science for Social Good*, Skopje, Macedonia, 2017.
77. Rubin, V.; Conroy, N.; Chen, Y.; Cornwell, S.: Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. In *Proceedings of the Second Workshop on Computational Approaches to Deception Detection*, pages 7–17, Stroudsburg, PA, USA, 2016. Association for Computational Linguistics.
 78. Pérez-Rosas, V.; Kleinberg, B.; Lefevre, A.; Mihalcea, R.: Automatic Detection of Fake News. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3391–3401 (2017)
 79. Samir, B.: “The Pope Has a New Baby!” Fake News Detection Using Deep Learning, 2017.
 80. Miraj, P.: Detection of Maliciously Authored News Articles. Msc. thesis, The Cooper Union For The Advancement of Science and Art, 2017.
 81. Ågren, A. and Ågren, C.: Combating Fake News with Stance Detection using Recurrent Neural Networks. Msc. thesis, University of Gothenburg, 2018.
 82. Rajendran, G.; Chitturi, B.; Poornachandran, P.: Stance-In-Depth Deep Neural Approach to Stance Classification. *Procedia Comput. Sci.* **132**, 1646–1653 (2018)
 83. Bhatt, G.; Sharma, A.; Sharma, S.; Nagpal, A.; Raman, B.; Mittal, A.: On the Benefit of Combining Neural, Statistical and External Features for Fake News Identification (2017)
 84. Ruchansky, N.; Seo, S.; Liu, Y.: CSI: A Hybrid Deep Model for Fake News Detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 797–806, New York, NY, USA, 2017. ACM.
 85. Singhania, S.; Fernandez, N.; Rao, S.: 3HAN: A Deep Neural Network for Fake News Detection. In *24th International Conference on Neural Information Processing (ICONIP 2017)*, pages 572–581, 2017.
 86. Volkova, S.; Shaffer, K.; Jang, J. Y.; Hodas, N.: Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 647–653, Stroudsburg, PA, USA, 2017. Association for Computational Linguistics.
 87. Wang, W. Y.: “Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 422–426, Stroudsburg, PA, USA, 2017. Association for Computational Linguistics.
 88. Girgis, S.; Amer, E.; Gadallah, M.: Deep Learning Algorithms for Detecting Fake News in Online Text. In *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, pages 93–97. IEE (2018)
 89. Fang, Y.; Gao, J.; Huang, C.; Peng, H.; Runpu, W.: Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection. *PLoS ONE* **14**(9), e0222713 (2019)