Research Article

The Use of LSTM Neural Network to Detect Fake News on Persian Twitter

Mohammad Mohsen Sadr [a]*, Afshin Mousavi Chalak [b], Soraya Ziaei [c] and Jafar Tanha [d]

[a] Faculty member of Payame Noor University, Tehran, Iran (ORCID: 0000-0003-0241-8163)
[b] Faculty member of Payame Noor University, Tehran, Iran (ORCID: 0000-0002-5614-1334)
[c] Faculty member of Payame Noor University, Tehran, Iran (ORCID: 0000-0003-3134-1468)
[d] Faculty member of Tabriz University, Tabriz, Iran (ORCID: 0000-0002-0779-6027)

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract: The spread of new Internet-based technologies has caused various news and events to be widely disseminated on various networks and quickly made available to the public. This has diminished the boundaries between news production and information creation and sharing in online media and the social media environment. The spread of fake news is a serious problem that has been created by the open nature of the web and social media, resulting in an increasing trend of fake news creation and publication. Meanwhile, the volume of fake news produced is very large and is adjusted in a way that easily deceives the audience. Machine learning and neural networks algorithms are commonly used to prevent the spread of fake news. In this study, a hybrid model of long short term memory (LSTM) and 14-layer bidirectional long short term memory (BLSTM) neural network has been used to identify the fake news on Persian texts and tweets. Based on the obtained results, the proposed model has the ability to identify fake news and rumors with 91.08% accuracy. According to the confusion matrix, it has been determined that the performance capability of the proposed model is 92.05%, its recall is 91.10%, and its f1 criterion is 91.57%. We further compare the results to Bayesian, k-NN, random forest, linear regression, perceptron neural network, SVM, decision tree, probabilistic gradient, Adaboost, gradient boost, and extra tree. The results show that the proposed approach outperforms the other algorithms.

Keywords: Fake news, rumors, artificial intelligence, deep learning, BLSTM neural network, Persian tweets

1. Introduction

The boundaries between news production, information creation, and sharing have been blurred in online media and social media environments (Chen et al. 2015). Fake news is a serious problem that is compounded by the open nature of the web and social media. Recent advances in new computer technology have simplified the process of creating and publishing fake news. But in the meantime, because the volume of published news is very large and is hand-picked quickly, the amount of fake and false news produced is very large and is adjusted in such a way that easily deceives the audience. In fact, glamorous news is part of the realm of information that attracts audiences. In 2016, the issue of fake news attracted a lot of attention, especially after the last US presidential election (Ahmed 2012). Due to the widespread production of fake news worldwide, the term "Fake News" was recognized as the word of year by macquuire dictionary in 2016.

In the 2020 perspective, big data is an important factor in the evolution of our understanding of the world. But there is also the risk of misinformation or so-called loss distribution (Kumar et al. 2019). This is a case of fake news. There are various definitions for fake news, but perhaps the definition of Ashida and Koraya (2018) is one of the most general definitions that stated "fake news is intentionally wrong and can mislead readers." Many news sites or individuals use compelling words to headline their news to engage the audience. Using this method provides more clicks for sites or channels to make more money. Fake news can upset the balance of the news ecosystem, and usually the popularity of fake news, especially in cyberspace, is much higher than real news (Shu et al. 2017). Fake news is deliberately used for political, financial and propaganda purposes to deceive the audience. But there are many challenges to detect fake news, because fake news is deliberately produced and attempts to mislead the audience. Detecting fake news based on content is mostly difficult (Kai et al. 2017). Fake news detection strategies are based on several factors. For example, this issue can be examined both from the content dimension of the news and from the social dimension of the news to ensure the authenticity of the news (Shu et al. 2017).

In some cases, some people or media outlets that broadcast fake news will continue to design fake pages similar to the main pages or fake addresses for sites so that the audience can better believe this fake news by visiting those sites. The biggest challenge is the lack of an efficient way to tell the difference between real and fake news, as even humans are often unable to express differences (Ahmed 2012). At present, the audience or, in a way, the customers of the companies care about the news published by them and are attracted to them; this shows that the audience is very impressed with the news, even fake news. According to statistics published by Kumar et al. (2019), 88% of customers rely on personal surveys and 72% of them blindly trust the company and news published by its name, which is why the spread of fake news under advertising titles has intensified to attract audiences and customers. On the other hand, this kind of news or propaganda may be used to tarnish the name of companies, the media, or individuals and politicians. Another important point is the ability to post comments in online news. Therefore, some companies and stakeholders try to better introduce themselves and create an unreal positive stream by hiring people to publish unrealistic positive comments on the news to convince customers. Others also try to tarnish the company or person by hiring someone to share negative comments on posts. In 2013, for example, the BBC published a report on Taiwanese authorities' investigation against Samsung for hiring students to post negative comments on the HTC website. According to what has been said, some believe that one of the biggest sources of fake news or rumors is social media websites, such as Google Plus, Facebook, Twitters, and so on. In the meantime, it can be stated that the only way to identify fake news is to gain extensive knowledge about the topics covered (Ahmed 2012). Social networks are also very popular among Iranians. Telegram, Instagram, WhatsApp, Facebook and Twitter have a special place among Iranian users, and Persian-speaking users use these networks more (TechRasa 2017; Azarafza et al. 2020). Discovering fake news is a complex and much more difficult task than identifying fake comments about a product and its spread on social media.

In this article, rumors have been extracted from the collection of rumors in Persian tweets related to the "DataHeart" database in combination with the collected data. The DataHeart database contains 650 healthy tweets and 645 tweets containing rumors. To improve the comparison and performance of DataHeart database, it has been upgraded to 7650 healthy tweets and 6370 tweets containing fake rumors with the data collected in this study. Due to the fact that a database suitable for rumors on Persian tweets as the main source of data is generally very limited due to the lack of data, news tweets and user comments on rumors and real news collected during this study can be very effective in improving the implementation process. In this article, a hybrid model of long short term memory (LSTM) and 14-layer bidirectional long short term memory (BLSTM) is used to identify fake news in Persian tweets. After pre-processing and deleting phrases and additional words and specifying the training and test data, the input text is evaluated to distinguish fake news from real. Persian words with Word2vec model receive basic training as Skip-gram. In the following, the features in Persian sentences will be converted to vector. Then different layers of LSTM and BLSTM are applied to improve the distinction of rumor from non-rumor and fake news from real news. Finally, with a single full connection, the network output is announced. According to the comparisons and results obtained, the proposed model has the ability to identify fake news and rumors with 91.08% accuracy. Also, according to the confusion matrix, it has been determined that the performance capability of the proposed model is 92.05%, its recall is 91.10%, and its f1 criterion is 91.57%. In order to measure the performance and validity of the model, the results obtained by artificial intelligence-based approaches, including one versus rest, Bayesian, k-NN, random forest, linear regression, perceptron neural network, SVM, decision tree, probabilistic gradient, Adaboost, gradient boost and extra tree, have been evaluated. The results of the performance appraisal assessment show the success of the proposed model.

2. Fake news classification

Jindal and Liu (2008) categorized fake content into three groups of wrong reviews (whose main purpose is to provide incorrect information about the product or increase its reputation or damage it), brand reviews (which target the brand, but do not share experiences with a particular product), and incorrect ads (which are invisible and do not contain text related to the product). According to these categories, appropriate solutions are predicted and presented. In general, fake news can be divided into three main categories of fake news (news that is completely fake and made by writers), humor news (fake news whose main purpose is to provide humor to readers) and weak news report (news that have a certain degree of real news, but they are not completely accurate and definitive) (Jindal and Liu 2008). For example, quotes of political figures are often used to report a completely fake story, which is designed to promote a particular poll or bigotry. In fact, a system for identifying fake news aims to help users identify and filter out all kinds of misleading news (Chen et al. 2015).

Research on fake news is a new topic, and many researchers around the world are currently working to improve ways to detect it. It is one of the new branches in LIS and NLP natural language processing that seeks to identify and label examples of deliberately misleading news that can be conveyed to the audience (Chen et al. 2015). In detecting fake news, it is generally sought to discover lies and false news and increase man's ability to discover false news. Using NLP, a variety of unrealistic and fake data can be collected, some of the most important of which include fake product reviews (Mukherjee et al. 2013), online disaster detection (Guillory and Hancock 2012), fake comments (Jindal and Liu 2008), fake social media profiles (Kumar et al. 2019), fake dating profiles (Toma and Hancock 2012), spam and phishing (Toolan and Carthy 2010), and fake scientific work (Meel and Vishwakarma 2019). Humans are usually not very good at detecting lies and fake news, so using a machine is beneficial (Rubin and Conroy 2012). Rubin and Conroy (2012) state the requirements for detecting fake news as follows:

- Access to honest and deceptive cases: The predicted methods should be able to find patterns and theories in the positive and negative points of the data (in this case, the counterparts of valid and reliable news is the main challenge);
- Access to digital text format: NLP prefers text and images are accompanied by text (in this case, audio and video must be sent properly);

- Trust the foundation of truth: In the texts, the reliability of the news base must always be evident. Reliable expert resources that have been tested over time and are based on the "check and balance" system are recommended as a basis for this;
- Homogeneity in length: For example, a short tweet with a title, a summary of a Facebook paragraph, and a long formula do not form a homogeneous data set;
- Homogeneity in writing: This set should be compared using new genres (e.g. news, editorials, and files) and topics (such as business, politics, science, and health), different types of authors (e.g. professional training, mainstream to citizens), journalists and in the news media.

According to the above statements, the news can be generally placed in the following categories (Chen et al. 2015):

- General News: Including an honest report published in digital journalism and newspapers or radio or blogs of reputable journalists, who have a lot of credibility as reputable citizens. This news is usually valid, unless proven otherwise. Predicting the correctness of a news story may be different from genres and topics, and so from real news to deceptive news should be kept in classifications;
- Allusive news: In political events, territorial disputes, wars or other current disputes, news channels or journalists may be accused of bias, blindness or lying. This news has its own sources to confirm;
- Unofficial news: Misleading news that can be picked up by someone or imitated by some researchers. So this news group is divided into three groups of serious fake news, large-scale scam, and humor.

3. Research Background

Ott et al. (2011) used the n-gram frequency model to identify fake ideas. They created the standard "golden" data by collecting deceptive comments and honest opinions from Amazon hotels' travel consultants. They divided all opinions (fake and truthful) into two categories of positive and negative, and then obtained 86% accuracy using SVM classification. When they removed the positive and negative separation, the model's accuracy was reduced from 86% to 84%, indicating that data separation in both the negative and positive categories improved performance. The researchers also used human opinions to judge fake or honest opinions, and the best statistic was 65%, indicating that machine performance was better (Ott et al. 2011).

By studying social media, Marchi (2012) examined adolescents' behaviors and attitudes toward various news items. By interviewing 61 high school students from different high schools, the researchers evaluated and prioritized the preference for specific news formats. The results of his study suggested a change in ways of accessing news information, new semantic approaches to information, and young people's preference for positive rather than objective news. This does not mean that young people ignore the basic ideals of professional journalism, but rather indicate that they are more realistic (Marchi 2012).

Shojaee et al. (2013) created a model based on stylometry for review classification. Using the standard golden data set created by Ott et al., the researchers extracted 234 stylometry features, which were divided into lexical and syntactic features and classified for SVM and Newbies network modeling. In this regard, they first used F-measure calculations to test the lexical and syntactic properties separately and then a combination of both. SVM went beyond combined or detached features from the Newbies network. However, the highest measured F score was 84% using both lexical and syntactic features (Shojaee et al. 2013).

Jiwei et al. (2014) used linguistic methods to categorize and identify fake news. These researchers used Linguistic Inquiry and Word Count (LIWC) and part of speech (POS) along with n-gram for fake news or comments detection. LIWC is software used to identify cognitive and emotional characteristics in speech. Using this method, words are categorized according to concepts. For example, "crying" falls into the category of grief and eventually leads to the production of an internal dictionary (Jiwei Li et al. 2014).

Tanhan et al. (2014) presented a boosting algorithm for multi-class semi-monitoring learning called Multi-SemiAdaBoost, which aims to minimize marginal cost on tagged data and minimize inconsistency on tagged and non-tagged data. This algorithm uses a loss exponential function for semi-monitored learning and can reinforce any type of basic fire class. An algorithm has been developed for semi-monitored multiclass learning. Most methods use one versus rest to convert multi-class problems to multiple binary classification problems (Tanha et al. 2014).

Liu (2015) focused on building its classifier using uncontrolled techniques to address the unavailability of tagged data and situations where outstanding features are not available. They developed a method for calculating the review factor by estimating the semantic overlap of content between games (Liu 2015).

Conroy et al. (2015) provided a comprehensive study on approaches to adopting and detecting fake news. They described the discovery of fake news as a major task in classifying news in terms of the accuracy of the news and the source, and by measuring the reliability of the source and the news, he identified the accuracy of the deliberate deception and the scale of its impact and risk (Conroy et al. 2015).

Khaldarova and Pantti (2016) presented an article assessing the role of news and television in the crisis in Ukraine and Russia, and assessed the role of fake news in the audience's mental and emotional deviation from social realities. They stated that television is a powerful weapon of Russia in emotional destruction and the government's main asset in the information warfare, which has shown distorted information in line with the Kremlin's narrative by presenting counterfeit stories (Khaldarova and Pantti 2016).

Rubin et al. (2016) studied the effect of deception and fake news called Satire, which causes "illusion of reality" in readers by creating false feelings of truth and is an inductive feeling. The researchers said that the general view of satire, interpretation and illustration is one of the most important features of emotional deception that can be used in civil society, science, business, soft news (illusion of reality) and .... In this regard, they used SVM-based classification and proposed that the algorithm used is based on 5 main features, including absurdity, humor, grammar, negative impact and writing. They tried to predict and test fake news. The algorithm has been implemented on 360 articles. The results of these researchers showed that the algorithm has an accuracy of 90%, recall of 84, and F1 criteria of 87% (Rubin et al. 2015).

Assessing the impact of the 2016 US presidential election on voting and the upcoming election in the United Kingdom, Kucharski (2016) focused on different news and market trends of British users and the emotional atmosphere of this election. The researcher said that information published on social media, especially the social and economic aspects, was a major lever in political victories (Kucharski 2016).

Allcott and Gentzkow (2017) reviewed news and comments on social media related to the 2016 US presidential election, examining changes and effects related to the publication of fake stories and false news in this regard. The researchers conducted a comprehensive online assessment and created a large database of online news websites that identified the four most important features, including "very important" news sources, findings, and repetitions. In the three months since the election, 14% of Americans followed very important news directly from Facebook and Twitter, of which 30 million times Trump and 8 million times Clinton have been repeated and reposted on Facebook. By evaluating fake or semi-fake news, it has been found that 12% of people believe this news to some extent and considered the candidacy of these people to be ideal (Allcott and Gentzkow 2017).

Wang (2017) conducted a study on the automatic discovery of fake news as a challenging problem in detecting deception and political and social effects in the real world. He has been severely restricted by using statistical approaches to counter fake news due to a lack of tagged data sets. In his paper, he presents a new and publicly available data set to detect fake news (Wang 2017).

Lazer et al. (2018) assessed the impact of the spread of fake news in various fields on the Internet, describing it as a global concern that can affect various social, economic, cultural and industrial aspects and play a deviant role. The researchers said that a new system of protection measures is needed (Lazer et al. 2018).

In an article, Bakir and McStay (2018) assessed the interactions between the 2016 US presidential election and the country's economic situation. The researchers examined the effects of fake news on the country's electoral conditions and the country's social and political climate, and considered the scientific, social, journalistic, media and public aspects. They stated that fake news and the phenomenon of fake news and public deception can have negative effects on all dimensions (Bakir and McStay 2018).

Clayton et al. (2019) studied the impact of fake news on the 2016 US presidential election on social media, noting the extent of this impact for Facebook. Demonstrating the Democrats' deliberate move to change the electoral situation, the researchers assessed the creation of fake stories and social deception. Based on the results of these researchers, the use of false news is always accompanied by special titles with a label "disputable" or "false" (Clayton et al. 2019).

Reis et al. (2019) stated that a large part of recent work has focused on understanding and discovering fake news published on social media. To achieve this goal, the present work explored several types of news-derived features, including social media resources and posts. In addition to examining the main features presented in the literature to detect fake news, the researchers used monitored algorithms such as SVM to detect fake news. Introducing a new collection in this field, they showed that the results of the research are useful in recognizing the importance of the features of false news discovery (Reis et al. 2019).

Monti et al. (2019) used geometric deep learning (GDL) to identify and extract fake news on social networks and cyberspace. Using natural language processing (NLP) and deep convolution neural network (CNN), the researchers identified general and counterfeit deception news. The accuracy of the algorithm was 92.7% (Monti et al. 2019).

4. Theoretical Framework

Rumors can be thought of as an active form of possible communication. Rumor is the expression of concerns and stress of some people about information fraud. Rumors do not convince anyone, they say something that the public is willing to believe. It's easily entered into the news and is generated with the help of ambiguity. The rumor can be confirmed by the authorities; In this case, it will become news and if it is denied, it will remain a rumor. In some cases, even when the news is denied by informed authorities, it may still remain a rumor, which happens when people do not trust the media. In the social sciences culture, written by Alan Bieru, rumors are defined as follows: "Rumors in French, like the Latin language, convey a news story, an unfounded commotion, a form of public opinion, and even some kind of fame. Sociologically, the rumor phenomenon is the process by which news is broadcasted without going through the usual channels. The rumor may be based on incorrect information, or the source of the information may be correct but exaggerated and misleading. The news may be passed from one person to another, from one group to another without a clear source or definite reasons for its accuracy" (Lippmann 2016).

People and governments have long been involved in conspiracies of rumors. The history of rumor dates back to the time of the Roman emperors and ancient Greece. As the city of Rome was set on fire in 12 AD, rumors spread among the people that the emperor Nero was the perpetrator, but Nero took the initiative to defend himself and attributed it to Christians by spreading an anti-rumor. This rumor spread well because Christians did not have a good reputation with the people. In such a way that people temporarily forgot their hatred and enmity towards Nero. In the late nineteenth century, while examining the Paris riots, Gustave Le Bon came up with a theory. He believed that riots were the result of group thinking harmony and that they were contagious. In the early twentieth century, with the use of scientific methods to study group behavior, "rumor" in the form of group behavior has been the subject of study and research by sociologists, psychologists, psychiatrists and anthropologists and researchers of popular culture. Some believed that creating rumors during this period depended on the entry of the press and radio into society; because during this period, due to the increase in social consciousness and political awareness of the people, they want to provide accurate and timely information about various events by these all-encompassing and widespread media. Therefore, if these media do not disseminate information in a timely manner or their information is contradictory, they will inadvertently provoke more rumors (Lippmann 2016). The proliferation and diversity of rumors in different situations indicates that rumors are made and addressed with specific attitudes and motivations. The following are some of the most important motivations for processing and spreading rumors (Bernays and Ewen 2011):

- Attracting social attention: Since one of the characteristics of a rumor is its special importance for a particular group, and therefore the ground for its dissemination is provided, it should be acknowledged that the rumor carrier also feels that he himself is an important and noteworthy person since his existence is the source of such important news. So one way to get people's attention in different places is exaggerated feeling of self-importance through gossip. Many people with low self-esteem and lack of social skills use this method to focus on themselves.

- Projection: Whenever an individual's emotional state is reflected in his or her interpretation of the environment without his or her knowledge, we call this state "projection." He has failed to use purposeful and completely neutral evidence to describe the facts around him. These people subconsciously express or disseminate their desires, inner tendencies, hopes, and aspirations in the form of "rumors" in the third-person quote: "They say," "They said," and "Someone said." That is, they want to show an interpretation of reality and in accordance with our private lives, which will be our tendency to believe and spread it.

5. Research methodology

In the proposed model, the first step is related to the use of Word2vec to increase and improve accuracy. The Word2vec algorithm is a group of text processing-related models used to generate embedding words. These models are neural networks used to teach the reconstruction of linguistic concepts. The Word2vec algorithm, as an input, takes up a large portion of the text and produces a vector space, typically a few hundred dimensions, with a unique word in the text body that is assigned to a corresponding vector in space. Using the Word2vec algorithm can be helpful in improving accuracy due to the semantic similarity, closeness of comments, the meaning of words and comments in tweets and sentences to distinguish rumors from non-rumors and fake news from non-fake. So in the first step, a pre-trained model for Persian is used. Word2vec uses both model structures to generate a distributed display of words: the bag of words model or N-gram. In the bag of continuous words, the model predicts the present word from a window of textual words around it. The order of the words in the text does not affect the prediction. In continuous skip architecture, this model uses the current word to predict the window around text words. In this architecture, the weight of these words is far greater than the word text distance. According to research, CBOW is faster, while skip-gram is slower, but it does better for incorrect words. In this study, the skip-gram method was used, reducing the speed and increasing the calculations; however, more precision can be expected, so more accuracy can be sacrificed to the calculations. In the proposed model, Word2vec is used on the input after the described pre-processing so that there is a better concept of input for fake news. After using the pre-trained model, they are applied as input to the proposed 14-layer model. It is a 14-layer model, consisting of 2 input layers, an embedded layer, 3 random deletion layers, an integration layer used to connect the two sides of the network, a LSTM layer, 2 bidirectional LSTM layers, 2 dense layers, and the last layer, which is a fully connected layer used to distinguish fake news from non-fake news, as shown in Fig. 1.

Layer (type)	Output Shape	Param #	Connected to
input_2 (InputLayer)	(None, 34)	0	
embedding_1 (Embedding)	(None, 34, 300)	2100	input_2[0][0]
input_1 (InputLayer)	(None, 34)	0	
dropout_2 (Dropout)	(None, 34, 300)	0	embedding_1[0][0]
dropout_1 (Dropout)	(None, 34)	0	input_1[0][0]
lstm_1 (LSTM)	(None, 34, 256)	570368	dropout_2[0][0]
dense_1 (Dense)	(None, 300)	10500	dropout_1[0][0]
time_distributed_1 (TimeDistrib	(None, 34, 300)	77100	lstm_1[0][0]
add_1 (Add)	(None, 34, 300)	0	dense_1[0][0] time_distributed_1[0][0]
dropout_3 (Dropout)	(None, 34, 300)	0	add_1[0][0]
bidirectional_1 (Bidirectional)	(None, 34, 256)	439296	dropout_3[0][0]
bidirectional_2 (Bidirectional)	(None, 512)	1050624	bidirectional_1[0][0]
dense_3 (Dense)	(None, 1028)	527364	bidirectional_2[0][0]
dense_4 (Dense)	(None, 7)	7203	dense_3[0][0]

Figure 1. Proposed model layering

After pre-processing and deleting phrases, additional words and specifying the training and test data, the input text is evaluated to distinguish fake news from real news. Persian words receive basic training as Skip-gram with Word2vec model. In the following, the features in the news, tweets and Persian sentences will become a vector. Then, different layers of LSTM and bidirectional LSTM are applied to improve the detection of rumor from non-rumor and fake news from real news. Finally, with a complete connection layer, the network output is announced. Fig. 2 shows the layout and input and output values of the proposed model. It should be noted that due to the high calculations of Word2vec model as n-gram and the use of Google Colab hardware that has limited memory (12 GB of RAM), there is a memory problem. Fig. 3 shows the error rate of the proposed model after 300 iterations. However, the error has been decreasing reduced until the 150 iteration. After the 200 iteration, the error decreases, but not like the previous iterations. However, the increase continues continuously, and model error is fixed in approximately 0.2, which is a very good value.

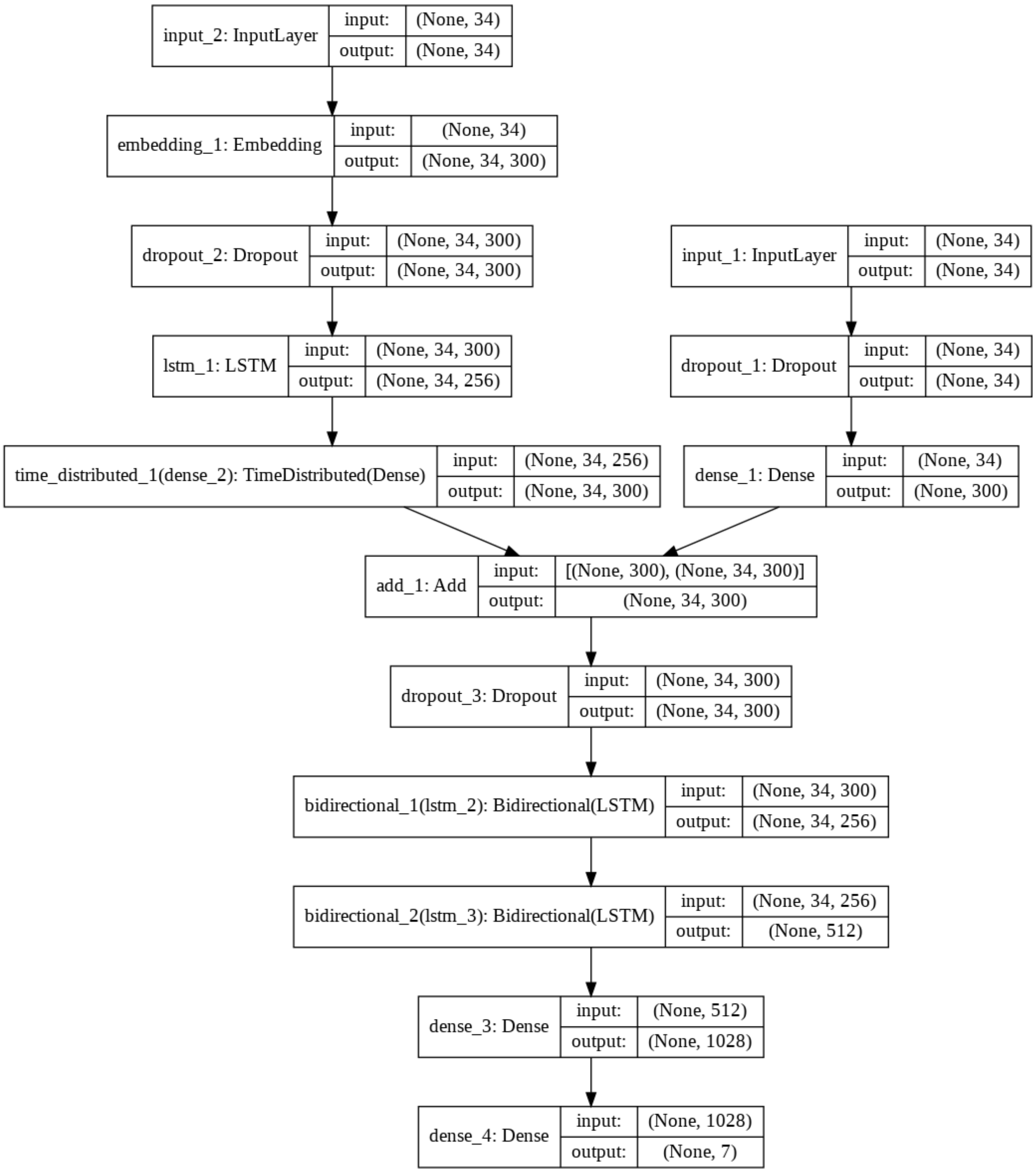

Figure 2. Layout and values of the proposed 14-layer model

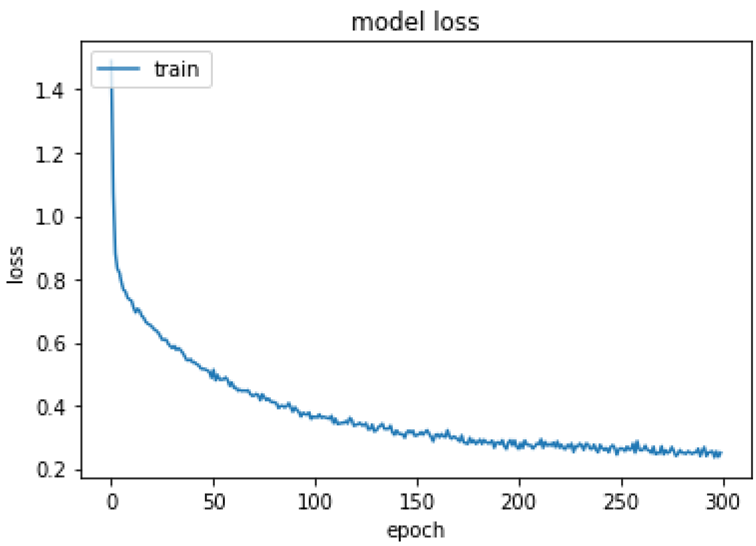

Figure 3. Tthe error rate of the proposed model

6. Research findings

In the proposed model, the whole network is taught using the necessary functions and layers. Accuracy is a criterion that expresses what percentage of this data is properly categorized. The confusion matrix shows how the categorization algorithm works according to the input data set by different categories of categorization problems. The relationships are described below.

$$\text{Accuracy} = \frac{TP_i + TN}{TP_i + FN + TN + FP_i}$$

$$\text{Precision} = \frac{TP_i}{TP_i + FP_i}$$

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

In order to measure the performance and validity of the model, the results obtained by artificial intelligence-based approaches, including one versus rest, Bayesian, k-NN, random forest, linear regression, perceptron neural network, SVM, decision tree, probabilistic gradient, Adaboost, gradient boost and extra tree, have been evaluated. The results of the performance appraisal assessment show the success of the proposed model. Accuracy can be expressed as the most important factor in the analysis of classifications that is the degree of proximity of a measurement to the actual value. In other words, it shows at what distance a set of measurements is closest to the actual value. Figs. 4 and 5 as well as Table 1 show the results of standard evaluation and validation. Based on these figures, it can be stated that the proposed method in the non-conversion mode for comparative algorithms has reached an accuracy of 91.08 for detecting fake news, which is almost one percent higher than the nearest algorithm (decision tree). Other algorithms have done almost the same thing, but the Adaboost algorithm has the lowest value among the 13 algorithms with a value of 21.916. Also, the proposed method in the state without converting words to vectors for comparative algorithms is more accurate than all methods. The closest to the proposed method is related to the basic tree algorithms, decision tree, and random forest.

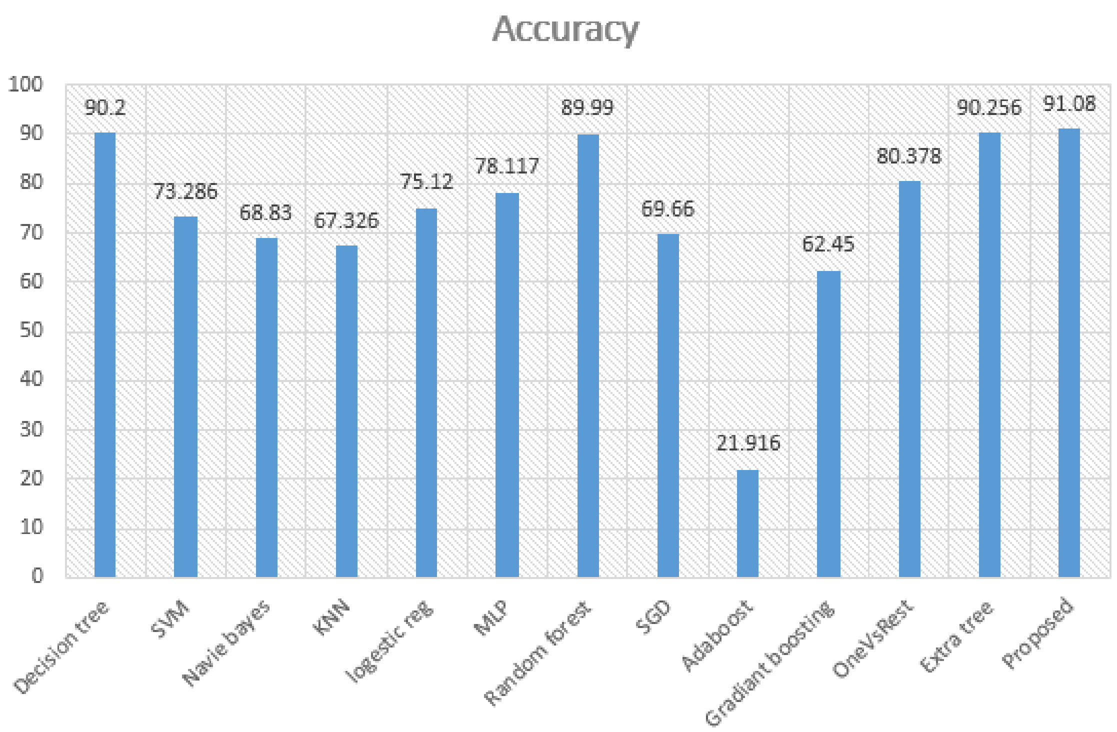

Figure 4. Comparison of the accuracy value of the proposed method with the other twelve algorithms in the state without word conversion

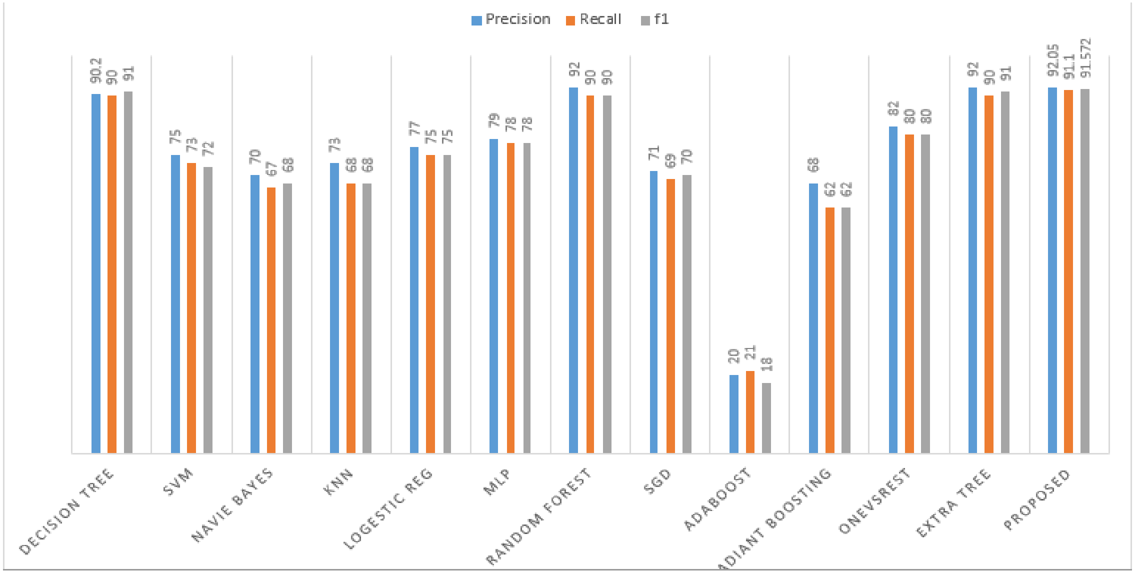

Figure 5. Comparison of the proposed model classification evaluation criteria with the other twelve algorithms

Table 1. Analysis results by different algorithms.

Algorithm	Precision	recall	Accuracy	F1-score	Error function
Proposed model	92.05	91.10	91.08	91.57	0.2373
One versus rest	82	80	80.378	80	-
Naïve Bayesian	70	67	68.183	68	-
KNN	73	68	67.326	68	-
Random forest	92	90	89.99	90	-
Linear regression	77	75	75.120	75	-
Perceptron	79	78	78.117	78	-
SVM	75	73	73.286	73	-
Decision tree	92	90	90.2	90	-
probabilistic gradient	71	69	69.66	70	-
Adaboost	20	21	21.916	18	-
Gradient boost	68	62	62.45	62	-
Extra tree	92	90	91	90.256	91

7. Discussion and Conclusion

The results of the evaluation of the proposed method by other methods are as follows: in the proposed method of LSTM and bidirectional LSTM network with word to vector conversion pre-processing, accuracy is 0.9205, recall is 0.911, and F1 criterion is 0.9108. In comparative mode without converting words to vectors for comparative algorithms, the linear regression algorithm has an accuracy value of 0.77, recall value of 0.75, and F1 value of 0.7512. In the Perceptron neural network algorithm, the value of the accuracy is 0.79, the value of the recall is 0.78, and the value of the F1 criterion is 0.7811. In the Naïve Bayesian categorization method, the value of accuracy is 0.70, the value of recall is 0.67, and the value of F1 criterion is 0.6818. In the case of random forest, the accuracy is 0.92, recall is 0.90, and the value of F1 is 0.899. In the KNN algorithm, the value of the accuracy is 0.73, recall is 0.68, and the value of F1 is 0.6732. In the SVM method, the value of the accuracy is 0.75, recall is 0.73, and the value of F1 is 0.7328. In the decision tree method, the accuracy is 0.92, recall is 0.90, and the value of F1 is 0.902. For classification algorithm of one versus rest, the value of accuracy is 0.82, recall is 0.80, and the value of F1 is 0.8037. For the probabilistic gradient classification algorithm, the value of the accuracy is 0.71, recall is 0.69, and the value of F1 is 0.6966. For the Adaboost classification algorithm, the value of the accuracy is 0.20, recall is 0.20, and the value of F1 is 0.2191. For Gradient Boosting classification algorithm, the value of accuracy is 0.68, recall is 0.62, and the value of F1 is 0.6245. For extra tree classification algorithm, the value of accuracy is 0.92, recall is 0.90, and the value of F1 is 0.9025.

Funding: no

Conflict of interest: The authors declare that they have no conflict of interest.

References

1. Ahmed H (2012) Detecting Opinion Spam and Fake News Using N-gram Analysis and Semantic Similarity. Dissertation, Department of Electrical and Computer Engineering, University of Ahram Canadian
2. Allcott H, Gentzkow M (2017) Social Media and Fake News in the 2016 Election J Econ Perspect 31(2):211-236
3. Azarafza M, Feizi-Derakhshi MR, Bagheri Shendi M (2020) TextRank-based Microblogs Keyword Extraction Method for Persian Language. In: Proceedings of the 3rd International Congress on Science and Engineering, Hamburg, Germany
4. Bakir V, McStay A (2017) Fake News and The Economy of Emotions. Digit Journal 6(2):154-175
5. Bernays E, Ewen S (2011) Crystallizing Public Opinion. Ig Publishing
6. Chen Y, Conroy NJ, Rubin VL (2015) News in an Online World: The Need for an “Automatic Crap Detector”. In: The Proceedings of the Association for Information Science and Technology Annual Meeting (ASIST2015), Nov. 6-10, St. Louis
7. Clayton K, Blair S, Busam JA, Forstner S, Glance J, Greem G, Bright A, Welch AT, Wolff AG, Zhou A, Nyhan B (2019) Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media. Polit Behav. https://doi.org/10.1007/s11109-019-09533-0
8. Conroy NJ, Rubin VL, Chen Y (2015) Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology 52(1):1-4
9. Guillory J, Hancock JT (2012) The Effect of Linkedin on Deception in Resumes. Cyberpsychol Behav Soc Netw 15(3):135-140
10. Jindal N, Liu B (2008) Opinion Spam and Analysis. ACM press
11. Jiwei Li MO, Cardie C, Hovy E (2014) Towards a general rule for identifying deceptive opinion spam. In: Proceedings of the 52nd AnnualMeeting of the Association for Computational Linguistics
12. Kai S, Sliva A, Wang S, Tang J, Liu H (2017) Fake news detection on social media: A data mining perspective. SIGKDD Explor Newsl 19:22-36
13. Khaldarova I, Pantti M (2016) Fake News. Journalism Practice 10(7):891-901
14. Kucharski A (2016) Study epidemiology of fake news. Nature 540:525
15. Kumar S, Asthana R, Upadhyay S, Upreti N, Akbar M (2019) Fake news detection using deep learning models: A novel approach. Trans Emerg Telecom Technol. https://doi.org/10.1002/ett.3767
16. Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Miriam JM (2018) The science of fake news. Science 359(6380):1094-1096
17. Lippmann W (2016) Public Opinion. CreateSpace Independent Publishing Platform
18. Liu B (2015) Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press
19. Marchi R (2012) With Facebook, Blogs, and Fake News, Teens Reject Journalistic “Objectivity”. J Commun Inq 36(3):246-262
20. Meel P, Vishwakarma D (2019) Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst Appl 2019:112986
21. Monti F, Frasca F, Eynard D, Mannion D, Bronstein MM (2019) Fake News Detection on Social Media using Geometric Deep Learning. Soc Inform Networks arXiv:1902.06673
22. Mukherjee A, Venkataraman L, Glance A (2013) Fake Review Detection: Classification & Analysis of Real & Pseudo Reviews. Technical Report, Dept. of CompSci, University of Illinois, Chicago & Google Inc
23. Ott M, Choi Y, Cardie C, Hancock JT (2011) Finding Deceptive Opinion Spam by Any Stretch of the Imagination. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, P11-1032, 309-319
24. Reis JCS, Correia A, Murai F, Veloso A, Benevenuto F (2019) Supervised Learning for Fake News Detection. IEEE Intell Syst 34(2):76-81
25. Rubin VL, Conroy N (2012) Discerning truth from deception: Human judgments & automation efforts. First Monday 17(3). https://doi.org/10.5210/fm.v17i3.3933
26. Rubin VL, Conroy N, Chen Y (2015) Towards News Verification: Deception Detection Methods for News Discourse. The Rapid Screening Technologies. Deception Detection and Credibility Assessment Symposium, Hawaii International Conference on System Sciences (HICSS48)
27. Shojaee S, Murad MAA, Azman AB, Sharef NM, Nadali S (2013) Detecting deceptive reviews using lexical and syntactic features. In: 13th International Conference on Intellient Systems Design and Applications. https://doi.org/10.1109/ISDA.2013.6920707

28. Shu K, Sliva A, Wang S, Tang J, Liu H (2017) Fake News Detection on Social Media: A Data Mining Perspective. Soc Inform Networks arXiv:1708.01967
29. Tanha J, van Someren M, Afsarmanesh H (2014) Boosting for multiclass semi-super- vised learning. Pattern Recognit Lett 37:63–77
30. TechRasa (2017) Infographic: Telegram Usage Statistics in Iran. Iran Startup & Tech Media. http://bit.do/techrasa20170906/
31. Toma CL, Hancock JT (2012) What Lies Beneath: The Linguistic Traces of Deception in Online Dating Profiles. J Commun 62(1):78-97
32. Toolan F, Carthy J (2010) Feature Selection for Spam and Phishing Detection. E-Crime Researchers Summit, Dallas, 1-12
33. Wang WY (2017) Liar, liar pants on fire: A new benchmark dataset for fake news detection. CoRR abs/1705.00648.