Deep learning algorithms for detecting Fake News in Online Text

Sherry Girgis
Faculty of Computer Science
Modern Academy for Computer Science and Management Technology
Cairo, Egypt
sherrygirgis91@gmail.com

Eslam Amer
Faculty of Computer Science
Misr International University
Cairo, Egypt
eslam.amer@miuegypt.edu.eg

Mahmoud Gadallah
Faculty of Computer Science
Modern Academy for Computer Science and Management Technology
Cairo, Egypt
mgadallah1956@gmail.com

***Abstract*— Spreading of fake news is a social phenomenon that is pervasive at the social level between individuals, and also through social media such as Facebook and Twitter. Fake news that we are interested in is one of many kinds of deception in social media, but it's more important one as it is created with dishonest intention to mislead people. We are concerned about this issue because we have noticed that this phenomenon has recently caused through the means of social communication to change the course of society and peoples and also their views, for example, during revolutions in some Arab countries have emerged some false news that led to the absence of truth and stirs up public opinion and also fake of news is one of the factors Trump successes in the presidential election. So we decided to face and reduce this phenomenon, which is still the main factor to choose most of our decisions. Techniques of fake news detection varied, ingenious, and often exciting. In this paper our objective is to build a classifier that can predict whether a piece of news is fake or not based only its content, thereby approaching the problem from a purely deep learning perspective by RNN technique models (vanilla, GRU) and LSTMs. We will show the difference and analysis of results by applying them to the dataset that we used called LAIR. We found that the results are close, but the GRU is the best of our results that reached (0.217) followed by LSTM (0.2166) and finally comes vanilla (0.215). Due to these results, we will seek to increase accuracy by applying a hybrid model between the GRU and CNN techniques on the same data set.**

Keywords— Deception detection; Deep Learning; Artificial Intelligence; RNN (Recurrent Neural Network); LSTM (long short-term memories) ; Vanilla ; GRU (Gated Recurrent Unit) ; CNN(Convolutional Neural Networks).

Keywords— Deception detection; Deep Learning; Artificial Intelligence; RNN (Recurrent Neural Network); LSTM (long short-term memories); Vanilla; GRU (Gated Recurrent Unit); CNN(Convolutional Neural Networks).

I. INTRODUCTION

Social media for news consumption, such as Facebook and Twitter is a double-edged sword [1]. On the one hand, low cost and easy access to information and dissemination quickly to push people to search for news and know what is happening at the beginning of events with details and updates at the moment unlike newspapers or magazines in the old days, on the other hand, it enables the widespread of 'fake news' because of its accessibility and lack of cost and control of the Internet. Recent reports suggest that the outcome of the U.S. Presidential Elections is due to the rise of online fake news.

Reports indicate that the human ability to detect deception without special assistance is only 54%. So we need to use machine learning for classifying texts automatically [2].

Fake news detection is considered one of the most dangerous types of deception because it recently caused deceiving many people, Fake news defined as the prediction of the chances of a particular news article (news report, editorial, expose, etc.) being intentionally deceptive (Rubin, Conroy & Chen, 2015) [3].

We concerned about the fake news because of the problem of fake news detection is more challenging than detecting deceptive reviews. A recent report by the Jump-shot Tech Blog1found that Facebook referrals accounted for 50% of the total traffic to fake news sites and 20% of the total traffic to reputable websites [4]. Since the majority of U.S. Adults – 62%– gets news on social media (Jeffrey and Elisa, 2016) [5] so that the ability to identify fake content in online sources is, therefore, an urgent need.

Our system establishes for detecting fake news by using deep learning technique that shows an improvement than linguistic cues recently, we use RNN models (vanilla, and GRU) and LSTM technique with the LAIR dataset (discuss it in details in next sections).

The organization of the paper is as follows: The next section we will talk about the fake news problem in general and changelings that faced us and the researchers when trying to reach the best result, then will presenting the previous findings of the researchers in this field to classify news passages as fake or not, next we will mention our experiments in details that we reached with the dataset that we use, and it's preparation, next we will mention our view of the steps of the system and proposed our model, Then will clarify our results and analysis of each result, Finally will presenting conclusions of all sections.

II. PROBLEM DEFINITION

Fake news becomes a major issue for the public and government. Fake news can take advantage of multimedia content to mislead readers and get published, which can lead to negative effects or even manipulation of public events. One of the unique challenges of detecting fake news on social media is how to identify fake news about recent events.

The task of detecting fake news has tested a variety of labels, from misinformation to rumors; to spam. There has been a large body of work surrounding text analysis of fake news and similar topics such as rumors or spam. We have tried to mention some papers that interested in this subject also and each paper will recall it's experience and results and is based on its own concept of these words deceptive [6].
There are two directions to detect text deception. The first of which is based on separate handheld features, which can capture linguistic and psychological causes, however, these features failed to classify text well, which limits performance. The second approach based on a neural network model learns document-level representation to discover deception text.

Neural network models have been used to learn semantic representations for NLP tasks and it reaches the highest competitive result According to (Le and Mikolov, 2014; Tang et al., 2015) [7] and also in NLP, fake data have been collected by crawling the web or crowdsourcing: fake product reviews (Mukherjee, Venkataraman, Liu, & Glance, 2013) [8].

Most researchers used many deep learning popular algorithms such as CNN, Bidirectional-LSTM, and RNN that we will mention in detail later, but datasets was a reason that why the fake news detection was not successful in the past because it was small and included unrealistic news, In 2017 William created a new benchmark dataset called LAIR which Collected 12.8K short data is labeled in different Contacts from POLITIFACT.COM, which Provides detailed analysis report and links to source documents for each case. William works on LAIR dataset with many techniques such as logistic regression, support vector machines, and (Bidirectional-LSTM and CNN) models for deep learning and CNN results is the best [9].

Our work in this paper will implement RNN models (Vanilla RNN, GRU) and LSTMs on LAIR dataset to determine if the news is truthful or deceptive and will show the results compared to William's results with analysis.

The following part, we represent some of the previous practical neural network models and their datasets to detect deception on a text.

III. RELATED WORK

Deception detection (in the framework of computational linguistics) is a text classification problem where our system should classify an unseen document as either truthful or deceptive. Such a system is first trained on known instances of deception. Used features are token unigrams and linguistic cues derived from the classes of words from the Linguistic Inquiry and Word Count (LIWC) [10]. This was a psychological experiment with the analysis of the participants' writings, focusing on the connection between deception and fantasy proneness. The task of detection that was performing is thus opinion spam detection or fake news detection, which is a more variant of deception detection.

In the past in deception detection mostly relied on manual feature selection based on, for example, psycholinguistic theories of deception and/or computational linguistics, followed by supervised machine learning to build a classifier [11], but recent NLP researches are now increasingly focusing on the use of new deep learning methods shown in Fig.1 [12].

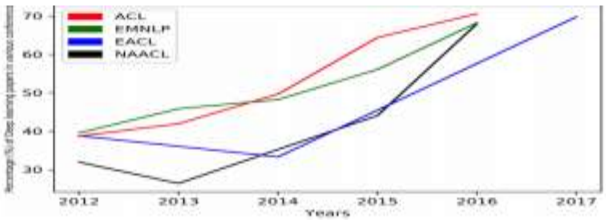

Fig. 1. Percentage of deep learning papers in ACL, EMNLP, EACL, NAACL over the last 6 years (long papers)

Here we mention some of the previous works used many techniques for fake news detection.

(Niall J. Conroy, 2015) [4] Using linguistic cues approaches and network analysis approaches to design a basic fake news detector which provides high accuracy in terms of classification tasks. They propose a hybrid system whose features like multi-layer linguistic processing, the addition of network behavior, they propose a method to detect online deceptive test by using a logistic regression classifier which is based on POS tags extracted from a corpus deceptive and truthful texts and achieves an accuracy of 72% which could be further improved by performing cross-corpus analysis of classification models and reducing the size of the input feature vector ,it's one of the best results that these features can reach so as to improve the results the next researchers have to use neural networks.

(Samir Bajaj, 2017)[13] used many techniques for neural networks and machine learning to determine which algorithm reach the best result among of them, he apply these algorithms to his dataset which collect it two different sources; from an open Kaggle dataset sold 13,000 Fake news articles and 50,000 authentic news articles (negative examples for the classifier) were extracted from the Signal Media News dataset. Split all of these data (63000 articles) into 60% training, 20% dev/validation, and 20% test sets. He used many techniques from machine learning and neural network such as Logistic Regression, Feedforward Network, RNN (Vanilla, GRU), LSTMs, Bi- LSTMs, CNN with Max-Pooling and CNN with Max-Pooling and Attention, The results proved that GRU gets the better F1 score and best results overall.

Other research by (Natali Ruchansky et. Al, 2017) [14], for getting better result more than previous one; she proposed a model called CSI, it is built up from deep neural networks which can extract information from different domains and capture temporal dependencies in user engagement with articles, and also select important features. CSI (which is composed of three modules: Capture, Score, and Integrate) evades the cost of manual features election by incorporating neural networks. The features they use to capture the temporal behavior and textual content in a general way that does not depend on the data context nor requires distributional assumptions. They created this technique to solve three main problems in fake text, first of them evaluate the matching score for the headlines and body of an article, second the emotion that reach to readers from an article and how make them feel, the last characteristic is knowing the source of the article by checking the structure of the URL, to the credibility of the media source. They used two datasets from Twitter and Weibo (real-world social media datasets), CSI gives the best performance overall comparison models and versions. We see that integrating user features boosts the overall numbers up to 4.3% from GRU-2. Put together, these results demonstrate that CSI successfully captures and leverages all three characteristics of text, response, and source, for accurately classifying fake news (Shown in table 1).

Other Works by (William Yang et.al, 2017) [12] Due to lack of fake news datasets and lack of their efficiency, he decided to present a new benchmark dataset called LIAR (will mention it in details in next sections): it's a new, publicly available dataset for fake news detection. Use it with many techniques such as logistic regression, support vector machines, and (Bidirectional-LSTM and CNN) models for deep learning. The results proved that CNN (Convolutional Neural Networks) models are the best.

As we showed that William's works did not recover using RNN technique, although GRU offered the best results, therefore we use this technique in LAIR dataset and will compare our works with his results and give a brief analysis.

TABLE I. COMPARISON OF DETECTION ACCURACY ON TWO DATASETS

	Twitter		**WEIBO**	
	Accuracy	***F-score***	***Accuracy***	***F-score***
DT-RANK	0.624	0.636	0.732	0.726
DTC	0.711	0.702	0.831	0.831
SVM-TS	0.767	0.773	0.857	0.861
LSTM-1	0.814	0.808	0.896	0.913
GRU-2	0.835	0.830	0.910	0.914
CI	0.847	0.846	0.928	0.927
CI-T	0.854	0.848	0.939	0.940
CSI	0.892	8.894	0.953	0.954

IV. EXPERMINENTS

A. Dataset

There are very useful datasets to study fake detection but the positive training data are collected from a tested out (in a way that was close to the real thing) (surrounding conditions). More importantly, these datasets are not good for fake statements detection; since the fake news on TVs and social media are much shorter than customer reviews.

According to (William Yang et.al, 2017) presents a new benchmark dataset called LIAR: it's a new, publicly available dataset for fake news detection. Collected a decade-long, 12.8K manually labeled short statements in various contexts from POLITIFACT.COM, Which provides a detailed analytical report and a link to its source level for each case. This dataset can be used for fact-checking research as well. They investigate the automatic detection of fake news based on surface-level linguistic patterns [12].

The LIAR dataset includes 12,836 short statements labeled for truthfulness, subject, context/venue, speaker, state, party, and prior history. With this size and time span of ten years, LIAR cases are collected in more natural context, such as political debate, TV ads, Facebook posts, tweets, interview, news release, etc. In each case, the labeler provides a lengthy analysis report to the ground each judgment [12].

They have evaluated several popular learning based methods on this dataset. The baselines include logistic regression, support vector machines, LSTM and the CNN model. We will present some examples of LIAR dataset in Fig.2 [12].

Statement: "Newly Elected Republican Senators Sign Pledge to Eliminate Food Stamp Program in 2015." **Speaker:** Facebook posts **Context:** social media posting **Label:** Pants on Fire **Justification:** More than 115,000 social media users passed along a story headlined, "Newly Elected Republican Senators Sign Pledge to Eliminate Food Stamp Program in 2015." But they failed to do due diligence and were snookered since the story came from a publication that bills itself (quietly) as a "satirical, parody website." We rate the claim Pants on Fire.	**Statement:** "Under the health care law, everybody will have lower rates, better quality care, and better access." **Speaker:** Nancy Pelosi **Context:** on 'Meet the Press' **Label:** False **Justification:** Even the study that Pelosi's staff cited as the source of that statement suggested that some people would pay more for health insurance. Analysis at the state level found the same thing. The general understanding of the word "everybody" is every person. The predictions don't back that up. We rule this statement False.

Fig.2. depicts some example from LIAR dataset.

B. *Data Preparation*

After we have LIAR data set as we mention in the previous section, we must preprocess it to be suitable for our system that can work with. Preprocessing means that data set is clearer to our algorithm by removing dummy characters, string, and Impurities. Preprocessing Works on three steps :

- Splitting: Separate each sentence from the next sentences to deal with them individually.
- Stop words removing: remove UN-important words from each sentence.
- Stemming: Returns each word to its origin.

C. *Deception System*

Due to the deception phenomenon that spread nowadays through the traditional media and the social media platform especially the Facebook and Twitter, which controls the users' selections in relation to the economic part such as (buying products, business, etc…) also in relation to the political life, so we decided to create s system which solves this problem not only fake reviews but we concerned about fake news by using preprocessing LAIR data which has been prepared based on what we have cleared in details in the previous part we apply it to Word embedding (or word vector) that gives each word a vector and each vector represents a latent feature of a word , then the result of the word vector apply to RNN models (Vanilla, GRU) and LSTM , then we will get the results that determine that a piece of news is deceptive or not.

We will explain in details each step of our system in the next section.

V. PROPOSED MODEL

Our work clarified in the following steps as follows:

First step: Preparing LIAR dataset in four levels:

- The first level is splitting each sentence to deal with separately.
- The second level is removing stop words and that includes identifying the useless words in each statement like (the, a, an, etc).
- The third level is stemming which every word return to its infinitive.

Second step: Output of stemming will be the input to word embedding which played an important role in deep learning based on deception analysis that includes representing each single word in each sentence by dimensional vector and get the relation between two words not only syntactic but also the same (as ‘see’ and ‘watch’ are very different in syntactic, but their meaning is somewhat related) [15]. Another benefit is that the algorithm detects the words that appear mostly together (like ‘wear’ and ‘clothes’) and it shows their relationship and then this is able to predict the next word [16].

Third step: Results of word embedding level will be the input to the RNN models (vanilla, GRU) and LSTMs technique.

Fourth step: the output of step four will Getting final result determining if the piece of news is truthful or deceptive. As is common in data mining problems, once the models are built, the process might be repeated with new data and new features.

VI. RESULTS AND ANALYSIS

We constructed three different experiments.

- Vanilla: Is the first model of RNN that used from 1980, and it’s just Single Layer Network (with feedback).
- GRU (Gated Recurrent Unit). researchers used to use it (from 2014) because of avoiding vanilla issues which filter the information flow to enable the modeling of long-term dependencies
- LSTM (long short-term memories) Behave like RNNs, but LSTMs have a different function of computing the hidden state by introducing input, forget, and output gate mechanism and an additional memory cell state (can store information for a longer time).

In the first experiment we used Vanilla RNN, Then we used GRU model in the second experiment, finally, in the last experiment we used the LSTM technique. The following table illustrates the accuracy of each model.

TABLE II. COMPARISON OF OUR RESULTS AND WILLIAM'S ACCURACY

Model	**Test Accuracy**
SVMs	0.255
Logistic Regress0ion	0.247
Bi-LSTMs	0.233
CNN	0.270
Vanilla	0.215
GRU	0.217
LSTM	0.2166

We compared our results with (William Yang et.al, 2017) in the previous table II.

We found the worst result of our experiments gets from vanilla because of its failure to solve complex tasks that have a practical application, It also changed the format of the original information, which meant that it was unable to hold the important memory content for more than a few time steps, And also Gradient vanishing Is one of its disadvantages.

LSTM also showed inefficiency compared to GRU and CNN because its two main drawbacks, first it is more expensive to calculate the network output and apply back propagation. we simply have more maths to do because of the complex activation. However this is not as important as the second point, second, the explicit memory adds several more weights to each node, all of which must be trained. This increases the dimensionality of the problem and potentially makes it harder to find an optimal solution.

The best result of our experiments is GRU, we have seen a slight improvement in its results than vanilla and LSTM because of solving gradient vanishing problem which is a problem in vanilla and it is easy to modify and doesn't need memory units, therefore, faster to train than LSTM and give as per performance.

As we compared our results with William Yang's result we found that CNN is the best among all the results as CNN tend to be much faster (~5 times faster) than RNN and more efficient depends on our implementation and because of Nvidia has historically focused much more on CNN than RNN, as computer vision mostly employs CNN.

VII. CONCLUSION

In recent years, deception detection in online reviews & fake news has an important role in business, law enforcement, national security, political due to the potential impact fake reviews can have on consumer behavior and purchasing decisions. Researchers used deep learning with the large dataset to increase in learning and thus get the best results by using word embedding for extract features or cues that distinguish relations between words in syntactic and semantic. In this paper we cover implementation of RNN technique models (Vanilla, GRU) and LSTMs that have been proposed for the detection of online fake news after we prepare our LAIR dataset applying to prepare data to word embedding to get vectors of words then entering this vectors to our deep learning technique, we found that the results of our experiments are close but GRU(Gated Recurrent Unit) is the best because it's solving the problems of Vanilla that popular of gradient vanishing problem and LSTMs (long short-term memories) which GRU is easy to modify and doesn't need memory units , so, we gain faster training than LSTM that effect in our performance ,but for comparing our results with William's results we found that CNN's(Convolutional Neural Networks) is the best from other models due to its speed and its best results and performance for windows .our future works will increase this accuracy by merging GRU and CNN's to get the best result .

REFERENCES

[1] Erlich, Aaron, et al. "The double-edged sword of mobilizing citizens via mobile phone in developing countries." Development Engineering 3, 2018, 34-46.

[2] Krishnamurthy, Gangeshwar, et al. "A deep learning approach for multimodal deception detection." arXiv preprint arXiv: 1803.00344, 2018.

[3] Conroy, Niall J., Victoria L. Rubin, and Yimin Chen. "Automatic deception detection: Methods for finding fake news." Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact:

[4] Pérez-Rosas, Verónica, et al. "Automatic Detection of Fake News." arXiv preprint arXiv:1708.07104 ,2017.

[5] Allcott, Hunt, and Matthew Gentzkow. "Social media and fake news in the 2016 election." Journal of Economic Perspectives 31.2 ,2017, 211-36.

[6] Kumar S, Shah N." False Information on Web and Social Media: A Survey". arXiv preprint arXiv: 1804.08559, 2018 Apr 23.

[7] Lopez MM, Kalita J." Deep Learning applied to NLP". arXiv preprint arXiv:1703.03091,2017 Mar 9.

[8] Rubin, Victoria L., Yimin Chen, and Niall J. Conroy. "Deception detection for news: three types of fakes." Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community. American Society for Information Science, 2015.

[9] Wang WY." " liar, liar pants on fire": A new benchmark dataset for fake news detection". arXiv preprint arXiv:1705.00648. 2017, May 1

[10] Verhoeven, Ben, and Walter Daelemans. "CLiPS Stylometry Investigation (CSI) corpus: a Dutch corpus for the detection of age, gender, personality, sentiment and deception in text." In LREC 2014-NINTH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, pp. 3081-3085,2014.

[11] Hagiwara, Masato. "A supervised learning approach to automatic synonym identification based on distributional features." In Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Student Research Workshop, pp. 1-6. Association for Computational Linguistics, 2008

[12] K Young T, Hazarika D, Poria S, Cambria E. "Recent trends in deep learning based natural language processing." arXiv preprint arXiv:1708.02709. 2017 Aug 9.

[13] Bajaj S. "The Pope Has a New Baby! Fake News Detection Using Deep Learning".2017 .

[14] Ruchansky N, Seo S, Liu Y. Csi." A hybrid deep model for fake news detection". InProceedings of the 2017 ACM on Conference on Information and Knowledge Management ,2017 .Nov 6. pp. 797-806, ACM.

[15] Lepping, Joachim. "Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.",2018.

[16] Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. "Distributed representations of words and phrases and their compositionality. InAdvances in neural information processing systems", 2013 ,pp. 3111-3119.

[17] www.kdnuggets.com.

[18] www.wikipedia .com