IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Study on Analyzing Fake news through Neural Network Models

1Ms.Mrunali Sanjay Amate, 2Prof.T.N.Ghorsad

1Student, 2Professor

1Sant Gadge Baba Amravati University,

2Sant Gadge Baba Amravati University

Abstract

Fake news is defined as a made-up story with an intention to deceive or to mislead. In this paper we present the solution to the task of fake news detection by using Deep Learning architectures. Gartner research [1] predicts that "By 2022, most people in mature economies will consume more false information than true information". The exponential increase in production and distribution of inaccurate news presents an immediate need for automatically tagging and detecting such twisted news articles. However, automated detection of fake news is a hard task to accomplish as it requires the model to understand nuances in natural language. Moreover, majority of the existing fake news detection models treat the problem at hand as a binary classification task, which limits model's ability to understand how related or unrelated the reported news is when compared to the real news. To address these gaps, we present neural network architecture to accurately predict the stance between a given pair of headline and article body. Our model outperforms existing model architectures by 2.5% and we are able to achieve an accuracy of 94.21% on test data.

Keywords: Deep Learning, Natural Language, Detection Models, predict, binary

1. Introduction

"Fake News" is a term used to represent fabricated news or propaganda comprising misinformation communicated through traditional media channels like print, and television as well as non-traditional media channels like social media. The general motive to spread such news is to mislead the readers, damage reputation of any entity, or to gain from sensationalism. It is seen as one of the greatest threats to democracy, free debate, and the Western order [3].

Fake news is increasingly being shared via social media platforms like Twitter and

Facebook [2]. These platforms offer a setting for the general population to share their opinions and views in a raw and un-edited fashion. Some news articles hosted or shared on the social media platforms have more views compared to direct views from the media outlets' platform. Research that studied the velocity of fake news concluded that tweets containing false information reach people on Twitter six times faster than truthful tweets [3]. The adverse effects of inaccurate news range from making people believe that Hillary Clinton had an alien baby, trying to convince readers that President Trump is trying to abolish first amendment to mob killings in India due to a false rumor propagated in WhatsApp. Technologies such as Artificial Intelligence (AI) and Natural Language Processing (NLP) tools offer great promise for researchers to build systems which could automatically detect fake news. However, detecting fake news is a challenging task to accomplish as it requires models to summarize the news and compare it to the actual news in order to classify it as fake. Moreover, the task of comparing proposed news with the original news itself is a daunting task as its highly subjective and opinionated. A different way to detect fake news is through stance detection which will be the focus of our study. Stance Detection is the process of automatically detecting the relationship between two pieces of text. In this study, we explore ways to predict the stance, given a news article and news headline pair.

Depending on how similar the news article content and headlines are, the stances between them can be defined as 'agree', 'disagree', 'discuss' or 'unrelated'. We experimented with several traditional machine learning models to set a baseline and then compare results to the state-of-the- art deep networks to classify the stance between article body and headline.

Considering these difficulties involved in detection of fake news, a good first step is to detect the stance between the body of text and the entity it's describing. The task of stance detection can be described as the process of automatically predicting if the news article or social media content is agreeing, disagreeing or unrelated to the entity it's describing.

Misusing the Data *"Have a Beer, It's Good for Your Brain,"* reported Inc. But you should wait a minute before you grab a pint (or two). The study was done on mice — not people. And the amount of beer was the equivalent of 28 kegs in humans.

Imprecise and Sloppy *"1 in 5 CEOs are Psychopaths, Study Finds."* But the headline is wrong. The research was based on a survey of professionals in the supply chain industry, not CEOs.

The above examples capture the complex nature of detection and classification of fake news. To rightly classify the above types of fake news, our language model needs to understand the subtleties involved in conveying messages through text.

Detecting fake news is hard for many reasons. First, manual task of identifying fake news is very subjective. Assessing the veracity of a news story is a complex and cumbersome task, even for trained experts.

News is not only spread through traditional media outlets anymore but also through various social media channels. Automated solution requires understanding the natural language processing which is difficult and complex. These complexities make it a daunting task to classify text as fake news.

1.1 Dataset Description: Fake News Challenge (FNC-1) Data

Fake News Challenge [7] opened to the public on December 1, 2016 as a competition. The data used for this competition were derived from the Emergent Dataset created by Craig Silverman. Emergent Research [8] is a research and consulting firm focused on the most dynamic sector of the global economy

Dataset includes body of the news article, the headline of the news article, and the label for relatedness (stance) of an article and headline. The data set is split into train, validation and test splits based on the methodologies outlined in Section 4. As an initial stage towards building an automated solution for identifying fake news, FNC-I targeted to solve the stance detection problem. The aim of solving stance detection is to determine the relatedness of news article and headline of the news article. Distribution across 4 stances is presented in table 1.

Table 1. Stance labels in FNC-1 training dataset

Stance category	Percentage	Description
Agree	7.36%	Headline agrees with the claim made in the news article
Disagree	1.68%	Headline disagrees with the claim made in the news article
Discuss	17.82%	Headline discusses same topic as news article
Unrelated	73.13%	Headline does not discuss same topic as news article

1.2 Introduction to Neural Network Architectures:

Given the significant breakthrough in neural network research, we used various versions of deep neural nets detailed in further sections. In this section we discuss different versions of neural network architectures used in common research.

1.2.1 Dense Neural Network (DNN)

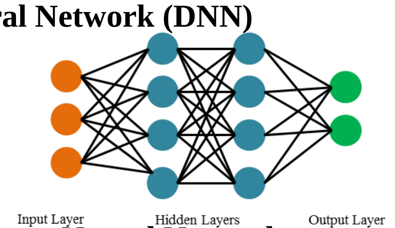


Fig. 1 Dense Neural Network

The fully connected dense neural network allows us to pass the input as sequence of words. The layered architecture allows us to experiment with the right depth that is needed for our task. The network consists of an input layer, an output layer and can consist of series of hidden layers.

1.2.2 Convolution Neural Network (CNN)

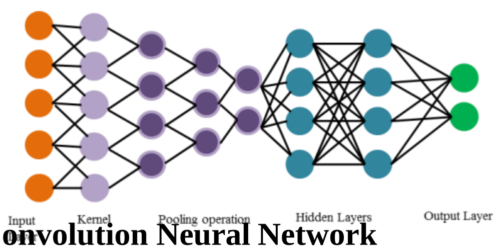


Fig. 2 Convolution Neural Network

These are very similar to ordinary Neural Networks: they are made up of neurons that have learnable weights and biases. A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers

1.2.3 Recurrent Neural Network (RNN)

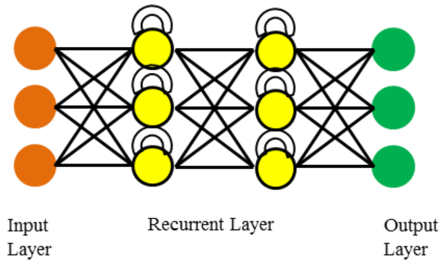


Fig.3 Recurrent Neural Network

Its popular with sequential data as each unit can have memory about the state of previous unit. This is particularly useful in natural language processing because it helps gain deeper understanding of language. RNN's have an input layer, output layer and can have a number of hidden recurrent units which have memory gates.

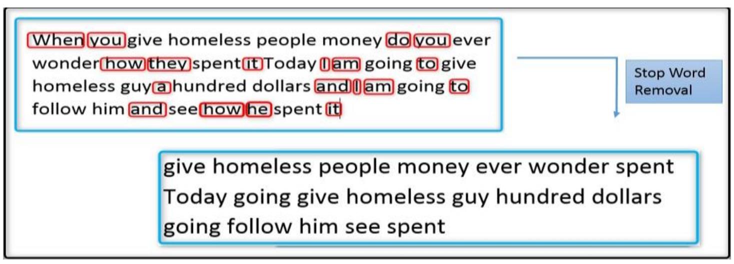
1.3 Data Preprocessing

Text data requires special preprocessing to implement machine learning or deep learning algorithms on them. There are various techniques widely used to convert text data into a form that is ready for modeling. The data preprocessing steps that we outline below are applied to both the headlines and the news articles. We also provide insights into different word vectors representations we used as part of our analysis.

1.3.1 Stop Word Removal

We start with removing stop words from the text data available. Stops Words (most common words in a language which do not provide much context) can be processed and filtered from the text as they are more common and hold less useful information. Stop words acts more like a connecting part of the sentences, for example, conjunctions like "and", "or" and "but", prepositions like "of", "in", "from", "to", etc. and the articles "a", "an", and "the". Such stop words which are of less importance may take up valuable processing time, and hence removing stop words as a part of data preprocessing is a key first step in natural language processing. We used Natural Language Toolkit – (NLTK) library to remove stop word.

Fig. 4. Example for Stop Word removal



1.3.2 Punctuation Removal

Punctuation in natural language provides the grammatical context to the sentence. Punctuations such as a comma, might not add much value in understanding the meaning of the sentence. Figure 5 shows an example of Punctuation removal process

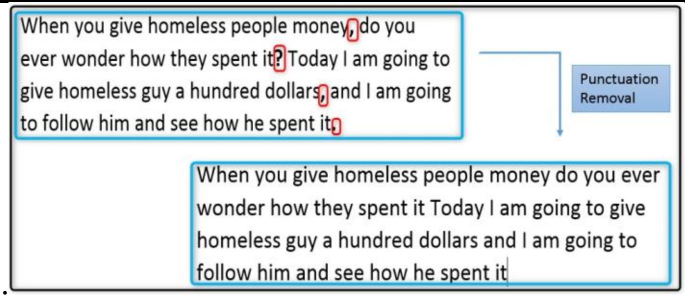


Fig. 5 . Example for punctuation removal

1.3.3 Stemming

Stemming is a technique to remove prefixes and suffixes from a word, ending up with the stem. Using stemming we can reduce inflectional forms and sometimes derivationally related forms of a word to a common base form. Figure 6 shows the example of stemming technique.

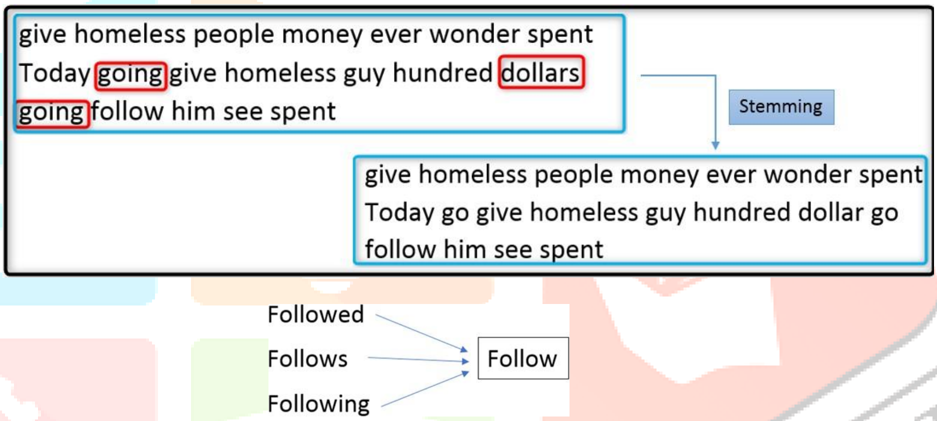


Fig. 6. Example Stemming

1.3.4 Word Vector Representation

Preparing the text from the body and headline of the news article for modeling is quite challenging. To perform text analytics, we need to convert raw text into numerical features. We experimented with two techniques to transform the raw text and feature extraction: Bag of Words and TF-IDF.

1.3.5 Bag of Word

The Bag of Words (BoW) technique processes each news article as a document and calculates the frequency count of each word in that document, which is further used to create numerical representation of the data, also called as vector features of fixed length. Bag of Words converts raw text to word count vector with CountVectorizer function for feature extraction. CountVectorizer splits the text form content, builds the vocabulary and encodes the text into a vector. This encoded vector will have a count for occurrences of each word that appears more like a frequency count as a key value pair. This methodology has drawbacks in terms of information loss. The relative position of the words is not considered, and the information about the context is lost. This loss can be expensive sometimes, compared to the gain in computing simplicity with the ease

2. Related Work

Mykhailo Granik et. al. in their paper [3] shows a simple approach for fake news detection using naive Bayes classifier. This approach was implemented as a software system and tested against a data set of Facebook news posts. They were collected from three large Facebook pages each from the right and from the left, as well as three large mainstream political news pages (Politico, CNN, ABC News). They achieved classification accuracy of approximately 74%. Classification accuracy for fake news is slightly worse. This may be caused by the skewness of the dataset: only 4.9% of it is fake news.

Himank Gupta et. al. [10] gave a framework based on different machine learning approach that deals with various problems including accuracy shortage, time lag (BotMaker) and high processing time to handle thousands of tweets in 1 sec. Firstly, they have collected 400,000 tweets from HSpam14 dataset. Then they further characterize the 150,000 spam tweets and 250,000 non- spam tweets. They also derived some lightweight features along with the Top-30 words that are providing highest information gain from Bag-of-Words model. 4. They were able to achieve an accuracy of 91.65% and surpassed the existing solution by approximately18%.

Marco L. Della Vedova et. al. [11] first proposed a novel ML fake news detection method which, by combining news content and social context features, outperforms existing methods in the literature, increasing its accuracy up to 78.8%. Second, they implemented their method within a Facebook Messenger Chabot and validate it with a real-world application, obtaining a fake news detection accuracy of 81.7%. Their goal was to classify a news item as reliable or fake; they first described the datasets they used for their test, then presented the content-based approach they implemented and the method they proposed to combine it with a social-based approach available in the literature. The resulting dataset is composed of 15,500 posts, coming from 32 pages (14 conspiracy pages, 18 scientific pages), with more than 2, 300, 00 likes by 900,000+ users. 8,923 (57.6%) posts are hoaxes and 6,577 (42.4%) are non-hoaxes.

Cody Buntain et. al. [12] develops a method for automating fake news detection on Twitter by learning to predict accuracy assessments in two credibility-focused Twitter datasets: CREDBANK, a crowd sourced dataset of accuracy assessments for events in Twitter, and PHEME, a dataset of potential rumors in Twitter and journalistic assessments of their accuracies. They apply this method to Twitter content sourced from BuzzFeed's fake news dataset. A feature analysis identifies features that are most predictive for crowd sourced and journalistic accuracy assessments, results of which are consistent with prior work. They rely on identifying highly retweeted threads of conversation and use the features of these threads to classify stories, limiting this work's applicability only to the set of popular tweets. Since the majority of tweets are rarely retweeted, this method therefore is only usable on a minority of Twitter conversation threads.

Shivam B. Parikh et. al. [13] aims to present an insight of characterization of news story in the modern diaspora combined with the differential content types of news story and its impact on readers. Subsequently, we dive into existing fake news detection approaches that are heavily based on text-based analysis, and also describe popular fake news datasets. We conclude the paper by identifying 4 key open research challenges that can guide future research. It is a theoretical Approach which gives Illustrations of fake news detection by analyzing the psychological factors.

Our news article data and headlines are in text format. Building automated machine learning models on text data involves cleaning and converting textual data to machine readable format. Natural-language processing (NLP) is an area of computer science and artificial intelligence concerned with the interactions between computers and human (natural) languages, and how to program computers to fruitfully process large amounts of natural language data [19]. In this project, we are leveraging textual data from FNC-1.

Natural language processing (NLP) models have made significant progress in automatically detecting sentiment and in mining opinion from text. A wide variety of benchmark datasets and techniques have emerged due to significant research conducted in this space [11]. Most of the pre-neural network era NLP techniques focused on developing extensive domain specific features. These techniques often involved manual feature engineering and they require linguists and semantic experts to parse and curate the text. However, the NLP landscape has evolved at great pace. Collobert et al. (2011) [12] introduced Natural Language Processing nearly from scratch which introduces unified neural network architecture and algorithms that can be applied to various NLP tasks. This paper [12] describes how we can learn word and sentence level representations instead of exploiting man-made input features carefully optimized for each task. This breakthrough research paved way for researchers to represent words and sentences as vectors, which can understand the context in which they are being used.

2.1 Word Vector Representation

The method of representing words as vectors is commonly referred to as word vector representations or word embeddings. In this paper, we experiment with the below mentioned word vector representations: Bag of Words: In a Bag of Words (BoW) model, sentences are represented as multiset of its words. BoW model disregards the order and hence it also disregards the context in which the word occurred.

Tf-Idf: Tf-idf stands for term frequency-inverse document frequency. Tf-idf weight gives us an indication on how important a given word is for a sentence or document in relation to the entire corpus.

GloVe: Glove representation are learnt by first constructing a co-occurrence matrix of all the words in the corpus and the dimensions are then reduced by matrix factorization methods. Researchers have trained and created GloVe models on multiple corpora and made them available to the public. The details can be found at [20].

3. Analysis of Problem

As the challenge of fake news detection is gaining leverage, the research community is starting to gather data from various sources to try to define the bounds and goals of this wide research. Some well-structured datasets have emerged and baselines are getting more and more compared across the most common ones. In the past decade, social media networks have become a valuable resource for data of different types, such as texts, photos, videos etc. On social media Fake news has experienced a resurgence of interest primarily due to the recent political climate and the growing concern around its negative effect. Even though the problem of fake news is not a new issue, detecting fake news is a complex task, and it is also very relevant given that humans tend to believe misleading informations. Nevertheless, for news that spread through social networks, there are three commonly accepted characteristics, content features, user features, and social features. In the case of textual content, in order to capture the text features, several existing work utilized both common linguistic features and syntactic features.

Additionally, user features and social features can also be derived from the user-driven social engagements of news consumption on social media platform. Capturing user-based features both at the individual level (registration age, number of followers/followees, number of tweets the user has authored) and group level user features (percentage of verified users and average number of followers' can provide useful information for fake news detection.

4. Proposed Methods

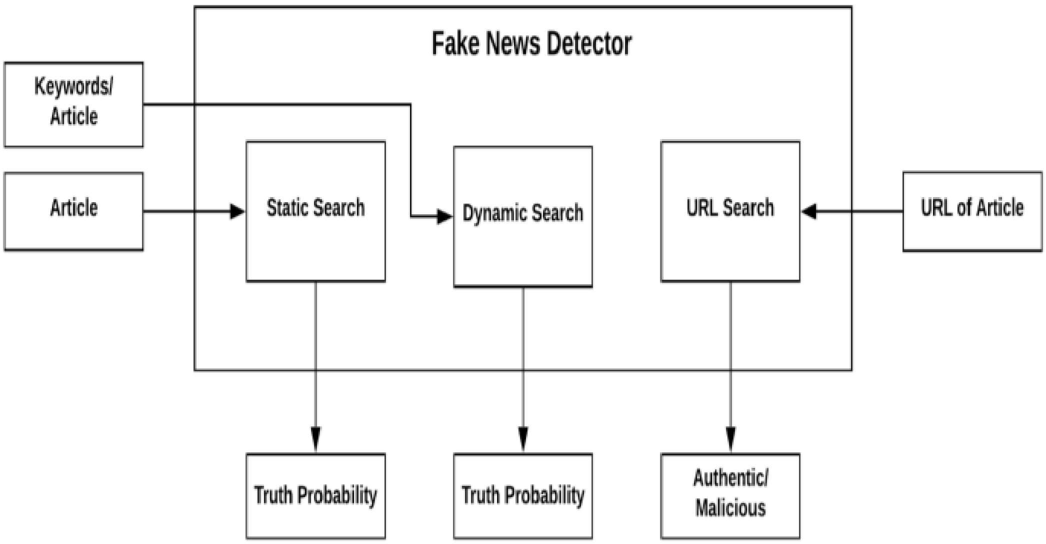


Figure 7: System Design

We can get online news from different sources like social media websites, search engine, homepage of news agency websites or the fact-checking websites. On the Internet, there are a few publicly available datasets for Fake news classification like Buzzfeed News, LIAR

[15], BS Detector etc. These datasets have been widely used in different research papers for determining the veracity of news. In the following sections, I have discussed in brief about the sources of the dataset used in this work.

Online news can be collected from different sources, such as news agency homepages, search engines, and social media websites. However, manually determining the veracity of news is a challenging task, usually requiring annotators with domain expertise who performs careful analysis of claims and additional evidence, context, and reports from authoritative sources. Generally, news data with annotations can be gathered in the following ways: Expert journalists, Fact-checking websites, Industry detectors, and Crowd sourced workers. However, there are no agreed upon benchmark datasets for the fake news detection problem. Data gathered must be pre-processed- that is, cleaned, transformed and integrated before it can undergo training process [16]. The dataset that we used is explained below:

LIAR: This dataset is collected from fact-checking website PolitiFact through its API [15]. It includes 12,836 human labelled short statements, which are sampled from various contexts, such as news releases, TV or radio interviews, campaign speeches, etc. The labels for news truthfulness are fine-grained multiple classes: pants-fire, false, barely-true, half-true, mostly true, and true.

4.1 Static Search Implementation-

In static part, we have trained and used 3 out of 4 algorithms for classification. They are Naïve Bayes, Random Forest and Logistic Regression.

Step 1: In first step, we have extracted features from the already pre-processed dataset. These features are; Bag-of-words, Tf-Idf Features and N-grams.

Step 2: Here, we have built all the classifiers for predicting the fake news detection. The extracted features are fed into different classifiers. We have used Naive-bayes, Logistic Regression, and Random forest classifiers from sklearn. Each of the extracted features was used in all of the classifiers.

Step 3: Once fitting the model, we compared the f1 score and checked the confusion matrix.

Step 4: After fitting all the classifiers, 2 best performing models were selected as candidate models for fake news classification.

Step 5: We have performed parameter tuning by implementing GridSearchCV methods on these candidate models and chosen best performing parameters for these classifier.

Step 6: Finally selected model was used for fake news detection with the probability of truth.

Step 7: Our finally selected and best performing classifier was Logistic Regression which was then saved on disk. It will be used to classify the fake news.

It takes a news article as input from user then model is used for final classification output that is shown to user along with probability of truth.

Conclusion

In the 21st century, the majority of the tasks are done online. Newspapers that were earlier preferred as hard-copies are now being substituted by applications like Facebook, Twitter, and news articles to be read online. Whatsapp's forwards are also a major source. The growing problem of fake news only makes things more complicated and tries to change or hamper the opinion and attitude of people towards use of digital technology. When a person is deceived by the real news two possible things happen- People start believing that their perceptions about a particular topic are true as assumed. Thus, in order to curb the phenomenon, we have developed our Fake news Detection system that takes input from the user and classify it to be true or fake. To implement this, various NLP and Machine Learning Techniques have to be used. The model is trained using an appropriate dataset and performance evaluation is also done using various performance measures. The best model, i.e. the model with highest accuracy is used to classify the news headlines or articles. As evident above for static search, our best model came out to be Logistic Regression with an accuracy of 65%. Hence we then used grid search parameter optimization to increase the performance of logistic regression which then gave us the accuracy of 75%. Hence we can say that if a user feed a particular news article or its headline in our model, there are 75% chances that it will be classified to its true nature.

The user can check the news article or keywords online; he can also check the authenticity of the website. The accuracy for dynamic system is 93% and it increases with every iteration. We intend to build our own dataset which will be kept up to date according to the latest news. All the live news and latest data will be kept in a database using Web Crawler and online database.

References

[1] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective" arXiv:1708.01967v3 [cs.SI], 3 Sep 2017

[2] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective" arXiv:1708.01967v3 [cs.SI], 3 Sep 2017

[3] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 2017, pp. 900-903.

[4] Fake news websites. (n.d.) Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Fake_news_website. Accessed Feb. 6, 2017 [5] Cade Metz. (2016, Dec. 16). The bittersweet sweepstakes to build an AI that destroys fake news.

[5] Conroy, N., Rubin, V. and Chen, Y. (2015). "Automatic deception detection: Methods for finding fake news" at Proceedings of the Association for Information Science and Technology, 52(1), pp.1-4.

[6] Markines, B., Cattuto, C., & Menczer, F. (2009, April). "Social spam detection". In Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web (pp. 41-48)

[7] Rada Mihalcea , Carlo Strapparava, The lie detector: explorations in the automatic recognition of deceptive language, Proceedings of the ACL-IJCNLP

[8] Kushal Agarwalla, Shubham Nandan, Varun Anil Nair, D. Deva Hema, "Fake News Detection using Machine Learning and Natural Language Processing," International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6, March 2019

[9] H. Gupta, M. S. Jamal, S. Madisetty and M. S. Desarkar, "A framework for real-time spam detection in Twitter," 2018 10th International Conference on Communication Systems &

Networks (COMSNETS), Bengaluru, 2018, pp. 380-383 [11] M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro and L. de Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals," 2018 22nd Conference of Open Innovations Association (FRUCT), Jyvaskyla, 2018, pp. 272- 279.

[10] C. Buntain and J. Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads," 2017 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, 2017, pp. 208-215.

[11] S. B. Parikh and P. K. Atrey, "Media-Rich Fake News Detection: A Survey," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, 2018, pp. 436-441

[12] Scikit-Learn- Machine Learning In Python [15] Dataset- Fake News detection William Yang Wang. " liar, liar pants on _re": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648, 2017.

[13] Titcomb, J., Carson, J.: www.telegraph.co.uk. Fake news: What exactly is it – and how can you spot it?

[14] Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election Technical report, National Bureau of Economic Research (2017)

[15] Langin, K.: http://www.sciencemag.org. Fake news spreads faster than true news on Twitter—thanks to people, not bots (2018)

[16] Wardle, C.: Fake News. It's complicated. First Draft New (2017). https://firstdraftnews.com/fake-news-complicated/

[17] Throne, J., Chen, M., Myrianthous, G., Pu, J., Wang, X., Vlachos, A.: 2017. Fake News Detection using Stacked Ensemble of Classifiers. In ACL.

[18] Davis, R., Proctor, C.: 2017. Fake News, Real Consequences: Recruiting Neural Networks for the Fight against Fake News. https://web.stanford.edu/class/cs224n/reports/2761239.pdf

[19] SemEval-2016: Semantic Evaluation Exercises International Workshop on Semantic Evaluation (SemEval-2016).

[20] Sobhani, P., Inkpen, D., Zhu, X.: A Dataset for Multi-Target Stance Detection. http://www.aclweb.org/anthology/E17-2088.

[21] Pang, B., Lee, L.: 2008. Opinion mining and sentiment analysis. http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf

[22] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, M., Kuksa, P.: 2011. Natural Language Processing (Almost) from Scratch. Journal of Machine Learning Research 12 (2011) 2493-2537.

[23] Johnson, J.: 2016. The Five Types of Fake News. https://www.huffingtonpost.com/dr-john- johnson/the-five-types-of-fake-ne_b_13609562.html

[24] As fake news spreads lies, more readers shrug at the truth. In https://www.nytimes.com/ 2016/12/06/us/ fake-news-partisan-republican-democrat.html.