Ministry of Science and Higher education
of the Russian Federation
ITMO University

Faculty of Digital Transformations
Educational program Big Data and Extreme Computing
Subject area (major) 01.04.02. Applied mathematics and informatics

REPORT
"Identification of fake news with machine learning"

Student: Zhumabaev Sultan J41321c

Supervisor: Gladilin P.E.

Date 04.03.2021

St. Petersburg
2021

TABLE OF CONTENTS

1 Introduction

The easy access and exponential growth of information available on social media has complicated the distinction between false and true information. The easy dissemination of information through sharing has exacerbated the exponential growth of falsification. At stake is also the credibility of social media where false information is disseminated. Thus, the research challenge is to automatically verify information, namely its source, content and publisher, to classify it as false or true.

Machine learning has played a vital role in classifying information, albeit with some limitations. This paper discusses various machine learning approaches to detect fake and fabricated news. The limitation of such approaches and improvisations by implementing deep learning is also discussed.

The term "*fake news*", a Collins' Word of the Year 2017, has gained popularity as a result of Donald Trump's election campaign, which refers to deliberately false and inaccurate mainstream media coverage. Interestingly, deliberately false information is much more popular with new media users: a Twitter study concluded that fake news stories are 70% more likely to be retweeted, while true stories are on average six times less likely to be viewed than fake news. Creating a fake news story that is likely to be popular is much easier than finding a really worthwhile true news story. In addition, given that despite attempts by states to combat the flow of false information, it is impossible to fully control new media, unlike traditional media, fake news is published and distributed freely in the media space, contributing to the increasing gap between hyper-reality and physical reality, as they form media models that distort reality. There are a number of classic tricks to draw attention to a false infopoint and hook the audience, such as the use of children in negative connotations (death of a child, illness of a child, etc.), fabricated visual evidence and scientific sensationalism, use of false testimony, demonisation of the enemy, etc. All of these techniques have one common flaw: they require a specific context for their use.

2 Literature review

Various approaches to detecting fake news have been proposed over the recent years. Before going on to an overview of methods to combat the spread of fake news, it is necessary to define what fake news is and what it can be.

The article "*Fake News: A Definition*" denotes fake news as a sociological phenomenon [1]. According to the article, fake news is false, often sensational, information disseminated under the guise of news reporting.

In the article "*Deception detection for news: three types of fakes*" by V. L. Rubin, Y. Chen, and N. J. Conroy presents the following classification of fake news [2].

The first type, serious fabrications, includes publications in tabloids. These outlets use eye-catching headlines, exaggerations, scandal-mongering, or sensationalism to increase traffic or profits. Yellow press focuses on breaking news, especially crime, gossip columns and astrology.

The second type, hoaxes, is social media or journalistic news. Unlike jokes, hoaxes are "relatively complex and large-scale fabrications" that may involve deceptions beyond mere gambling and "cause material damage or harm to the victim".

A third type, humorous news, includes news sites that specialise in fake information and do not pretend to be a true source, games to identify fake news, and satire and parody. The characteristics of this type include the fact that, due to their humorous tone, they can be easily identified.

In the article "*Fake News Detection Using Machine Learning approaches: A systematic Review*", another classification of fake news is highlighted [3]:

1. Visual-based: These fake news posts use graphics a lot more in as content, which may include morphed images, doctored video, or combination of both.

2. User-based: This type of fabricated news is generated by fake accounts and is targeted to specific audience which may represent certain age groups, gender, culture, political affiliations.

3. Knowledge-based: these types posts give scientific explanation to the some unresolved issues and make users to believe it is authentic. For example natural remedies of increased sugar level in human body.

4. Style-based posts are written by psedojournalists who pretend and copy style of some accredited journalists.

5. Stance-based: It actually is representation of truthful statements in such a way which changes its meaning and purpose.

The article "*Media-Rich Fake News Detection: A Survey*" provides sources of fake news [4].

Standalone website:

1. Popular news sites.

2. Blog sites: Blog sites are big on user-generated content and heavily rely on unsupervised content, also considered the best place to get wrong information.

3. Media sites.

Social media: Sharing is the most common way of circulating the content on these sites. More than 70% of its users use them for their daily news source.

1. Facebook (Status —Wall Posts —Dedicated Page —Ad): Users can make a Facebook wall post and/or create Facebook pages and produce/share content using these pages. Better yet, it is concerning to see Facebook allowing users to create paid ads for pretty much any post, which could very well be fake news and can reach larger audience.

2. Twitter (Tweet —Re-Tweet): Twitter is also a social media site, allows you to create a tweet (limited character) and retweet (share a popular tweet with other).

3. Emails: Emails (news) are also a great way for consumers to receive news, it is really challenging to validate the authenticity of news emails.

4. Broadcast networks (PodCast): Podcasts are an audio multimedia category, very small number of users still use this service and consume their news.

5. Radio service: Radio talk shows are a popular source of news and it is challenging to validate the truthfulness of the audio.

The article "*Fake news detection on social media: A data mining perspective*" provides basic methods for analyzing fake news [10]:

1. Linguistic Features based methods.

2. Deception Modeling based methods.

3. Clustering based methods.

4. Predictive Modeling based methods.

5. Content Cues based methods.

6. Non-Text Cues based methods.

The article "*Automatic deception detection: Methods for finding fake news*" by Conroy, N. J., Rubin, V. L., & Chen, Y. (2015, November) describes the typologies of several varieties of credibility assessment methods arising from two main categories: linguistic replication approaches and network analysis approaches [5]. A fake news detection method that is a hybrid of these two approaches is also presented.

The article "*Detecting Fake News using Machine Learning: A Systematic Literature Review*" provides the most common methods of machine learning [9]. These methods include the following approaches.

Support Vector Machine: This algorithm is mostly used for classification. This is a supervised machine learning algorithm that learns from the labeled data set. Researchers in (Singh et al., 2017) used various classifiers of machine learning and the support vector machine have given them the best results in detecting the fake news.

Naive Bayes: Naive Bayes is also used for the classification tasks. This can be used to check whether the news is authentic or fake. Researchers in (Pratiwi et al., 2017) used this classifier of machine learning to detect the false news.

Logistic Regression: This classifier is used when the value to be predicted is categorical. For example, it can predict or give the result in true or false. Researchers in (Kaur et al., 2020) have used this classifier to detect the news whether it is true or fake.

Random Forests: In this classifier, there are different random forests that give a value and a value with more votes is the actual result of this classifier. In (Ni et al., 2020) researchers have used different machine learning classifiers to detect the fake news. One of these classifiers is the random forest.

Recurrent Neural Network: This classifier is also helpful for detecting the fake news. Researchers in (Jadhav & Thepade, 2019) have used the recurrent neural network to classify the news as true or false.

Neural Network: There are different algorithms of machine learning that are used to help in classification problems. One of these algorithms is the neural network. Researchers in (Kaliyar et al., 2020) have used the neural network to detect the fake news.

K-Nearest Neighbor: This is a supervised algorithm of machine learning that is used for solving the classification problems. This stores the data about all the cases to classify the new case on the base of similarity. Researchers (Kesarwani et al., 2020) have used this classifier to detect fake news on social media.

Decision Tree: This supervised algorithm of machine learning can help to detect the fake news. It breaks down the dataset into different smaller subsets. Researchers in (Kotteti et al., 2018) have used different machine learning classifiers and one of them is the decision tree. They have used these classifiers to detect the fake news.

The article "*A Benchmark Study on Machine Learning Methods*" for Fake News Detection compares deep learning methods and neural networks with the classical machine learning methods presented above [8].

In article "*Automatic Online Fake News Detection Combining Content and Social Signals*" is presented by the novel ML fake news detection method which uses combination of news content and social context features [6]. By combining these approaches, the authors of the article managed to increase the accuracy of the method.

In the article "*Some Like it Hoax: Automated Fake News Detection in Social Networks*", the authors hypothesize the relationship between the many users who interact with the news and the fact that the news is fake or not [7].

3 Models, Algorithms and Datasets

This research compares the following 5 classification methods:

1. Support Vector Machine.

2. Naive Bayes.

3. Random Forests.

4. K-Nearest Neighbor.

5. Decision Tree.

The methods presented above will classify fake news by the following three datasets:

1. First dataset: 17903 fake news and 20826 true ones. This dataset does not require preprocessing.

2. Second dataset: 2137 fake news and 1872 true ones. This dataset also does not require preprocessing.

3. Third dataset: 10387 fake news and 10413 true ones. In this dataset, you need to delete empty rows and duplicates.

Classifiers are evaluated by the following metrics:

1. Accuracy is a parameter that can be used to determine the number of correctly classified forecasts and is expressed as:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

. This can be further developed by using the results of a confusion matrix that includes TP, TN, FP, which FN, and is defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

.

2. The F1 score is a number between 0 and 1 that represents the weighted average of precision and recall. F1 score is a better efficiency measure than accuracy, and it is defined:

$$F1_{score} = \frac{2 \cdot (recall \cdot precision)}{recall + precision}$$

.

3. The likelihood of the desired result is used to calculate the success of a model using log loss. The higher the log deficit, the higher the chance of the real class. The lower the ranking, the more the model has done. When the number of potential classes $(M) = 2$, the log loss can be expressed as follows:

$$-(y_i \log(p_i) + (1 - y_i)) \log(1 - p_i)$$

.

4 Experimental research

The estimated metrics were calculated for each method relative to the three data sets. The initial data were divided into training and test samples, 75% and 25%, respectively.

The results of the measurements are shown in the following tables 1- 3:

Table 1: Results of evaluation of classification methods for the first dataset

	accuracy	f1	log_loss
SVM	0.92	0.91	0.15
Naive Bayes	0.89	0.85	2.16
Random Forests	0.92	0.90	0.13
KNN	0.87	0.89	0.54
Decision Tree	0.91	0.88	5.68

Table 2: Results of evaluation of classification methods for the second dataset

	accuracy	f1	log_loss
SVM	0.89	0.93	0.27
Naive Bayes	0.89	0.93	3.57
Random Forests	1.0	1.0	0.059
KNN	0.86	0.91	0.19
Decision Tree	1.0	1.0	9.99

Table 3: Results of evaluation of classification methods for the third dataset

	accuracy	f1	log_loss
SVM	0.73	0.71	0.34
Naive Bayes	0.69	0.65	4.08
Random Forests	0.89	0.88	0.17
KNN	0.75	0.73	0.43
Decision Tree	0.87	0.87	7.58

On the presented datasets, methods with trees show better results than other methods considered.

5 Conclusion

While doing this course work, I studied the phenomenon of fake news. The analysis of the subject area was carried out, the most frequently used algorithms in this problem were identified. Five algorithms were applied in practice to determine the type of news. From the results, I can make an assumption that on the selected datasets, the Random Forests method and Decision Tree show the best result. Of these methods, the method of Random Forests proved to be the best in terms of quality metrics. However, it should be borne in mind that this method is prone to retraining, which was demonstrated by the example of the second dataset.

As for the future work I would like to evaluate these methods on whether it is possible to replace the real data with synthetic in order to solve a Machine Learning Problem by learning a Machine Learning model on the sampled data and then evaluating the score which it obtains when evaluated on the real data.

References

[1] GelferT, A. Fake News: A Definition. Informal Logic, Vol. 38, No.1 (2018), pp. 84-117.

[2] Rubin, Victoria & Chen, Yimin & Conroy, Nadia. (2015). Deception Detection for News: Three Types of Fakes.

[3] S. I. Manzoor, J. Singla and Nikita, "Fake News Detection Using Machine Learning approaches: A systematic Review," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 230-234, doi: 10.1109/ICOEI.2019.8862770.

[4] Parikh, S. B., & Atrey, P. K. (2018, April). Media-Rich Fake News Detection: A Survey. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 436–441). IEEE.

[5] Conroy, N. J., Rubin, V. L., & Chen, Y. (2015, November). Automatic deception detection: Methods for finding fake news. In Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community (p. 82). American Society for Information Science.

[6] Della Vedova, M. L., Tacchini, E., Moret, S., Ballarin, G., DiPierro, M., & de Alfaro, L. (2018, May). Automatic Online Fake News Detection Combining Content and Social Signals. In 20182018–22 Conference of Open Innovations Association (FRUCT) (pp. 272-279). IEEE.

[7] Tacchini, E., Ballarin, G., Della Vedova, M. L., Moret, S., & de Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. arXiv preprint arXiv:1704.07506.

[8] Khan, Junaed Younus & Khondaker, Md. Tawkat Islam & Iqbal, Anindya & Afroz, Sadia. (2019). A Benchmark Study on Machine Learning Methods for Fake News Detection.

[9] Ahmed, Alim & Aljabouh, Ayman & Donepudi, Praveen & Choi, Myung. (2021). Detecting Fake News Using Machine Learning : A Systematic Literature Review.

[10] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective", ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.