

Fake News Detection Models and Performances

Pramod Sekharan Nair, Kefyalew Aragaw Hailekidan, Vinitha Gangadharan Nair



Abstract: Fake News detection is a hard problem for decades after the advent of social media. As misinformation, so called fake news continues to be rapidly distributing on internet, the reality has becoming increasingly shaped by false information. Time after time we have consumed or being exposed to inaccurate information. The last few years have been talking about guarding against misinformation and not progressed much in this direction.

The social media is one of the medium where the fake news spreads so rapidly and impact many in a lesser span of time. Machine Learning and Natural Language processing are the core techniques to detect the fake news and stopping from spreading on social media. Many researchers putting their effort in this new challenge to curb down. This paper provides an insight on feature extraction techniques used for fake news detection on soft media. Text feature extraction works with extracting the document information which represent the whole document without loss of the sole information but words which are considered irrelevant were ignored for the purpose of improving the accuracy. Term Frequency Inverse Document Frequency (TF-IDF), BoW(Bag of Words) are some of the important techniques used in text feature extraction. These techniques are discussed with their significance in this paper. One of the important approach, Automated Readability Index is used to test the readability of the text to build the model also discussed in this paper. This paper will play a significant role for the researchers who are interested in the area of fake news Identification.

Keywords: Fake News, Social Media, Fake News Detection, Fake News Identification, TF-IDF Approach, Bag of Words.

I. INTRODUCTION

Fake news is connected with two words that are misinformation and disinformation. The misinformation is inadvertent sharing of misinformation. The disinformation is a deliberate creation and sharing of information known to be false. The latter is a dangerous threat to the society and the world as a whole. This is one of the main threats to many counties and industries today [8]. Fake news has been existing before technology and technological advancements were introduced. In fact, the way they spreading out to information seekers may differ as of traditional and modern information age. However, unlike the advantages, the wide spread usage and acceptance of internet around the globe, has created a significant reason in increasing fake news or hoaxes

[14] especially social media networks like facebook, twitter etc. Social media news feeds which are designed knowingly to attract a targeted people reach the destination in no time. Hence, a mechanism that filters out the truth from the daily exploiting numerous false news feed is very demanding. Fake news on the social media has made negative impact on people and society.

In one of the studies [9] reported reveals that, twenty nine percent of respondents agreed that social media sites should be the most responsible to take necessary steps to eliminate or protect the social media users from fake news. At the same time a majority of people think other medias are also responsible for spreading fake news. Fourteen percent of the total involved admitted the deliberate sharing of political fake news on social media sites. The fake news and the way it spread will continue to grow as long as there are no preventive mechanisms.

One of the wide spread social media, Facebook has encountered this problem many times via fake accounts. Almost 800 Unites States accounts have been removed by Facebook in recent past that spreads fake news. The website states that, the social media network is suspected as played a major role in facilitating the interference of external entity in the U.S politics, the threat is also from the domestic players in creating and spreading fake news. Many at times the motivation is financial or political gains [10].

The properties of the fake news [15,16] identification can be made by spreading behavior, user properties, ratings, network properties, references, shares, textual content, audio content, video content, temporal features and mitigation behavior. The fake news spread by a very few people but in the very initial stage of the share itself spreads rapidly and the outreach will be enormous before it debunked.

II. DATA TYPES IN NEWS

There are four main data types of formats [7] that found in news. The text news is more popular as compared to other formats. The social media increases the usage of other formats in spreading the news.

A. Text Based Content

Text content is not only simple words, this has several hidden characteristics as grammar, tone and helps to analyze the discourse pragmatically in the given context. The text is analyzed based on the linguistic [17].

B. Multimedia Based Content

The multiple forms of medium as images, videos, audios, and graphics are multiple mediums and while it integrated called as multimedia. This is always very attracting the readers' attention as compared to do any communication in full text.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Pramod Sekharan Nair*, Computer Science and Engineering, Medi-Caps University, Indore, India. Email: pramodsnair@yahoo.com

Kefyalew Aragaw Hailekidan, Computing Technology, Aksum University, Axum, Ethiopia. Email: kefyalewaragaw@gmail.com

Vinitha Gangadharan Nair, Management Studies, Medi-Caps University, Indore, India. Email: vinithapramod2009@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

C. Hyperlinks or Embedded Based Contents

The hyperlinks help the news writers to redirect to different sources to justify or to provide more proof related to the news story. The social media helps the news writers to embed the snapshots to provide the proof to the given news, like the posts on Facebook, tweets on twitter, videos on YouTube, sound clips on cloud, Instagram post etc.

D. Audio Based Contents

Audio based contents are considered as part of multimedia format. In many cases it has its own standalone existence in spreading the news. This format includes audio clips, podcast, broadcast networks and radio services. This medium has very important role especially in social media for spreading the fake news.

III. PROBLEM STATEMENT

Common belief that fake news exists solely for political advantages, but this is far from the actual reason. In fact, might not even be the main one. The reasons behind generation and propagation of fake news includes manipulation of media and propaganda, influence, both political as well as social provocation, social unrest and even financial profit.

Our diverse culture and immense population with people from different communities has historically been used as strategy to perpetrate communal violence for political as well as financial gain and doing that is the way easier in today's age through the medium of social media, messaging applications etc. At these trying times a fake news article that is allowed to exist without any check or repercussion could be reason for a clash which would benefit only the political powers and the capitalists who have stakes in the situation.

IV. MOTIVATION

Fake news spreads the roots by making social medias or other media channels, which are very hard to control. Millions of people, organizations and other bodies that are the victims of fake news irrespective of the number of occurrences. Different researchers are following different approaches to identify the fake news from social media. There are different feature extraction methods adopted by different researchers in fake news identification. This is an emerging area of research due to the high impact that fake news creates in the society. We have made an effort to discuss and analyse some of the important works in this field.

V. FAKE NEWS IDENTIFICATION METHODS

Most of the fake news identification techniques are using the feature extraction [7] extract the relevant content from the news to prepare the data for identification. The researchers have been used different feature extraction methods depends up on the type of news data. The news data can be as text, video, audio or as the combination of all or any of the said. The selection of useful features is important in the accuracy of the detection.

Non-Text Content Based: This method follows a non-text based [18] approach where the belief is that the non-text plays a vital role in convincing the contaminated news. The way the images are used in a news having a great role to play in catching the news by the readers. So, the strategic use of images is one of the methods to spread the fake news.

Text Content Based; Fake news spreads mainly based on text messages. This leads to text content analysis as one of the leading approaches to identify the fake news. The research works on content-based approach largely dependent on feature extraction [27, 28,16,14].

Text Review Content Based; There are algorithms to identify duplicate reviews [19,20,21,22,23,24,25] as by assuming the fake reviewers will give similar of the genuine review or almost similar to the genuine reviews. The Table 2 shows the claimed accuracy of some of the text review-based algorithms.

A. Feature Extraction Methods

Fake news detection is much dependent on feature extraction. The identified and extracted features are important to decide the accuracy of the classification algorithm for fake news identification. There are three steps involved in this approach; Text Pre-processing, Word weight calculation and Feature word selection. Some of the important feature extraction methods are discussed below.

In addition, linguistic based features also extracted [4] to increase the accuracy of identification algorithms. Linguistic features are filtered from the text document based on the organization of documents. The filtering process as in the order as characters then the words, sentences and finally the document. The conversations that are driven by user on social media can also be identified and extracted as features.

Term Frequency-Inverse Document Frequency; One of the important approaches of feature extraction is Term Frequency-Inverse Document Frequency (TF-IDF) [12,13, 29]. This is based on the weight calculation of features in the text document by identifying occurrence or frequency of words in a document. However, Inverse Document Frequency (IDF) works through identifying the word which is not occurred more often throughout the document. Based on the TF-IDF value the word significance can be identified. The TF-IDF high value of a particular word represents the importance of that word in the document. TABLE.3 shows some of the important methods and the accuracy. As an example, the common words from the news articles are extracted and then the words that helps to identify the classes given better weight. Elaborate mathematically as:

$$weight\ w, e = \log(tfidfw, e + 1) * \log\left(\frac{n}{x_w}\right)$$

Where

$tfidf\ w, e$:- is the value of the word w in the given article.

n :- is the number of articles in the given training set

x_w :- is the number of articles where the word w occurs.

- **n-Gram;** n-grams are one way to help machines understand a word in its context to get a better understanding of the meaning of a word. N-grams look up at what comes after and before the interested word. The 'n' represents the number of words comes together in a sequence, the 'n' very as bigram, trigram etc [2]. In baseline model, works simply with predicting a news item is fake depending on the presence or absence of words, characteristic to it. Whereas, bigram works with determining a set of unique pairs of words and their probabilities in the news provided. And trigram does the same to the bigrams but it looks pair of 3(three) words.



- Bag of Words [11]; This is a method widely used in the field of Natural Language Processing for the purpose of retrieving specific information. The ‘Bag’ is the multi set of words in a sentence or document without considering the order or grammar. Sometimes this Bag of Words are called as histogram of words or to be converted to histogram, that will be order less but will give the frequency. In each different news documents, the distribution of interested words can be identified and to be normalized [1]. This will help to identify which all the documents match the best from the millions of documents on the web. In similar approach the Bag of Visual words can also be done by converting in to vectors.
- Semantic Similarity; The semantic similarity plays a major role in natural language processing. The two different approaches to detect the semantic similarity are knowledge based and topological based. The topological based similarity identifies the similarity by using ontologies (WordNet [30]). The knowledge based identifies with the help of expert knowledge, which uses a text corpus.
- Part-of-Speech; One of the most used syntactic approach. This is a tagging approach, labelling each word of a text to a particular part of speech based on definition and surrounding context. Many of the text analysis approaches are dependent on part of speech.

Table 1: Some of the Text Content based algorithms and accuracy

Reference Number	Methods/Models Used	Claimed Accuracy
[27]	Unigrams, Bigrams, Part-of-speech, Past posted false information, Hash tags, URLs	Mean average precision score 95%
[16]	n-grams, Syntax, Punctuations, LIWC, Readability,	Average accuracy 74%
[14]	Support, Appearance, Network, Creator	98%
[28]	User details, Metadata, Tweet details	97.7%

Table 2: Some of the Text Review based algorithms and accuracy

Reference Number	Approach/ Model	Claimed Accuracy
[19]	Logistic Regression	78%
[22]	Bayesian Model	86%
[24]	Part-of-Speech Syntactic. Semantic similarity	F1-Score 0.5-0.7
[25]	Positive Sentiment Deceptive with SVM Classifier with feature sets, n-grams and LIWC[26]	F1-Score 0.9
[23]	Negative Sentiment Deceptive with SVM Classifier	F1-Score 0.86
[21]	Sparse Adaptive Generative Model and SVM Classifier	65%
[20]	Iterative Learning Model	F1-Score 0.63

Table 3: Performance Evaluation of some of the Fake News detection methods

Reference Number	Methods Used	Algorithms Used	Accuracy
[1]	Tf-Idf and Bag of Words	Tf-Idf Vectors with Dense Neural Network	94.31%
[3]	Tf-Idf	Stochastic Gradient Descent, Support Vector Machines, Linear Support Vector Machines, K-Nearest Neighbour, LR, and Decision Trees	LSVM 92%
[5]	n-grams, Automated readability Index, Syntactic information using POS, TF-IDF, Online Relevance	SVM classifier, Random Forest Classifier	Random Forest 94.87%
[6]	Crawl to collect uniform resource locators (URLs) for the click baits.	Bayes Net, Logistic, Random Tree	BayesNet: 99.4%

other available medias. The popularity and fast accessibility are being misused by many to propagate the fake news.

VI. CONCLUSION

Due to the increasing acceptance of internet and social media, individuals consume news from social media than any



Creating and propagating fake news make negative impact in the society. In this paper, we have discussed some of the important fake news content types and the detection techniques and algorithms used to identify the fake news. The main challenge of fake news detection researches is the unavailability of fake news data set in different domains. Most of the cases the data set to be generated accordance to the requirements of the experiments. This will affect the performance of the algorithm. This research topic opens to find a generic approach to fake news detection.

REFERENCES

1. A. Thota, S. Ahluwalia, N. Lohia "Fake News Detection: A Deep Learning Approach". SMU Data Science Review: Vol. 1: No. 3, Article 10, 2018.
2. D. Jurafsky, J. H. Martin "Speech and Language Processing", 2018.
3. H. Ahmed, I. Traore, S. Saad, "Detecting opinion spams and fake news using text classification", Security and Privacy, John Wiley & Sons, Ltd, 29 December 2017.
4. K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu "Fake News Detection on Social Media: A Data Mining Perspective", Newsletter ACM SIGKDD Explorations Newsletter, Volume 19 Issue 1, June 2017.
5. K. Gunasekaran, G. Ganesan, S. S. Ramanujan, B. Srinivasan "Fake News Detection in Social Media", Vol. 1, No. 1, Article 1, 2016.
6. M. Aldwairi, A. Alwahedi. "Detecting Fake News in Social Media Networks", Elsevier: The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2018).
7. S. Parikh, P. K. Atrey "Media-Rich Fake News Detection: A Survey, Fake Multimedia" 2018.
8. <http://www.fakenewschallenge.org>
9. <https://www.statista.com>
10. <http://time.com>
11. Y. Zhang, R. Jin, Z. Zhou, "Understanding bag-of-words model: A statistical framework", International Journal of Machine Learning and Cybernetics, December 2010.
12. L. Hua-Meng, L. Hai-Rui, X. Liang. "TFIDF Algorithm Based on Information Gain and Information Entropy". Computer Engineering, 38(08): 37-40, 2012.
13. Hui H C, S. Y. Sen University, Guangzhou. "A Text Similarity Measurement Combining Word Semantic Information with TF-IDF Method". Chinese Journal of Computers, 34(5):856-864, 2011.
14. S. Kumar, R. West, J. Leskovec. "Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes". Proceedings of the 25th International Conference, World Wide Web, 2016.
15. B. D. Horne, S. Adali. "This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news". 2nd International Workshop on News and Public Opinion, March 2017.
16. V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea. Automatic detection of fake news. arXiv preprint arXiv:1708.07104, 2017.
17. D. S. K. R, V. Singh, R. Dasgupta, I. Ghosh, "Automated fake news detection using linguistic analysis and machine learning.". International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (SBPBRiMS), 2017.
18. Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading online content: Recognizing clickbait as false news.", Proceedings, 2015 ACM on Workshop on Multimodal Deception Detection, 2015.
19. N. Jindal, B. Liu. "Opinion spam and analysis". Proceedings, 1st ACM International Conference on Web Search and Data Mining. ACM, 2008.
20. F. Li, M. Huang, Y. Yang, X. Zhu. "Learning to identify review spam". Proceedings, International Joint Conference on Artificial Intelligence, 2011.
21. J. Li, M. Ott, C. Cardie, E. Hovy. "Towards a general rule for identifying deceptive opinion spam". In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 2014.
22. A. Mukherjee, B. Liu, N. Glance. "Spotting fake reviewer groups in consumer reviews". Proceedings, 21st International Conference on World Wide Web. ACM, 2012.
23. M. Ott, C. Cardie, J. T. Hancock. "Negative deceptive opinion spam". Proceedings, 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2013.
24. V. Sandulescu, M. Ester. "Detecting singleton review spammers using semantic similarity". Proceedings, 24th international conference on World Wide Web. ACM, 2015.

25. M. Ott, Y. Choi, C. Cardie, J. T Hancock. "Finding deceptive opinion spam by any stretch of the imagination". Proceedings, 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 2011.
26. J. W. Pennebaker, M. E. Francis, R. J Booth. "Linguistic inquiry and word count": Liwc 2001. Mahway: Lawrence Erlbaum Associates, 71(2001):2001, 2001
27. V. Qazvinian, E. Rosengren, D. R. Radev, Q. Mei. "Rumor has it: Identifying misinformation in microblogs". Proceedings, Conference on Empirical Methods in Natural Language Processing. ACL, 2011.
28. A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi. "Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy". Proceedings, 22nd International Conference on World Wide Web. ACM, 2013.
29. Salton G, Yu C T. On the construction of effective vocabularies for information retrieval[M]// Operator algebras, unitary representations, enveloping algebras, and invariant theory. Birkhauser, 1990:48-60.
30. Miller, G.A.: WordNet: A Lexical Database for English. Communications of the ACM, Vol. 38, No. 11, 39-41. (1995)

AUTHORS PROFILE



Pramod Sekharan Nair is a Professor at the Department of Computer Science and Engineering, University of Medi-Caps, Indore, India. He has received B.Tech, M.Tech. and Ph.D in Computer Science. His research interest spanned over Business Intelligence, Data Mining, Machine Learning, Data Science, Big Data, Computer Networks, Artificial Intelligence and IoT. He has the experience of 21 years in academia and Research. Published many research papers in peer reviewed journals.



Kefyalew Aragaw Hailekidan Is an M.Sc. student at Department of Computing Technology, University of Aksum, Axum, Ethiopia. He is having a research interest on Business Intelligence, Data Science and Natural Language Processing.



Vinitha Gangadharan Nair M. Com, MBA is an Assistant Professor at Department of Management Studies, University of Medi-Caps, M.P, India. Her research interest is on Business Intelligence and Data Analysis. She is having the Industry, academia and research experience of more than 10 years.