Malevolent Application of Bots-Detecting and Forecasting the Prospect of Bots Triggering Misinformation Spread Using the Machine Learning Approach

Monikka Reshmi Sethurajan[1]and Natarajan K.[2]
[1]Research Scholar, Christ (Deemed to be University), Bangalore

[2]Associate Professor, Christ (Deemed to be University), Bangalore

Corresponding Author's Email- monikkareshmi.sethurajan@res.christuniversity.in, natarajan.k@christuniversity.in

Abstract. Despite the fact that social media platforms like Facebook and Twitter, offer implausible opportunities to their users, social media remains a major part of our everyday life. Because they are immensely popular among a wide range of types of users, social media platforms, such as Facebook, are increasingly used by automated accounts, also acknowledged as bots. The key work of these bots is disseminating fake news, promoting specific ideas and products, manipulating the stock market, and even distributing sexually explicit materials. On the other hand, machine learning models have successfully detected fake accounts created by bots or computers in many instances. Machine Learning algorithms such as SVM, Naïve Bayes, Bayes Network, Neural Network, and other classifiers that were used for Bots detection were compared for precision, readability, accuracy. An in-depth review of systematic literature discusses the supervised machine learning classifiers that make use of labeled data to train. From the perspective of the context, we examine the previous literature examining the role of bot-generated accounts that spreads such misinformation also known as fake news, similarities to the COVID-19 pandemic. By examining known bots' prediction through machine learning approaches, we conclude by stating the prevalence of bots in misinformation spread. The majority of Binary classification models are used to detect bots and the major limitation is the model's capacity to figure out the correspondence between the reported news and the actual news, so we suggest certain ideas to sort this out.
Keywords: Misinformation, Fake news, Social Bots, spambots, SMP, Machine Learning, Bot prediction, COVID-19 pandemic.

1 Introduction

Facebook and Twitter are online social networking services like opinion-sharing, video-sharing, and news-sharing is done [1-2]. The use of SN may be enjoyed by many, but deceptive activity like fake news might lead users to believe misinformation [3]. MID has emerged as a top research topic in recent years, since SN have gained so much attention [4-5]. In practice, Automated misinformation detection, however, is challenging because it requires advanced models to analyze the relationship between reported and genuine information [4]. In addition, ML has been applied to a variety of applications by industry and academia to solve many complex problems in MID [6-8]. An overview of the MID literature will be provided in this study using ML tools, this survey aims to collect as much information as possible. As a dynamic platform, social media sites (SN) are being used for various purposes, from educa-

tion to business, from medical to telemarketing, but unfortunately, they can also be used for illegal purposes [4,9-10].

2 Review of Literature

Media coverage of the corona virus rose around the world along with the increase in online information. COVID-19 misinformation on social media continues to emerge despite the pandemic being ongoing. As rumor spreads, speculations spread about the virus's origin, potential treatments, and the disease's severity [11] and prevalence. Twenty-four percent of tweets related to COVID-19 contained misinformation, and seventeen percent contained unverified info [12]. According to the authors, misinformation and verified information engage Twitter users equally, suggesting that myths as well as facts about the virus are shared on the social network. An earlier study showed that completely false claims about the virus have a higher rate of propagation and are more likely to be liked. Those who tweeted misinformation likewise used not as much of tentative language than those who tweeted effective ones [13].

During humanitarian crises and propagated through SMPs, misinformation has become an established trend. Researchers who examined social media after the 2010 Haiti earthquake observed an increase in misinformation, rumors, and conspiracies [14]. In the aftermath of the Sandy Hook Elementary School shooting of 2012, rumor theory and twitter are examined [15], 2012 Hurricane Sandy [16], and the bombing of the Boston Marathon [17-18] and the Ebola epidemic of 2013 [19].

People can spread misinformation directly, as well as using accounts that are automated, commonly stated as bots. A social bot, which masquerades as a real user on a platform such as Twitter, will make excessive posts, retweeting news items on a regular basis, and tagging influential figures in an attempt to spread the content [20]. The amount of time spent by bots during Twitter discussions related to controversial political issues and public health issues is disproportionate, although their bias appears to be less evident [21-23].

An initial scoping review was combined with a secondary analysis to provide an overview [24-25]. The first step in our analysis is a literature review on bots: what bots are and how are these bots distinguishes, and how to detect bots using machine learning. Furthermore, our study examines the effects of the COVID-19 pandemic on bots and how bots spread misinformation.

3 Methodology

As far as we are aware, this is the first systematic review of social media bots detection methods that is based on a predefined search strategy and includes works published between 2010 and 2021. An in-depth look at detection methods for bots was conducted in this review, which focused on highlighting the techniques used to detect bots in social media and comparing them with current methods. It was identified that there are several gaps in the literature, including: there is a disproportionate amount of literature on Twitter and rarely use methods other than supervised machine learning, most public datasets are not large enough or accurate, integrated systems and real-time detection are necessary, and spreading awareness of these issues is necessary to protect legitimate users.

4 Technical Features of Bots

Despite the difficulty in detecting fake social engagement [26], this vulnerability is systematically exploited [27]. The following factors are attributed to the creation of these fake accounts, among others:

- SMPs do not expect individuals to divulge their true identity in accordance with their privacy policies [28]. It is increasingly difficult to trust people, and this results in detrimental outcomes [29]. [30] someone who has been falsely accused or misinformed. As a case in point, consider cyberbullying [31]. The spreading of false rumors causes children to be bullied online.
- The spread of chaos and pandemonium on SMPs is the fault of malicious individuals and groups. One recent example was false stories spread in the US about Hurricane Sandy [32]. People affected by the storm became aware of the hurricane through false reports.
- As a result of gamification, sites tend to become more popular and have higher social ratings when you have more "likes" or "followers" [30]. In order to remain competitive, people tend to find new ways to stay ahead of their competition [30]. The candidate with the most votes usually wins a political election [33].
- Obtaining false accounts and taking false actions is easy. In one example, false accounts are being purchased online on a marketplace [34], at minimal cost, or delivered through crowdsourcing platforms [35]. There is even the option of buying Twitter followers and Facebook "likes" online [26].

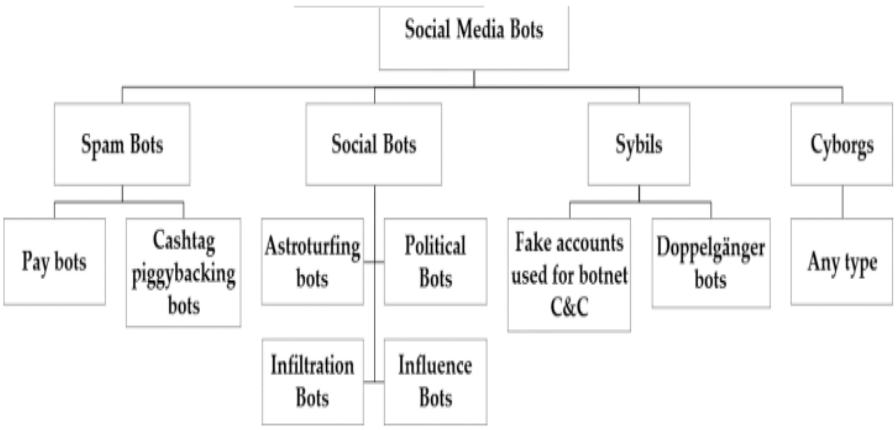

Fig.1.Malicious social media bots typology

5 Overview on Bots

There are roughly three characteristics that distinguish bots from humans:

- This includes Hashtags and connections which are network properties with friends and followers
- Activity and temporal patterns on accounts
- Content in profiles and tweets.

It is advantageous to categorize features according to the platform they are applicable to [36].

Network Properties

Studies such as the one mentioned above have used networks based on connections between friends and followers, hashtag usage, retweets, and mentions to identify social bots [37-40]., taking advantage of network homophily (i.e., A human tends to follow another human, and a bot tends to follow another bot). The characteristics of bot networks change as bots become more sophisticated, and some studies found that groups of bots can build social networks that mimic human ones [41].

Account Activity and Temporal Patterns

Bots can be identified by patterns of content generation. A bot composes rarer novel tweets than a human, Retweets other people's tweets more frequently than his own, and tweets less frequently between time intervals [38]. [42] revealed that humans are more likely to be retweeted than bots, pointing to the fact that bots have difficulty composing convincing or interesting tweets. In other studies, however, this has not always been true [21-22; 43]. There is a decrease in the number of new tweets as the session progresses, because humans every time they are online, they need to modify their behavior. [44] found that bots are not participating during these social media sessions use, and that their behavior does not differ as a result.

Profile and Tweet Content

It is possible to identify social bots using metadata from profiles, such as account age and username. In 2016 [42], they found that bots had a shorter age (i.e., they had been created within the past few years) and longer usernames. Researchers have also looked at automated sentiment analysis of tweet content to distinguish humans from bots. Humans are more often flip-flopping in their sentiment than bots, according to a study [33].

6 Development of Machine learning techniques for Detecting Bots spreading Misinformation

Information that's inaccurate and created to mislead the reader [45-46] is misinformation. In the sphere of misinformation, there are numeric, categorical, textual, graphic, etc., information-packed sources of fake news, rumors, spam, and misinformation that are intended to harm the public. [47-49]. To this end, various techniques have been developed over the years to distinguish legitimate from fraudulent information or users [2; 50-51). These types of misinformation are extremely difficult to analyze using traditional methods. Therefore, a variety of MID features can be detected with ML approaches.

Industry and research communities alike have shown considerable interest in developing ML and DL approaches for a variety of applications [52]. The increasing prevalence of MID has been mainly due to ML-based detection methods. The automatic MID has been the topic of a large number of research works [53-55], along with connected terms, e.g., rumor [56-57], misinformation [47; 57-59], and detection of spam [60-63]. Research issues and existing scenarios of the current problem are reviewed systematically is necessary to optimize ML for MID in academia as well as in industry.

Since the beginning of the 20th century, numerous methods have been proposed to deal with various issues like Online social networks can be compromised by fake news and other misinformation (misinformation, anomalous behavior, etc.). Researchers identify and investi-

gate research gaps in numerous fields, using a variety of techniques to attempt to fill these gaps. A huge number of domains are now being explored with deep learning techniques, including (CNN), (RNN) [64], and (LSTM) [65], all of which can be applied to various areas of research. To perform complex tasks at a high level of sophistication, machine learning is crucial.

Although data scientists and engineers around the world have been working on them for years [8; 66], It remains challenging to build and deploy them effectively. Tests can be done in a very short period of time, even after data training is completed. The ML framework integrates modularized ML algorithms, techniques for optimization and distribution, and infrastructure sustenance to speed up ML processing [67; 68]. System-level development and research are boosted by them in order to simplify the implementation process.

The 45th United States Presidential Election was followed by another notifying survey. To prevent fake news from spreading about the competing representatives, controlling fake news on the Internet was essential. We evaluated several learning-based methods using LR, SVM, LSTM, and CNN. Metadata and text were incorporated into a neural network architecture. Through these techniques, LIAR, a dataset that set benchmarks for fake news detection, was formulated, making the detection of false information and fact-checking easier. As well as statements posted to social media and online, the dataset included those made on television and subsequently criticized by critics and journalists. The problem is a six-way classification problem. Final predictions are generated using a randomly initialized matrix of vectors, in which metadata embeddings are performed and a standard max-pooling operation is applied to it, and then a SoftMax activation function.

According to [62], their fake-news detection technique is based on feedback from microblogs for the exact news. Support vector machines and Naive Bayes classifiers are both employed in these systems to detect deception, respectively. The information is obtained by asking people directly about abortion, execution, and friendship as well as whether they believe that information is true or false. Around 70% of the detections are accurate. It describes a method to detect fake news that uses naive Bayes classifiers, Random Forests, and Logistic Regression, among other artificial intelligence algorithms. Based on a manually labeled news dataset, this study examines how these methods work to detect fake news, and supports (or does not support) the idea of implementing AI to detect fake news. In contrast to similar articles, Logistic Regression is used in the current paper for fake news detection; in addition, the developed system is tested on a relatively new dataset, allowing for an accurate assessment of its performance.

It is shown in [69-70] that fake news can be detected using a naive Bayes classifier in their paper. An application based on this approach was developed and tested using Facebook news posts. These figures were gathered from three major right-wing and left-wing Facebook pages, as well as three major mainstream news sites (Politico, CNN, ABC News). Approximately 74% of their predictions were accurate. There was a slight variance in the classification of fake news. Only 4.9% of the dataset is fake news, which may be due to its skewness. It is based on different machine learning approaches. They give a framework based on 400,000 tweets from the HSpam14 dataset to handle thousands of tweets in 1 second. The framework uses different machine learning approaches to deal with various problems like accuracy shortage, time lag (BotMaker) and high processing time. They then classify the spam tweets

as 150,000 and the non-spam tweets as 250,000. In addition, the researchers derived some lightweight features along with the Top-30 words that provide the highest levels of information from the Bag-of-Words model. The accuracy of their solution was 91.65%, which is 17% better than the existing solution. This research first established a ML fake news detection method that combined news content and social context features to outperform existing methods, increasing its accuracy to 78.8% [71].

7 Discussion

Table.1.An overview of existing techniques.

Source	Process involved	Inference
[83]	Explains how fake news is defined, that it is detected and identified, and the different types of false information	Sorted fake news according to serious fabrications, hoaxes, and humorous fabrications.
[84]	Malicious content is spread through various methods including Click-Bait, Rumours, Spamming, and Bots. The article explores how such content can be countered.	Clickbait can be identified by variations between the amount of character in the title and the number of characters in the article keyword. The percentage of text in an image compared to that of the title can also be used to identify clickbait.
[57],[59],[85].	Datasets for fake news are not sufficient. A number of datasets were explored as well as numerous feature extraction principles.	A benchmark for diagnosing fake news is provided by the LIAR dataset. It is important to preprocess content and tag it properly to conduct an analysis.
[86]	Examining the various available datasets and assessing their relevance to social media and internet data as measured by the PHEME and LIAR benchmarks. Formulating a multiclass classification for six classes of texts. SoftMax is used in final predictions	For the detection of fake news, a multi-modal dataset is necessary. Dataset size determines methods such as clustering.
[87]	This study describes the types of fake news sources and defines a 3-tiered classification system based on contextual, social, and temporal references. Examines challenges of implementation of source reliability.	Using network analysis to characterize fake news trends. The dissemination of fake news can be mitigated through network propagation.
[74]	Bayesian influence is used to detect fake news by using crowd signals	The flagging of fake news in social media is an effective ap-

		proach to identifying them.
[88], [89]	Methods such as bag of words and n-grams can be used for labeling and estimating word occurrence probabilities.	The approaches involved in lexical disambiguation serve to analyze various features relevant to analysis.

Machine learning techniques are considered effective for detecting fake news on social media. According to [72], the following procedures were used:

- a dataset of labeled tweets was obtained;
- The users specified a set of features (followers, followers, URLs, spam words, and hashtags) to be extracted;
- In addition, they chose the testing methodology (Nave Bayes, Unsupervised Clustering, and Decision Trees).

Legitimate users and fake news writers behave differently by definition. As many people as possible are looking to be noticed by the first category of users. There are many followers and few followers on their accounts. If a user tweets too many URLs, it may indicate that they are writing fake news. The dataset can automatically detect spam words. We need to remove stop words (prepositions, determinants, etc.) from vocabularies in order to create one for fake news writers and legitimate users. The misinformation spreaders' vocabulary must be stripped of a group of words that are common to both vocabularies. Posts of misinformation are usually replied to by their authors. A lower number of replies is normally received by their messages since they post false information. As a final step, misinformation spreaders' use a number of hashtags in their posts so that they will be seen by an increasing number of readers. Combining learning methods might result in a detection accuracy at a high level, according to many researchers. Culprits of fake news claim that legitimate users behave differently from writers of fake news. The spammer category includes social bots as well. The difference between them and genuine users is difficult to discern.

For [58] An article's content, the responses it receives from readers, are all factors that can determine if it is genuine or not, and the sources from which readers share it are the most important features. The vector space model represents text as a vector. Users' responses include the frequency and duration of their responses. User characteristics are represented as feature vectors. In order to analyze these Twitter and Weibo data [73], a recurrent neural network was applied. Based on the investigation of the obtained results, it was found that the most suspicious users promoted fake news first in each dataset. In contrast, users of both types act similarly when confronted with actual news.

Using a new tool that allowed users to report fake news on Facebook, researchers looked at how users used the tool [74]. When they see fake news appearing on the network, they can report it. A flagged article is sent to an expert if its flagging score exceeds a threshold. A third-party fact-checking organization may be contacted to examine this article. A fake article may be marked as such by the expert, and it may then be removed. In general, users can be categorized into good (those who flag), spammers, and indifferent (those who do not flag). In order to detect misinformation and valid user performance, Bayesian interference was used. In addition, even large proportions of adversarial users were able to leverage flags.

There are two existing solutions listed in [75]: B.S. Detector is a browser extension, and Fake-News-Detection is an open-source software system. The methodologies they employ and the way they implement them differ from each other. For example, one employs a direct comparison of links. Second, a Machine Learning model that uses SVMs, Gradient Boosting, Bounded Decision Trees, and Random Forests. During the process of running a check, the user can specify which method to use. For measuring the quality of Web pages, both packages use open-source resources. The list includes unreliable or otherwise questionable sources that have been professionally curated. Researchers tested whether the software was accurate at detecting if the articles were from reliable sources. Despite their high accuracy, they are not suitable for ordinary users.
In the following section, some of these popular classification methods are discussed.

Support Vector Machine: Classification is the primary purpose of this algorithm. Machine learning is a supervised process that can be applied to labeled data. [76] used a variety of classifiers for ML, but SVM proved to be the most effective in detecting fake news.

Naïve Bayes: The classification tasks are also carried out using Naive Bayes. By using this, one can determine if the news is authentic. In [77] ML was employed to detect fake news using this classifier.

Logistic Regression: When predicting categorical values, this classifier is used. A true or false result can be predicted or given by its use. This classifier has been used by researchers in [78] for detecting whether news is true or false.

Random Forests: A value with more votes is what the classifier actually results in. There are different random forests used in this classifier. Several machine learning classifiers have been used to detect fake news by researchers [79]. A random forest classifier is one of these.

Recurrent Neural Network: Fake news can also be detected using this classifier. [80] have classified news as true or false by using a recurrent neural network.
Neural Network: To help with classification problems, different algorithms of machine learning can be used. A neural network is one such algorithm. A neural network has been used to detect fake news by researchers at [81].

K-Nearest Neighbor: Classification problems are solved with this supervised algorithm of machine learning. To categorize the new case according to similarity, all the data from previous cases is stored here. In detecting fake news on social media, this classifier has been used by researchers (Kesarwani et al., 2020).

Decision Tree: Detecting fake news can be made easier with this supervised learning algorithm. Subsets of the dataset can be broken down to identify fake news. In their research [82] researchers used a decision tree as one of the machine learning classifiers. Using these classifiers, they detected fake news.
As a first step towards ensuring and maintaining the integrity of the information shared through the platform, Since May 2020, Twitter has begun labeling rumors about COVID-19

as fake and misleading. The social network introduced more controls, including showing the top news feed includes vetted articles, disabling anti-mask clusters, and distribution of anti-misinformation posts to users who share fake news. The methods of combating bots and misinformation have had mixed results despite social media, citizens, and policymakers agreeing that they are urgent problems. In order to combat bots and misinformation, social media companies need to use automated methods or manual moderation to avoid censoring online speech. It is possible for moderation and promotion of content to result in inconsistent policy decisions across platforms. Machine learning models may also suffer from biases with unintended downstream consequences, as multiple studies have shown.

Finally, bot detection algorithms and bot makers are in a continuous arms race. Identifying bots will become more difficult as their intelligence and humanlike characteristics increase. In a sample of publicly available datasets, we found that most recognized bots are now Tweets related to COVID-19. In the aftermath of the COVID-19 pandemic, it is possible that social bots will latch onto future global issues, perhaps even using the same accounts. Also, bot generation technologies will continue to evolve, particularly as the capabilities of ML methods improve for tasks such as the generation of text and images. Creators of bots will continue to use techniques like these, possibly fooling detection algorithms as well as human users. As a result, the government, social media platforms, and detection techniques alone cannot stop cybercrime will be able to address issues related to bots and misinformation in today's world.

8 Conclusion

Fake news has become easier to spread due to the increased use of the internet. Social media platforms and the internet are regularly being used by a large number of people. Posting news on these platforms is not subject to any restrictions. Thus, some people and organizations have taken advantage of these platforms and spread false information about them. A person's reputation can be damaged or a business can suffer. A political party can also use fake news to influence the public's opinion. A way must be found to detect this type of false information. In addition to the application of ML classifiers goes far beyond the identification of fake news. Data sets called training data sets are used to train the classifiers. Afterwards, the fake news is automatically detected by the classifiers. An analysis of supervised machine learning classifiers is presented in this systematic literature review, which require labeled training data. Detecting fake news requires labeled data, which isn't easily accessible. We can use machine learning to detect fake news in the future through the use of unsupervised classifiers.

Today, individuals tend to use only one device and one application when receiving information in a mobile environment. Initially, search engines were competing for customers in this manner. Social networking tools were not available. Today, social networking fills that niche. Internet users have been inundated with all kinds of information, like fake news and hoaxes, as a result of fools' freedom in posting information. A major concern is that people might be influenced by political fake news. The main focus of this paper was to find out how to identify fake news on the Internet. Semantic approaches are currently not capable of providing universally reliable results. There is a need for human expertise. Human decision-makers cannot, however, make independent judgments in sensitive situations. Using multiple filtering layers to reduce fake news can be a compromise.

References

[1]. Gao, Huiji, and Huan Liu. "Data analysis on location-based social networks." In Mobile social networking, pp. 165-194. Springer, New York, NY, 2014.

[2]. Islam, Md Rafiqul, Muhammad Ashad Kabir, Ashir Ahmed, Abu Raihan M. Kamal, Hua Wang, and AnwaarUlhaq. "Depression detection from social network data using machine learning techniques." Health information science and systems 6, no. 1 (2018): 1-12.

[3]. Kumar, Srijan, Robert West, and Jure Leskovec. "Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes." In Proceedings of the 25th international conference on World Wide Web, pp. 591-602. 2016.

[4]. Wu, Liang, Fred Morstatter, Kathleen M. Carley, and Huan Liu. "Misinformation in social media: definition, manipulation, and detection." ACM SIGKDD Explorations Newsletter 21, no. 2 (2019): 80-90.

[5]. Goswami, Anuradha, and Ajey Kumar. "A survey of event detection techniques in online social networks." Social Network Analysis and Mining 6, no. 1 (2016): 1-25.

[6]. Lin, Xiang, Xiangwen Liao, Tong Xu, Wenjing Pian, and Kam-Fai Wong. "Rumor detection with hierarchical recurrent convolutional neural network." In CCF International Conference on Natural Language Processing and Chinese Computing, pp. 338-348. Springer, Cham, (2019): 338–348

[7]. Yenala, Harish, Ashish Jhanwar, Manoj K. Chinnakotla, and Jay Goyal. "Deep learning for detecting inappropriate content in text." International Journal of Data Science and Analytics 6, no. 4 (2018): 273-286.

[8]. Yin, Jun, Qian Li, Shaowu Liu, Zhiang Wu, and Guandong Xu. "Leveraging Multi-level Dependency of Relational Sequences for Social Spammer Detection." arXiv preprint arXiv:2009.06231 (2020).

[9]. Vartapetiance, Anna, and Lee Gillam. "Deception detection: dependable or defective?" Social Network Analysis and Mining 4, no. 1 (2014): 166.

[10]. Islam, Md Rafiqul, Shaowu Liu, Xianzhi Wang, and Guandong Xu. "Deep learning for misinformation detection on online social networks: a survey and new perspectives." Social Network Analysis and Mining 10, no. 1 (2020): 1-20.

[11]. Gozzi, Nicolò, Michele Tizzani, Michele Starnini, Fabio Ciulla, Daniela Paolotti, André Panisson, and Nicola Perra. "Collective response to media coverage of the COVID-19 pandemic on Reddit and Wikipedia: mixed-methods analysis." Journal of medical Internet research 22, no. 10 (2020): e21597.

[12]. Kouzy, Ramez, Joseph Abi Jaoude, AfifKraitem, Molly B. El Alam, Basil Karam, Elio Adib, Jabra Zarka, Cindy Traboulsi, Elie W. Akl, and Khalil Baddour. "Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter." Cureus 12, no. 3 (2020).

[13]. Shahi, Gautam Kishore, Anne Dirkson, and Tim A. Majchrzak. "An exploratory study of covid-19 misinformation on twitter." Online social networks and media 22 (2021): 100104.

[14]. Oh, O., Kwon, K. H., & Rao, H. R. An exploration of social media in extreme events: Rumor theory and Twitter during the Haiti earthquake 2010.

[15]. Williamson, E. "How Alex Jones and Infowars helped a Florida man torment Sandy Hook family." The New York Times 29 (2019).

[16]. Wang, B., & Zhuang, J. Rumor response, debunking response, and decision makings of misinformed Twitter users during disasters. Natural Hazards, 93(3), (2018):1145-1162.

[17]. Gupta, Aditi, Hemank Lamba, and PonnurangamKumaraguru. "$1.00 per rt# bostonmarathon# prayforboston: Analyzing fake content on twitter." In 2013 APWG eCrimeresearchers' summit, pp. 1-12. IEEE, 2013.

[18]. Starbird, Kate, Jim Maddock, Mania Orand, Peg Achterman, and Robert M. Mason. "Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing." IConference 2014 Proceedings (2014).

[19]. Jin, Fang, Wei Wang, Liang Zhao, Edward Dougherty, Yang Cao, Chang-Tien Lu, and Naren Ramakrishnan. "Misinformation propagation in the age of twitter." Computer 47, no. 12 (2014): 90-94.

[20]. Shao, Chengcheng, Giovanni Luca Ciampaglia, OnurVarol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. "The spread of low-credibility content by social bots." Nature communications 9, no. 1 (2018): 1-9.

[21]. Bessi, Alessandro, and Emilio Ferrara. "Social bots distort the 2016 US Presidential election online discussion." First monday 21, no. 11-7 (2016).

[22]. Badawy, Adam, Emilio Ferrara, and Kristina Lerman. "Analyzing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign." In 2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp. 258-265. IEEE, 2018.

[23]. Yuan, Xiaoyi, Ross J. Schuchard, and Andrew T. Crooks. "Examining emergent communities and social bots within the polarized online vaccination debate in Twitter." Social media+ society 5, no. 3 (2019): 2056305119865465.

[24]. Leggio, Lorenzo, James C. Garbutt, and Giovanni Addolorato. "Effectiveness and safety of baclofen in the treatment of alcohol dependent patients." CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders) 9, no. 1 (2010): 33-44.

[25]. Zhu, Yi, Sharath Chandra Guntuku, Weisi Lin, GheorghitaGhinea, and Judith A. Redi. "Measuring individual video qoe: A survey, and proposal for future directions using social media." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 14, no. 2s (2018): 1-24.

[26]. Li, Yixuan, Oscar Martinez, Xing Chen, Yi Li, and John E. Hopcroft. "In a world that counts: Clustering and detecting fake social engagement at scale." In Proceedings of the 25th International Conference on World Wide Web, pp. 111-120. 2016.

[27]. Thomas, Kurt, Chris Grier, Dawn Song, and Vern Paxson. "Suspended accounts in retrospect: an analysis of twitter spam." In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, pp. 243-258. 2011.

[28]. Tuna, Tayfun, EsraAkbas, Ahmet Aksoy, Muhammed Abdullah Canbaz, UmitKarabiyik, Bilal Gonen, and Ramazan Aygun. "User characterization for online social networks." Social Network Analysis and Mining 6, no. 1 (2016): 1-28.

[29]. Gurajala, S., White, J. S., Hudson, B., Voter, B. R., & Matthews, J. N. Profile characteristics of fake Twitter accounts. Big Data & Society, 3(2), (2016): 2053951716674236.

[30]. Xiao, Cao, David Mandell Freeman, and Theodore Hwa. "Detecting clusters of fake accounts in online social networks." In Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security, pp. 91-101. 2015.

[31]. Galán-García, P., Puerta, J. G. D. L., Gómez, C. L., Santos, I., &Bringas, P. G. Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying. Logic Journal of the IGPL, 24(1),(2016):42-53.

[32]. Gupta, Aditi, Hemank Lamba, PonnurangamKumaraguru, and Anupam Joshi. "Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy." In Proceedings of the 22nd international conference on World Wide Web, pp. 729-736. 2013.

[33]. Dickerson, John P., Vadim Kagan, and V. S. Subrahmanian. "Using sentiment to detect bots on twitter: Are humans more opinionated than bots?" In 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), pp. 620-627. IEEE, 2014.

[34]. Gurajala, Supraja, Joshua S. White, Brian Hudson, and Jeanna N. Matthews. "Fake Twitter accounts: profile characteristics obtained using an activity-based pattern detection approach." In Proceedings of the 2015 International Conference on social media& Society, pp. 1-7. 2015.

[35]. Viswanath, Bimal, M. Ahmad Bashir, Mark Crovella, Saikat Guha, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. "Towards detecting anomalous user behavior in online social networks." In 23rd {USENIX} Security Symposium ({USENIX} Security 14), pp. 223-238. 2014.

[36]. Ahmed, Faraz, and Muhammad Abulaish. "A generic statistical approach for spam detection in online social networks." Computer Communications 36, no. 10-11 (2013): 1120-1129.

[37]. Varol, Onur, Emilio Ferrara, Clayton Davis, Filippo Menczer, and Alessandro Flammini. "Online human-bot interactions: Detection, estimation, and characterization." In Proceedings of the international AAAI conference on web and social media, vol. 11, no. 1. 2017.

[38]. Davis, Clayton Allen, OnurVarol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. "Botornot: A system to evaluate social bots." In Proceedings of the 25th international conference companion on world wide web, pp. 273-274. 2016.

[39]. Dickerson, John P., Vadim Kagan, and V. S. Subrahmanian. "Using sentiment to detect bots on twitter: Are humans more opinionated than bots?." In 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014), pp. 620-627. IEEE, 2014.

[40]. Stephens, Monica, and Ate Poorthuis. "Follow thy neighbor: Connecting the social and the spatial networks on Twitter." Computers, Environment and Urban Systems 53 (2015): 87-95.

[41]. Yang, Zhi, Christo Wilson, Xiao Wang, Tingting Gao, Ben Y. Zhao, and Yafei Dai. "Uncovering social network sybils in the wild." ACM Transactions on Knowledge Discovery from Data (TKDD) 8, no. 1 (2014): 1-29.

[42]. Ferrara, Emilio, OnurVarol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. "The rise of social bots." Communications of the ACM 59, no. 7 (2016): 96-104.

[43]. Cresci, Stefano, Roberto Di Pietro, MarinellaPetrocchi, Angelo Spognardi, and Maurizio Tesconi. "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race." In Proceedings of the 26th international conference on world wide web companion, pp. 963-972. 2017.

[44]. Pozzana, Iacopo, and Emilio Ferrara. "Measuring bot and human behavioral dynamics." Frontiers in Physics 8 (2020): 125.

[45]. Fernandez M, Alani H. Online misinformation: challenges and future directions. Companion Proc Web Conf (2018):595–602

[46]. Zhang, Huiling, Alan Kuhnle, J. David Smith, and My T. Thai. "Fight under uncertainty: Restraining misinformation and pushing out the truth." In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 266-273. IEEE, 2018.

[47]. Ma, Jing, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J. Jansen, Kam-Fai Wong, and Meeyoung Cha. "Detecting rumors from microblogs with recurrent neural networks." (2016): 3818.

[48]. Bharti, Santosh Kumar, Ramkrushna Pradhan, Korra Sathya Babu, and Sanjay Kumar Jena. "Sarcasm analysis on twitter data using machine learning approaches." Trends in Social Network Analysis (2017): 51-76.

[49]. Helmstetter, Stefan, and Heiko Paulheim. "Weakly supervised learning for fake news detection on Twitter." In 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 274-277. IEEE, 2018.

[50]. De Choudhury, Munmun, Scott Counts, and Eric Horvitz. "Predicting postpartum changes in emotion and behavior via social media." In Proceedings of the SIGCHI conference on human factors in computing systems, pp. 3267-3276. 2013.

[51]. De Choudhury, Munmun, Michael Gamon, Aaron Hoff, and AstaRoseway. "“Moon Phrases”: A social media faciliated tool for emotional reflection and wellness." In 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, pp. 41-44. IEEE, 2013.

[52]. LeCun, Yann, YoshuaBengio, and Geoffrey Hinton. "Deep learning." nature 521, no. 7553 (2015): 436-444.

[53]. Jain, Suchita, Vanya Sharma, and Rishabh Kaushal. "Towards automated real-time detection of misinformation on Twitter." In 2016 International conference on advances in computing, communications and informatics (ICACCI), pp. 2015-2020. IEEE, 2016.

[54]. Qazvinian, Vahed, Emily Rosengren, Dragomir Radev, and Qiaozhu Mei. "Rumor has it: Identifying misinformation in microblogs." In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pp. 1589-1599. 2011.

[55]. Zhang, Huiling, Md Abdul Alim, Xiang Li, My T. Thai, and Hien T. Nguyen. "Misinformation in online social networks: Detect them all with a limited budget." ACM Transactions on Information Systems (TOIS) 34, no. 3 (2016): 1-24.

[56]. Sampson, Justin, Fred Morstatter, Liang Wu, and Huan Liu. "Leveraging the implicit structure within social media for emergent rumor detection." In Proceedings of the 25th ACM international on conference on information and knowledge management, pp. 2377-2382. 2016.

[57]. Yu, Feng, Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. "A Convolutional Approach for Misinformation Identification." In IJCAI, pp. 3901-3907. 2017.

[58]. Ruchansky, Natali, SungyongSeo, and Yan Liu. "Csi: A hybrid deep model for fake news detection." In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 797-806. 2017.

[59]. Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. "Fake news detection on social media: A data mining perspective." ACM SIGKDD explorations newsletter 19, no. 1 (2017): 22-36.

[60]. Hu, Xia, Jiliang Tang, Yanchao Zhang, and Huan Liu. "Social spammer detection in microblogging." In Twenty-third international joint conference on artificial intelligence. 2013.

[61]. Yin, Jun, Zili Zhou, Shaowu Liu, Zhiang Wu, and Guandong Xu. "Social spammer detection: a multi-relational embedding approach." In Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 615-627. Springer, Cham, 2018.

[62]. Markines B, Cattuto C, Menczer F. Social spam detection. In: Proceedings of the 5th international workshop on adversarial information retrieval on the web, (2009):41–48

[63]. Wang, De, Danesh Irani, and Calton Pu. "A social-spam detection framework." In Proceedings of the 8th annual collaboration, electronic messaging, anti-abuse and Spam conference, pp. 46-54. 2011.

[64]. Cho, Kyunghyun, Bart Van Merriënboer, CaglarGulcehre, DzmitryBahdanau, FethiBougares, Holger Schwenk, and YoshuaBengio. "Learning phrase representations using RNN encoder-decoder for statistical machine translation." arXiv preprint arXiv:1406.1078 (2014).

[65]. Sun, Xiao, Chen Zhang, Shuai Ding, and Changqin Quan. "RETRACTED ARTICLE: Detecting anomalous emotion through big data from social networks based on a deep learning method." Multimedia Tools and Applications 79, no. 13 (2020): 9687-9687.

[66]. Hardy, William, Lingwei Chen, Shifu Hou, Yanfang Ye, and Xin Li. "DL4MD: A deep learning framework for intelligent malware detection." In Proceedings of the International Conference on Data Science (ICDATA), p. 61. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2016.

[67]. Chalapathy, Raghavendra, and Sanjay Chawla. "Deep learning for anomaly detection: A survey." arXiv preprint arXiv:1901.03407 (2019).

[68]. David, Omid E., and Nathan S. Netanyahu. "Deepsign: Deep learning for automatic malware signature generation and classification." In 2015 International Joint Conference on Neural Networks (IJCNN), pp. 1-8. IEEE, 2015.

[69]. Granik, Mykhailo, and Volodymyr Mesyura. "Fake news detection using naive Bayes classifier." In 2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON), pp. 900-903. IEEE, 2017.

[70]. Gupta, Himank, MohdSaalim Jamal, Sreekanth Madisetty, and Maunendra Sankar Desarkar. "A framework for real-time spam detection in Twitter." In 2018 10th International Conference on Communication Systems & Networks (COMSNETS), pp. 380-383. IEEE, 2018.

[71]. Della Vedova, Marco L., Eugenio Tacchini, Stefano Moret, Gabriele Ballarin, Massimo DiPierro, and Luca de Alfaro. "Automatic online fake news detection combining content and social signals." In 2018 22nd Conference of Open Innovations Association (FRUCT), pp. 272-279. IEEE, 2018.

[72]. Patil S.M., Malik A.K. Correlation Based Real-Time Data Analysis of Graduate Students Behaviour. In: Santosh K., Hegadi R. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science, vol 1037. Springer, Singapore (2019)

[73]. Ma, Jing, Wei Gao, Zhongyu Wei, Yueming Lu, and Kam-Fai Wong. "Detect rumors using time series of social context information on microblogging websites." In Proceedings of the 24th ACM international on conference on information and knowledge management, pp. 1751-1754. 2015.

[74]. Tschiatschek, Sebastian, Adish Singla, Manuel Gomez Rodriguez, Arpit Merchant, and Andreas Krause. "Fake news detection in social networks via crowd signals." In Companion Proceedings of the The Web Conference 2018, pp. 517-524. 2018.

[75]. Qiu, Chen, VeranikaMikhailava, and Vitaly Klyuev. "Comparing existing solutions for detecting fake news." Процессыуправления и устойчивость 5, no. 1 (2018): 420-425.

[76]. Singh, Vivek, Rupanjal Dasgupta, Darshan Sonagra, Karthik Raman, and Isha Ghosh. "Automated fake news detection using linguistic analysis and machine learning." In International conference on social computing, behavioral-cultural modeling, & prediction and behavior representation in modeling and simulation (SBP-BRiMS), pp. 1-3. 2017.

[77]. Pratiwi, InggridYanuarRisca, Rosa Andrie Asmara, and Faisal Rahutomo. "Study of hoax news detection using naïve bayes classifier in Indonesian language." In 2017 11th International Conference on Information & Communication Technology and System (ICTS), pp. 73-78. IEEE, 2017.

[78]. Kaur, Sawinder, Parteek Kumar, and PonnurangamKumaraguru. "Automating fake news detection system using multi-level voting model." Soft Computing 24, no. 12 (2020): 9049-9069.

[79]. Ni, Bo, Zhichun Guo, Jianing Li, and Meng Jiang. "Improving Generalizability of Fake News Detection Methods using Propensity Score Matching." arXiv preprint arXiv:2002.00838 (2020).

[80]. Jadhav, Shrutika S., and Sudeep D. Thepade. "Fake news identification and classification using DSSM and improved recurrent neural network classifier." Applied Artificial Intelligence 33, no. 12 (2019): 1058-1068.

[81]. Kaliyar, Rohit Kumar, Anurag Goswami, Pratik Narang, and Soumendu Sinha. "FNDNet–a deep convolutional neural network for fake news detection." Cognitive Systems Research 61 (2020): 32-44.

[82]. Kotteti, Chandra Mouli Madhav, XIshuang Dong, Na Li, and Lijun Qian. "Fake news detection enhancement with data imputation." In 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), pp. 187-192. IEEE, 2018.

[83]. Conroy, Nadia K., Victoria L. Rubin, and Yimin Chen. "Automatic deception detection: Methods for finding fake news." *Proceedings of the association for information science and technology* 52, no. 1 (2015): 1-4.

[84]. Kadian, Arjun, Vivek Singh, and AnolBhattacherjee. "Detecting Clickbait Using User Emotions and Behaviors on Social Media." (2018).

[85]. Buntain, C., &Golbeck, J. (2017, November). Automatically identifying fake news in popular twitter threads. In 2017 IEEE International Conference on Smart Cloud (SmartCloud) (pp. 208-215). IEEE.

[86]. Buntain, Cody, and Jennifer Golbeck. "Automatically identifying fake news in popular twitter threads." In *2017 IEEE International Conference on Smart Cloud (SmartCloud)*, pp. 208-215. IEEE, 2017.

[87]. Parikh, Shivam B., and Pradeep K. Atrey. "Media-rich fake news detection: A survey." In *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*, pp. 436-441. IEEE, 2018.

[88]. Shu, Kai, H. Russell Bernard, and Huan Liu. "Studying fake news via network analysis: detection and mitigation." In *Emerging research challenges and opportunities in computational social network analysis and mining*, pp. 43-65. Springer, Cham, 2019.

[89]. Davis, Richard, and Chris Proctor. "Fake news, real consequences: recruiting neural networks for the fight against fake news." (2017).

[90]. Long, Yunfei. "Fake news detection through multi-perspective speaker profiles." Association for Computational Linguistics, 2017.