

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/349829903>

Detecting Fake News using Machine Learning: A Systematic Literature Review

Article in *Psychology (Savannah, Ga.)* · January 2021

DOI: 10.17762/pae.v58i1.1046

CITATIONS

27

READS

3,494

4 authors, including:



Ayman Aljarbouh

Central Asian University

32 PUBLICATIONS 146 CITATIONS

[SEE PROFILE](#)



Praveen Donepudi

UST-Global

24 PUBLICATIONS 269 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Developing EFL Learners' Knowledge [View project](#)



Measurement of the Extent of Accounting Disclosure: A Study on the Big Four Banks of China (会计信息披露程度的计量：对中国四大银行的研究) [View project](#)

Detecting Fake News using Machine Learning: A Systematic Literature Review

¹Alim Al Ayub Ahmed, ²Ayman Aljarboubh, ³Praveen Kumar Donepudi, ⁴Myung Suh Choi

¹School of Accounting, Jiujiang University, Jiujiang, Jiangxi, CHINA

²Department of Computer Science, University of Central Asia, 310 Lenin Street, 722918 Naryn, KYRGYZSTAN

³Enterprise Architect, Information Technology, UST-Global, Inc., Ohio, USA

⁴Monta Vista High School, 21840 McClellan Rd, Cupertino, CA 95014, USA

ABSTRACT

Internet is one of the important inventions and a large number of persons are its users. These persons use this for different purposes. There are different social media platforms that are accessible to these users. Any user can make a post or spread the news through these online platforms. These platforms do not verify the users or their posts. So some of the users try to spread fake news through these platforms. These fake news can be a propaganda against an individual, society, organization or political party. A human being is unable to detect all these fake news. So there is a need for machine learning classifiers that can detect these fake news automatically. Use of machine learning classifiers for detecting the fake news is described in this systematic literature review.

Keywords

Online fake news, Machine learning, fake news, Text Classification, social media

1. Introduction

World is changing rapidly. No doubt we have a number of advantages of this digital world but it also has its disadvantages as well. There are different issues in this digital world. One of them is fake news. Someone can easily spread a fake news. Fake news is spread to harm the reputation of a person or an organization. It can be a propaganda against someone that can be a political party or an organization. There are different online platforms where the person can spread the fake news. This includes the Facebook, Twitter etc. Machine learning is the part of artificial intelligence that helps in making the systems that can learn and perform different actions (Donepudi, 2019). A variety of machine learning algorithms are available that include the supervised, unsupervised, reinforcement machine learning algorithms. The algorithms first have to be trained with a data set called train data set. After the training, these algorithms can be used to perform different tasks. Machine learning is using in different sectors to perform different tasks. Most of the time machine learning algorithms are used for prediction purpose or to detect something that is hidden.

Online platforms are helpful for the users because they can easily access a news. But the problem is this gives the opportunity to the cyber criminals to spread a fake news through these platforms. This news can be proved harmful to a person or society.

Readers read the news and start believing it without its verification. Detecting the fake news is a big challenge because it is not an easy task (Shu et al., 2017). If the fake news is not detected early then the people can spread it to others and all the people will start believing it. Individuals, organizations or political parties can be effected through the fake news. People opinions and their decisions are affected by the fake news in the US election of 2016 (Dewey, 2016).

Different researchers are working for the detection of fake news. The use of Machine learning is proving helpful in this regard. Researchers are using different algorithms to detect the false news. Researchers in (Wang, 2017) said that fake news detection is big challenge. They have used the machine learning for detecting fake news. Researchers of (Zhou et al., 2019) found that the fake news are increasing with the passage of time. That is why there is a need to detect fake news. The algorithms of machine learning are trained to fulfill this purpose. Machine learning algorithms will detect the fake news automatically once they have trained.

This literature review will answer the different research questions. The importance of machine learning to detect fake news will be proved in this literature review. It will also be discussed how machine learning can be used for detecting the false news. Machine learning algorithms that are used to

detect false news will be discussed in the literature review.

The structure of the rest of paper is as Methodology in section two, section three shows the research questions, section four is showing the search process model that is followed for this literature review, result and discussion is given in section five, the conclusion is presented in section six. In the last, references are given for the papers that are discussed in this literature review.

2. Methodology

This literature review is written for answering some research questions. So the methodology that

is used is the systematic literature review. This methodology helps in answering the research questions. The papers were collected from various databases to be discussed in this literature review. To answer the research questions, different research papers are discussed and cited in this literature review.

Exclusion and Inclusion

A number of papers are published every day. So when a string is searched a number of papers are presented in the result. Not all the papers are relevant to that string. This means there is a need for the criteria. The criteria for inclusion and exclusion that is followed in this literature review is given in the below table.

Exclusion Criteria	Inclusion Criteria
The language of the paper is not the English language.	Papers that are written in the English language.
The complete paper is not accessible.	Paper can be accessed completely.
Paper is not related to machine learning and fake or false news detection.	Paper showing content related to machine learning and fake or false news detection.

Table 1: Exclusion and Inclusion Criteria

Papers that fulfilled the above mentioned inclusion criteria were included in the literature review.

Quality Assessment

Quality of all included papers was assessed on the basis of the research work presented in those papers. The papers in which the researchers have discussed the machine learning use for fake or false news detection were considered as good quality papers to be included in this literature review.

3. Research Question

A SLR has to answer some RQs. In this literature review, three research questions will be answered on the basis of valid arguments. These two research questions are given below.

RQ₁: Why machine learning is required to detect the fake news?

RQ₂: Which machine learning supervised classifiers can be used for detecting fake news?

RQ₃: How classifiers of machine learning are trained to detect fake news?

These research questions will be answered in the result and discussion section of this literature review.

4. Search Process

A search process is followed to collect the papers that can be discussed in this literature review. This search process can easily be understood through the below given diagram.

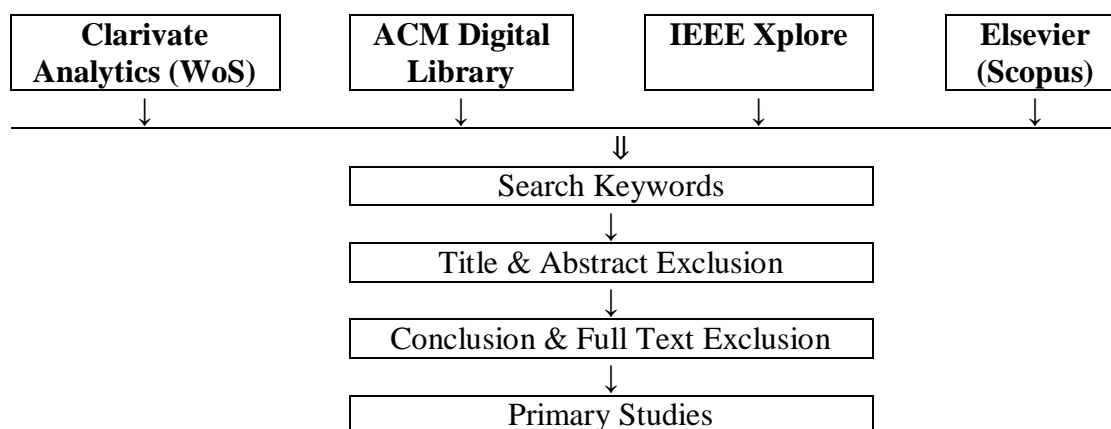


Figure 1: Search Process Model Diagram

Papers were collected from different databases. But not all of them were relevant to the topic. So first of all, the papers were excluded on the basis of their titles and abstract. An abstract is a kind of short summary of the whole paper that can give the idea about the contents presented in the paper. In the next phase, the further part of the papers was studied against the inclusion and exclusion criteria. Seventy three papers were collected from different databases against the search keyword. After the exclusion, there were twenty six papers were remaining that are discussed in this literature review.

5. Result and Discussion

Internet is one of the great sources of information for its users (Donepudi, 2020). There are different social media platforms that includes Facebook or Twitter that helps the people to connect with other people. Different kind of news are also shared on these platforms. People nowadays prefer to access the news from these platforms because these are easy to use and easy to access platforms. Another advantage to the people is that these platforms provide options of comments, reacts etc. These advantages attract people to use these platforms (Donepudi et al., 2020b). But as like their advantages, these platforms are also used as the best source by the cyber criminals. These persons can spread the fake news through these platforms. There is also a feature of sharing the post or news on these platforms and this feature also proves helpful for spreading such fake news. People start believing in such news as well as shares the news with other peoples. Researchers in (Zubiaga et al., 2018) said that it is difficult to control the false

news from spreading on these social media platforms.

Anyone can be registered on these platforms and can start spreading news. A person can create a page as a source of news and can spread the fake news. These platforms do not verify the person whether he is really reputable publisher. In this way, anyone can spread news against a person or an organization. These fake news can also harm a society or a political party. The report shows that it is easy to change people opinions by spreading fake news (Levin, 2017). Therefore, there is a need for detecting these fake news from spreading so that the reputation of a person, political party or an organization can be saved.

RQ1: Why machine learning is required to detect the fake news?

Increasing use of internet has made it easy to spread the false news. Different social media platforms can be used to spread fake news to a number of persons. With the share option of these platforms, the news spread in a fast way. Fake news just not only affects an individual but it can also affect an organization or business (Donepudi et al., 2020a). So controlling the fake news is mandatory. A person can know the news is fake only when he knows the complete story of that topic. It is a difficult task because most of the people do not know about the complete story and they just start believing in the fake news without any verification.

The question arises here how to control fake news because a person cannot control the fake news. The answer is machine learning. Machine learning can help in detecting the fake news (Khan et al., 2019). Through the use of machine learning these fake

news can be detected easily and automatically (Della Vedova et al., 2018). Once someone will post the fake news, machine learning algorithms will check the contents of the post and will detect it as a fake news. Different researchers are trying to find the best machine learning classifier to detect the fake news (Kurasinski, 2020). Accuracy of the classifier must be considered because if it failed in detecting the fake news then it can be harmful to different persons. The accuracy of the classifier depends on the training of this classifier. A model that is trained in a good way can give more accuracy. There are different machine learning classifiers are available that can be used for detecting the fake news that will be answered in the next question.

RQ 2: Which machine learning supervised classifiers can be used for detecting fake news?

Detecting the fake news is one of the most difficult tasks for a human being. The fake news can easily be detected through the use of machine learning. There are different machine learning classifiers that can help in detecting the news is true or false. Nowadays, the dataset can easily be collected to train these classifiers. Different researchers used machine learning classifiers for checking the authenticity of news. Researchers in (Abdullah-All-Tanvir et al., 2019) used the machine learning classifiers for detecting the fake news. According to the experiments of the researchers the SVM and Naïve Bayes classifiers are best for detecting fake news. These two are better than other classifiers on the basis of accuracy they provide. A classifier with more accuracy is considered as a better classifier. The major thing is the accuracy that is provided by any classifier. Classifier with more accuracy will help in detecting more fake news. Researchers in (Kudarvalli & Fiaidhi, 2020) said that detection of false news is necessary because many persons spread the fake news of social media to mislead the people. To save the individuals or organizations from losing their reputation because of false news it is necessary to detect it (Rahman et al., 2020). They have said that the machine learning is very helpful in this regard. They used the different machine-learning algorithms and they also found that the Logistic regression is a better classifier because it gives more accuracy.

Researchers in (Aphiwongsophon & Chongstitvatana, 2018) said that the social media

produce a large number of posts. Anyone can register on these platforms can do any post. This post can contain false information against a person or business entity. Detecting such false news is an important and also a challenging task. For performing this task the researchers have used the three machine learning methods. These are the Naïve Bayes, Neural network and the SVM. The accuracy provided by the Naïve Bayes was 96.08%. On the other hand, the other two methods that are neural network and SVM provided the accuracy of 90.90%.

According to the researchers of (Ahmed et al., 2017), false news has major impact on the political situation of a society. False news on the social media platforms can change opinions of peoples. People change their point of view according to a fake news without verifying it. There is a need for a way that can detect such news. The researchers have used classifiers of machine learning for this purpose. The classifiers that are used by different researchers are the K-Nearest Neighbor, Support Vector Machine, Logistic Regression, Linear Support Vector Machine, Decision tree, Stochastic Gradient Descent. According to results, linear support vector machine provided the good accuracy in detecting the false news.

Researchers (Reis et al., 2019) have used the machine-learning classifiers for the detection of fake news. They have used different features to train these classifiers. Training of the classifiers is an important task because a trained classifier can give the more accurate results. According to the researchers of (Granik & Mesyura, 2017), artificial intelligence is better to detect the fake news. They have used Naïve Bayes classifier to detect fake news from Facebook posts. This classifier has given them the accuracy of 74% but they said the accuracy can be improved. To improve the accuracy different ways are also described by these researchers in that paper. There are classifiers of machine learning that are used for detecting fake news.

Some of these popular classifiers are given below that are used for this purpose.

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.

Naïve Bayes: Naïve 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.

RQ3: How machine learning classifiers are trained for detecting fake news?

Training of the classifiers of machine learning is an important task. This plays an important role for the accuracy of results of these classifiers. A classifier must have to be trained in a proper way with proper data set. Different researchers have trained the machine learning classifiers to detect the fake news. The main problem that occurs while training these classifiers is that mostly the training data set in an imbalanced form (Wang et al., 2020). Researchers in (Al Asaad & Erascu, 2018) have used the supervised machine learning classifiers for fake news detection. To train these classifiers they have used the three different models for feature extraction. Actually, these features are used to train the classifiers. These models are the TF-IDF Model, N-Gram Model, Bag of Words Model. These models extract the features from the training data set and then the classifier is trained through these features. Researchers in (Ahmed et al., 2018) has trained some machine learning classifiers to detect the fake news. For the training purpose, they have used a training data set. They have first removed the unnecessary words and the words are transformed to its single form. So that the training dataset that is given to these classifier should only have the valuable data.

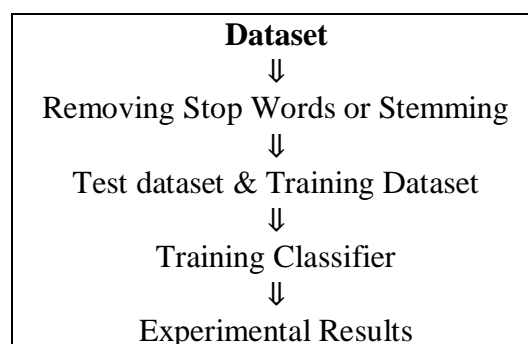


Figure 2: Training Dataset

Figure 3 shows the steps that are used while training a classifier. After the training a classifier is then used for experiments.

6. Conclusion

Due to increasing use of internet, it is now easy to spread fake news. A huge number of persons are regularly connected with internet and social media

platforms. There is no any restriction while posting any news on these platforms. So some of the people takes the advantage of these platforms and start spreading fake news against the individuals or organizations. This can destroy the reputes of an individual or can affect a business. Through fake news, the opinions of the people can also be changed for a political party. There is a need for a way to detect these fake news. Machine learning classifiers are using for different purposes and these can also be used for detecting the fake news. The classifiers are first trained with a data set called training data set. After that, these classifiers can automatically detect fake news.

In this systematic literature review, the supervised machine learning classifiers are discussed that requires the labeled data for training. Labeled data is not easily available that can be used for training the classifiers for detecting the fake news. In future a research can be on the use of the unsupervised machine learning classifiers for the detection of fake news.

References

1. Abdullah-All-Tanvir, Mahir, E. M., Akhter S., & Huq, M. R. (2019). Detecting Fake News using Machine Learning and Deep Learning Algorithms. 7th International Conference on Smart Computing & Communications (ICSCC), Sarawak, Malaysia, Malaysia, 2019, pp.1-5, <https://doi.org/10.1109/ICSCC.2019.8843612>
2. Ahmed, H., Traore, I., & Saad, S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques. *Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments*, 127–138, Springer, Vancouver, Canada, 2017. https://doi.org/10.1007/978-3-319-69155-8_9
3. Ahmed, H., Traoré, I., & Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Secur. Priv.*, 1(1), 1-15. <https://doi.org/10.1002/spy2.9>
4. Al Asaad, B., & Erascu, M. (2018). A Tool for Fake News Detection. 2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Timisoara, Romania, 2018, pp.379-386. <https://doi.org/10.1109/SYNASC.2018.00064>
5. Aphiwongsophon, S., & Chongstitvatana, P. (2018). Detecting Fake News with Machine Learning Method. 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 528-531. <https://doi.org/10.1109/ECTICon.2018.8620051>
6. Della Vedova, M. L., Tacchini, E., Moret, S., Ballarin, G., DiPierro, M., & de Alfaro, L. (2018). Automatic online fake news detection combining content and social signals. FRUCT'22: Proceedings of the 22st Conference of Open Innovations Association FRUCT. Pages 272–279. <https://dl.acm.org/doi/10.5555/3266365.3266403>
7. Dewey, C. (2016). Facebook has repeatedly trended fake news since firing its human editors. The Washington Post, Oct. 12, 2016.
8. Donepudi, P. K. (2019). Automation and Machine Learning in Transforming the Financial Industry. *Asian Business Review*, 9(3), 129-138. <https://doi.org/10.18034/abr.v9i3.494>
9. Donepudi, P. K. (2020). Crowdsourced Software Testing: A Timely Opportunity. *Engineering International*, 8(1), 25-30. <https://doi.org/10.18034/ei.v8i1.491>
10. Donepudi, P. K., Ahmed, A. A. A., Saha, S. (2020a). Emerging Market Economy (EME) and Artificial Intelligence (AI): Consequences for the Future of Jobs. *Palarch's Journal of Archaeology of Egypt/Egyptology*, 17(6), 5562-5574. <https://archives.palarch.nl/index.php/jae/article/view/1829>
11. Donepudi, P. K., Banu, M. H., Khan, W., Neogy, T. K., Asadullah, ABM., Ahmed, A. A. A. (2020b). Artificial Intelligence and Machine Learning in Treasury Management: A Systematic Literature Review. *International Journal of Management*, 11(11), 13-22.
12. Granik, M., & Mesyura, V. (2017). Fake news detection using naive Bayes classifier. 2017 IEEE First Ukraine

- Conference on Electrical and Computer Engineering (UKRCON)*, Kiev, pp.900-903,
<https://doi.org/10.1109/UKRCON.2017.8100379>
13. Jadhav, S. S., & Thepade, S. D. (2019). Fake News Identification and Classification Using DSSM and Improved Recurrent Neural Network Classifier, *Applied Artificial Intelligence*, 33(12), 1058-1068, <https://doi.org/10.1080/08839514.2019.1661579>
 14. Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2020). FNDNet-A deep convolutional neural network for fake news detection. *Cognitive Systems Research*, 61, 32-44.
<https://doi.org/10.1016/j.cogsys.2019.12.005>
 15. Kaur, S., Kumar, P. & Kumaraguru, P. (2020). Automating fake news detection system using multi-level voting model. *Soft Computing*, 24(12), 9049–9069.
<https://doi.org/10.1007/s00500-019-04436-y>
 16. Kesarwani, A., Chauhan, S. S., & Nair, A. R. (2020). Fake News Detection on Social Media using K-Nearest Neighbor Classifier. *2020 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, Las Vegas, NV, USA, pp.1-4,
<https://doi.org/10.1109/ICACCE49060.2020.9154997>
 17. Khan, J. Y., Khondaker, M., Islam, T., Iqbal, A., & Afroz, S. (2019). A benchmark study on machine learning methods for fake news detection. *Computation and Language*.
<https://arxiv.org/abs/1905.04749>
 18. Kotteti, C. M. M., Dong, X., Li, N., & Qian, L. (2018). Fake news detection enhancement with data imputation. *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)*. Athens, 2018, pp.187-192.
<https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00042>
 19. Kudarvalli, H. & Fiaidhi, J. (2020). Experiments on Detecting Fake News using Machine Learning Algorithms. *International Journal of Reliable Information and Assurance*, 8(1), 15-26.
<https://dx.doi.org/10.21742/IJRIA.2020.8.1.03>
 20. Kurasinski, L. (2020). Machine Learning explainability in text classification for Fake News detection. Malmö University Publications.
<https://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-20058>
 21. Levin, S. (2017). Pay to sway: report reveals how easy it is to manipulate elections with fake news. *The Guardian*. 13 Jun 2017
 22. Ni, B., Guo, Z., Li, J., & Jiang, M. (2020). Improving Generalizability of Fake News Detection Methods using Propensity Score Matching. *Social and Information Networks*.
<https://arxiv.org/abs/2002.00838>
 23. Pratiwi, I. Y. R., Asmara, R. A., & Rahutomo, F. (2017). Study of hoax news detection using naïve bayes classifier in Indonesian language. *2017 11th International Conference on Information & Communication Technology and System (ICTS)*, Surabaya, pp.73-78.
<https://doi.org/10.1109/ICTS.2017.8265649>
 24. Rahman, M. M., Chowdhury, M. R. H. K., Islam, M. A., Tohfa, M. U., Kader, M. A. L., Ahmed, A. A. A., & Donepudi, P. K. (2020). Relationship between Socio-Demographic Characteristics and Job Satisfaction: Evidence from Private Bank Employees. *American Journal of Trade and Policy*, 7(2), 65-72.
<https://doi.org/10.18034/ajtp.v7i2.492>
 25. Reis, J. C. S., Correia, A., Murai, F., Veloso A., & Benevenuto, F. (2019). Supervised Learning for Fake News Detection. *IEEE Intelligent Systems*, 34(2), 76-81,
<https://doi.org/10.1109/MIS.2019.2899143>
 26. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*,

- 19(1), 22-36.
<https://doi.org/10.1145/3137597.3137600>
27. Singh, V., Dasgupta, R., Sonagra, D., Raman, K., & Ghosh, I. (2017). 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 (SBP-BRIMS)*.
<http://doi.org/10.13140/RG.2.2.16825.67687>
28. Wang, W. Y. (2017). "Liar, liar pants on fire": A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 2(Short Papers), 422–426.
<http://dx.doi.org/10.18653/v1/P17-2067>
29. Wang, W., Wang, S., Fan, W., Liu, Z., & Tang, J. (2020). *Global-and-Local Aware Data Generation for the Class Imbalance Problem*. *Proceedings of the 2020 SIAM International Conference on Data Mining*.
<https://doi.org/10.1137/1.9781611976236.35>
30. Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). Fake News: Fundamental theories, detection strategies and challenges. In *WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining* (pp. 836-837). (WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining). Association for Computing Machinery, Inc. <https://doi.org/10.1145/3289600.3291382>
31. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys*, 51(2), 1-36.
<https://doi.org/10.1145/3161603>