iJRASET
International Journal For Research in
Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** I **Month of publication:** January 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.48911

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

# A Literature Survey on "Misinformation Flagging System"

Kritik Modi[1], Shubham Shreshth[2], Karan V[3], Himanshu Yadav[4], Sarala D V[5]
[1, 2, 3, 4]*Student in Dept. of CSE, Dayananda Sagar College of Engineering, Bangalore-560078, Karnataka, India*
[5]*Asst. Prof in Dept. of CSE, Dayananda Sagar College of Engineering, Bangalore-560078, Karnataka, India*

***Abstract: Many people rely on social media as their main information source these days. But it has two sides to it. The cheap cost, simple access, and quick information distribution encourage consumers to read news on social media. But it also facilitates the dissemination of incorrect information without any malicious intent and disinformation can go unnoticed. As publishing news online is inexpensive and spreading it through social media is much simpler and faster than through traditional channels, a lot of misinformation is created on the internet for a number of reasons, such as financial benefits and political benefits of many different parties. The ecology of the news industry's authenticity check might be upset by fake news. Consumers are purposefully led to believe inaccurate or biased information via fake news. Propagandists typically use fake news to spread political ideas or exert influence. Some fake news is simply designed to incite people's mistrust and confound them, making it difficult for them to tell what is genuine from what is not.***

## I. INTRODUCTION

Misinformation or "manipulated news" plays a great role in shaping public opinion. Misinformation can have harmful effects on both individuals and society in the long term. One of the prominent and crucial aspects of news from social media is that anyone can pretend as a news publisher and without any financial investment. Firstly, misinformation is written intentionally to mislead its readers and accept false beliefs and this changes(indirectly) the way people react to factual news. Propagandists frequently use manipulated news to spread or influence political messages. Secondly, data or information on social media is very large or huge, mostly user generated and sometimes noisy as well. Third, the mass circulation of misinformation breaks the credibility of factual news' entire ecosystem.

There are no studies or research available that provides an understanding of - (i) What are the characteristics that can be used to describe or identify a user; (ii) can the characteristics described be utilized for detection of misinformation; and (iii) how effective are the characteristics in distinction of real or fake news. So, an automatic system for news classification is needed. To give an understanding, we try to answer the following three research questions: Q1) Which type or group of users are more likely to propagate or disseminate false information or real information? Q2) What are the attributes or characteristics that distinguish users who are more prone to spreading false information from those who tend to share real information? Are there any discernible distinctions between these two groups of users? Q3) Can characteristics or attributes that describe a user be used to detect misinformation, and if so, what is the approach?

Users always prefer to stay close to people or group of people with same or similar interest. To develop a successful and feasible system for detecting misinformation, it is necessary and rational to analyze supplementary data from various perspectives. The social media-based news environment consists of three primary elements, which are publishers, news stories, and social media users. Investigations in the field of journalism has suggested a connection between the political inclinations of publishers and truthfulness of news. The political inclinations denote the perceived bias of the publisher in their decision on how news is reported and covered. Recognizing user's interactions with news on social media platforms can aid in the detection of misinformation. On social media platforms, there are wide range of users with varying levels of credibility. The credibility score states "The trustworthiness".

By incorporating the credibility levels of users in analysing the interactions between users and news, the prediction of misinformation can be improved. Currently available manual designed and context specific textual features are not enough or are typically insufficient for identifying misinformation. In order to improve detection, it is necessary to incorporate additional data sources, such as knowledge based and the social interactions of users.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

## II. LITERATURE REVIEW

[1] Deals with Misinformation classification using the WOA-xgbTree algorithm and content features, these features define the news article, i.e., useful, and unique features. These features are extracted and fed into the xgbTree algorithm, powered up by the Whale Optimization Algorithm (WOA). XgbTree algorithm stands for Extreme Gradient Boosting Tree algorithm and is an optimized tree boosting ensemble technique, i.e., it takes inputs from multiple ML models to produce the most efficient results. The main feature of this algorithm is that it produces decision trees sequentially. It is a scalable implementation of an old algorithm and is useful for multiple Supervised Learning problems involving mainly classification and regression. Moreover, this algorithm shows good compatibility with the WOA. The main highlight of this method is that it requires only the news content to work, unlike other methods which require the source and its characteristics as well. According to "Saeid Sheikhi", a dataset (ISOT Fake News Dataset) of nearly 44000 recently obtained news articles, both authentic and false, was used for testing and training this model, and the model achieved a good accuracy and F1 score and was able to successfully classify over 91% of the articles.

[6] The motivation for this task was to learn more about the various already existing methods of flagging misinformation. This paper deals with various already existing methods for misinformation detection compares them and explores the opportunities those comparison results provide. Already existing fake news classification methods can be classified into four main categories: Style-based, Knowledge-based, Source-based, and Propagation-based. There are various fact-checking websites as well which work on different algorithms and analyse different types of content. Some of the popular ones are Politifact, The Washington Post Fact Checker, FactCheck, Snopes, TruthOrFiction, FullFact, HoaxSlayer, and GossipCop. Most misinformation detection techniques have been seen to use the datasets obtained from Politifact and GossipCop since these are the closest ones to the subject of "problem-causing misinformation." Apart from automatic fact-checking methods, manual fact-checking also exists. This is done in cases where accurate classification is the top priority and is generally conducted by experts in that news domain. The chances of excluding manual inspection are very low since the tools used to identify fake news are not particularly accurate. Methods for fake news detection include Supervised Learning, Decision Trees, Graph Neural Networks, Propagation Networks, Deep Learning, and multiple other ML techniques. New methods are formulated as the input type changes. Most of the methods consider only the news content as the input, but to obtain more accurate results the news source cannot be ignored. Moreover, the most recent developments around this problem statement have seen the rise of graph propagation networks which track the spread of fake news across social media and convert them into a numerical representation which is then fed into the ML models.

[3] The research looks at integrating the user credibility levels (credibility score denotes "the quality of being trustworthy") to track the user-news interactions and has the potential to better misinformation prediction. Framework TriFN for modelling tri-relationship for misinformation detection comprises of five major components: News contents embedding, user embedding, user-news interaction embedding, publisher-news relation embedding, and a semi-supervised classification component. Methods for detecting misinformation mostly rely on social contexts and news content. Real news may be distinguished from false information using clues found in news content. In news content-based techniques, features are extracted as linguistic and visual-based features. Linguistic-based features identify various writing styles and astonishing news reports that occur in misinformation, such as lexical and syntactic features. Visual-based features attempt to identify fake images that are deliberately created or identify characteristics of images that are generally present in misinformation. In social context-based approaches, there are three main features: User-based features, post-based features, and network-based features. With the purpose of figuring out their characteristics. and integrity, user-based features are taken from user profiles.

[9] The objective of this study is to understand the relationship between users' profiles on social media and misinformation. Firstly, we measure users sharing behaviour i.e., users who share true news and misinformation, along with it we also gather implicit and explicit user behaviour and perform a comparative analysis on them. Through feature importance analysis, we further validate these features' efficacy. We collect and analyse user profile features from implicit and explicit aspects. There are no implicit features readily available but are derived from users' online behaviours, such as tweets. Age, personality, region, profile image, and political bias are some of the implicit features. Explicit features are directly extracted from the meta-data given by social media API queries. We try to identify differentiation between users who share real news and misinformation, which further is used to characterize discriminative features for misinformation detection (Filtering Bot Users, Identifying User Groups) By using the Gini impurity to create a feature significance score, we examine feature importance in the Random Forest. (Register Time, Verified, Political Bias, Personality, Status Count).

[2] This paper provides a thorough examination of how to spot fake news on social media, including explanations of fake news based on psychological and sociological theories, current data mining methods, evaluation criteria, and representative datasets.

The detection of fake news is said to as a binary classification problem since it is simply a bias on information that has been skewed by the publisher. Next, a general framework for data mining consists of the phases of model construction and feature extraction for the detection of fake news. Different types of feature representations can be constructed based on these content attributes to extract distinguishing traits of fake news. We classify current techniques into News Content Models and Social Context Models based on their main input sources. In the News Content Models method, we use existing factual sources and news content features primarily to identify fake news. Available methods can be divided into Knowledge-based (Knowledge-based strategies try to leverage outside sources to verify assertions made in news material.) and Style-based (Style-based approaches aim to identify manipulators in the writing of news content in order to identify fake news.) Social media's characteristics give academics new tools to complement and improve News Content Models. Social context models incorporate pertinent user social interactions into the study, capturing this supplemental data from various angles.

*[5]* The primary objective of this study is to study the methods for misinformation detection. Many of them depend on identifying user traits, content, and context that point to misinformation. We also examine datasets that have previously been employed to classify fake news. Analyzing false news information alone will not be enough to develop a dependable and efficient detecting system. Thus, in order to have a thorough understanding of online social data, additional essential significant factors are explored in this study, including author and user analysis and news context. Resources for fact-checking are often used by major media outlets. Real-time news is often a combination of facts, therefore sometimes the whole issue cannot be properly explained by a binary classification result. In the most recent fact-checking resources, a wide range of evaluation criteria or visual metrics is utilised to assess the news's level of veracity. Fact-checking is a useful tool for spotting fake news since it informs readers of what is accurate, false, or in-between. For the classification of online hoaxes, frauds, and misleading information, supervised machine learning algorithms such as Random Forest, Logistic Regression, Support Vector Machine (SVM), Decision Trees, and K-nearest Neighbour have been widely used in the past. Unsupervised learning is more realistic and practical for tackling problems in the real world. However, there aren't many studies that focus specifically on identifying bogus news on the internet without supervision. The majority of them concentrate on sentiment analysis or semantic similarity analysis. By integrating word similarity and word-order similarity, an unsupervised similarity assessment for online fraudulent reviews may successfully identify almost identical internet assessments. The correlations between visual information and important contextual data may be used to predict the sentiment of social photographs from two large-scale datasets in an unsupervised sentiment analysis framework for social media photos.

*[4]* Through the use of a framework (SpotFake) for a multimodal false news detection system, this research presents a method to identify fake news. The SpotFake framework learns textual features from the provided article or post using language models like BERT (Bidirectional Encoder Representations from Transformers).and VGG-19 pre-trained on the ImageNet dataset to incorporate image features. After this, the desired news vector is then created by concatenating the representations of both. Finally, this news vector is used for classification. Two publicly accessible datasets, the Twitter MediaEval Dataset and the Weibo Dataset, are used to train SpotFake. There are three sub-modules in SpotFake. An extractor for textual features is the first sub-module that uses the BERT language model to extract contextual text features. The visual feature extractor, which is the second sub-module, uses pre-trained VGG-19 to extract the visual features from a post. A multimodal fusion module, which forms the last sub-module, integrates the representations received from various modalities to create a news feature vector. SpotFake extracts data using pre-trained ImageNet models and a language transformer model, then classifies data using a fully connected layer. It performs better than the baselines on average by 6% accuracy.

*[8]* The purpose of this study is to recognise and comprehend social media posts that include fake news. In this research paper, the features are extracted from news stories, including sources and social media posts. Features for fake news detection are extracted from news content, news source, and the environment. From news content, Textual Features are extracted that include language Features (Syntax), Lexical Features, Psycholinguistic Features like Linguistic Inquiry and Word Count (LIWC), Semantic Features, and Subjectivity. From News Source Features, this set is composed of three features also called domain localization Bias, Credibility and Trustworthiness, and Domain Location. Environment Features contains user statistics about the user engagement involvement with the social media handles. The Environment features contain two features. Engagement, Temporal Patterns. After the extraction of the features, the researchers used various classifiers to check the most suitable classifier among them, which includes Naive Bayes (NB), K-Nearest Neighbours (KNN), Support Vector Machine with RBF kernel (SVM), Random Forests (RF), and XGBoost (XGB). Since they had already extracted the features already, they didn't use any neural network. They measured the effectiveness of each classifier using the area under the ROC curve (AUC) and the Macro F1 score. The trade-off between true and false positive rates can be controlled using the decision threshold; hence the AUC is employed.

The F1 score combines recall and precision for each class into a single metric, and the Macro F1 score shows how well the classifier performed overall. Examining the ROC for XGB classifier, they found that it can categorise nearly all of the false information while incorrectly categorising 40% of the actual information.

## III. DATASETS

### *A. Benjamin Political News Dataset*

For the purpose of identifying political and satirical tales online, Horne and Adali (2017) developed this dataset. The dataset includes 75 articles from the news categories listed below: false, satirical, and real. Real sources were collected from "most trusted" list of Business Insider, while fake sources were gathered from fake and misleading news sites list managed by Zimdar (Fake misleading clickbait or satirical information sources).

### *B. Burfoot Satire News Dataset*

Burfoot and Baldwin (2009) manually collected this dataset, which includes 233 satirical news items and 4000 real news samples. The Genuine news items are collected from English Gigaword Corpus, using newswire documents samples and the satire news stories closely related in subject to the existent ones are chosen (Horne & Adali,2017).

### *C. BuzzFeed News*

More than 2000 news samples from Facebook are included in this dataset, which was uploaded in September 2016. (Horne & Adali, 2017; BuzzFeed news).

All news samples have been rated as mainly true, not factual, conflicting between right and wrong content, and mostly false by professional BuzzFeed reporters.

This dataset also offers other pertinent information for each news sample, including the URLs of the news posts, published data, the amount of shares, responses, and comments.

### *D. Cred Bank Dataset*

Streaming tweets from the period of October 2014 to February 2015 are included in this corpus. The whole collection consists of exceeding 60 million tweets that cover 1049 real-world occurrences. Thirty annotators assess the veracity of the tweets.

### *E. Fake News Challenge Dataset*

Streaming tweets from the period of October 2014 to February 2015 are included in this corpus. The whole collection consists of exceeding 60 million tweets that cover 1049 real-world occurrences. Thirty annotators assess the veracity of the tweets.

### *F. FakeNewsNet*

This dataset, which includes 211 false reports and 211 verified reports that were gathered from BuzzFeed.com and PolitiFact.com, is made available by Shu et al. (2017). Additionally, it includes crucial elements for each news sample, such as user details, news content and details on social media activities. In Shu, Wang, and Liu, the dataset has a more in-depth examination (2017).

### *G. LIAR*

For the purpose of identifying bogus news online, the data is put forth and provided in Wang (2017). It includes 12,800 hand labelled brief remarks from PolitiFact.com in a variety of circumstances. The six ratings for each data sample are true, mostly true, half true, barely true, false, and pants-fire. For each case, a thorough analytical report and references to the original papers are also provided.

### *H. ISOT Fake News Dataset*

Contains about 44,000 news stories that have been categorised as both factual news and fake news. The dataset used consists of aggregated news from various sources, cleaned and processed for reliability.

While the true news was obtained from his website at Reuters, the false news was obtained from dubious websites as stated by PolitiFact and Wikipedia.

### *I. Twitter MediaEval Dataset*

As a result of the MediaEval challenge to verify multimedia usage, this data collection was made available. It consists about 17,000 unique tweets on various events. The training set consists of 9,000 and 6,000 fake news and actual news tweets respectively, while the test set consists of 2000 news tweets.

### *J. Weibo Dataset*

China's Xinhua News Agency and Weibo, a microblogging platform, are reputable news sources where real news is obtained. Weibo gathered fake news from May 2012 to June 2016. The official Weibo mechanism for disproving rumours verifies the messages that have been collected.

## IV. METHODOLOGIES

Our main aim is to classify a news article as true or false. In the previous sections, we have discussed fake news, its meaning, effects and why we need an accurate misinformation classification system. We also went through the various different datasets used in different approaches. Now, in this section, we will cover the actual methodologies used for classifying the news articles.

*[1]* This paper deals with Misinformation classification using WOA-xgbTree algorithm and content features. Content features are those features which define the news article, i.e., the useful and unique features. These features are extracted and fed into the xgbTree algorithm, powered up by the Whale Optimization Algorithm (WOA). The main highlight of this method is that it requires only the news content to work, unlike other methods which require the source and its characteristics as well. According to Saeid Sheikhi, a dataset (ISOT Fake News Dataset) of nearly 44000 recently obtained news articles, both authentic and false, was used for testing and training this model, and the model achieved a good accuracy and F1 score and was able to successfully classify over 91% of the articles. The content features are extracted using NLP techniques. The features extracted are: special symbols, articles, keywords, numbers, special keywords, offensive or swear words, abbreviations, ellipsis, uppercased words, length and number of words. As the name suggests, the Whale Optimization Algorithm is inspired by a whale, more specifically the Humpback whale's feeding behaviour. The hunting behaviour of these whales is termed "bubble mesh feeding" since it involves a loop path or an involute shaped path in the search space. This method is found to improve the search results and hence is taken into consideration as a good efficiency boosting algorithm. XgbTree algorithm stands for Extreme Gradient Boosting Tree algorithm and is an optimized tree boosting ensemble method, i.e., it takes inputs from multiple ML models to produce the most efficient results. The main feature of this algorithm is that it produces decision trees sequentially. It is a scalable implementation of an old algorithm and is useful for multiple Supervised Learning problems involving mainly classification and regression. Moreover, this algorithm shows good compatibility with the WOA.

*[4]* This research introduces a solution to detect Fake News using a framework (SpotFake) for multimodal fake news detection system. The SpotFake framework uses language models like BERT (Bidirectional Encoder Representations from Transformers) to learn textual features in the given article or post and VGG-19 to incorporate image features. After this, the representations of the both are then concatenated together to produce the desired news vector. Finally, this news vector is used for classification. SpotFake uses language transformer models and pre-trained ImageNet models for extraction and classifies using fully connected layers.

*[2]* Problem Definition: Given the social news engagements E among n users for news article a, the task of fake news detection is to predict whether the news article a is a fake news piece or not, i.e., F: E → {0, 1} such that,

$$\mathcal{F}(a) = \begin{cases} 1, & \text{if } a \text{ is a piece of fake news,} \\ 0, & \text{otherwise.} \end{cases}$$

where F is the prediction function, which we want to study. Because fake news is essentially a bias on information that has been manipulated by the publisher, fake news detection is defined as a binary classification problem. Next, a general framework for data mining that consists of the phases of model construction and feature extraction for the detection of fake news. The goal of the feature extraction phase is to formalize the representation of news content and related supporting data into a mathematical structure. The model construction phase then develops machine learning models to more effectively distinguish between fake and legitimate information based on the feature representations. Feature Extraction: Features in news content explain the metadata associated with a piece of news. Some typical news content features are source, headline, body text and image/video etc. Different types of feature representations can be constructed based on these raw content attributes to extract distinguishing traits of fake news. Typically, the news content we are looking at will mostly be Linguistic-based and/or visual-based.

The user-driven social interactions of news consumption on social media platforms may also be used to infer additional social context elements for the news items. Social interactions depict how news spreads over time, which offers helpful supplementary data to estimate the accuracy of news items. Users, generated posts, and networks are the three main facets of the social media setting that we wish to depict.

*[7]* Features for fake news detection are taken from: 1) news content (e.g., language processing techniques). 2) news source (e.g., reliability and trustworthiness). 3) environment (e.g., social network structure). From news content, the textual features extracted include : i) Language Features (Syntax) ii) Lexical Features iii) Psycholinguistic Features: Linguistic Inquiry and Word Count (LIWC) iv) Semantic Features v) Subjectivity From News Source Features, this set is composed of three features and is also called domain localisation : i) Bias ii) Credibility and Trustworthiness iii) Domain Location Environment Features contains user statistics about the user engagement involvement with the social media handles. The Environment features contains two features: i) Engagement ii) Temporal Patterns After the extraction of the features, the researchers used various classifier to check the most suitable classifier among them, which includes k-Nearest Neighbours (KNN), Naive Bayes (NB), Random Forests (RF), Support Vector Machine with RBF kernel (SVM), and XGBoost (XGB). Since they had already extracted the features already, they didn't use any neural network. They measured the effectiveness of each classifier using the area under the ROC curve (AUC) and the Macro F1 score.

*[6]* Already existing fake news classification methods can be classified into four main categories: Knowledge-based, Style-based, Propagation-based and Source-based. The ML models or NLP techniques used for each of them heavily depend on these styles and the type of input data or fake news features accumulated. The propagation of fake news has been studied scientifically as well as psychologically and many theories and “effects” have been identified. The common ones include Bandwagon Effect, Normative Influence Theory, Social Identity Theory, Availability Cascade, Validity Effect, Echo Chamber Effect, Confirmation Bias and the Selective Exposure Effects. These terms are quite common in human psychology and the studies of public behaviour and the spread of information among masses. There are various fact-checking websites as well which work on different algorithms and analyze different types of content. Some of the popular ones are PolitiFact, The Washington Post Fact Checker, FactCheck, Snopes, TruthOrFiction, FullFact, HoaxSlayer and GossipCop. Most of the misinformation detection techniques have been seen to use the datasets obtained from Politifact and GossipCop since these are the closest ones to the subject of “problem causing misinformation”. Methods for fake news detection include Supervised Learning, Decision Trees, Graph Neural Networks, Propagation Networks, Deep Learning and multiple other ML techniques. New methods are formulated as the input type changes. Most of the methods consider only the news content as the input, but to obtain more accurate results the news source cannot be ignored. Moreover, the most recent developments around this problem statement have seen the rise of graph propagation networks which track the spread of fake news across social media and convert them into a numerical representation which is then feeded into the ML models. Research and development around this method is still under progress but it has seen good results up till now. Moreover, it makes perfect sense to analyze the spread of fake news if we want to track down the source and the users creating and spreading it.

*[5]* For the classification of online hoaxes, frauds, and misleading information, supervised machine learning algorithms have been widely used in many past fake news classification models. A labeled dataset's quality has a significant impact on how well a supervised learning model performs. For the following reasons, it is challenging to produce a comprehensive, high-quality dataset for false news identification. The real-world online dataset is often vast, imperfect, unstructured, unlabeled, and noisy. Every day, social media generates a significant volume of misleading information with various purposes and various linguistic features. Getting the data's ground truth label is challenging. Moreover, unsupervised learning is more realistic and practical for tackling problems in the real world.

## V. CONCLUSION

In this survey, we looked at multiple possibilities to correctly verify fake news with different methodologies and also calculating the effectiveness against given testing datasets. Using these techniques, we can reduce the spread of fake news at the earliest. As there are also, many news reliability firms present but in the process of manually dealing with the details it takes too much time. So, to reduce this time gap we need to implement different automated tools, machine learning or deep learning models which can precisely compute the reliability of the news.

## REFERENCES

[1] Saeid Sheikhi. “An effective fake news detection method using WOA-xgbTree algorithm and content-based features.” at Applied Soft Computing 109 (2021): 107559

[2] Shu, Kai, et al. "Fake news detection on social media: A data mining perspective." ACM SIGKDD explorations newsletter 19.1 (2017): 22-36.

[3] Kai Shu, Suhang Wang, Huan Liu. “Beyond News Contents: The Role of Social Context for Fake News Detection.” at Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining (2019): 312-320

[4] Shivangi Singhal, Rajiv Ratn Shah, Tanmoy Chakraborty, Ponnurangam Kumaraguru, Shin'ichi Satoh. "SpotFake: A Multi-modal Framework for Fake News Detection." at IEEE Fifth International Conference on Multimedia Big Data (2019)
[5] Zhang, Xichen, and Ali A. Ghorbani. "An overview of online fake news: Characterization, detection, and discussion." Information Processing & Management 57.2 (2020): 102025.
[6] Zhou, Xinyi, and Reza Zafarani. "A survey of fake news: Fundamental theories, detection methods, and opportunities." ACM Computing Surveys (CSUR) 53.5 (2020): 1-40.
[7] Reis, Julio CS, et al. "Supervised learning for fake news detection." IEEE Intelligent Systems 34.2 (2019): 76-81.
[8] Shu, Kai, et al. "The role of user profiles for fake news detection." Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining. 2019.
[9] Pérez-Rosas, Verónica, et al. "Automatic detection of fake news." arXiv preprint arXiv:1708.07104 (2017).
[10] Jain, Akshay, and Amey Kasbe. "Fake news detection." 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, 2018.
[11] Oshikawa, Ray, Jing Qian, and William Yang Wang. "A survey on natural language processing for fake news detection." arXiv preprint arXiv:1811.00770 (2018).
[12] Shu, Kai, Suhang Wang, and Huan Liu. "Exploiting tri-relationship for fake news detection." arXiv preprint arXiv:1712.07709 8 (2017).
[13] Kaliyar, Rohit Kumar, et al. "FNDNet–a deep convolutional neural network for fake news detection." Cognitive Systems Research 61 (2020): 32-44.
[14] Zhou, Xinyi, and Reza Zafarani. "Network-based fake news detection: A pattern-driven approach." ACM SIGKDD explorations newsletter 21.2 (2019): 48-60.
[15] Ruchansky, Natali, Sungyong Seo, and Yan Liu. "Csi: A hybrid deep model for fake news detection." Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017.
[16] Karimi, Hamid, et al. "Multi-source multi-class fake news detection." Proceedings of the 27th international conference on computational linguistics. 2018.
[17] Thota, Aswini, et al. "Fake news detection: a deep learning approach." SMU Data Science Review 1.3 (2018): 10.
[18] Shu, Kai, Suhang Wang, and Huan Liu. "Understanding user profiles on social media for fake news detection." 2018 IEEE conference on multimedia information processing and retrieval (MIPR). IEEE, 2018.
[19] Shu, Kai, et al. "defend: Explainable fake news detection." Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019.
[20] Monti, Federico, et al. "Fake news detection on social media using geometric deep learning." arXiv preprint arXiv:1902.06673 (2019).
[21] Granik, Mykhailo, and Volodymyr Mesyura. "Fake news detection using naive Bayes classifier." 2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON). IEEE, 2017.
[22] Parikh, Shivam B., and Pradeep K. Atrey. "Media-rich fake news detection: A survey." 2018 IEEE conference on multimedia information processing and retrieval (MIPR). IEEE, 2018.
[23] Wang, William Yang. "" liar, liar pants on fire": A new benchmark dataset for fake news detection." arXiv preprint arXiv:1705.00648 (2017).
[24] Khattar, Dhruv, et al. "Mvae: Multimodal variational autoencoder for fake news detection." The world wide web conference. 2019.
[25] Gravanis, Georgios, et al. "Behind the cues: A benchmarking study for fake news detection." Expert Systems with Applications 128 (2019): 201-213.
[26] Zhou, Xinyi, and Reza Zafarani. "Fake news: A survey of research, detection methods, and opportunities." arXiv preprint arXiv:1812.00315 2 (2018).
[27] Faustini, Pedro Henrique Arruda, and Thiago Ferreira Covões. "Fake news detection in multiple platforms and languages." Expert Systems with Applications 158 (2020): 113503.
[28] Singhania, Sneha, Nigel Fernandez, and Shrisha Rao. "3han: A deep neural network for fake news detection." International conference on neural information processing. Springer, Cham, 2017.
[29] Karimi, Hamid, and Jiliang Tang. "Learning hierarchical discourse-level structure for fake news detection." arXiv preprint arXiv:1903.07389 (2019).
[30] Sitaula, Niraj, et al. "Credibility-based fake news detection." Disinformation, Misinformation, and Fake News in Social Media. Springer, Cham, 2020. 163-182.