Fake News Detection on Social Media: A Systematic Survey

Mohamed K. Elhadad
Dept. of Elec. and Comp. Eng.
University of Victoria
Victoria, Canada
melhaddad@uvic.ca

Kin Fun Li
Dept. of Elec. and Comp. Eng.
University of Victoria
Victoria, Canada
kinli@uvic.ca

Fayez Gebali
Dept. of Elec. and Comp. Eng.
University of Victoria
Victoria, Canada
fayez@uvic.ca

***Abstract*—These days there are instabilities in many societies in the world, either because of political, economic, and other societal issues. The advance in mobile technology has enabled social media to play a vital role in organizing activities in favour or against certain parties or countries. Many researchers see the need to develop automated systems that are capable of detecting and tracking fake news on social media. In this paper, we introduce a systematic survey on the process of fake news detection on social media. The types of data and the categories of features used in the detection model, as well as benchmark datasets are discussed.**

***Index Terms*—Fake News Detection, Social Media, Misinformation, Disinformation, Mal-information**

I. Introduction.

Nowadays all over the world, especially in developing countries online social media platforms (Facebook, Twitter, Snapchat, YouTube, etc.) has become the main source for people to get news from, accounted for almost 62 percent of the users [1]. Moreover, these online social media platforms are suitable places where users can share their emotions, stories and concerns, and provide better means to get quick response and feedback on different global issues, even though some untruthful thoughts may be put forward purposely [2].

Social media play a vital role in economics, politics and social developments, for example, the mobilization of pro-democratic movements (as in Egypt in 2011 and 2013, or Ukraine in 2013–2014), however, it also allows misleading information to travel more broadly, inexpensively and lethally than traditional news sources. The spread of misleading information on the web and social media could affect the operation of the financial markets, the response time during critical situations, and terrorist attacks, etc. [3]. The main goal of such misleading information is to cause financial or political damage to public figures, agencies, or even countries. By making some interesting and often fabricated headline, viewers are attracted to it thus the headline becomes a trending item [4]. Furthermore, misleading information could be used in information warfare between countries, companies and institutions.

The term fake news is very old. It gained a bad reputation and became a popular term during the 2016 US presidential election campaign [5]. Fake news has been widely used to indicate information pollution in the form of false news, hoaxes, propaganda, rumours, and junk news [6]. Although all these terms indicate deceiving information, we believe that there are still no agreed definition and categorization for it. Hence, in this paper, we introduce a general categorization of fake news into Disinformation, Misinformation, and Mal-information. Disinformation is defined as false information created and shared by people with harmful intent. Misinformation is defined as some kind of false information disseminated online by people who do not have an ill intent, while Mal-information is defined as the sharing of genuine information with the intent to cause harm. [7–11]. In TABLE I we make a comparison of the different types of fake news and the number of publication since 2004 till now, as well as the corresponding category or categories assigned. The information provided is obtained from queries to Google Scholar.

Active research has been ongoing for the development of automated, reliable, and accurate techniques for detecting fake news on social media. The detection of fake news could be defined as the process of estimating whether a particular news article of any topic, from any domain, is being intentional or unintentionally misleading. Most of the fake news detection systems deploy machine learning techniques to assist users in filtering the news they are viewing, and detecting whether a particular news article is deceptive or not. This classification and analysis are done based on comparing a given news article with some pre-known news corpora that contain both misleading and truthful news articles [18].

In this paper, we present a systematic survey on fake/false news on social media and their detection focusing on the research from 2017 till date. In section II, we introduce the research methodology that we follow. In section III, we introduce and discuss our research findings. Finally, section IV concludes and suggests future work.

II. Research Methodology

We conduct a systematic review on Fake/False news detection on social media, based on the research protocol Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) [19]. This systematic process is divided into five stages, as follows:

(i) Stage 1: Defining Article Eligibility Criteria determined by the Inclusion/Exclusion Criteria (IEC), which are:

978-1-7281-2794-1/19/$31.00 ©2019 IEEE

TABLE I
COMPARATIVE DESCRIPTION OF TERMINOLOGIES FOR FAKE NEWS AND THEIR RELATED REFERENCES IN SOCIAL MEDIA

Term	Definition	research count related to social media	Category—		
			Disinfo.	Misinfo.	Malinfo.
Fake/False news	fabricated news articles that could be potentially or intentionally misleading for the readers, as they mimic traditional news content in form but not in the intent or the organizational process.[6]	269100	✓	✓	✓
Hoax	a fiction intentionally fabricated to masquerade as the truth[12]	18300	✓	✓	
Propaganda	news stories which are created to influence the emotions, opinions and actions of the target audiences by means of deception, in selectively omitting or providing one–sided messages for political, ideological or religious purposes.[13]	142000	✓		✓
Satire/Parody	a form of news which is written with the goal of entertaining or criticizing the readers, and it could mimic genuine news; this is harmful when shared out of context. [13]	61400			✓
Rumors	originated from a Latin word that means noise, and has been identified by some scholars as a subset of propaganda [14]. As unverified claim that did not originate from news events, it could spread from one user to another [15]	39900	✓	✓	
Click-bait	low quality journalism which is intended to attract traffic and benefit from advertising revenue[16]	2110	✓		
Junk news	more generic and aggregates several types of information; it usually refers to the overall content that pertains to a publisher rather than a single article[17]	28300	✓		✓

a) IEC-1: the research paper must be written in English.
b) IEC-2: the research paper must be original work from a very technical perspective.
c) IEC-3: the research paper must focus on the detection and identification of fake/false news in social media platforms.
d) IEC-4: the research paper must be published between 2017 and 2019 to cover the period after the 2016 US presidential election.
e) IEC-5: the article has a purpose to analyze a method or approach to perform fake news detection in social media.

(ii) Stage 2: Defining Information Resource
a) The literature could be searched on any online database with a significant repository for an academic study, in our case the Google Scholar scientific database, as it contains the results from other scientific databases such as IEEE Xplore, ScienceDirect, ACM Digital Library, etc.
b) The IEC qualified articles are used to find further related research.

(iii) Stage 3: Literature Selection, as follows:
a) Use the set of key words as shown in TABLE II.
b) Apply IECs on the resulting papers to eliminate non-eligible papers.
c) Read the title and abstract of the articles that are not eliminated from the previous stage, in order to determine whether the items are eligible for the next review.
d) Short-listed articles are re-assessed to find related studies. The articles that reference-listed and associated with the short-listed articles are re-assessed by repeating stage 3 and stage 4

(iv) Stage 4: Data Collection: Data are queried manually, using the keywords determined in stage 3 part (a), by creating a data extraction form. As shown in TABLE II, the queries result in a total of 152640 items based on different keywords used. Some of results are redundant and need to be investigated using the IEC as in stage 1.

(v) Stage 5: Data Item Selection: Data are obtained from short-listed articles that consist of method or approach used for detecting fake/false news on social media, which satisfy the constraints in stage 1 and stage 3. After refining the used queries for all keywords, the candidate articles are obtained. After reviewing the titles and abstracts, 196 articles are selected as shown in TABLE II.

TABLE II
COLLECTED DATA

No.	Used Key Words	number of studies found	Candidate research	Selected
1	fake news detection on social media	10500	488	182
2	detection of fake news on social media	17000	146	2
3	identify fake news on social media	17100	251	4
4	identifying fake news on social media	16800	162	3
5	fake news detection on twitter	5640	431	5
6	false news detection on social media	17000	10	0
7	detection of false news on social media	16900	9	0
8	identify false news on social media	17300	18	0
9	identifying false news on social media	17100	10	0
10	false news detection on twitter	17300	6	0

Then the content of these articles are reviewed and finally 47 research papers are fully investigated in this research.

III. RESEARCH FINDINGS AND DISCUSSION

Fake news detection is not simple but rather is a complicated problem [3, 20–22]. The detection task requires several steps to classify a given set of news articles (text documents) [21, 23, 24]. The preprocessing of the collected news documents will vary depending on the type of data and the language used in the documents. Fig. 1 shows the typical types of data in news documents. Generally, most of news articles contain textual data. Therefore, the documents must be transformed to another representation, in order to be able to extract feature vectors that contain enough information to ensure accurate classification, and suitable to be maintained by machines.

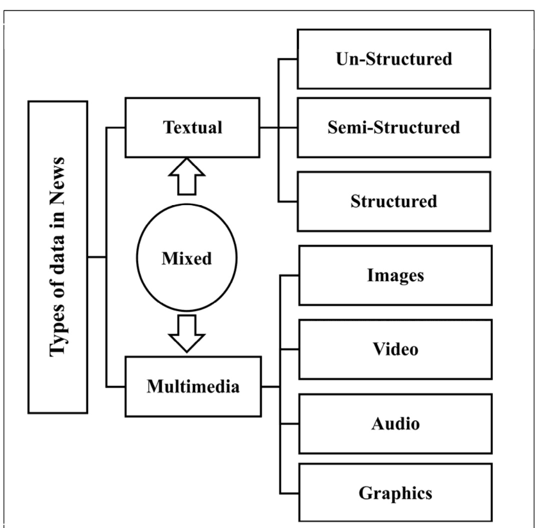

Fig. 1. Types of data in News.

Carlos Castillo et. al. (2011) [25], Kai Shu et. al. (2017) [4], and Ugur Kursuncu et. al. (2019) [26], all gave a categorization for the types of extracted features according to their perspective. From our point of view, the extracted features could be categorized into content-based features, context-based features, and domain-based features, depending on the application and the type of data in the news document as shown in Fig. 2. The content dependent features are meta-information that represent the raw content of a news document including the author, editor, publisher, news title, the body of the news itself, and any attached multimedia. So, depending on the content that is available in our data, we could extract different kinds of features. Lexical and syntactic features, such as number of URLS, word length, hashtag, retweet, and term frequency, can be extracted from the title and the body of the news document. While from any attached multimedia, we could extract visual and statistical features with information such as clarity score, coherence score, similarity distribution histogram, image ratio, multi-image ratio, etc. [4].

The author/editor/publisher of the news document beside being a part of the news content, they are also highly related to the news context. News context is the social engagements of the news article consumption on social media platform. This social engagement represents the news propagation over time and the group of users that engaged with this news document. Hence, we could extract from individuals, group, and postings, social based features such as number of followers, friends count, registration age, number of authored posts/tweets, related social groups, demographic information, user stance, average credibility scores, etc. Moreover, we could extract features about the domain that the news article belongs to, by extracting propagation features that consider characteristics related to the propagation tree, that can be built from the retweets of a message in a certain domain. These include features such as the depth of the re-tweet tree and the number of initial tweets of a topic. Some of the studies on text classification in general, used the weighted feature vectors to improve classification results [27–35]. The weight of each feature indicates the feature's importance thus enhancing the classification results.

After an in-depth review of the collected papers, our findings are summarized in TABLE III, with respect to the referenced work, publishing year, research type, the publisher, approach used, the social media platform, and types of extracted features, ordered by publishing year. It can be observed that Twitter is the most popular target for recent research on social media networks. This could be attributed to the extraction of

TABLE III
SUMMARIZATION OF RESEARCH FINDINGS

Reference	Year	Type	Publisher	Approach	Social Media Platform	Features Type
[36]	2017	Conference	NOBIDS	Survey	-	Content, Context, Domain
[4]	2017	Journal	ACM	Survey	Multiple	Content, Context, Domain
[37]	2017	Conference	AIS	Supervised Learning, Deep Learning	Facebook	Content
[38]	2017	Conference	IEEE	Supervised Learning, Other	Twitter, Facebook	Content, Context, Domain
[39]	2017	Journal	arXiv	Other	Facebook, Twitter	Content, Context
[21]	2018	Journal	arXiv	Survey	Multiple	Content
[40]	2018	Journal	ELSEVIER	Supervised Learning	Multiple	Content
[41]	2018	Conference	IEEE	Supervised Learning	Twitter	Content
[42]	2018	Conference	Springer	Supervised Learning	Multiple	Content
[43]	2018	Journal	arXiv	Supervised Learning, Deep Learning, Other	Multiple	Content
[44]	2018	Conference	IEEE	Supervised Learning	Twitter, Facebook	Content, Context
[45]	2018	thesis	ProQuest	Supervised Learning	Twitter, Facebook	Content
[46]	2018	Journal	ELSEVIER	Other	Twitter	Content, Context
[47]	2018	Conference	Springer	Deep Learning	Twitter	Content, Context
[48]	2018	Conference	IEEE	Survey	Twitter, Facebook	Content, Context
[49]	2018	Conference	ACM	Deep Learning	Twitter, Sina Weibo	Content
[20]	2018	Journal	ACM	Survey	Multiple	Content, Context, Domain
[50]	2018	Conference	IEEE	Supervised Learning, Other	Twitter, Facebook	Content, Context, Domain
[51]	2018	Conference	Springer	Survey	-	Content, Context, Domain
[3]	2018	Journal	arXiv	Survey	Multiple	Content, Context, Domain
[52]	2018	Journal	arXiv	Supervised Learning	Twitter	Content, Context
[53]	2018	Conference	IEEE	Supervised Learning, Deep Learning, Other	Twitter, Facebook	Content, Context
[54]	2018	Journal	arXiv	Deep Learning	Twitter	Content, Context
[55]	2018	Conference	ACM	Supervised Learning, Other	Twitter	Content
[56]	2018	Conference	Springer	Deep Learning	Multiple	Content, Context
[57]	2018	Conference	ACL	Supervised Learning, Deep Learning, Other	Multiple	Content, Context
[58]	2018	Conference	ACM	Deep Learning, Other	Twitter	Content, Context, Domain
[59]	2018	Conference	IEEE	Other	Twitter	Content
[60]	2018	Journal	arXiv	Supervised Learning, Deep Learning	Multiple	Content, Context, Domain
[61]	2019	Journal	arXiv	Supervised Learning	Multiple	Content, Context
[62]	2019	Conference	HICSS	Supervised Learning	Multiple	Content, Context
[6]	2019	Journal	arXiv	Survey	Multiple	Content, Context
[63]	2019	Journal	arXiv	Supervised Learning, Deep Learning	Twitter	Content, Context
[64]	2019	Journal	arXiv	Supervised Learning	Multiple	Domain
[65]	2019	Journal	Springer	Supervised Learning, Deep Learning	Twitter, Facebook	Content, Context
[66]	2019	Conference	IEEE	Supervised Learning	Multiple	Content
[67]	2019	Journal	HAL	Supervised Learning, Deep Learning	Multiple	Content
[68]	2019	Journal	arXiv	Supervised Learning, Deep Learning, Other	Twitter, Facebook	Content, Context, Domain
[69]	2019	Journal	Nature	Supervised Learning	Twitter	Content, Context
[70]	2019	Conference	AAAI	Other	Multiple	Content
[71]	2019	Conference	AAAI	Deep Learning	Twitter, Sina Weibo	Content
[72]	2019	Conference	IEEE	Supervised Learning	Twitter	Content
[73]	2019	Journal	IEEE	Supervised Learning	Facebook	Content, Context, Domain
[74] [75]	2019 2019	Conference Conference	HICSS ACM	Survey Other	Multiple Twitter, Facebook	Content, Context Content, Context
[76]	2019	Conference	ACM	Supervised Learning	Multiple	Content

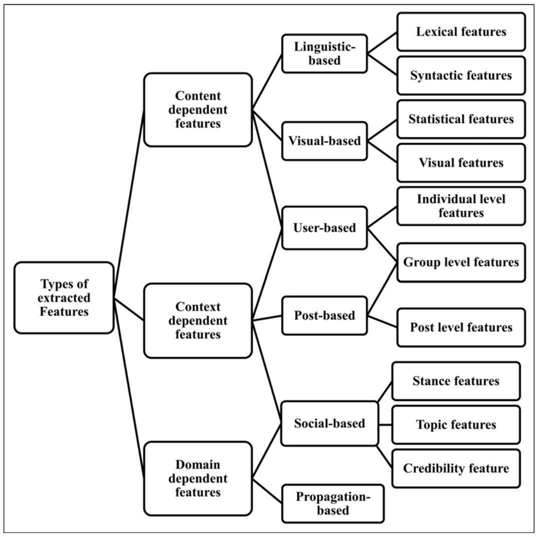

Fig. 2. Types of extracted features.

data from it using its API is much easier than from Facebook and other platforms. Based on our categorization shown in Fig. 2, the following observations can be made:

(i) Nearly 26 percent of the investigated papers extract content, context, and domain dependent features in their detection systems. Techniques used include supervised learning such as Support Vector Machine and Logistic/Linear Regression, deep learning such as recurrent and convolutional neural network, and other techniques such as Long Short Term Memory.

(ii) Nearly 36 percent of the investigated papers extract content and context dependent features in building their detection systems as follows:

  a) Almost 28 percent of them are using the supervised learning techniques such as Support Vector Machine, logistic/linear regression, random forest, etc.

  b) Almost 17 percent of them are using deep learning techniques including recurrent and convolutional neural networks. The same percentage are using other techniques such as Linguistic Inquiry and Word Count (LIWC), Rhetorical Structure Theory (RST), etc.

  c) Almost 21 percent of them are using a combination of supervised learning, deep learning, and other techniques in their experiments.

(iii) Nearly 36 percent of the investigated papers extract only the content dependent features in their detection systems as follows:

  a) Almost 50 percent of them are using the supervised learning techniques such as Support Vector Machines, logistic/linear regression, and Random Forest.

  b) Almost 12.5 percent of them are using deep learning techniques such as recurrent and convolutional neural networks, and the same percentage use other techniques including Long Short Term Memory, and TraceMiner, etc.

  c) Almost 25 percent of them are using in their experiments a combination of supervised learning, deep learning, and other techniques.

Moreover, it is clear that content dependent features are the most commonly used features, since most of the available data is textual data in either structured, semi-structured or unstructured formats. Almost 80 percent of the available textual data are in unstructured format [23], and the supervised learning techniques are extensively used in building the detection models for them. Finally, the main and general challenge for building and testing the effectiveness of any detection system is the lack of a generalized and standard dataset for fake news detection [4, 20–22, 77, 78].

Additionally, we observe from the investigated research that the main challenge facing the researchers in implementing new techniques for fake news detection is the poor quality of the existing data. The most commonly used datasets are shown in TABLE IV. Among these, LIAR, BuzzfeedNews, and the PolitiFact dataset are the most popular, probably due to the variety of features types that could be extracted from them.

IV. Conclusions and Future work

In this paper, we present a systematic survey on fake news detection. Different forms of fake news are discussed: misinformation, disinformation, and mal-information. We give a general overview on the methodology that we follow in conducting our survey. Some of the covered research papers are introudced in details. Furthermore, the types of data in news documents are summarized and the types of features that could be extracted from each type are discussed. Moreover, the most commonly used bench mark datasets are provided. It is noticed that there is still a shortage in fake news detection systems for non-English news. Also, there is a lack of work done in making real-time detection systems. In addition, there is not much research in big fake news data, as well as how to minimize the feature vector size. In our current and future work, we are investigating the impact of using semantic information in enhancing the efficiency of the detection system, and will make experimental comparison between different classification techniques, with different datasets using different evaluation metrics.

References

[1] H. Allcott and M. Gentzkow, “Social media and fake news in the 2016 election,” *J. of economic perspectives*, vol. 31, no. 2, pp. 211–36, 2017.

[2] D. Westerman, P. R. Spence, and B. Van Der Heide, “Social media as information source: Recency of updates and credibility of information,” *J. of Comp. Mediated Comm.*, vol. 19, no. 2, pp. 171–183, 2014.

[3] S. Kumar and N. Shah, “False information on web and social media: A survey,” *arXiv:1804.08559*, 2018.

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

TABLE IV
FAKE NEWS DETECTION DATASETS

No.	Dataset	Purpose	Size	References
1	FEVER	Fact Extraction	185K	[20, 21]
2	KaggleFN	Fake News Detection	13K	[20, 42, 67]
3	FNC-1	Stance Detection	50K	[20, 40, 43]
4	LIAR	Fake News Detection	12.8K	[4, 6, 20, 21, 41, 47, 48, 56, 57, 62, 70, 74, 76]
5	PHEME	Fake News and Rumor Detection	6425	[20, 21, 38, 48, 58, 74]
6	Cred-1	Fact Extraction	4856	[20]
7	Weibo	Fake News and Rumor Detection	4664	[20, 45, 49, 63, 71, 74]
8	GossipCop	Fake News Detection	3570	[6, 20, 21, 60, 61, 64]
9	BuzzfeedNews	Fake News Detection	2282	[4, 6, 20, 21, 37–39, 44, 45, 48, 50, 53, 56, 65, 68, 70, 73, 75]
10	PolitiFact	Fake News Detection	488	[6, 20, 39, 44, 50, 53, 54, 60, 61, 64, 65, 68, 75]
11	FakevsSatire	Fake News Detection	486	[20]
12	NewsFN-2014	Fake News Detection	221	[20]
13	Cred-2	Fact Extraction	157	[20]
14	BuzzfeedPolitical	Fake News Detection	120	[3, 20, 44]
15	Twitter15	Fake News and Rumor Detection	1,490	[20, 71]
16	Twitter	Fake News and Rumor Detection	992	[20]
17	Twitter16	Fake News and Rumor Detection	818	[20, 71]

[5] K. Niklewicz, "Weeding out fake news: an approach to social media regulation," *European View*, vol. 16, no. 2, pp. 335–335, 2017.

[6] F. Pierri and S. Ceri, "False news on social media: A data-driven survey," *arXiv:1902.07539*, 2019.

[7] C. Ireton and J. Posetti, *Journalism, fake news and disinformation: handbook for journalism education and training*. UNESCO Publ., 2018.

[8] A. Karduni, "Human-misinformation interaction: Understanding the interdisciplinary approach needed to computationally combat false information," *arXiv:1903.07136*, 2019.

[9] V. Bakir and A. McStay, "Fake news and the economy of emotions: Problems, causes, solutions," *Digi. JRNL*, vol. 6, no. 2, pp. 154–175, 2018.

[10] A. Yap, L. G. Snyder, and S. Drye, *Academy of Social Science J.*, vol. 3, no. 7, pp. 1214–1221, 2018.

[11] M. F. Noordin, R. Othman, and A. H. R. Rassa, "Social media and knowledge management disruptive technology," in *Knowledge Management Int. Conf.*, 2018, pp. 6–11.

[12] C. D. MacDougall, *Hoaxes*. Dover Publ., 1958, vol. 465.

[13] E. C. Tandoc Jr, Z. W. Lim, and R. Ling, "Defining "fake news" a typology of scholarly definitions," *Digi. JRNL*, vol. 6, no. 2, pp. 137–153, 2018.

[14] J.-N. Kapferer, *Rumors: Uses, interpretations, and images*. Trans. Pub., 2013.

[15] S. M. Alzanin and A. M. Azmi, "Detecting rumors in social media: A survey," *Proc. comp. sci.*, vol. 142, pp. 294–300, 2018.

[16] S. Volkova, K. Shaffer, J. Y. Jang, and N. Hodas, "Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter," in *Proc. of the 55th Annual Meeting of the Association for Comp. Linguistics*, 2017, pp. 647–653.

[17] S. C. Woolley and P. N. Howard, *Computational propaganda: political parties, politicians, and political manipulation on social media*. Oxf. Univ. Press, 2018.

[18] V. L. Rubin, Y. Chen, and N. J. Conroy, "Deception detection for news: three types of fakes," in *Proc. of the 78th ASIS&T Annual Meeting: Infor. Sci. with Impact: Research in and for the Community*. American Society for Info. Sci., 2015, p. 83.

[19] D. Moher, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement,"

Annals of internal medicine, vol. 151, no. 4, p. 264, 2009.

[20] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, and Y. Liu, "Combating fake news: A survey on identification and mitigation techniques," *arXiv:1901.06437*, 2019.

[21] R. Oshikawa, J. Qian, and W. Y. Wang, "A survey on natural language processing for fake news detection," *arXiv:1811.00770*, 2018.

[22] S. Ghosh and C. Shah, "Towards automatic fake news classification," *Proc. of the Association for Info. Sci. and Tech.*, vol. 55, no. 1, pp. 805–807, 2018.

[23] M. K. Elhadad, K. M. Badran, and G. I. Salama, "A novel approach for ontology-based dimensionality reduction for web text document classification," *Int. J. of SW Innovation (IJSI)*, vol. 5, no. 4, pp. 44–58, 2017.

[24] ——, "A novel approach for ontology-based feature vector generation for web text document classification," *Int. J. of SW Innovation (IJSI)*, vol. 6, no. 1, pp. 1–10, 2018.

[25] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proc. of the 20th Int. Conf. on WWW*. ACM, 2011, pp. 675–684.

[26] U. Kursuncu, M. Gaur, U. Lokala, K. Thirunarayan, A. Sheth, and I. B. Arpinar, "Predictive analysis on twitter: Techniques and applications," in *Emerging Research Challenges and Opportunities in Computational SNA and Mining*. Springer, 2019, pp. 67–104.

[27] A. H. Mohammad, T. Alwada'n, and O. Al-Momani, "Arabic text categorization using support vector machine, na¨ıve bayes and neural network," *GSTF J. on Comp.*, vol. 5, no. 1, 2018.

[28] Y. Liu and X. Zhu, "Short text sentiment classification based on feature extension and ensemble classifier," in *AIP Conf. Proc.*, vol. 1967, no. 1. AIP Publ., 2018, p. 020051.

[29] W. Yun, W. X. An, Z. Jindan, and C. Yu, "Combining vector space features and convolution neural network for text sentiment analysis," in *Conf. on Complex, Intel., and SW Intensive Sys.* Springer, 2018, pp. 780–790.

[30] M. Jiang, Y. Liang, X. Feng, X. Fan, Z. Pei, Y. Xue, and R. Guan, "Text classification based on deep belief network and softmax regression," *Neural Computing and Apps.*, vol. 29, no. 1, pp. 61–70, 2018.

[31] H. Chen, Y. Hou, Q. Luo, Z. Hu, and L. Yan, "Text feature selection based on water wave optimization algorithm," in *The 10th Int. Conf. on Adv. Computational Intel.* IEEE, 2018, pp. 546–551.

[32] E. K. Shimomoto, L. S. Souza, B. B. Gatto, and K. Fukui, "Text classification based on word subspace with term-frequency," in *Int. Joint Conf. on NN*. IEEE, 2018, pp. 1–8.

[33] D. S. Sachan, M. Zaheer, and R. Salakhutdinov, "Investigating the working of text classifiers," *arXiv:1801.06261v2*, 2018.

[34] W. Zhang, L. Yu, T. Yoshida, and Q. Wang, "Feature weighted confidence to incorporate prior knowledge into support vector machines for classification," *Knowledge and Info. Sys.*, vol. 58, no. 2, pp. 371–397, 2019.

[35] M. K. Elhadad, K. F. Li, and F. Gebali, "Sentiment analysis of arabic and english tweets," in *Workshops of the Int. Conf. on Adv. Info. Net. and Apps.* Springer, 2019, pp. 334–348.

[36] Ö. Özgöbek and J. A. Gulla, "Towards an understanding of fake news," in *Proc. of the 3rd Norwegian Big Data Symposium, in conjunction with NxtMedia Conf.* NOBIDS, 2017, pp. 35–42.

[37] C. Janze and M. Risius, "Automatic detection of fake news on social media platforms." in *PACIS*, 2017, p. 261.

[38] C. Buntain and J. Golbeck, "Automatically identifying fake news in popular twitter threads," in *2017 IEEE Int. Conf. on Smart Cloud*. IEEE, 2017, pp. 208–215.

[39] K. Shu, S. Wang, and H. Liu, "Exploiting tri-relationship for fake news detection," *arXiv:1712.07709*, 2017.

[40] M. Aldwairi and A. Alwahedi, "Detecting fake news in social media networks," *Procedia Computer Science*, vol. 141, pp. 215–222, 2018.

[41] C. M. M. Kotteti, X. Dong, N. Li, and L. Qian, "Fake news detection enhancement with data imputation," in *2018 IEEE 16th Int. Conf. on Dependable, Autonomic and Secure Comp.* IEEE, 2018, pp. 187–192.

[42] J. Z. Pan, S. Pavlova, C. Li, N. Li, Y. Li, and J. Liu, "Content based fake news detection using knowledge graphs," in *Int. Semantic Web Conf.* Springer, 2018, pp. 669–683.

[43] A. Hanselowski, A. PVS, B. Schiller, F. Caspelherr, D. Chaudhuri, C. M. Meyer, and I. Gurevych, "A retrospective analysis of the fake news challenge stance detection task," *arXiv:1806.05180*, 2018.

[44] M. L. Della Vedova, E. Tacchini, S. Moret, G. Ballarin, M. DiPierro, and L. de Alfaro, "Automatic online fake news detection combining content and social signals," in *2018 22nd Conf. of Open Innovations Association*. IEEE, 2018, pp. 272–279.

[45] L. F. Meyn, "Fake news prediction on facebook: Design and implementation of a fake news prediction tool," Ph.D. dissertation, Univ. of Nebraska, 2018.

[46] C.-S. Atodiresei, A. Tănăselea, and A. Iftene, "Identifying fake news and fake users on twitter," *Proc. Comp. Sci.*, vol. 126, pp. 451–461, 2018.

[47] F. C. Fernández-Reyes and S. Shinde, "Evaluating deep neural networks for automatic fake news detection in political domain," in *Ibero-American Conf. on AI.* Springer, 2018, pp. 206–216.

[48] S. B. Parikh and P. K. Atrey, "Media-rich fake news detection: A survey," in *2018 IEEE Conf. on Multimedia Info. Processing and Retrieval*. IEEE, 2018, pp. 436–441.

[49] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao, "Eann: Event adversarial neural networks for multi-modal fake news detection," in *Proc. of the 24th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*. ACM, 2018, pp. 849–857.

[50] K. Shu, S. Wang, and H. Liu, "Understanding user profiles on social media for fake news detection," in *2018 IEEE Conf. on Multimedia Info. Processing and Retrieval*. IEEE, 2018, pp. 430–435.

[51] J. Meinert, M. Mirbabaie, S. Dungs, and A. Aker, "Is it really fake?–towards an understanding of fake news in social media communication," in *Int. Conf. on Social Comp. and Social Media*. Springer, 2018, pp. 484–497.

[52] L. de Alfaro, M. Di Pierro, R. Agrawal, E. Tacchini, G. Ballarin, M. L. Della Vedova, and S. Moret, "Reputation systems for news on twitter: A large-scale study," *arXiv:1802.08066*, 2018.

[53] S. Gupta, R. Thirukovalluru, M. Sinha, and S. Mannarswamy, "Cimtdetect: A community infused matrix-tensor coupled factorization based method for fake news detection," in *2018 IEEE/ACM Int. Conf. on Adv. in SNA and Mining*. IEEE, 2018, pp. 278–281.

[54] J. Zhang, L. Cui, Y. Fu, and F. B. Gouza, "Fake news detection with deep diffusive network model," *arXiv:1805.08751*, 2018.

[55] L. Wu and H. Liu, "Tracing fake-news footprints: Characterizing social media messages by how they propagate," in *Proc. of the 11th ACM Int. Conf. on Web Search and Data Mining*. ACM, 2018, pp. 637–645.

[56] M. Dong, L. Yao, X. Wang, B. Benatallah, Q. Z. Sheng, and H. Huang, "Dual: A deep unified attention model with latent relation representations for fake news detection," in *Int. Conf. on Web Info. Sys. Eng.* Springer, 2018, pp. 199–209.

[57] H. Karimi, P. Roy, S. Saba-Sadiya, and J. Tang, "Multi-source multi-class fake news detection," in *Proc. of the 27th Int. Conf. on Comp. Linguistics*, 2018, pp. 1546–1557.

[58] O. Ajao, D. Bhowmik, and S. Zargari, "Fake news identification on twitter with hybrid cnn and rnn models," in *Proc. of the 9th Int. Conf. on Social Media and Society*. ACM, 2018, pp. 226–230.

[59] G. B. Guacho, S. Abdali, N. Shah, and E. E. Papalexakis, "Semi-supervised content-based detection of misinformation via tensor embeddings," in *2018 IEEE/ACM Int. Conf. on Advances in SNA and Mining*. IEEE, 2018, pp. 322–325.

[60] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media," *arXiv:1809.01286v3*, 2019.

[61] K. Shu, X. Zhou, S. Wang, R. Zafarani, and H. Liu, "The role of user profile for fake news detection," *arXiv:1904.13355v1*, 2019.

[62] A. Olivieri, S. Shabani, M. Sokhn, and P. Cudré-Mauroux, "Creating task-generic features for fake news detection," in *Proc. of the 52nd Hawaii Int. Conf. on Sys. Sci.*, 2019.

[63] C. Guo, J. Cao, X. Zhang, K. Shu, and M. Yu, "Exploiting emotions for fake news detection on social media," *arXiv:1903.01728*, 2019.

[64] K. Shu, D. Mahudeswaran, S. Wang, and H. Liu, "Hierarchical propagation networks for fake news detection: Investigation and exploitation," *arXiv:1903.09196*, 2019.

[65] K. Shu, D. Mahudeswaran, and H. Liu, "Fakenewstracker: a tool for fake news collection, detection, and visualization," *Comp. and Math. Org. Theory*, vol. 25, no. 1, pp. 60–71, 2019.

[66] T. Traylor, J. Straub, N. Snell *et al.*, "Classifying fake news articles using natural language processing to identify in-article attribution as a supervised learning estimator," in *2019 IEEE 13th Int. Conf. on Semantic Comp.* IEEE, 2019, pp. 445–449.

[67] P. Bharadwaj and Z. Shao, "Fake news detection with semantic features and text mining," 2019.

[68] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani, "Fake news early detection: A theory-driven model," *arXiv:1904.11679*, 2019.

[69] A. Bovet and H. A. Makse, "Influence of fake news in twitter during the 2016 us presidential election," *Nature comm.*, vol. 10, no. 1, p. 7, 2019.

[70] S. Yang, K. Shu, S. Wang, R. Gu, F. Wu, and H. Liu, "Unsupervised fake news detection on social media: A generative approach," in *Proc. of 33rd AAAI Conf. on AI*. AAAI, 2019, pp. 1–8.

[71] Y. Liu and Y.-F. B. Wu, "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks," in *32nd AAAI Conf. on AI*, 2018.

[72] S. Helmstetter and H. Paulheim, "Weakly supervised learning for fake news detection on twitter," in *2018 IEEE/ACM Int. Conf. on Adv. in SNA and Mining*. IEEE, 2018, pp. 274–277.

[73] J. C. Reis, A. Correia, F. Murai, A. Veloso, F. Benevenuto, and E. Cambria, "Supervised learning for fake news detection," *IEEE Intelligent Sys.*, vol. 34, no. 2, pp. 76–81, 2019.

[74] F. Cardoso Durier da Silva, R. Vieira, and A. C. Garcia, "Can machines learn to detect fake news? a survey focused on social media," in *Proc. of the 52nd Hawaii Int. Conf. on Sys. Sci.*, 2019.

[75] K. Shu, S. Wang, and H. Liu, "Beyond news contents: The role of social context for fake news detection," in *Proc. of the 12th ACM Int. Conf. on Web Search and Data Mining*. ACM, 2019, pp. 312–320.

[76] T. Rasool, W. H. Butt, A. Shaukat, and M. U. Akram, "Multi-label fake news detection using multi-layered supervised learning," in *Proc. of the 2019 11th Int. Conf. on Comp. and Automation Eng.* ACM, 2019, pp. 73–77.

[77] D. Król and K. Wiśniewska, "on rumor source detection and its experimental verification on twitter," in *Asian Conf. on Intel. Info. and DB Sys.* Springer, 2017, pp. 110–119.

[78] J. Cao, J. Guo, X. Li, Z. Jin, H. Guo, and J. Li, "Automatic rumor detection on microblogs: A survey," *arXiv:1902.07539*, 2018.