Journal on Interactive Systems, 2023, 14:1, doi: 10.5753/jis.2023.3020

Fake news detection: a systematic literature review of machine learning algorithms and datasets

Humberto Fernandes Villela [Universidade FUMEC | humberto.villela@gmail.com]
Fábio Corrêa [Universidade FUMEC | fabiocontact@gmail.com]
Jurema Suely de Araújo Nery Ribeiro [Universidade FUMEC | jurema.nery@gmail.com]
Air Rabelo [Universidade FUMEC | air@fumec.br]
Dárlinton Barbosa Feres Carvalho [Universidade Federal de São João del-Rei | darlinton@acm.org]

Abstract

Fake news (i.e., false news created to have a high capacity for dissemination and malicious intentions) is a problem of great interest to society today since it has achieved unprecedented political, economic, and social impacts. Taking advantage of modern digital communication and information technologies, they are widely propagated through social media, being their use intentional and challenging to identify. In order to mitigate the damage caused by fake news, researchers have been seeking the development of automated mechanisms to detect them, such as algorithms based on machine learning as well as the datasets employed in this development. This research aims to analyze the machine learning algorithms and datasets used in training to identify fake news published in the literature. It is exploratory research with a qualitative approach, which uses a research protocol to identify studies with the intention of analyzing them. As a result, we have the algorithms Stacking Method, Bidirectional Recurrent Neural Network (BiRNN), and Convolutional Neural Network (CNN), with 99.9%, 99.8%, and 99.8% accuracy, respectively. Although this accuracy is expressive, most of the research employed datasets in controlled environments (e.g., Kaggle) or without information updated in real-time (from social networks). Still, only a few studies have been applied in social network environments, where the most significant dissemination of disinformation occurs nowadays. Kaggle was the platform identified with the most frequently used datasets, being succeeded by Weibo, FNC-1, COVID-19 Fake News, and Twitter. For future research, studies should be carried out in addition to news about politics, the area that was the primary motivator for the growth of research from 2017, and the use of hybrid methods for identifying fake news.

Keywords: *Algorithms, datasets, accuracy, fake news, artificial intelligence.*

1 Introduction

Currently, the term fake news is on the rise as this type of news can remarkably influence society, promoting significant political, economic, or social impacts (Zhang *et al.*, 2016; Islam *et al.*, 2020). They are false news created to be highly broadcastable, usually with malicious intent (i.e., to deceive, cause ambiguity, or falsehood).

In the political context, specifically, Almeida *et al.* (2021) point out the impacts of fake news on the US presidential elections in 2016, having President Donald Trump elected. In Brazil, similar impacts were imputed to the election of President Jair Bolsonaro in 2018. Due to the behavior of many Brazilian voters relying on social media as the primary source to access news, this channel is fruitful for the proliferation of fake news (ALMEIDA *et al.*, 2021).

Due to the relevant impact this type of news has caused on society, researchers have been seeking to develop ways to detect them. Using algorithms for the automated identification of fake news presents itself as a promising line of research. The quality of these algorithms is commonly verified by accuracy, which is the measure of correctness in classifying whether a news item is true or false (Chapra & Canale, 2016). That is, accuracy represents the assertiveness characteristic of the algorithm in detecting fake news.

However, the quality of algorithms is directly related to their specific purpose, restricted to a language or news style, that is, according to the datasets used for training (Ahuja & Kumar, 2020). For this reason, Jiang *et al.* (2021) and Ahuja and Kumar (2020) recommend continuing research using varied datasets, such as in languages other than English, given that this is the most used.

Nevertheless, recurring studies to detect fake news have achieved relevant accomplishments. The research by Burfoot and Baldwin (2009) reached 71% accuracy in detecting fake news, while Ahmed, Traore, and Saad (2017) raised this metric to 87%, the same percentage achieved by Low *et al*. (2022).

Medeiros and Braga (2020) carried out research aiming, among others, to identify the accuracy achieved by algorithms in detecting fake news. From the 32 studies examined, the proposed algorithms show an accuracy of 73.7% to 98.0%. This result highlights the relevance of this theme and the constant search to promote a more assertive detection of fake news, to minimize its impacts in the aforementioned political, economic, or social contexts, among others.

Accordingly, this research is grounded on the recommendations of Ahmed, Traore, and Saad (2017), Ahmad *et al.* (2020), Agarwal *et al.* (2020), Aslam *et al.* (2021) e Jiang *et al.* (2021) regarding the use of algorithms to identify fake news. Therefore, the objective of this investigation is to analyze the accuracy obtained and the datasets used in fake news identification algorithms.

This study intends to contribute by highlighting more successful algorithms and datasets in identifying fake news to better support researchers' understanding of the state-of-the-art, given the variety of options mentioned above. Besides, it consolidates recommendations for future studies. Thus, the theoretical contribution stands out for presenting algorithms and datasets with expressive accuracy, providing a means for their identification and practical application for the identification of fake news, and continuing the research by Medeiros and Braga (2020), carried out in August of 2019. In addition, by exploring fake news, a phenomenon with relevant political, economic, and social impacts, we seek to encourage discussion on this topic and shed light on the importance of computing research in tackling this matter.

Thus, this article is subdivided into parts to present the work carried out. In addition to this introduction, the following section explains the theoretical and practical aspects of fake news, contextualizing this phenomenon for the purposes of this research. The following section explains the methodological procedures used to achieve the desired objective. Accordingly, the results are discussed in the subsequent section, and so on, promoting this study's conclusion. The references that support this investigation end it.

2 Fake news detection

The endless need to be informed in a progressively connected and ubiquitous world is already part of most people's routines. The ease of communicating through the internet is ever more simplified and decentralized. People and devices connected, through the Internet of Things (IoT), communicate through the global network at all times, culminating in an increasingly high volume of shared communications.

The increasing expressive volume of communications shared through the network, often treated as Big Data, composed of textual and multimedia data, has produced valuable information. Through this perspective, Agarwal *et al.* (2020) advocate that data become an increasingly valuable asset and gain greater relevance when transformed into information and assimilated as knowledge by users.

In this scenario, Ali *et al.* (2021) express the idea of hyper-connectivity, which consists of the high connection of users and machines, high transmission speed, ease of communication, and access to real-time information worldwide. It is a vibrant path in technological evolution, full of opportunities as well as struggles, such as facilitating fake news.

Fake news is usually a modified version of plausible news to mislead, cause ambiguity or falsehood, and be widely disseminated. It is mainly propagated through social media, with intentional use and challenging identification. In this way, fake news is a specific type of disinformation, like rumor and SPAM. They can generate political, economic, and social impacts (Zhang *et al.*, 2016; Islam *et al.*, 2020).

It is noteworthy that fake news is not new. In 1835, The Sun, a New York newspaper, published a series of fake news about the supposed discovery of life on the moon, being this case nowadays referenced as the Great Moon Hoax (Shabani & Sokhn, 2018). However, in the modern context of hyper-connectivity, social media has propelled the spread of this type of news.

Social media platforms have a plethora of misinformation, which has caught the attention of researchers in developing mechanisms to detect them (Jiang *et al.*, 2021, 2020; Goldani, Momtazi & Safabakhsh, 2021; Birunda & Devi, 2021; Sahoo & Gupta, 2021; Goel *et al.*, 2021; Pardamean & Pardede, 2021). The European Commission has established a group of experts to advise and discuss policy initiatives to combat fake news and the spread of disinformation online (Assad & Erascu, 2018). Usually, the detection approaches consist of classifying (fake) news in a binary form (i.e., true or false) (Zhang *et al.*, 2019). However, the subtleties of human language add high complexity to the detection algorithms, even for this binary classification.

Nevertheless, Collins *et al.* (2021) subdivide fake news into clickbait, propaganda, satire and parody, hoaxes, and others (e.g., name theft, journalistic fraud). Medeiros and Braga (2020) split into conspiracy theories, hoaxes, rumors, biased news, and satires. Identifying these categories is a complex task due to the nuances of human language. For example, since satire and parody employ sarcasm and humor, the algorithmic analysis must consider this feature to call a news story true or false.

In the meantime, it is notable that social media are the biggest target of fake news due to their ease and breadth of disseminating information nowadays. Another relevant feature that is being widely used and favors the spread of fake news is the resemblance to plausible news, confusing people and making it challenging to combat false information dissemination (Islam *et al.*, 2020). Furthermore, it is noteworthy that the absence or even only a delay in a disclaimer by official entities involved in some rumor, which is false but convincing, favors its dissemination (Abouzeid *et al.*, 2019).

Ruchansky, Seo, and Liu (2017) claim that many stakeholders profit from the publication of fake news online because the more provocative the news is, the greater the response and the greater its yield. A rumor can be vastly profitable for someone or some organization. Given the sophistication of fake news due to its creator's intentions, Asaad and Erascu (2018) point out that critical thinking is an important ally in combating the spread of disinformation.

Although recognizing fake news by algorithms is tricky, like in satire and parody, the increasing volume of communications shared on the network is expressive; therefore, it is unfeasible to attribute to humans the responsibility of classifying them (Agarwal *et al.*, 2020). Ahmad *et al.* (2020) articulate that numerous techniques help in the classification of articles as fake based on their textual content. Many of them make use of verification sites such as PolitiFact[1] and Snopes[2]. Curators also maintain repositories, with lists of sites considered false or ambiguous. Medeiros and Braga (2020) point out several mechanisms for detecting fake news, divided into automatic and semi-automatic/manual (Figure 1); however, human critical thinking, in these cases, is necessary to classify the news as false or true (Ahmad *et al.*, 2020).

1 https://www.politifact.com/

2 https://www.snopes.com/

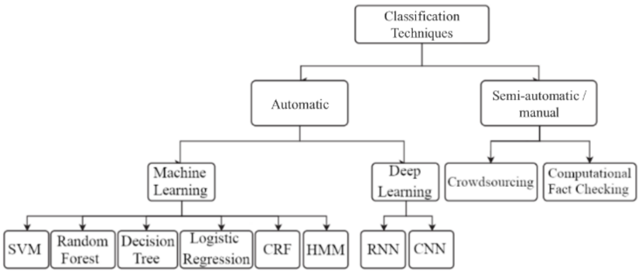

Figure 1. Different approaches proposed for detecting fake news found in the literature. Source: translated from Medeiros & Braga (2020, p. 3)

Ahmed, Traore, and Saad (2017) and Ahmad *et al*. (2020) recommend the use of artificial intelligence (AI) algorithms to extract linguistic features from textual articles through machine learning. Agarwal *et al*. (2020) proposed, as another way to identify fake news, labeling or classifying a particular news or article on a defined scale, thus giving the reader an idea about the credibility of that published text. Another AI technique used in the process of identifying disinformation is deep learning. To Agarwal *et al.* (2020), the nature of self-learning and resource maps gave deep learning a significant advantage compared to other statistical modeling and learning methods.

Since research has shown that only 54% of humans can detect fraud without special assistance, Aslam *et al*. (2021) claim that efforts should be made to build an automated system to classify news as real or fake, aiming at greater classification accuracy. Jiang *et al*. (2021) agree with Aslam *et al*. (2021), remarking on the essentiality of a machine-driven approach when they state that the use of automated tools to detect fake news has become an essential requirement to tackle the issue.

The accuracy for identifying fake news by an algorithm is defined as the measure of correct answers attributed to a group of news stories as true or false (Chapra & Canale, 2016). This result is intrinsically related to the features (e.g., language, style) of the news stories of interest, which are called the dataset (Ahuja & Kumar, 2020).

Accordingly, given the impacts of fake news in society, it is relevant to pinpoint the algorithmic means developed and the accuracy achieved by them, as well as the datasets used, to provide a better understanding regarding the computational state-of-the-art at tackling the issue. To this end, the methodological procedures are outlined as follows.

3 Methodology

This research has an exploratory nature and a qualitative approach since this investigation aims to analyze the accuracy obtained by algorithms and the datasets used in fake news identification. For this, a research protocol is employed, based on Dresch, Lacerda, and Antunes Júnior (2015), to specify the research carried out in this study. Table 1 presents the planned research protocol, which sets a systematic review of scientific publications.

Table 1. Research Protocol. Source: The authors.

Item	Description
Research Question (RQ)	RQ1. What is the accuracy of the main algorithms used to identify fake news? RQ2. Which datasets are used? RQ3. What are the top recommendations for future research?
Exclusion Criteria (EC)	EC1. Articles in languages other than English or Portuguese; EC2. Abstracts, technical reports, secondary studies, presentations, or systematic reviews; EC3. Incomplete or unavailable articles for download; EC4. It does not address the use of algorithms to identify fake news; EC5. It does not present methods or accuracy in the identification of fake news; EC6. Article duplicates.
Search fields	Title, abstract and keywords.
Temporal space	Publications between 2010 and 2021.
Languages	English and Portuguese.
Databases	ACM Digital Library, IEEE Xplore, Scopus, Science Direct, Springer, Web of Science, EBSCO
Descriptors (Query String)	QS1. ("computational techniques") AND (("fake news") OR (disinformation) OR (misinformation) OR (malinformation)) QS2. (("fake news detection") OR ("disinformation detection") OR ("misinformation detection") OR ("malinformation detection"))

The performed research design included articles from the last ten years (from 2010 to 2021) to identify the current state-of-the-art about algorithms and datasets used to identify fake news and the accuracy achieved. It also looked at the recommendations for future research in the field.

As a source for publications of computational techniques in the context of fake news, the research considered seven databases. This choice is based on the criterion of scope. The ACM Digital Library and IEEE Xplore are fruitful bases on the subject, with the other bases and the Portuguese language being added to expand the search.

The descriptors were defined based on the authors' prior knowledge of the subject and a trial-and-error process using the databases' search engines. The authors performed and discussed the analysis together until reaching an agreement.

4 Results and discussion

The application of the proposed research protocol, through the search employing the query strings QS1 and QS2 (Table 1), was carried out on 06/10/2021 and retrieved 507 articles from the following databases: ACM Digital Library (ACM), IEEE Xplore (IEEE), Scopus (SCO), Science Direct (ScD), Springer (Spr), Web of Science (WoS), EBSCO (EBS). Of this amount, three articles were disregarded by the EC1 exclusion criterion, 82 by EC2, 80 related to EC3, 150 related to EC4, 51 related to EC5, and 80 related to EC6, totaling 446 articles not consistent with the intent of this research. Table 2 details the whole process providing the specific number of publications retrieved from each source and the application of each exclusion criteria along the analysis. Thus, 61 publications comprise the sample considered for analysis by this research.

Table 2. Application of the Research Protocol. Source: The authors.

	ACM	IEEE	SCO	ScD	Spr	WoS	EBS	Σ
QS1	44	0	2	2	8	6	1	63
QS2	16	188	91	36	36	45	32	444
ΣS	60	188	93	38	44	51	33	507
EC1	0	0	1	1	0	1	0	3
EC2	3	45	11	5	7	7	4	82
EC3	12	42	11	5	1	4	5	80
EC4	36	37	32	6	29	9	1	150
EC5	2	23	12	5	2	5	2	51
EC6	4	13	13	12	2	16	20	80
ΣEC	57	160	80	34	41	42	32	446
Σ	3	28	13	4	3	9	1	61

The 61 scientific articles were fully read to answer the research questions (RQ1, RQ2, and RQ3 in Table 1). Regarding RQ1 (What is the accuracy of the main algorithms used to identify fake news?), Table 3 presents the results obtained through the analyzed articles and the accuracy (Acc) reported by them.

Table 3. Accuracy of the analyzed algorithms used to identify fake news. Source: The authors.

Author	Algorithm	Acc
1) Jiang *et al*. (2021)	*Stacking Method*	99.9%
2) Jiang *et al*. (2020)	*Bidirectional Recurrent Neural Network (BiRNN)*	99.8%
3) Goldani, Momtazi e Safabakhsh (2021)	*Convolutional Neural Network (CNN)*	99.8%
4) Birunda e Devi (2021)	*Gradient Boosting*	99.5%
5) Sahoo e Gupta (2021)	*Long short-term memory (LSTM)*	99.4%
6) Goel *et al*. (2021)	*Robustly Optimized BERT Pretraining Approach (RoBERTa)*	99.3%
7) Pardamean e Pardede (2021)	*Bidirectional encoder representation of transformers (BERT)*	99.2%
8) Kaliyar *et al*. (2020a)	*Convolutional Neural Network (CNN)*	99.1%
9) Albahr e Albahar (2020)	*Naive Bayes*	99.0%
10) Gereme *et al*. (2021)	*Convolutional Neural Network (CNN)*	99.0%
11) Nasir, Khan e Varlamis (2021)	*Hybrid CNN-RNN*	99.0%
12) Kaliyar, Goswami e Narang (2021)	*Fake news detection in social media with a BERT-based (FakeBERT)*	98.9%
13) Dadkhah *et al*. (2021)	*AWD-LSTM*	98.8%
14) Goldani, Safabakhsh e Momtazi (2021)	*CNN com marginloss*	98.4%
15) Sridhar e Sanagavarapu (2021)	*BiLSTM-CapsNet*	98.0%
16) Umer *et al*. (2020)	*CNN-LSTM*	97.8%
17) Thakur *et al*. (2020)	*Gradient Boosting (GB)*	97.6%
18) Agarwal *et al*. (2020)	*CNN+RNN*	97.2%
19) Ayoub, Yang e Zhou (2021)	*DistilBERT*	97.2%
20) Yu *et al*. (2020)	*IARNet*	96.9%
21) Xie *et al*. (2021)	*Stance Extraction and Reasonic Network (SERN)*	96.6%
22) Ozbay e Alatas (2019)	*Grey Wolf Optimization (GWO)*	96.5%
23) Fang *et al*. (2019)	*SMHA-CNN*	95.5%
24) Kaliyar *et al*. (2020b)	*DeepNet*	95.2%
25) Faustini e Covões (2020)	*Random Forest (RF)*	95.0%
26) Varshney e Vishwakarma (2020)	*Random Forest*	95.0%
27) Ahuja e Kumar (2020)	*S-HAN*	93.6%
28) Ivancov, Sarnovsk e Maslej-kre (2021)	*LSTM*	93.6%
29) Verma *et al*. (2021)	*WELFake*	92.6%
30) Wang *et al*. (2021)	*SemSeq4FD*	92.6%
31) Kumar, Anurag e Pratik (2021)	*EchoFakeD*	92.3%
32) Song *et al*. (2021)	*CARMN*	92.2%

33) Sharma, Garg e Shrivastava (2021)	*BiLSTM*	91.5%
34) Albahar (2021)	*SVM-RNN-GRUs bidirecionais*	91.2%
35) Bahad, Saxena e Kamal (2020)	*BiLSTM-RNN*	91.1%
36) Torgheh *et al.* (2021)	*GRU-LSTM-CNN*	90.8%
37) Lakshmanarao, Swathi e Kiran (2019)	*Random Forest*	90.7%
38) Li *et al.* (2020)	*MCNN-TFV*	90.1%
39) Aslam et al. (2021)	*Bi-LSTM-GRU-dense*	89.8%
40) Kumar et al. (2020)	CNN + BiLSTM	88.8%
41) Lin et al. (2020)	BERT	88.7%
42) Kaliyar et al. (2020b)	DeepFakE	88.6%
43) Wang et al. (2020)	Knowledge-driven Multi-modal Graph Convolutional Networks (KMGCN)	88.6%
44) Najar et al. (2019)	Bayesian inference algorithm	87.9%
45) Chen et al. (2018)	AERNN	87.6%
46) Alanazi e Khan (2020)	SVM	87.1%
47) Mugdha et al. (2020)	Gaussian Naive Bayes	87.0%
48) Kaliyar et al. (2019)	Gradient Boosting	86.0%
49) Ajao et al. (2019)	LSTM HAN	86.0%
50) Lin et al. (2019)	XGBoost	85.5%
51) Qawasmeh et al. (2019)	FND-Bidirectional LSTM concatenated	85.3%
52) Qi et al. (2019)	Multi-domain Visual Neural Network (MVNN)	84.6%
53) Shabani e Sokhn (2018)	CROWDSOURCING	84.0%
54) Barua et al. (2019)	LSTM+GRU (Recurrent Neural Networks)	82.6%
55) Khattar et al. (2016)	MVAE (Multimodal Variational Autoencoder)	82.4%
56) Gangireddy et al. (2020)	GTUT	80.0%
57) Kesarwani et al. (2020)	K-Nearest Neighbor	79.0%
58) Ren et al. (2020)	AA-HGNN (Adversarial Active Learning based Graph Neural Network)	73.5%
59) Jardaneh et al. (2019)	Random Forest	76.0%
60) Al-Ahmad et al. (2021)	Algoritmos genéticos	75.4%
61) Konkobo et al. (2020)	SSLNews	72.3%

Ten studies employed the Convolutional Neural Network (CNN) algorithm (Table 3). Three used typical CNN (articles 3, 8, and 10), and the other seven as a hybrid model, as follows: CNN-RNN (articles 11 and 18), CNN with marginloss (article 14), CNN-LSTM (article 16), SHMA-CNN (article 23), GRU-LSTM-CNN (article 36), CNN-BiLSTM (article 40). Goldani, Momtazi, and Safabakhsh (2021) achieved the highest accuracy of the CNN approaches, reaching 99.8% with deep learning (article 3). Through the articles analyzed (Table 2), science's growth in identifying fake news is underlined. (Figure 2).

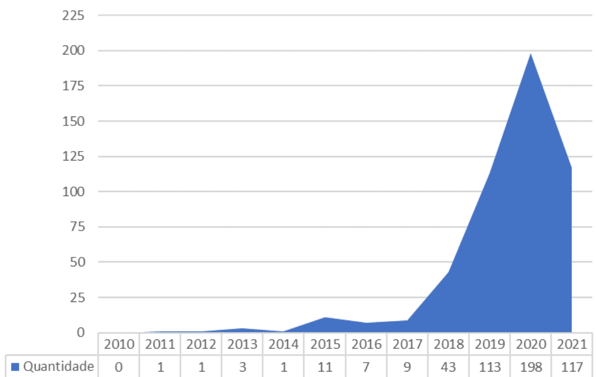

Figure 2. Evolution over the years of the number of publications of algorithms for fake news identification. Source: The authors.

Accuracy greater than 90% was only achieved in 2019 by Lakshmanarao, Swathi, and Kiran (2019) (article 37). Supporting Almeida *et al.* (2021), Goldani, Momtazi, and Safabakhsh (2021) and Kumar, Anurag, and Pratik (2021) cite the 2016 US presidential elections as the biggest motivator for research applied to the identification of disinformation or fake news.

The word cloud of the most recurrent keywords (Figure 3) in the analyzed articles showed that the terms "fake news detection", "fake news", "deep learning", "machine learning", and "feature-extraction", with 24, 15, 16, 13 and 12 occurrences, respectively, were the words with the highest recurrence in the 61 articles analyzed by this research.

Figure 3. Word cloud of most recurrent keywords in articles
Source: The authors.

Regarding RQ2 (Which datasets are used?), the investigation reveals that many datasets were being used to develop fake news identification methods. Table 4 presents them with the respective amount of use in the publications.

Table 4. Datasets used in the analyzed studies of fake news detection.
Source: The authors.

Dataset	Amount	Frequency
Kaggle	39	63.9%
Weibo	6	9.8%
FNC-1	3	4.9%
COVID-19 Fake News	2	3.3%
Twitter	2	3.3%
NewsFN	1	1.6%
Bengali Language	1	1.6%
btvlifestyle	1	1.6%
Slovak language	1	1.6%
Fake vs Satire	1	1.6%
fake news Amharic	1	1.6%
LUN	1	1.6%
Fakeddit	1	1.6%
Facebook	1	1.6%
Total	61	100.0%

The datasets are primarily related to a language (Ahuja & Kumar, 2020), English being the most used and fostering most algorithms for such language. Among those used, Kaggle - a Google platform used by data scientists) - stands out, which includes several datasets for studies of artificial intelligence. Kaggle was the biggest dataset provider (ISOT, Kaggle Fake News, LIAR, Kaggle, PolitiFact, BuzzFeed, Kaggle Indonesia data) of the analyzed articles (Table 4). After the Kaggle datasets, the most used was Weibo (a Chinese microblog similar to Twitter), with six occurrences; FNC-1, with three occurrences, COVID-19 Fake News and Twitter, both with two occurrences, and the others with only one occurrence.

Additionally, we sought to understand: a) Is there any relationship between algorithms results and the datasets used? b) Is there any pattern in the results provided by the algorithms that had Kaggle as a database?

All the 61 analyzed papers presented a different combination of datasets and algorithms, except for surveys by Goldani, Momtazi and Safabakhsh (2021) and Kaliyar *et al.* (2020), who used the Kaggle dataset with the CNN algorithm and obtained an accuracy of 99.8% and 99.1%, respectively. Thus, it is not possible to infer whether there is any relationship between the results of the algorithms and the database used (a). This yield may be due to the character of originality and fast evolution in the field since, in principle, better algorithms or datasets are required for better performance, and the domain of fake news identification presents a vast and increasing range of applications.

In view of the above, it was also not possible to verify whether there is any pattern in the results provided by the algorithms that had Kaggle as a database (b), given that it is not possible to compare results further considering the different experimental setups. The data that support these perceptions are expressed in Table 5.

Table 5. Algorithm and dataset used for each of the analyzed studies.
Source: The authors.

Algorithm – Author	Dataset
1) Stacking Method – Jiang *et al.* (2021)	Kaggle
2) Bidirectional Recurrent Neural Network (BiRNN) – Jiang *et al.* (2020)	Kaggle
3) Convolutional Neural Network (CNN) – Goldani, Momtazi and Safabakhsh (2021)	Kaggle
4) Gradient Boosting – Birunda and Devi (2021)	Kaggle
5) Long short-term memory (LSTM) – Sahoo and Gupta (2021)	Facebook
6) Robustly Optimized BERT Pretraining Approach (RoBERTa) – Goel *et al.* (2021)	Kaggle
7) Bidirectional encoder representation of transformers (BERT) – Pardamean and Pardede (2021)	Kaggle
8) Convolutional Neural Network (CNN) – Kaliyar *et al.* (2020a)	Kaggle
9) Naive Bayes – Albahr and Albahar (2020)	Kaggle
10) Convolutional Neural Network (CNN) – Gereme *et al.* (2021)	fake news Amharic
11) Hybrid CNN-RNN – Nasir, Khan and Varlamis (2021)	Kaggle
12) Fake news detection in social media with a BERT-based (FakeBERT) – Kaliyar, Goswami and Narang (2021)	Kaggle
13) AWD-LSTM – Dadkhah *et al.* (2021)	Kaggle
14) CNN com marginloss – Goldani, Safabakhsh and Momtazi (2021)	Kaggle

15) BiLSTM-CapsNet – Sridhar and Sanagavarapu (2021)	Kaggle
16) CNN-LSTM – Umer *et al*. (2020)	FNC-1
17) Gradient Boosting (GB) – Thakur *et al*. (2020)	Kaggle
18) CNN+RNN – Agarwal *et al*. (2020)	Kaggle
19) DistilBERT – Ayoub, Yang and Zhou (2021)	COVID-19 Fake News
20) IARNet – Yu *et al*. (2020)	Weibo
21) Stance Extraction and Reasonic Network (SERN) – Xie *et al*. (2021)	Fakeddit
22) Grey Wolf Optimization (GWO) – Ozbay and Alatas (2019)	Kaggle
23) SMHA-CNN – Fang *et al*. (2019)	Kaggle
24) DeepNet – Kaliyar *et al*. (2020b)	Kaggle
25) Random Forest (RF) – Faustini and Covões (2020)	btvlife-style
26) Random Forest – Varshney and Vishwakarma (2020)	Kaggle
27) S-HAN – Ahuja and Kumar (2020)	Kaggle
28) LSTM – Ivancov, Sarnovsk and Maslej-kre (2021)	Slovak language
29) WELFake – Verma *et al*. (2021)	Kaggle
30) SemSeq4FD – Wang *et al*. (2021)	LUN
31) EchoFakeD – Kumar, Anurag and Pratik (2021)	Kaggle
32) CARMN – Song *et al*. (2021)	Weibo
33) BiLSTM – Sharma, Garg and Shrivastava (2021)	Kaggle
34) SVM-RNN-GRUs bidirecionais – Albahar (2021)	Kaggle
35) BiLSTM-RNN – Bahad, Saxena and Kamal (2020)	Kaggle
36) GRU-LSTM-CNN – Torgheh *et al*. (2021)	Twitter
37) Random Forest – Lakshmanarao, Swathi and Kiran (2019)	FNC-1
38) MCNN-TFV – Li *et al*. (2020)	NewsFN
39) Bi-LSTM-GRU-dense – Aslam *et al*. (2021)	Kaggle
40) CNN + BiLSTM – Kumar *et al*. (2020)	Kaggle
41) BERT – Lin *et al*. (2020)	FNC-1
42) DeepFakE – Kaliyar *et al*. (2020b)	Kaggle
43) Knowledge-driven Multimodal Graph Convolutional Networks (KMGCN) – Wang *et al*. (2020)	Weibo
44) Bayesian inference algorithm – Najar *et al*. (2019)	Kaggle
45) AERNN – Chen *et al*. (2018)	Weibo
46) SVM – Alanazi and Khan (2020)	Kaggle
47) Gaussian Naive Bayes – Mugdha *et al*. (2020)	Bengali Language
48) Gradient Boosting – Kaliyar *et al*. (2019)	Kaggle
49) LSTM HAN – Ajao *et al*. (2019)	Kaggle
50) XGBoost – Lin *et al*. (2019)	Kaggle
51) FND-Bidirectional LSTM concatenated – Qawasmeh *et al*. (2019)	FNC-1
52) Multi-domain Visual Neural Network (MVNN) – Qi *et al*. (2019)	Weibo
53) CROWDSOURCING – Shabani and Sokhn (2018)	Fake vs Satire
54) LSTM+GRU (Recurrent Neural Networks) – Barua *et al*. (2019)	Kaggle
55) MVAE (Multimodal Variational Autoencoder) – Khattar *et al*. (2016)	Weibo
56) GTUT – Gangireddy *et al*. (2020)	Kaggle
57) K-Nearest – Neighbor Kesarwani *et al*. (2020)	Kaggle
58) AA-HGNN (Adversarial Active Learning based Graph Neural Network) – Ren *et al*. (2020)	Kaggle
59) Random Forest – Jardaneh *et al*. (2019)	Twitter
60) Algoritmos genéticos – Al-Ahmad *et al*. (2021)	COVID-19 Fake News
61) SSLNews – Konkobo *et al*. (2020)	Kaggle

Regarding RQ3 (What are the top recommendations for future studies?), some perspectives are presented, such as using other languages for the dataset. The accuracy of identifying fake news for the authors does not only depend on the algorithm but also on the dataset language. Therefore, Jiang *et al.* (2021) and Ahuja and Kumar (2020) recommend extending studies by applying research in datasets from other languages.

Another research suggestion is classifying fake news using a scoring model, such as a credibility rate. The scoring model is justified by the difficulty in classifying information as only false or true (Agarwal *et al*., 2020), given the inherent complexity of human language and other aspects, such as the area of news (e.g., economics or politics).

The combination of models generating hybrid models is a recommendation highlighted by Jiang *et al*. (2020), Pardameanm, and Pardede (2021), and Kaliyar, Goswamim and Narang (2021). Another recommendation was to use algorithms based on deep learning in future research and understand how this technique can help identify fake news (Bahad, Saxena & Kamal, 2020).

Despite much research being directed toward textual information, Song *et al*. (2021) and Varshney and Vishwakarma (2020) recommended research on the exploitation of visual information in search of fake news. Goel *et al*. (2021) highlight the relevance of further expanding research on fake news in other areas, given that many were restricted to the identification of fake news in datasets exclusive to political news.

Fang *et al*. (2019) recommended a better understanding of how the classifier detects fake news. Thus, it allows modifying or replacing features to avoid detection method that relies on very specific semantics of fake news, which could be explored and generate misclassification (e.g., false positives, false negatives).

For Albahar (2021), the challenge is great, and researchers need to devote more attention to understanding the patterns of news structures and what is considered false in the digital universe. For the researcher, fake digital news continues to acquire new formats, making it difficult to distinguish fake news embedded in long news.

Finally, some limitations regarding the performed analysis threaten this study's validity. Although the number of publications retrieved from the scientific repositories is expressive, the method employed does not intend to be exhaustive. Accuracy comparison between different algorithms and datasets provides a limited view of the matter. A more accurate comparison between algorithms required a controlled environment and advanced (statistic) analysis (e.g., n-fold cross-validation, paired t-tests). The classification of the algorithms is also variable according to different authors' perspectives and theoretical backgrounds, especially regarding mixed approaches, called hybrid methods.

5 Conclusion

This research intended to investigate the computational techniques and datasets used in fake news identification, analyzing the accuracy reported in scientific literature. For this, three questions were investigated. Regarding the accuracy of the main algorithms used to identify fake news (RQ1), the top three approaches are as follows: the Stacking Method, with 99.9% accuracy, Bidirectional Recurrent Neural Network (BiRNN), with 99.8%, and the Convolutional Neural Network (CNN), also with 99.8%.

The most popular technique was CNN, being used in ten studies. The scientific evolution in the past years for fake news identification is remarkable. An accuracy superior to 90% was reached only in 2019, with the 21 highest accuracies, above 96.6%, dating between 2020 and 2021.

Regarding the datasets used for the identification of fake news (RQ2), Kaggle has a more significant predominance, probably due to its popularity and contemplating several datasets on its platform for studies of artificial intelligence. After Kaggle, Weibo (i.e., a Chinese microblog similar to Twitter), FNC-1, COVID-19 Fake News, and Twitter were found and are presented in order considering the highest number of occurrences in the analyzed studies.

The top recommendations for future research in fake news identification (RQ3) are pointed out as follows:

- the use of other languages in the datasets;
- the classification through a scoring model;
- development based on hybrid models;
- the use of algorithms based on deep learning;
- the exploration of visual information;
- expansion of research in other areas beyond politics;
- the replacement of keywords with synonyms;
- understanding the patterns of news structures.

It is emphasized that the accuracy of 90% is considered a relevant result in this complex process of identifying fake news. Most of the research used datasets in controlled environments (e.g., Kaggle) or without information updated in real-time (from social networks). Few studies were applied directly in social network environments (where there is greater dissemination of disinformation).

The results show a compelling development in computational techniques to identify fake news. However, considering the ongoing trend, more research is still demanded to tackle the increasing complexity of fake digital news on social media. Thus, we suggest for future research the need to extend research beyond political news, an area that was the primary motivator for the growth of research from 2017, and the use of hybrid methods for fake news classification.

References

Almeida, L.D., Fuzaro, V. Nieto, F., & Santana, A.L.M. (2021). Identificação de "Fake News" no contexto político brasileiro: uma abordagem computacional. In: *Proceedings Workshop sobre as Implicações da Computação na Sociedade (WICS)*, Porto Alegre, Brasil, pp. 78-89. https://doi.org/10.5753/wics.2021.15966

Abouzeid, A., Granmo, O.C., Webersik, C., & Goodwin, M. (2019). Causality-based Social Media Analysis for Normal Users Credibility Assessment in a Political Crisis. *Proceedings of 25th Conference of Open Innovations Association (FRUCT)*, Helsinki, Finland, pp. 1-14. https://doi.org/10.23919/FRUCT48121.2019.8981500

Agarwal, A., Mittal, M., Pathak, A., & Goyal, L.M. (2020). Fake News Detection Using a Blend of Neural Networks: An Application of Deep Learning. *SN Computer Science*, 1,3:1-9.

Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M.O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. *Hindawi*, 1-11. https://doi.org/10.1155/2020/8885861

Ahmed, H., Traore, I., & Saad, S. (2017). Detection of Online Fake News Using NGram Analysis and Machine Learning Techniques. *Proceedings of International conference on intelligent, secure, and dependable systems in distributed and cloud environments*, Springer, Cham, pp. 127-138

Ahuja, N., & Kumar, S. (2020). S-HAN: Hierarchical Attention Networks with Stacked Gated Recurrent Unit for Fake News Detection. *Proceedings 8th International Conference on Reliability, Infocom Technologies and Optimization*, Noida, India, pp. 873-877.

Ajao, O., Bhowmik, D., & Zargari, S. (2019). Sentiment aware fake news detection on online social networks. *Proceedings ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, pp. 2507-2511. https://doi.org/10.1109/ICASSP.2019.8683170

Al-Ahmad, B., Al-Zoubi, A.M., Abu Khurma, R., & Aljarah, I. (2021). An evolutionary fake news detection method for covid-19 pandemic information. *Symmetry*, 13,6:1-16.

Alanazi, S.S., & Khan, M.B. (2020). Arabic fake news detection in social media using readers' comments: Text mining techniques in action. *International Journal of Computer Science and Network Security*, 20,9:29-35.

Albahar, M. (2021). A hybrid model for fake news detection: Leveraging news content and user comments in fake news. *IET Information Security*, 15,2:169-177.

Albahr, A., & Albahar, M. (2020). An empirical comparison of fake news detection using different machine learning algorithms. *International Journal of Advanced Computer Science and Applications*, 11,9:146-152.

Ali, H., Khan, M. S., AlGhadhban, A., Alazmi, M., Alzamil, A., Al-Utaibi, K., & Qadir, J. (2021). All Your Fake Detector Are Belong to Us: Evaluating Adversarial Robustness of Fake-news Detectors Under Black-Box Settings. *IEEE Access*, 9:81678-81692.

Asaad, B., & Erascu, M. (2018). A tool for fake news detection. *Proceedings 2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, Timisoara, Romania, pp. 379-386. https://doi.org/10.1109/SYNASC.2018.00064

Aslam, N., Ullah Khan, I., Alotaibi, F.S., Aldaej, L.A., & Aldubaikil, A.K. (2021). Fake detect: A deep learning ensemble model for fake news detection. *Hindawi*, 1-8. https://doi.org/10.1155/2021/5557784

Ayoub, J., Yang, X.J., & Zhou, F. (2021). Combat COVID-19 infodemic using explainable natural language processing models. *Information Processing & Management*, 58,4:1-11. https://doi.org/10.1016/j.ipm.2021.102569

Bahad, P., Saxena, P., & Kamal, R. (2020). Fake News Detection using Bi-directional LSTM-Recurrent Neural Network. *Procedia Computer Science*, 165:74-82.

Barua, R., Maity, R., Minj, D., Barua, T., & Layek, A.K. (2019). F-NAD: An application for fake news article detection using machine learning techniques. Proceedings IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India, pp. 1-6). https://doi.org/10.1109/IBSSC47189.2019.8973059

Birunda, S.S., & Devi, R.K. (2021). A Novel Score-Based Multi-Source Fake News Detection using Gradient Boosting Algorithm. *Proceedings of International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, Coimbatore, India, pp. 406–414. https://doi.org/10.1109/ICAIS50930.2021.9395896

Burfoot, C., & Baldwin, T. (2009). Automatic satire detection: Are you having a laugh? *Proceedings of the ACL-IJCNLP 2009 conference short papers*, p. 161–164.

Chapra, S.C., & Canale, R.P. (2016). *Métodos Numéricos para Engenharia*-7ª Ediçao. McGraw Hill Brasil.

Chen, W., Yang, C., Cheng, G., Zhang, Y., Yeo, C.K., Lau, C.T., & Lee, B.S. (2018). Exploiting Behavioral Differences to Detect Fake News. *Proceedings 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, USA, pp. 879-884. https://doi.org/10.1109/UEMCON.2018.8796519

Collins, B., Hoang, D.T., Nguyen, N.T., & Hwang, D. (2021). Trends in combating fake news on social media–a survey. *Journal of Information and Telecommunication*, 5,2:247-266. https://doi.org/10.1080/24751839.2020.1847379

Dadkhah, S., Shoeleh, F., Yadollahi, M.M., Zhang, X., & Ghorbani, A.A. (2021). A real-time hostile activities analyses and detection system. *Applied Soft Computing*, 104:1-28. https://doi.org/10.1016/j.asoc.2021.107175

Dresch, A., Lacerda, D.P. & Antunes Júnior, J.A.V. (2015). *Design science research*: método de pesquisa para avanço da ciência e tecnologia. Bookman Editora.

Fang, Y., Gao, J., Huang, C., Peng, H., & Wu, R. (2019). Self multi-head attention-based convolutional neural networks for fake news detection. *PloS one*, 14,9:1-13. https://doi.org/10.1371/journal.pone.0222713

Faustini, P.H.A., & Covões, T.F. (2020). Fake news detection in multiple platforms and languages. *Expert Systems with Applications*, 158:1-9. https://doi.org/10.1016/j.eswa.2020.113503

Gangireddy, S.C.R., Long, C., & Chakraborty, T. (2020). Unsupervised fake news detection: A graph-based approach. *Proceedings of the 31st ACM conference on hypertext and social media*, pp. 75-83. https://doi.org/10.1145/3372923.3404783

Gereme, F., Zhu, W., Ayall, T., & Alemu, D. (2021). Combating fake news in "low-resource" languages: Amharic fake news detection accompanied by resource crafting. *Information*, 12,1:1-9. https://doi.org/10.3390/info12010020

Goel, P., Singhal, S., Aggarwal, S., & Jain, M. (2021). Multi Domain Fake News Analysis using Transfer Learning. *Proceedings 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, pp.

1230–1237. https://doi.org/10.1109/ICCMC51019.2021.9418411

Goldani, M.H., Momtazi, S., & Safabakhsh, R. (2021). Detecting fake news with capsule neural networks. *Applied Soft Computing*, 101:1-8. https://doi.org/10.1016/j.asoc.2020.106991

Goldani, M.H., Safabakhsh, R., & Momtazi, S. (2021). Convolutional neural network with margin loss for fake news detection. *Information Processing & Management*, 58,1:1-12. https://doi.org/10.1016/j.ipm.2020.102418

Islam, M.R., Liu, S., Wang, X., & Xu, G. (2020). Deep learning for misinformation detection on online social networks: a survey and new perspectives. *Social Network Analysis and Mining*, 10,1:1-20. https://doi.org/10.1007/s13278-020-00696-x

Ivancová, K., Sarnovský, M., & Maslej-Krcšňáková, V. (2021). Fake news detection in Slovak language using deep learning techniques. *Proceedings of 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Herl'any, Slovakia, pp. 255-260. https://doi.org/10.1109/SAMI50585.2021.9378650

Jardaneh, G., Abdelhaq, H., Buzz, M., & Johnson, D. (2019). Classifying Arabic tweets based on credibility using content and user features. *Proceedings IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, Amman, Jordan, pp. 596-601. https://doi.org/10.1109/JEEIT.2019.8717386

Jiang, T., Li, J.P., Haq, A.U., & Saboor, A. (2020). Fake News Detection using Deep Recurrent Neural Networks. *Proceedings of 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China, pp. 205-208. https://doi.org/10.1109/ICCWAMTIP51612.2020.9317325

Jiang, T.A.O., Li, J.P., Haq, A.U., Saboor, A. & Ali, A. (2021). A novel stacking approach for accurate detection of fake news. *IEEE Access*, 9:22626-22639. https://doi.org/10.1109/ACCESS.2021.3056079

Kaliyar, R.K., Goswami, A., & Narang, P. (2019). Multiclass Fake News Detection using Ensemble Machine Learning. *IEEE 9th International Conference on Advanced Computing (IACC).* Tiruchirappalli, India, pp. 103-107. https://doi.org/10.1109/IACC48062.2019.8971579

Kaliyar, R.K., Goswami, A., & Narang, P. (2021). FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimedia tools and applications*, 80,8:11765-11788. https://doi.org/10.1007/s11042-020-10183-2

Kaliyar, R.K., Goswami, A., Narang, P., & Sinha, S. (2020a). 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

Kaliyar, R.K., Kumar, P., Kumar, M., Narkhede, M., Namboodiri, S., & Mishra, S. (2020b). DeepNet: an efficient neural network for fake news detection using news-user engagements.*Proceedings of 5th International Conference on Computing, Communication and Security (ICCCS)*, Patna, India, pp. 1-6). https://doi.org/10.1109/ICCCS49678.2020.9277353

Kesarwani, A., Chauhan, S.S., & Nair, A.R. (2020). Fake news detection on social media using k-nearest neighbor classifier. *Proceedings 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

Khattar, D., Goud, J.S., Gupta, M., & Varma, V. (2019). Mvae: Multimodal variational autoencoder for fake news detection. *Proceedings The world wide web conference*, pp. 2915-2921). https://doi.org/10.1145/3308558.3313552

Konkobo, P.M., Zhang, R., Huang, S., Minoungou, T.T., Ouedraogo, J.A., & Li, L. (2020). A deep learning model for early detection of fake news on social media. *Proceedings 7th International Conference on Behavioural and Social Computing (BESC)*, Bournemouth, United Kingdom, pp. 1-6). https://doi.org/10.1109/BESC51023.2020.9348311

Kumar, S., Asthana, R., Upadhyay, S., Upreti, N., & Akbar, M. (2020). Fake news detection using deep learning models: A novel approach. *Transactions on Emerging Telecommunications Technologies*, 31,2:1-23. https://doi.org/10.1002/ett.3767

Kumar, R., Anurag, K., & Pratik, G. (2021). EchoFakeD: improving fake news detection in social media with an efficient deep neural network. *Neural Computing and Applications*, 33,14:8597–8613. https://doi.org/10.1007/s00521-020-05611-1

Lakshmanarao, A., Swathi, Y., & Kiran, T.S.R. (2019). An effecient fake news detection system using machine learning. *International Journal of Innovative Technology and Exploring Engineering*, 8,10:3125-3129.

Li, Q., Hu, Q., Lu, Y., Yang, Y., & Cheng, J. (2020). Multilevel word features based on CNN for fake news detection in cultural communication. *Personal and Ubiquitous Computing*, 24,2:259–272.

Lin, J., Tremblay-Taylor, G., Mou, G., You, D., & Lee, K. (2019). Detecting fake news articles. *Proceedings IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, pp. 3021-3025, https://doi.org/10.1109/BigData47090.2019.9005980

Lin, S.X., Wu, B.Y., Chou, T.H., Lin, Y.J., & Kao, H.Y. (2020). Bidirectional perspective with topic information for stance detection. *Proceedings 2020 International Conference on Pervasive Artificial Intelligence (ICPAI)*, Taipei, Taiwan, pp. 1-8. https://doi.org/10.1109/ICPAI51961.2020.00009

Low, J. F., Fung, B. C., Iqbal, F., & Huang, S. C. (2022). Distinguishing between fake news and satire with transformers. *Expert Systems with Applications*, 187, 115824.

Medeiros, H.I., & Braga, R. B. (2020). Fake News detection in social media: a systematic review. *Proceedings 16th Simpósio Brasileiro de Sistemas de Informação (SBSI)*, Porto Alegre, Brasil, pp. 1-8. https://doi.org/10.5753/sbsi.2020.13782

Mugdha, S.B.S., Ferdous, S.M., & Fahmin, A. (2020). Evaluating machine learning algorithms for bengali fake news detection. *Proceedings 23rd International Conference on Computer and Information Technology (ICCIT)*, DHAKA, Bangladesh, pp. 1-6. https://doi.org/10.1109/ICCIT51783.2020.9392662

Najar, F., Zamzami, N., & Bouguila, N. (2019). Fake news detection using bayesian inference. *Proceedings 20th International Conference on Information Reuse and Integration for Data Science (IRI)*, Los Angeles, CA, USA, pp. 389-394. https://doi.org/10.1109/IRI.2019.00066

Nasir, J.A., Khan, O.S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights*, 1,1:1-13. https://doi.org/10.1016/j.jjimei.2020.100007

Ozbay, F.A., & Alatas, B. (2019). A novel approach for detection of fake news on social media using metaheuristic optimization algorithms. *Elektronika ir Elektrotechnika*, 25,4:62–67.

Pardamean, A., & Pardede, H.F. (2021). Tuned bidirectional encoder representations from transformers for fake news detection. *Indonesian Journal of Electrical Engineering and Computer Science*, 22,3:1667-1671.

Qawasmeh, E., Tawalbeh, M., & Abdullah, M. (2019). Automatic identification of fake news using deep learning. *Proceedings 6th international conference on social networks analysis, Management and Security (SNAMS)*, Granada, Spain, pp. 383-388. https://doi.org/10.1109/SNAMS.2019.8931873

Qi, P., Cao, J., Yang, T., Guo, J., & Li, J. (2019). Exploiting multi-domain visual information for fake news detection. *Proceedings IEEE international conference on data mining (ICDM)*, Beijing, China, pp. 518-527. https://doi.org/10.1109/ICDM.2019.00062

Ren, Y., Wang, B., Zhang, J., & Chang, Y. (2020). Adversarial active learning based heterogeneous graph neural network for fake news detection. *Proceedings IEEE International Conference on Data Mining (ICDM)*, Sorrento, Italy, pp. 452-461. https://doi.org/10.1109/ICDM50108.2020.00054

Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 797-806.

Sahoo, S.R., & Gupta, B.B. (2021). Multiple features based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, 100:1-16. https://doi.org/10.1016/j.asoc.2020.106983

Shabani, S., & Sokhn, M. (2018). Hybrid machine-crowd approach for fake news detection. *Proceedings of 4th International Conference on Collaboration and Internet Computing (CIC)*, Philadelphia, PA, USA, pp. 299-306. https://doi.org/10.1109/CIC.2018.00048

Sharma, D.K., Garg, S., & Shrivastava, P. (2021). Evaluation of tools and extension for fake news detection. *Proceedings of International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Noida, India, pp. 227-232. https://doi.org/10.1109/ICIPTM52218.2021.9388356

Song, C., Ning, N., Zhang, Y., & Wu, B. (2021). A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Information Processing and Management*, 58,1:1-14.

Sridhar, S., & Sanagavarapu, S. (2021). Fake news detection and analysis using multitask learning with BiLSTM CapsNet model. *11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, pp. 905-911. https://doi.org/10.1109/Confluence51648.2021.9377080

Thakur, A., Shinde, S., Patil, T., Gaud, B., & Babanne, V. (2020). MYTHYA: Fake News Detector, Real Time News Extractor and Classifier. *Proceedings of 4th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp. 982-987. https://doi.org/10.1109/ICOEI48184.2020.9142971

Torgheh, F., Keyvanpour, M.R., Masoumi, B., & Shojaedini, S.V. (2021). A Novel Method for Detecting Fake news: Deep Learning Based on Propagation Path Concept. *Proceedings of 26th International Computer Conference, Computer Society of Iran (CSICC)*, Tehran, Iran, pp. 1-5. https://doi.org/10.1109/CSICC52343.2021.9420601

Umer, M., Imtiaz, Z., Ullah, S., Mehmood, A., Choi, G.S., & On, B.W. (2020). Fake news stance detection using deep learning architecture (CNN-LSTM). *IEEE Access*, 8: 156695-156706.

Varshney, D., & Vishwakarma, D.K. (2021). Hoax news-inspector: a real-time prediction of fake news using content resemblance over web search results for authenticating the credibility of news articles. *Journal of Ambient Intelligence and Humanized Computing*, 12,9:8961-8974. https://doi.org/10.1007/s12652-020-02698-1

Verma, P.K., Agrawal, P., Amorim, I., & Prodan, R. (2021). WELFake: Word Embedding Over Linguistic Features for Fake News Detection. *IEEE Transactions on Computational Social Systems*, 8,4:881-893.

Wang, Y., Qian, S., Hu, J., Fang, Q., & Xu, C. (2020). Fake news detection via knowledge-driven multimodal graph convolutional networks. *Proceedings of the 2020 International Conference on Multimedia Retrieval*, pp. 540-547. https://doi.org/10.1145/3372278.3390713

Wang, Y., Wang, L., Yang, Y., & Lian, T. (2021). SemSeq4FD: Integrating global semantic relationship and local sequential order to enhance text representation for fake news detection. Expert Systems with Applications, 166:1-12. https://doi.org/10.1016/j.eswa.2020.114090

Xie, J., Liu, S., Liu, R., Zhang, Y., & Zhu, Y. (2021). SeRN: Stance extraction and reasoning network for fake news detection. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, pp. 2520-2524. https://doi.org/10.1109/ICASSP39728.2021.9414787

Yu, J., Huang, Q., Zhou, X., & Sha, Y. (2020). Iarnet: An information aggregating and reasoning network over heterogeneous graph for fake news detection. *Proceedings of 2020 International Joint Conference on Neural Networks (IJCNN)*, Glasgow, UK, pp. 1-9. https://doi.org/10.1109/IJCNN48605.2020.9207406

Zhang, H., Alim, M.A., Li, X., Thai, M.T., & Nguyen, H.T. (2016). Misinformation in online social networks: Detect them all with a limited budget. *ACM Transactions on Information Systems*, 34,3:1-24. https://doi.org/10.1145/2885494

Zhang, Q., Lipani, A., Liang, S., & Yilmaz, E. (2019). Reply-aided detection of misinformation via bayesian deep learning. *Proceedings of 19 the world wide web conference*, pp. 2333-2343. https://doi.org/10.1145/3308558.3313718