Social Network Analysis and Mining (2024) 14:120
https://doi.org/10.1007/s13278-024-01280-3

REVIEW PAPER

# A comprehensive overview of fake news detection on social networks

**Upasna Sharma**[1] · **Jaswinder Singh**[1]

Received: 11 February 2024 / Revised: 5 June 2024 / Accepted: 6 June 2024 / Published online: 24 June 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2024

**Abstract**
As social media and web-based forums have grown in popularity, the fast-spreading trend of fake news has become a major threat to the government and other agencies. With the rise of social media and internet platforms, misinformation may quickly spread over borders and language boundaries. Detecting and neutralizing fake news in several languages can help to protect the integrity of global elections, political discourse, and public opinion. The lack of a robust multilingual database for training the classification models makes detecting fake news a difficult task. This paper looks at it by describing several forms of fake news (like serious fabrications, large-scale hoaxes, stance news, deceptive news, satire news, clickbait, misinformation, rumour). This review paper includes different steps, features, tools for mitigating the scourge of information pollution, and different available datasets. This study presented a taxonomy for detecting fake news, which gives a comprehensive overview and analysis of existing DL-based algorithms focusing on diverse techniques. This paper also includes the monolingual and multilingual fake news detection models. Finally, this paper ends with the technical challenges.

**Keywords** Social media · Fake news detection · Deep learning · Natural language processing · Linguistic features · Crowdsource

## 1 Introduction

The concept of fake news existed long before the Internet and other computational technologies were developed. It is believed that fake news was effectively used for the first time long back in the 18th century. In 1779 during the American civil war, where Benjamin Franklin spread fake news for the first time by sending a letter to Capt Samuel Gerrish in which the cruelties by Britishers and their allied forces were mentioned in such a way in the newspapers that it should openly influence the public opinion. As a result, the dialogue for peace got stopped (Okoro et al. 2018). But with the advancement in techniques, the chances to spread fake news with the help of the internet, social media sites, and apps have increased exponentially. A growing number of so-called "fake news" stories have emerged in recent social and political events, such as the 2016 US presidential election (Bovet and Makse 2019). The authors discovered that 25% of the tweets transmitted are fake or extremely biased news. Fake pandemic news travels throughout the Internet during the Covid19, producing widespread psychological fear among citizens.

Fake news is a type of social media news report that contains purposefully inaccurate information (Rashkin et al. 2017). Fake news frequently employs multimodal information, such as text and images, to deceive readers and spread its effect. Fake news is becoming more common on social media sites such as Twitter and Facebook. Social media has become the primary platform for online social interaction and information transmission due to its ease of use, low cost, and rapid rate (Shu et al. 2017). According to statista.com, a study showed that 47% of polled U.S. adults said they had seen a lot or at least some fake news regarding the coronavirus, underlining the problem of fake news circulation and untrustworthy sources attempting to profit from the public's desire for news and updates during times of crisis. Detecting fake news is a difficult job because it necessitates models that summarise the news and equate it to real news

✉ Upasna Sharma
erupasna401@gmail.com

Jaswinder Singh
dr.jaswinder@pbi.ac.in

1 Department of Computer Science and Engineering, Punjabi University Patiala, Patiala, India

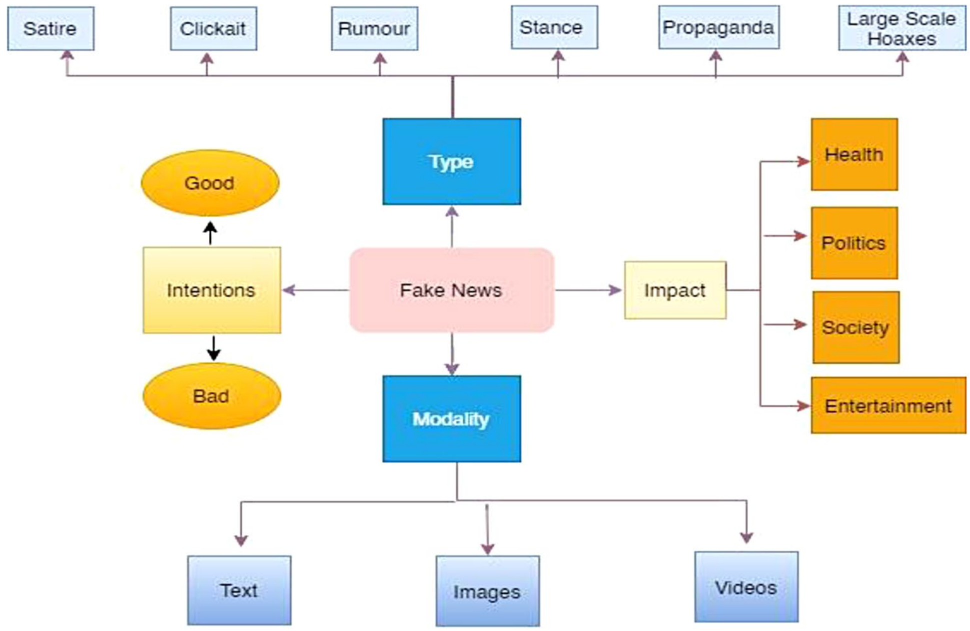



**Fig. 1** General structure of fake news

to identify it as fake (Thota 2018). Any news that is published with the intent to deceive or give the reader incorrect details is considered fake news. Detecting disinformation on social media is a critical issue that is often technically difficult to solve (Ruchansky 2017).

Several fact-checking websites have been deployed to reveal or validate reports in order to address the increasing misinformation and disinformation (e.g. snopes.com[1], politifact.com[2]). These websites are critical in the fight against fake news, but they need expert research, which delays responses.

The general structure of fake news in Fig. 1 which shows its various types, impact on the individual as well as the whole, intention behind spreading the fake news and modalities.

*Satire news* The news which criticizes ideas and people in a funny or entertaining way. It can be entrenched with humor, irony, or negative emotions (Rubin et al. 2015).

*Clickbait* The headlines whose main target is to attract users and inspire them to click on the link to reach the particular page(Chakraborty et al. 2016).

*Rumor* A rumor is an unsubstantiated assertion posted by users on social media platforms that has the potential to spread beyond their private network (Asghar et al. 2019).

*Stance news* This type of news tells us whether the two parts of the text are related to each other or not, which means the headline is related to the body of the news article or not.

Propaganda: Propaganda is defined as “news stories created by a political entity to influence public.” While governments are typically linked with propaganda, activist groups, corporations, religious organisations, the media, and individuals all engage in it.

*Large scale hoaxes* These are news articles having false information, but concealed as correct news in an attempt to mislead the viewers (Tacchini et al. 2017).

mislead individuals and convince them of false information (Wani et al. 2021). To address this, the authors have curated and released a manually annotated dataset of 10,700 real and false social media messages and articles on Covid19). Fake news can have an impact on individuals, organizations, and political parties. In the 2016 US presidential election, fake news influenced people’s opinions and decisions, and in 2019, when India launched an airstrike on Balakot (Kaliyar et al. 2020). Stock prices, stock buying decisions, investment plans, elections (Allcott and Gentzkow 2017), health, education, and even natural disaster reactions are all affected by false information (Hakak et al. 2021).

Figure 2 shows that the evaluation of social networks is important for the development of big data because it can generate a lot of data types every day (abundant), grow quickly in response to demand (alacrity), which has a very low degree of accuracy for data processing and analytics

[1] https://www.snopes.com.

[2] https://www.politifact.com.

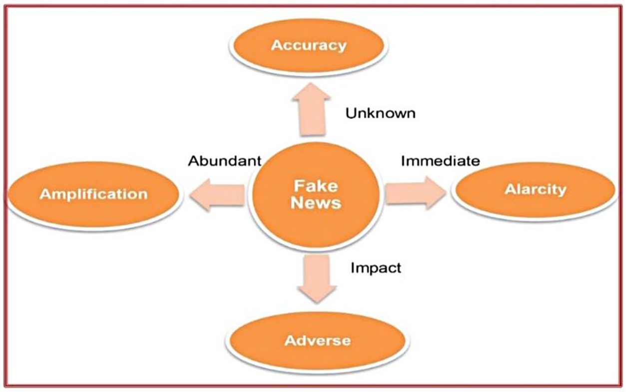



**Fig. 2** Fake News in terms of 4 A

(veracity), and have the potential to contribute adverse effect to society through spreading fake news. To identify false news, a powerful AI framework that can discriminate between "Genuine" and "Fake" news is required. Identifying real information among the vast volume of data (Pan et al. 2018) provided by users on social media platforms like Facebook, Twitter, Instagram, and YouTube is a challenging task (Kaliyar et al. 2020).

Misinformation can have an impact on all aspects of life, including social, political, economic, stock market, disaster response, and crises. Its goal is to deceive the public, influence political elections, education, online shopping behaviour, and health, and endanger public security and social stability, either purposefully or accidentally (Wu et al. 2019).

### 1.1 Motivation and contribution of this paper

Users on Facebook may be too quick to believe and spread false information. Multifarious public concerns and approaches to this issue have recently been articulated. Because of the volume of news published every day, many computer discipline studies have proposed deep learning-based models for detecting fake news. Fake news is one of the issues that the online world is dealing with these days. Fake news is the greatest threat to our ostensibly free press; in addition to distorting and corrupting beliefs, it has also resulted in practical consequences such as cybercrime, phishing, cyber-attacks, and so on.

The focus of this study is definitions, benchmarked datasets, taxonomy, features for fake news detection, open issues, and conclusion. Deep learning, which has recently captured the attention of practically everyone in this subject, is the greatest notion for addressing the problem. There are several sections to this survey paper. Section 2 discusses the different steps for creating fake news, available tools and services, and critical features such as the creator and target users of fake news. Section 3 presents the features of fake news detection. Section 4 categorizes and summarizes fact-checking datasets. Section 5 presents the taxonomy of Fake News Detection Techniques based on Deep Learning. Section 6 analysed the monolingual and multilingual fake news detection models. Section 7 discusses the comparisons of existing fake news detection model based. Section 8 finally, the paper is concluded with technical challenges, open issues, and conclusion (Figs. 3 and 4).

### 1.2 Chronological review

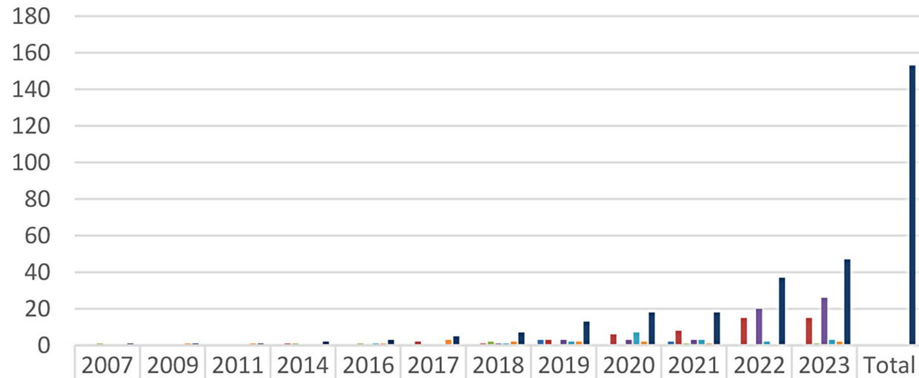
The papers for this literature review were gathered from a variety of databases. The chronological review of publication patterns from 2007 to 2023 is indicated in Fig. 3. The information pertains to the number of contributions made in the realm of identifying fake news with machine learning and deep learning algorithms. It is shown that interest in detecting fake news climbed from 1 in 2007, 2009, and 2011 to 2 in 2014. Figure 3 illustrates the growing popularity of deep learning approaches over the years 2021,2022, and 2023. Hybrid approaches may lead to more publications, but their usefulness is still being established in this emerging field. This inspires researchers to use more new approaches in the coming years.

#### 1.2.1 Exclusion and inclusion

A literature review was conducted to identify existing studies on detecting false news. The Table 1 below clearly illustrates the criteria for accepting and rejecting papers. Papers were gathered from several databases. But not all of them were germane to the subject. So, first and foremost, papers were rejected based on their titles and abstracts. An abstract is a brief overview of the entire work that provides insight into the information contained in it. In the following phase, the remaining papers were compared to the standards for exclusion and inclusion. Against the search phrase, fifteen hundred papers were gathered from various databases like PubMed, Google Scholar, IEEE Xplore, Elesvier, Arxiv, and ACM Digital Library. Following the elimination, one hundred fifty-three papers remained for discussion in this literature review.

## 2 Different steps of fake news

A limited number of people develop and initiate false information. People, relationships, content, and time are four key aspects of networked data that may be analysed multidimensionally using an iOLAP framework based on polyadic factorization (Chi et al. 2009). Fake news on the internet




| | 2007 | 2009 | 2011 | 2014 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Machine Learning | | | | | | | | 3 | | 2 | | | |
| Supervised Deep Learning | | | | 1 | | 2 | 1 | 3 | 6 | 8 | 15 | 15 | |
| Unsupervised Deep Learning | 1 | | | 1 | 1 | | 2 | | | 1 | | 1 | |
| Hybrid Deep Learning | | | | | | | 1 | 3 | 3 | 3 | 20 | 26 | |
| Monolingual & Multilingual | | | | | 1 | | 1 | 2 | 7 | 3 | 2 | 3 | |
| Fake News Datasets | | 1 | 1 | | 1 | 3 | 2 | 2 | 2 | 1 | | 2 | |
| Total Papers | 1 | 1 | 1 | 2 | 3 | 5 | 7 | 13 | 18 | 18 | 37 | 47 | 153 |



**Fig. 3** A chronological review of publications about existing fake news detection models

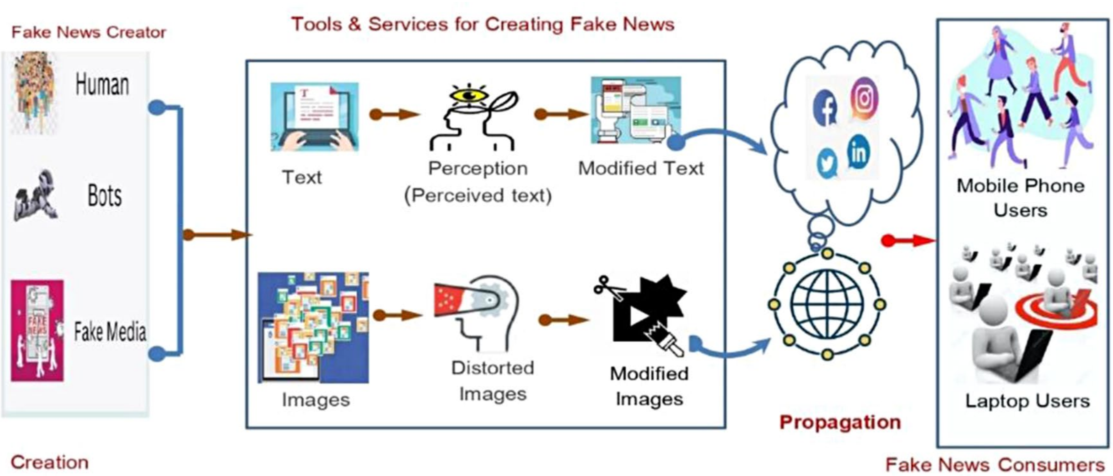



**Fig. 4** Different steps of fake news

**Table 1** Criteria for accepting and rejecting papers

| Inclusion criteria | Exclusion criteria |
|---|---|
| English-written papers. | The paper is written in another language than English. |
| Papers that can be fully accessed | Papers that are not fully accessible |
| Papers related to fake news detection using machine /deep learning techniques. | Papers unrelated to machine/deep learning. Papers related to stance, rumour, and spam detection |

nowadays is obtrusive and varied in terms of topics, styles, and platforms (Shu et al. 2017). According to this paper, fake news has reasonably manipulated, false, and misleading content that can drastically influence the passion and motive of journalism, may it be print media, electronic, and social media.

Fake news can be clearly defined and understood by using the following components Fig 4.

### 2.1 Fake news developer/publisher

Fake news can be created by humans, cyborgs, spammers and social bots also. The most prevalent non-human fake news developers are social bots and cyborgs. Bots on social networking sites post incessantly, disseminating uncoated information utilizing trending topics and hashtags as the primary techniques for reaching a larger audience, which, in many cases, aids the spread of fake news. Spammers are relentless and very inventive. To persuade users to open misleading news messages, they are continuously producing bait that is ever more appealing. (Stringhini et al. 2010) analyze the spread of spam on social networks. Spammers typically spread their spam by sending out numerous tweets at once and uploading identical false postings after a predetermined amount of time (Shahid et al. 2022). The publisher is defined as the entity that distributes the narrative to the general public. Trolls, in this context, are those who have accounts on social media sites for the sole goal of harassing others. To elicit arguments, insults, and slurs directed towards other users and public figures. The ultimate generators of untreatable information are real persons who seek to undermine the integrity of online social communities.

### 2.2 Tools and services

Many of these tools and services are sold in various online communities throughout the world for manipulating and propagating the message across relevant social media networks. There is a wide range of tools and services available, some of which are quite straightforward (for example likes/followers/shares/tweets/retweets). People are spending more time on these sites to stay up to date on the newest news and information, therefore their role in propagating fake news cannot be overstated. Fake news employs the use of accurate photographs or video to fabricate a story. With the introduction of digital photography, powerful photo editing tools, and technological competence, photo processing has become increasingly prevalent. Deepfake films employ AI algorithms to deceptively swap out a person's face or voice for someone else's, enabling the synthetic media to be used maliciously to propagate misinformation or for amusement.(Verdoliva 2020) also analysed the techniques for detecting faked photos and movies, often known as visual media integrity verification.

### 2.3 News content

The publisher's major information in the narrative is referred to as content. The authenticity of this material may be genuine, false, or unknown at the time of publication. News is made up of four components (Shu et al. 2017):

**Source** This tells us where the news has come from, who write it, and whether it is a reliable source or not.

**Headline** Main attraction of the news content, which trying to attract the reader.

**Body text** Which gives the actual information about the news.

**Images/Video** Audio, video, and images that are embedded in the news articles.

### 2.4 Social context

The intrinsic tri-relationship is formed by the social environment during the news transmission process on social media between publishers, news articles, and readers, with the potential to enhance the detection of fake news (Shu et al. 2017). The quality of social media news is far inferior to that of established news agencies. Fake news or news with purposely incorrect information is widely distributed on the internet for a variety of reasons, including financial and political benefits. Because of the low cost and simple accessibility to the Internet, a growing number of traditional media outlets, like NBC News, the New York Times, and The Washington Post, are making substantial shifts from mainstream to digital platforms. To detect fake news at an early stage, one has to primarily and efficiently rely on news content and limited social context information, which leads to facing multiple challenges (Zhou and Zafarani 2020). Social context models have been used by various existing fake news detection approaches as well as for rumor detection (Shu et al. 2017).

#### 2.4.1 Stance based

The task of automatically assessing whether a person is in favor of, neutral toward, or against a specific entity, event, or idea from a post is known as stance detection. Information retrieval, text summarization, and textual entailment all benefit from automatically recognizing stance. To learn latent stance from topics, topic model approaches such as Latent Dirichlet Allocation (LDA) can be used (Yin et al. 2017).

#### 2.4.2 Propagation-based

The task of investigating the information authenticity of news spread via Twitter, a popular microblogging platform. They concentrated on approaches for evaluating the credibility of a series of tweets that are automatically generated.

Microblog posts about "trending" subjects are evaluated and classified as credible or not (Castillo et al. 2011). We can have built homogeneous and heterogeneous credibility networks for the evaluation of the propagation process. The network was represented as a directed graph of the Flicker social network, which is characterized by a significant clustering in the nearby area (Cha et al. 2009).

#### 2.4.3 Target victims/consumers

The production of fake news with the intention of attacking unpopular, frequently oppositional social and political groups. Users of online social media or other online news platforms could be among them. Students, voters, parents, senior citizens, and other groups may be targeted depending on the news's goals. One of the most popular activities on social media, according to consumers, is reading or watching the news. Spending too much time on social media exposes users to incorrect or fraudulent content, especially if they have a strong political identity or participation (Halpern et al. 2017) (Fig. 4 and 5).

## 3 Features for detecting false news

The following section of the study presents a thorough, critical examination of DL-based methods for false news identification. Content-based approaches, on the one hand, rely on content features, which refer to information that may be derived directly from the text, such as linguistic features. User-based features give the characteristics of the user's information. Context-based techniques, on the other hand, are more diverse. Figure 6 outlines the various types of features for representing and identifying fake news.

### 3.1 Content-based features

To classify fake news contents of news and news sources need to be analysed carefully. Linguistic features are collected from text material at various levels, such as characters, words, phrases, features at the sentence level, such as the frequency of function words and phrases i.e., n-grams (Zhang et al. 2017), and approaches to a bag of words (Caropreso and Matwin 2006), in terms of document organization, Punctuation and Parts of Speech (POS) tagging. Natural Language Tool Kit (NLTK) NLP offers a collection of techniques and algorithms that facilitate the processing of multimodal data flowing on OSNs, turning unstructured data into structured data (Camacho et al. 2020). (Wagner 2010) is used to tokenize text. A stylistic distinction exists between authentic and false news pieces, according to a study of thousands of news articles. Genuine news stories use more language to differentiate themselves, whilst fraudulent news items use more certainty.

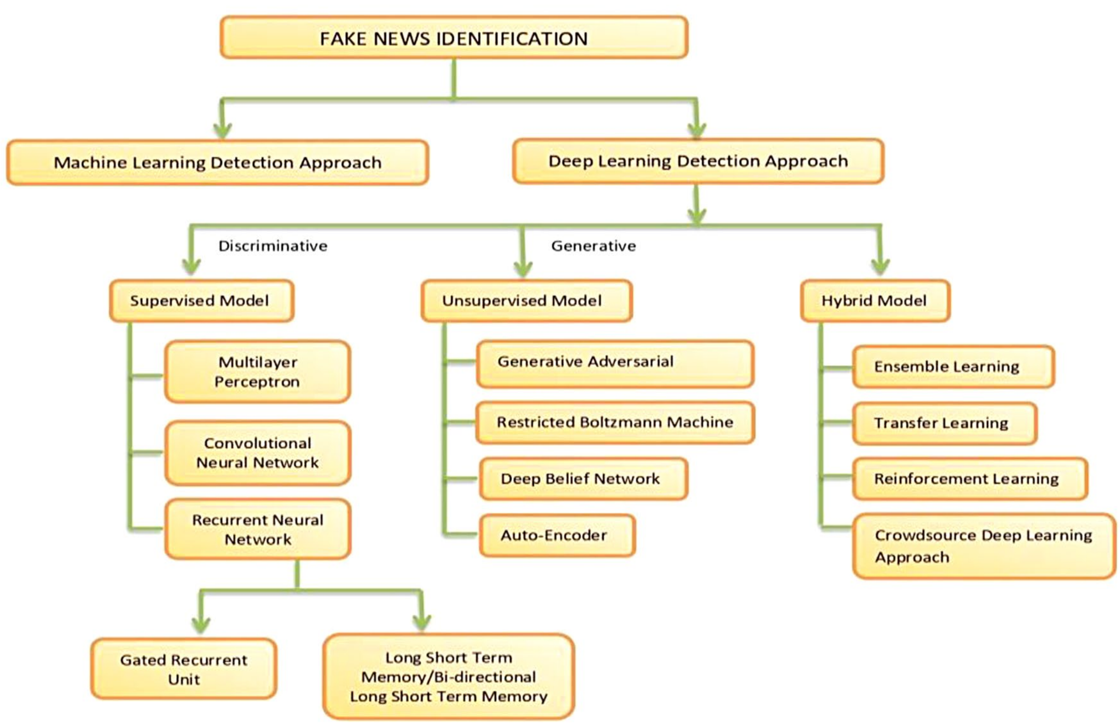



**Fig. 5** Taxonomy for identifying fake news using deep learning approaches


**Fig. 6** Some images of fake news spread over social media. Source: Twitter

#### 3.1.1 Syntactic features

Syntactic features mean the words and sentences such as the number of adjectives, nouns, and verbs. The occurrence of particular POS patterns has been assessed by part-of-speech taggers. Additionally, it has been discovered that sentence complexity serves as both a subjectivity index and a sign of the veracity of the information. Deep syntax analysis has been examined to envisage the instances of deception.

#### 3.1.2 Semantic features

The content of the headline and title derives some meaning or not. This gives the analysis of language patterns, structures, and meanings of the news. The title of the fake news is in conflict with or not related to the body of the news. Semantic characteristics are also calculated by including such as sentiment analysis, named entity recognition, linkages, and relations (Braşoveanu and Andonie 2019). Visual semantic features that describe the semantic level of textual and visual features are also analysed by authors(Xue et al. 2021). Bag of words, n-grams (Zhang et al. 2017), and many more features extracting techniques are used for detecting these types of features.

#### 3.1.3 Sentiment features

Sentiment Analysis, also known as opinion mining, is a method of consciously distinguishing, separating, valuing, and examining emotional states and abstract information that employs content inquiry and statistics. It aims to decipher the true meaning of phrases and sentences. It is also known as polarity, and it assigns classification levels to the authors' declarations. Classification could be positive, negative, or neutral. (Bhutani et al. 2019) developed a novel technique for detecting fake news that includes sentiment as a key characteristic to increase accuracy.

#### 3.1.4 Lexical features

It includes character, frequency of word usage in text, unique words doubt words, negation words, abbreviations, and vulgar expressions in the news text. N-gram techniques (Bondielli and Marcelloni 2019) can be employed for finding the salient features in the content and expressions.

#### 3.1.5 Style-based features

Style-based features try to capture the writing style of fake news authors containing different characteristics. Fake news creators aim to copy the writing style of actual news creators to mislead online users. Fake news creators have still left some clues, which helps in distinguishing the fake news content from the real news content. In order to detect the style-based patterns(Ahmed, 2017) studied the keystroke pattern and time span of the fake news creator and the real news creator.

### 3.2 User-based features

With the widespread of online users and social media websites, the number of online users is increasing day by day (Parikh and Atrey 2018). In order to identify the unique characteristics of real users, the suspicious user and the non-humans like social bots.

#### 3.2.1 User profiling features

These features capture the personal digital data associated with the user. It covers the account name, the date of registration of the user, how many accounts a particular user has, geolocation information, the account is verified or not.

#### 3.2.2 User credibility features

The credibility of the user is evaluated by the number of friends, followers, and status of the account. The number of followers of legitimate users is often close to the number of friends. But the number of followers is more than the actual friends in the case of social bots. This feature assessed the credibility of the online account.

### 3.3 Social context based features

Social context-based features help in evaluating the characteristics of the distribution pattern of online news. These features help in observing the reactions and emotions that certain topics generate. Users can acquire customized local coherent structures faster with community search, especially in large-size networks. So far, a great variety of community search methods have been proposed (C. Wang and Zhu 2019), including topology-based methods and semantics-enhanced methods.

#### 3.3.1 Network based features

The features of network based analysis helps in finding the unique characteristics of certain networks (Shu et al. 2018), and the likeness and unlikeness of different online accounts, gives the brief idea of different network analysis: friendship network which indicates the following/followee structure of users whose post related tweets or articles and diffusion network (Kwon et al. 2013) which helps in tracking the route of the spread of news.

#### 3.3.2 Distribution based features

These features aim to capture the distinct propagation patterns of spreading online news. It helps in getting the propagation-based features of particular most frequent URLs, user mention, hashtag, the total number of tweets, and retweets. A propagation tree can be used to determine the origin of the news, post, tweets, etc. It helps in generating the features associated with the propagation tree, which is containing the information on the number of subtrees, degree of the root node and depth of the tree, etc.

#### 3.3.3 Temporal based features

Temporal-based features aim to capture the online user creator commenting activities in a time-series manner. For detecting suspicious posting actions, temporal-based features are a good characteristic to identify the false level of news (Alzanin and Azmi 2018). The most frequently used temporal-based features include the time gap between the two posts and the rate at which posting, commenting, and replying for a certain post occurs. time, day, and week in which the particular post was published, shared and commented (Tables 2, 3, 4, 5, 6 and 7).

## 4 Datasets of fake news

Different existing benchmarked datasets are discussed in detail:

**Ott et al.'s dataset** Contains the dataset of fake reviews about the tourism news domain having 800 reviews in the English language (Ott et al. 2011). Samples were collected from TripAdvisor's[3] social media platform.

**Burfoot Satire News Dataset**[4] The genuine news items were gathered using newswire document samples from the English Gigaword Corpus, while the satire news stories were chosen based on their topical similarity to the real ones (Horne & Adali 2017).

**Fact-checking dataset** This dataset consists of a total of 221 statements from two popular fact-checking websites Channel 4 and Politifact (Vlachos and Riedel 2014).

**Benjamin Political News Dataset**[5] The fake news sources were gathered from Zimdars' list of fake and deceptive news websites. The true sources are well-established news media businesses from Business Insider's "Most-Trusted" list.
**BuzzFeed News dataset**[6] This dataset includes additional relevant information for each news sample, such as the URL of the news piece and the date it was published (Horne & Adali 2017).
**LIAR**[7] This data was gathered using the PolitiFact API (Wang 2017), a fact-checking website. News veracity is classified into several categories: false, barely true, half-true, mainly true, and true.
**CREDBANK**[8] Dataset with roughly 60 million tweets spanning 96 days, beginning in October 2015 were evaluated by 30 Amazon Mechanical Turk annotators (Mitra and Gilbert 2015).

[3] http://tripadvisor.com.

[4] https://github.com/rfong/satire/tree/master/corpus.

[5] https://github.com/rpitrust/fakenewsdata1.

[6] The dataset is analysed from: https://github.com/BuzzFeedNews/2016-10-facebook-fact-check/tree/master/data.

[7] The dataset is analysed from https://www.cs.ucsb.edu/~william/data/liar_dataset.zip.

[8] The dataset is analysed from: https://github.com/compsocial/CREDBANK-data.

**Table 2** list of the features that are utilised to detect fake information

| S. No. | Features extracted by classifiers | Citation |
|---|---|---|
| 1 | **Content/Text features** | (X. Zhang and Ghorbani 2019),(Caropreso and Matwin 2006), (Kwon et al. 2013). 2013, (Ahmed, 2017), (Castillo et al. 2011), (Zhang et al. 2017), (Ma et al. 2016), (Yang et al. 2018), (Reis et al. 2019), (Shu et al. 2018), (Giachanou et al. 2022) |
| | Average sentence length | (Verma et al. 2021) |
| | The average length of a post/ tweet | (Castillo et al. 2011), (Sahoo and Gupta 2019) |
| | Sentence with the phrase | (Castillo et al. 2011), (Sahoo and Gupta 2019) |
| | Number of uppercase letters | (Castillo et al. 2011), (Yang et al. 2018), (Sahoo and Gupta 2019) |
| | function words, complex words, long words | (Castillo et al. 2011), (Pérez-Rosas et al. 2018), (Yang et al. 2018) |
| | Textual sentiment analysis | (Sharma and Garg 2021), (Sahoo and Gupta 2019), (Kishwar and Zafar 2023) |
| | Punctuation, special symbols | (Yang et al. 2018), (Sahoo and Gupta 2019), (Pérez-Rosas et al. 2018), (Jin et al. 2016) |
| | Length of the news | (Pérez-Rosas et al. 2018), (Castillo et al. 2011), (Verma et al. 2021) |
| | Frequently occurring words | (Castillo et al. 2011), (Horne & Adali 2017), (Kwon et al. 2013) |
| | Negation words, use of pronouns, exclusive words | (Castillo et al. 2011), (Ahmed, 2017), (Kwon et al. 2013), (Yang et al. 2018) |
| | Unigram, bigram, ngram | (Ahmed et al. 2017), (Kaur et al. 2020), (Vereshchaka et al. 2020),(Kishwar and Zafar 2023) |
| 2 | **Social context based features** | (Shu et al. 2018) |
| | Tweets posted by users | (Shu et al. 2018), (Ma et al. 2016),(Alyoubi et al. 2023) |
| | URLs posted by users | (Grinberg et al. 2019), (Y. Li et al. 2020b) |
| | Number of pages that the user liked | (Sahoo and Gupta 2019) |
| | Stories shared in number | (Sahoo and Gupta 2019), (Shu et al. 2018) |
| | Amount of tags | (Sahoo and Gupta 2019), (Jarrahi and Safari 2021), (Shu et al. 2018), (Castillo et al. 2011) |
| | Hashtags shared | (Sahoo and Gupta 2019), (Shu et al. 2018), (Castillo et al. 2011), (Singhal et al. 2019) |
| | Number of post/ tweets | (Shu et al. 2018), (Sahoo and Gupta 2019), (Jarrahi and Safari 2021), (Castillo et al. 2011), (Indu and Thampi 2019), (Singhal et al. 2019), (Alrubaian et al. 2021) |
| | Current location | (Sahoo and Gupta 2019), (Jarrahi and Safari 2021), (Castillo et al. 2011), (Indu and Thampi 2019), (Singhal et al. 2019), (Alrubaian et al. 2021) |
| | Friendships count | (Alrubaian et al. 2021), (Indu and Thampi 2019), (Sahoo and Gupta 2019), (Jarrahi and Safari 2021), (Castillo et al. 2011), (Singhal et al. 2019) |
| | Time of posting | (Sahoo and Gupta 2019), (Alrubaian et al. 2021), (Castillo et al. 2011), (Jarrahi and Safari 2021), (Indu and Thampi 2019) |
| | Spam-filled messages | (Sahoo and Gupta 2019) |
| | Number of followers | (Alrubaian et al. 2021), (Jarrahi and Safari 2021), (Indu and Thampi 2019), (Castillo et al. 2011), (Shu et al. 2018), (Jin et al. 2016) |
| | Max. length of a tweet | (Wani et al. 2021) |
| | Retweets | (Sahoo and Gupta 2019), (Wani et al. 2021), (Castillo et al. 2011) |
| | Tweets with mention | (Shu et al. 2018), (Sahoo and Gupta 2019), (Kwon et al. 2013), (Alrubaian et al. 2021) |
| | The overall count of likes | (Sahoo and Gupta 2019), (Shu et al. 2018), (Singhal et al. 2019) |
| 3 | **Distribution features** | (Kwon et al. 2013), (Castillo et al. 2011), (Varol et al. 2017), (Shu et al. 2018), (Jin et al. 2016), (Vosoughi et al. 2017). (Wu and Liu 2018) |
| 4 | **Temporal features** | (Castillo et al. 2011), (Kwon et al. 2013), (Vosoughi et al. 2017), (Ma et al. 2016), (Chen et al. 2019), (Varol et al. 2017), (Jin et al. 2016), (Reis et al. 2019), (Habib et al. 2019), (Huang et al. 2020) |
| 5 | **Psycho-linguistic features** | (Verma et al. 2021) |
| 6 | **Linguistic features** | (Mahyoob et al. 2020), (Varol et al. 2017), (Verma et al. 2021), (Kwon et al. 2013), (Pérez-Rosas et al. 2018), (Vosoughi et al. 2017), (Reis et al. 2019) |
| 7 | **Visual Features** | (Sharma and Garg 2021), (Yang et al. 2018), (Jin et al. 2016), (Xue et al. 2021) |
| 8 | **Structural features** | (Kwon et al. 2013), (Verma et al. 2021), (Ma et al. 2016) |

**Table 3** Existing Fake News datasets

| Datasets | No. of samples | Domain | Feature type |
|---|---|---|---|
| Burfoot Satire News (Burfoot and Baldwin 2009) | 4000 real news, 233 fake news | Economy, Politics, Technology, Society | Linguistic |
| (Ott et al. 2011) | 800 | Tourism | Linguistic |
| Fact Checking (Vlachos and Riedel 2014) | 221 Real, 221 Fake | Politics, Society | Linguistic |
| CREDBANK (Mitra and Gilbert 2015) | 60 million tweets | Society | Linguistic, User-Based, Social Context Based |
| Media Eval2015 (Boididou et al. 2016) | 7032 Fake tweets 5008 Real tweets | Tweets related to 11 events like Hurricane Sandy | Linguistic, User-Based, Social Context Based |
| Buzzfeed News (Horne & Adali 2017) | 2283 Facebook post | Political | Linguistic, User-Based, Social Context Based |
| Benjamin Political News (Horne & Adali 2017) | 75 | Political, Satire | Linguistic |
| LIAR (Wang 2017) | 12,386 | Political | Linguistic |
| FEVER (Thorne et al. 2018) | 185,445 | Society | Linguistic |
| FA-KES (Abu Salem et al. 2019) | 804 articles | Syrian war | Linguistic |
| Spanish-v2 (Posadas-Durán et al. 2019) | 491 true, 480 Fake | Society, Sports, Environment, Science, Health, Security, Politics, International | Textual in Spanish Language |
| FakeNewsNet (Shu et al. 2018) | GossipCop 85crore (Fake) 20crore (Real)<br>Politifact 12 crore(Fake) 29 crore(Real) | Political, Entertainment | Linguistic, User-Based, Social Context Based |
| TDS2020 | 46,700 articles | | Textual |
| FakeCovid (Shahi and Nandini 2020) | 5182 articles in 40 Languages | Covid | Textual |
| IFND (Sharma and Garg 2021) | 37,809 Real, 19,059 Fake | Covid19,Politics, Election, Violence | Textual, Images ` |
| TALLIP (Mohawesh et al. 2023) | | Technology, Education, Business, Politics, Entertainment, Celebrity News | Textual |
| Dravidian_Fake News Data (Raja et al.,2023) | 27,000 news articles in 4 languages(Telugu, Tamil, Malyalam, Kannada) | | Textual |

**FakeNewsNet**[9] (Shu et al. 2018) Published a dataset that contains 211 fake news and 211 true news stories culled from BuzzFeed.com and PolitiFact.com.

**FEVER**[10] Fact Extraction and VERification (FEVER) proposed by (Thorne et al. 2018) consists of the fake news articles on society having 185,445 claims extracted from Wikipedia in the English language.

**BS detector** This dataset is gathered from a browser addon called BS detector, which was created to verify the accuracy of news stories.

**MediaEval 2015** Published a dataset collected from Twitter (Boididou et al. 2016). This collection is targeted for message-level verification.

[9] The dataset is analysed from: https://github.com/KaiDMML/FakeNewsNet.

[10] The dataset is analysed from: http://fever.ai/resources.html.

**Table 4** Relative comparisons of supervised algorithms for classifying fake news

| Citation | Supervised model | Feature extraction / word embedding technique | Dataset | Results | Data type |
|---|---|---|---|---|---|
| (Ruchansky 2017) | RNN | | Twitter<br>Weibo | 89.2%<br>95.3% | Text |
| (Long et al. 2017) | LSTM | Skip gram | LIAR | 41.5% | Text |
| (Yang et al. 2018) | CNN | | Kaggle, Weibo | 92.10% | Text & Image |
| (Abedalla et al. 2019) | LSTM | Glove | | 71.2% | Text |
| (Giachanou et al. 2020) | MLP | Bert Base | ImageNet | 76.28% F1 score | Text & Image |
| (Li et al. 2020a) | MCNN | Word2Vec, TFIDF | Weibo / NewsFN | 88.82%<br>90.10% | Text |
| (Kaliyar et al. 2020) | MLP, CNN | Glove | Kaggle | 98.36% | Text |
| (Vijayaraghavan et al. 2020) | LSTM | CV | FakeNewsNet | 94.88% | Text |
| (Vereshchaka et al. 2020) | LSTM | Ngram | FakeNewsNet | 75% | Text |
| (Jehad and Yousif 2021) | MLP, CNN | TFIDF, Bert, Roberta, GPT2 & Funnel | Kaggel, LIAR, ISOT | 95.74% | Text |
| (Jehad and Yousif 2021) | MLP | Bert | ISOT | 99.94% | Text |
| (Kaliyar et al. 2021b) | CNN | BERT | Kaggle | 98.90% | Text |
| (Goldani et al. 2021) | CNN | Static / Non Static WE | LIAR<br>ISOT | 99.1%<br>99.9% | Text |
| (Saleh et al. 2021) | CNN | NGram, TFIDF, Glove | Kaggle<br>FAkeNewsNet<br>FA-KES5<br>ISOT | 99.99%<br>98.65%<br>97.23%<br>100% | Text |
| (Madhubala et al. 2021) | CNN | Word2Vec | Kaggle | 98% | Text |
| (Sahoo and Gupta 2021) | LSTM | | Facebook Raw Data generated by the crawler | 99.42% | Text |
| (Giachanou et al. 2022) | CNN | BERT, Glove | | | Text |
| (Iwendi et al. 2022) | GRU | | COVID19 | 86.12% | Text |
| (Karnyoto et al. 2022) | Bi-GRU | BERT, Glove | COVID19 | 91.9% | Text |
| (Kishwar and Zafar 2023) | LSTM, CNN | BERT, Glove | Pakistani news | 94.4%<br>93% | Text |
| (Mallik & Kumar, 2024) | LSTM | Word2Vec | ISOT | 99.97% | Text |
| (Alyoubi et al. 2023) | CNN | MARBERT | Arabic Twitter dataset | 95.6% | Text |

**Table 5** Relative comparisons of unsupervised algorithms for classifying fake news

| Citation | Unsupervised model | Feature extraction / word embedding technique | Dataset | Results | Data type |
|---|---|---|---|---|---|
| (Tzortzis and Likas 2007) | DFN | | Spam Assassim<br>Enron Spam<br>Ling Spam | 97.5%<br>97.4%<br>99.45% | Text |
| (Alexandre et al. 2016) | RBM | | SpamBase<br>Ling Spam<br>CSDMC | Not more than 66% | Text (Spam) |
| (Aghakhani et al. 2018) | GAN | | Trip Advisor hotel reviews | 89.1% | Text |
| (Li et al. 2021) | AE | | Media Eval<br>Weibo<br>Twitter | Good positive results | Text |
| (Ali et al. 2023) | ICNNAEN-DM | Glove | LIAR | 89.59% | Text |

**Table 6** Relative comparisons of hybrid algorithms for classifying fake news

| Citation | Hybrid model | Feature extraction / word embedding technique | Dataset | Results | Data type |
|---|---|---|---|---|---|
| (Drif et al. 2019) | CNN-LSTM | Glove | LIAR<br>News Article | 62.5%<br>72.5% | Text |
| (Agarwal et al. 2020) | CNN-RNN | Glove | Kaggle | 97.21% | Text |
| (Cruz et al. 2020) | ULMFiT<br>BERT + Transfer Learning | | Fakenews | 91.59% | Text |
| (Kaliyar et al. 2021a) | CNN-LSTM | Glove | PHEME<br>FN-COV | 91.88%<br>98.62% | Text |
| (Nasir et al. 2021) | CNN-RNN | Glove | FA-KES<br>ISOT | 65%<br>~100% | Text |
| (Aslam et al. 2021) | BiLSTM-GRU | | LIAR | 89.8% | Text |
| (Mohapatra et al. 2022) | BiLSTM + Self Attention | Glove, Word2Vec | Kaggle | 98.65% | Text |
| (Kausar et al., 2022) | LSTM | Ngram, TFIDF | KaggleFakeNews<br>WELFAKE | 94%<br>96.8% | Text |
| (Lee and Kim 2022) | Transfer Learning + BiLSTM | | COVID19 | 78.8% | Text |
| (Palani and Elango 2023) | BERT + BiLSTM + CNN | | Kaggle | 99.06% | Text |
| (Zhang et al. 2023) | CNN (Conv-FFD) | | CHEF, Rumor | 87% | Text |

**IFND**[11]: Collected real news from a variety of reputable sources and fake news from the sources like Boomlive, AltNews, and Digit eye. Used a parsehub scrapper tool for scraping data from the websites.

## 5 Taxonomy of fake news detection techniques based on deep learning

We outline a taxonomy for identifying fake news in this section. The objectives of fake news detection are to discover the truthfulness of the piece of information which has been shared online. We examine various techniques for DL model-based fake news categorization. The recommended solution taxonomy for the same is shown in Fig. 7 with the division of supervised, unsupervised, and hybrid learning. The following is a full description of these AI methods .

### 5.1 Machine learning detection approach

For the refinement of algorithms, different types of training datasets are used to train the model. Supervised machine learning algorithms like Support Vector Machine(SVM) (Han and Mehta 2019), Decision Tree, Multinomial Naïve Bayes, and many more algorithms have been used for fake news detection. In some experiments, only content-based features have been exploited and in some experiments, both content and context-based features (Bondielli and Marcelloni 2019) have been evaluated using different datasets. For text classification, the SVM algorithm performed better for detecting fake news. The majority of fake news machine learning algorithms have used a supervised learning paradigm. Support Vector Machines (SVMs) are one of the most extensively utilized categorization algorithms in a variety of fields. Machine. (Elsaeed et al. 2021) proposed a fake news detection system by using machine learning algorithms as well as a set of voting classifiers. The proposed vote classifier outperforms existing machine learning algorithms on two datasets: Fake-or-Real-News and MediaEval. (Choraś et al. 2021) highlighted the most significant aspects of intelligent systems in the detection of misinformation sources. Also presented current and future developments in this critical area of computer science study due to the demands of societies from all around the world. A thorough and critical examination of previous machine learning-based techniques for fake news detection is offered. (BalaAnand et al. 2019) proposed a new graph-based semi-supervised learning technique for detecting bogus users in enormous amounts of Twitter data. Game theory, support vector machine (SVM), decision tree, and k-nearest neighbour (KNN) techniques were used to evaluate the performance of the suggested algorithm with the accuracy of 90.3% for detecting fake users.

### 5.2 Deep learning detection approach

Deep learning approaches have main advantages over machine learning approaches. Deep learning has

[11] The dataset is analysed from: https://github.com/sonalgarg174/Dataset.

**Table 7** Comparison of different fake news detection models based on monolingual and multilingual

| Citation | Dataset | Languages | Domain | No. of samples | Label |
|---|---|---|---|---|---|
| (Martínez-gallego 2016) | Fake and real news dataset, Spanish Fake News Corpus971, fake news in Spanish1600 | Spanish and English | | Spanish:51,233 English:2571 | 2 |
| (Monteiro et al. 2018) | Fake.Br corpus | Portuguese | General | 7200 | 2 |
| (Vogel and Jiang 2019) | GermanFakeNC | German | Refugee crisis | 490 | 2 |
| (Guibon et al. 2019) | Storyz | English and French | Vaccination | English:5105, French: 941 | 3 |
| (Santos et al. 2020) | Fake.Br corpus | Portuguese | General | 7200 | 2 |
| (Abonizio et al. 2020) | Fake News Corpus | English, Portuguese, Spanish | News articles | 6129, 2538, 1263 | 3 |
| (Henrique et al. 2020) | TwitterBR, FakeBrCorpus, FakeNewsData, FakeOrRealNews, btvlifestyle | Germanic, Slavic, and Latin | Politics | 8981, 7200, 150, 5903, 137 samples of datasets | 2, 3 |
| (Dementieva and Panchenko 2020) | FakeNewsDataset | English, Spanish, French, Russian, and, German | Political, Entertainment | 240(fake), 240(real) | 2 |
| (Y. Li et al. 2020b) | MM-COVID | English, Spanish, Portuguese, Hindi, French and Italian | COVID | > 52 million | 2 |
| (Kar et al. 2020) | COVID 19 Twitter dataset on Hindi and Bengali, Twitter dataset on English | Hindi, Bengali and English | Covid19 | Hindi:454, Bengali: 480, English: 504 | 2 |
| (Koloski et al. 2020) | PAN 2020 Twitter | English and Spanish | User profile | Fake news users' profiles on Twitter | 2 |
| (Meesad 2021) | COVID19 | Thai | Covid | 41,448 | 3 |
| (Majumder and Das 2021) | No. of real and fake tweets in every language | Bengali, English, Hindi, German, French, Italian, and Spanish | Covid 19 | Bengali:1500, English:13,587 Hindi:902, German: 14,457 Italian:13,618, Spanish:12,599 | 2 |
| (Fouad et al. 2022) | Real dataset, Benchmark dataset, Merged dataset | Arabic | Tweets | 1979, 2582, 4561 | 2 |
| (Ahuja and Kumar 2023) | Kaggle | English, German, and French | General news | 40,000 | 2 |
| (Mohawesh et al. 2023) | TALLIP | English, Hindi, Vietnamese, Swahili, and Indonesian | Technology, Education, Business, Politics, Entertainment, Celebrity News | | 3 |

revolutionized and changed the way artificial intelligence works. In the past few years, deep learning classifiers have seen an exceptional rise in research fields, including text mining, topic classification, and NLP.

The following is a brief explanation of each of these models of deep learning approach for identifying fake news.

#### 5.2.1 Supervised/discriminative model

In supervised learning or classification applications, this group of DL approaches is used to give a discriminative function. Generally, supervised techniques pick up on different features from labelled instances. Multi-Layer Perceptrons (MLP), Convolutional Neural Networks (CNN), Artificial Neural Network (ANN), Feed Forward networks

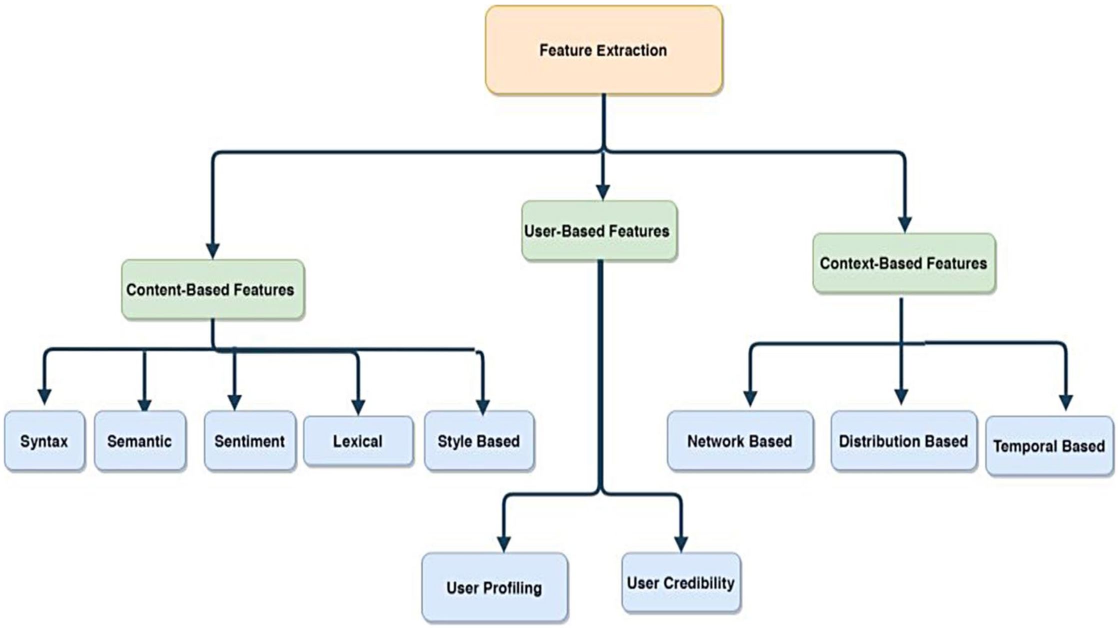



**Fig. 7** Features for fake news detection

(FNN), Recurrent Neural Networks (RNN), and their derivatives are the three primary types of discriminative architectures. We thoroughly explore these methods in the sections that follow.

**5.2.1.1** Multilayer perceptron (MLP) A supervised learning strategy is the multi-layer perceptron (MLP). A feed-forward artificial neural network that produces a set of outputs from a collection of inputs is called a multilayer perceptron (MLP). Backpropagation is used by MLP to train the network.

(Giachanou et al. 2020) proposed a technique for detecting fake news that was multimodal and made use of both text and image based features. They linked the LSTM model for picture representation to the VGG16. As a deep contextual text representation, they also employed BERTbase. Finally, the vector is sent into a multi-layer perceptron (MLP) for classification after joining all extracted feature vectors.

(Jehad and Yousif 2021) Used the Term Frequency-Inverse Document Frequency (TF-IDF) feature extraction approach and Multi-Layer Perceptron (MLP) algorithm were suggested for fake news detection. The classification accuracy of the suggested algorithm was 95.47%. offered an analysis of comparable models for contextualised text representation, such as BERT, RoBERTa, GPT2, and funnel transformer, within architectures using MLP, CNN, and single-layer perceptrons (SLP).

(Samadi et al. 2021) compared various models, including BERT, GPT2, RoBERTa, and funnel transformer, using single-layer perceptron (SLP), multi-layer perceptron (MLP), and CNN architectures.

**5.2.1.2** Convolutional neural network (CNN) A well-known discriminative deep learning architecture that learns directly from the input without the aid of human feature extraction is the convolutional neural network (CNN or ConvNet) (Sarker 2021). A Convolutional Neural Network (CNN) uses matrix multiplication to provide outputs that can be used in subsequent training stages. Convolution is the term for this technique. This is why this kind of neural network is known as a convolutional neural network (Nasir et al. 2021).

(Yang et al. 2018) Proposed TI-CNN (Text and Image information based Convolutional Neural Network) model. The text and image information is used to train TI-CNN at the same time. The convolutional neural network observes all of the input at the same time and can be trained much faster than the RNN and LSTM models. The collection contains 20,000 news pieces, with roughly 12,000 of them being fake news.

(Asghar et al. 2019) The suggested method uses a convolutional neural network and bidirectional long short-term memory to classify tweets into rumours and non-rumors.

(Qi et al. 2019) Using a multi-branch CNN-RNN model, authors were able to extract visual elements from false news photographs while simultaneously designing a CNN-based network to automatically detect their intricate patterns in the frequency domain.

(Li et al. 2020a) Created multilevel convolutional neural network MCNN-TFW and then used a technique for estimating the weight of highly sensitive terms (TFW), which has demonstrated their greater value to categorize the article as false or truthful on dataset in cultural subject.

Kaliyar et al. (2020) developed FNDNet deep learning-based convolutional neural network model. This model(FNDNet) achieved an accuracy rate of 98.36% and a false positive rate is 0.59%. The limitation of this model is that it worked for only two label datasets that is for real and fake values only.

(Kaliyar et al. 2021b) Used Bidirectional Encoder Representation Transformer(BERT) as a word embedding model for the designing of a deep convolutional neural network. The dataset trained for this model is the fake news dataset from Kaggle having 1000 fake and real news articles. The findings of the classification showed that FakeBERT produced more accurate findings with an accuracy of 98.90%.

(Alsaeedi and Al-Sarem 2020) Conventional neural network was presented to detect rumors propagating on Twitter (CNN). The authors examined the results of several comparable rumor detection methods and machine learning techniques that used the PHEME dataset to highlight the performance of the model in terms of accuracy, f-measure, recall, and precision. This model had an accuracy of 87%.

(Goldani et al. 2021) For the purpose of identifying bogus news, convolutional neural networks (CNN) with margin loss and other embedding models have been proposed. For the LIAR dataset, the authors employed static word embedding, while for the ISOT, they used non-static word embedding.

(Saleh et al. 2021) To identify fake news, authors suggested an improved convolutional neural network model (OPCNN-FAKE). For ML/DL models, the feature matrices N-gram, Term Frequency—Inverse Document Frequency (TF-IDF), and Glove word embedding had been utilised to represent features. Using four benchmark datasets for fake news, authors compared the OPCNN-FAKE with LSTM, RNN, and the six conventional ML techniques: DT, RF, LR, KNN, NB, and SVM.

(Madhubala et al. 2021) Used deep Convolutional Neural Networks (CNN) and NLP techniques for data preprocessing. The algorithm was compared to a baseline CNN model and a Naive Bayes machine learning model. Got the accuracy of 98%.

(Varlamis et al. 2022) Proposed a survey of works using Graph Convolutional Networks (GCNs) for the job of identifying false news, false accounts, and rumours that circulate on social networks is conducted in the current work, which focuses on the popular and promising graph representation techniques.

(Giachanou et al. 2022) Looked into the effect of several psycholinguistic variables in distinguishing between people who spread fake news. The authors proposed the CheckerOrSpreader model, which employed a Convolution Neural Network (CNN) to distinguish between fake news spreaders and checkers.

**5.2.1.3** Recurrent neural network (RNN) and its subtypes An additional common neural network is a recurrent neural network (RNN), which uses time-series or sequential data and inputs the results of the previous stage into the current stage (Sarker 2021). Recurrent networks, like CNN and feedforward, also learn from training data, but they differ in that they include a "memory" that allows them to use data from earlier inputs to influence their current input and output. It is difficult to learn long data sequences with ordinary recurrent networks because of the problem of vanishing gradients.

(Ruchansky 2017) suggested a creative solution to the problem of fake news detection. In order to determine the temporal pattern of user engagements on a particular article, RNN was used in the first module. Analyse the user's actions in the second portion by comprehending the source characteristic. The outcome of both modules' results was then combined to determine whether the item was authentic or fake.

The recurrent network has a number of well-liked variations that reduce the drawbacks and excel in a variety of real-world application domains. These are discussed below.

## Gated recurrent unit (GRU)

A Gated Recurrent Unit (GRU), a common variation of the Recurrent Network that uses gating techniques to control and manage information flow between neural network cells (Cho et al. 2014). Update gate and reset gate are the two gates of GRU. The structure of the GRU makes it possible to capture dependencies from lengthy data sequences in an adaptive fashion without losing data from previous portions of the sequence.

(Iwendi et al. 2022) Authors have suggested named entity-based features, sentiment features, and linguistic features. Modern deep learning models like GRU, RNN, and LSTM were used by the model to detect fake news about COVID-19. The accuracy achieved by GRU was 86.12% among all other algorithms.

(Karnyoto et al. 2022) In order to tackle the fake news identification problem, authors applied GPT2 and BERT as pre-trained models utilising the BiGRU-Attention-Capsule Network model. The BiGRU-CRF layer processes the output of the capsule network features by segmenting its characteristics. It proved that BERT produced a better outcome.

### Long short term memory (LSTM)/ Bidirectional-LSTM

To address the shortcomings of RNN, Hochreiter and Schmidhuber created the LSTM cell. Long-term data storage is possible in memory cells found in LSTM units. Three gates control the flow of data into and out of memory cells. The Forget gate is the first gate, the Input gate is the second, and the Output gate is the final gate. The Forget Gate determines whether or not we should retain the information from the previous time step. Input gates are used to rate the significance of newly transmitted information. “Output Gate” chooses and manages the outputs. In order to improve model performance on sequence classification problems, the regular LSTM is extended with the BiLSTM algorithm. The input travels in both ways in Bi LSTM in order to keep both the past and the future data.

(Long et al. 2017) In order to detect fake news, an attention-based LSTM model with speaker profiles was developed in this work as a novel approach. When speaker profiles are incorporated into an attention-based LSTM detection model, accuracy can rise by a net 14.5%, or over 41.5%, over the most recent model.

(Abedalla et al. 2019)The authors studied the issues of fake news by examining the prior research on identifying fake news utilizing machine learning and deep learning domains. The authors assembled models mainly on long short-term memory, bidirectional long short term memory, and convolutional neural network on FNC-1 dataset. LSTM networks that used the relationship between the article’s headline and body to identify fake news achieved a high accuracy of 71.2%.

(Vereshchaka et al. 2020) Analysed the features of false news, as well as the sociocultural and textual features of fake news, to address the problem of fake news identification. For the classification of bogus and true news, deep learning models such as LSTM, RNN, and GRU were used. The maximum level of accuracy is reached using LSTM.

(Vijayaraghavan et al. 2020) The effectiveness of various Natural Language Processing models including count vectorizer, TFIDF, and Word2Vec in solving the binary classification fake news problem is investigated by the authors. Then they trained the model on ANN, LSTM, random forest, logistic regression, and SVM. Out of those, the best result is obtained when CV and LSTM are combined with an accuracy of 94.88%.

(Sahoo and Gupta 2021) employed a variety of Facebook account-related features, some news content features, and then the features of account were analysed by machine learning algorithms like KNN, random forest, logistic regression, naive bayes, and deep learning algorithm like LSTM. Better accuracy was provided by LSTM at 99.42%.

#### 5.2.2 Unsupervised / generative model

Unsupervised learning uses data set without labels, which means that the classes are not labelled. New data points can be produced using generative models. Typically, this class of DL approaches is employed to characterise the high-order association features or characteristics for the analysis of patterns (Deng 2014). The essential principle of generative neural networks is that exact supervisory data, such as intended class labels, is unimportant throughout the learning process. Generative Adversarial Network (GAN), Restricted Boltzmann Machine (RBM), Autoencoder (AE), and Deep Belief Network (DBN) are popular deep learning algorithms for generative or unsupervised learning.

**5.2.2.1** Generative adversarial network Generative modelling is an unsupervised modelling problem that involves automatically recognising and understanding the patterns or regularities in input data such that the model may be used to produce new examples that might have been reasonably derived from the original set of data. The two neural networks that make up a GAN are the generator, which generates new data with attributes that are comparable to the original data, and the discriminator, which forecasts the likelihood that a following sample will be taken from genuine data rather than data produced by the generator. As a result, two models are trained in GAN modelling to compete with one another. The discriminator seeks to separate the real data from the fake data produced by the generator while the generator tries to deceive and confound the discriminator by producing more plausible data.

(Aghakhani et al. 2018) Proposed FakeGAN for detecting deceptive reviews using GAN. 800 TripAdvisor reviews from 20 hotels in Chicago were analysed, and the results showed that FakeGAN performed on par with the best methods with an accuracy of 89.1%.

**5.2.2.2** Restricted boltzmann machine The Boltzmann Machine is a generative unsupervised learning that uses a dataset of original data to teach a probability distribution that is then used to new data to draw conclusions. The input layer of a Boltzmann machine, also known as the visible layer, and one or more hidden layers, also known as the hidden layer. The restricted Boltzmann machine approach, which is utilised for feature extraction and feature selection, is crucial for dimensionality reduction, classification, and regression.

RBMs were used by the authors to automatically extract spam detection-related features. RBMs obtained good results in the datasets used in this work, hence the findings they demonstrated were suitable for learning features from email messages.

**5.2.2.3** Deep belief network Deep belief networks are probabilistic unsupervised models made up of several layers of random, hidden variables. The connections between the top two levels are undirected and symmetric. Directed connections from the upper layer are sent down to the lower layers.

(Tzortzis and Likas 2007) Authors defined spam as an unexpected communication that contains unsuitable content and explained how DBNs were initially used for spam identification.

Deep belief network for feature selection was used in the suggested methodology (Selvaganapathy et al. 2018) for the identification and classification of malicious URLs.

**5.2.2.4** Auto-encoder A common unsupervised learning method that uses neural networks to learn representations is called an auto-encoder (AE). When dealing with high-dimensional data, auto-encoders are frequently utilised, and dimensionality reduction describes the representation of a set of data.

(Li et al. 2021) Authors suggested autoencoder-based unsupervised false news detection technique (UFNDA). Author gathered and combined information from Twitter's content-based features, user, propagation, and image characteristics in social networks, and analysed it using the suggested approach UFNDA.

(Ali et al. 2023) In this paper, the ICNNAEN-DM Web-Informed-Augmented Fake News Detection Model was suggested. It used deep autoencoding and stacked layers of convolutional neural networks. For accurate identification and prevention of false news, the study emphasised the significance of effective news representational properties and appropriate model design.

### 5.2.3 Hybrid model

In most cases, hybrid deep learning approaches are made up of two or more basic deep learning models, where the basic model is one of the earlier stated discriminative or generative deep learning models.

**5.2.3.1** Ensemble learning Multiple separate models are combined using ensemble learning to improve generalisation performance. The benefits of both deep learning models and ensemble learning are combined in deep ensemble models for learning, which improves the generalisation performance of the final model. Despite the variety of deep learning architectures and their capacity to handle difficult issues and the capacity to automatically extract information, the main issue in deep learning is that for tuning the ideal hyper-parameters requires a lot of knowledge and experience, which makes it a laborious and time-consuming operation.

(Drif et al. 2019) A hybrid Convolutional Neural Network model (CNN) and a Long Short Term Memory (LSTM) recurrent neural network architecture is suggested, utilising local features produced by CNN and the long-distance relationships learned by LSTM. The accuracy of the CNN-LSTM was 62.34% on the Liar dataset, outperforming all other models. The achieved accuracy for the News Articles dataset was 72.50%.

(Bahad et al. 2019) presented a fake news detection method based on a Bidirectional LSTM recurrent neural network. The Bidirectional LSTM-RNN model is significantly more powerful than unidirectional models, according to the findings.

(Agarwal et al. 2020) Presented the GloVe model for text pre-processing to construct a vector space of words and establish a linguistic relationship that was described and tested. With the addition of word embeddings, the proposed model, which combines the architectures of convolutional neural networks and recurrent neural networks, has achieved benchmark results in the prediction of fake news, with precision values of 97.21%.

(Kaliyar et al. 2021a) Presented a hybrid model that combines numerous convolutional neural network (CNN) branches with layers of Long Short Term Memory (LSTM). The authors collected a dataset (FN-COV) of 69,976 fake and real news articles during the COVID-19 epidemic and achieved an accuracy of 98.62%.

(Nasir et al. 2021) Proposed a hybrid deep learning model which is the combination of convolutional and recurrent neural networks for the classification of fake news. FA-KES dataset consists of 804 articles and achieved an accuracy of 60% and the ISOT dataset consists of 45,000 articles and gains an accuracy of 99.96%. It has been observed that large-scale datasets achieved high accuracy as compared to small-scale datasets.

(Aslam et al. 2021) Developed a deep learning ensemble model to categorise news as fake or true using the LIAR dataset. Deep learning model Bi-LSTM-GRU-dense was utilised for textual attribute statement. According to experimental findings, the proposed study had an accuracy of 0.898.

(Mohapatra et al. 2022) The author suggested a deep learning method based on Bi-LSTM (Bidirectional long short term memory) and adding self-attention to it. With an accuracy score of 98.65%, the classification result showed that the proposed hybrid deep learning model beats existing methods.

(Kausar et al., 2022) The content-based features for the proposed hybrid model were first extracted using the N-gram with TF-IDF, and sequential features were then extracted using a deep learning model LSTM or BERT. For the WELFAKE and KaggleFakeNews datasets, the suggested technique had provided accuracy values of 96.8% and 94%, respectively.

The proposed model used an automated method for data augmentation known as auxiliary Generative Adversarial Networks to create additional samples of fake news. Convolutional and recurrent neural networks were then combined to efficiently detect fake news. When used with the Buzzfeed, FakeNewsNet, and FakeNewsChallenges datasets to detect fake news, the suggested model outperformed state-of-the-art algorithms with 93.87% accuracy.

**5.2.3.2** Transfer learning Transfer learning is a two-stage training method for DL models that entails pre-training and fine-tuning using training data from the target task. Natural language processing, visual and speech recognition, sentiment classification, spam filtering, and other fields can all benefit from DTL approaches.

(Cruz et al. 2020) The first dataset on fake news in Filipino language was created by authors and termed as fake news filipino. ULMFiT, BERT, and GPT-2 have all been benchmarked TL approaches that demonstrate a significant performance advantage over few-shot techniques. By utilising transfer learning approaches, authors were able to obtain an accuracy of 96%.

(Lee and Kim 2022) The proposed model first undergoes transfer learning using a BiLSTM model after being pre-trained with a sizable dataset on COVID19. The proposed model was demonstrated through testing to have an accuracy of 78.8%, an improvement of 8% over the linear model used as a baseline model.

**5.2.3.3** Reinforcement learning Reinforcement learning addresses the sequential making decisions problem differently from the other methodologies. Deep reinforcement learning is a kind of machine learning that allows an agent to behave in an interactive environment executing actions and observing the outcomes. The agent receives positive feedback for each excellent action, and negative feedback or a penalty for each bad action.

(Wang et al. 2020) Developed a method for detecting fake news that is reinforced and weakly supervised. The suggested architecture was made up of three major components: an annotator, a reinforced selector, and a fake news detector. Based on user reports, the annotator can automatically give weak labels to unlabeled news.

(Mosallanezhad et al. 2022) The author addressed the shortcomings of existing automated false news identification models by combining additional information (e.g., comments from users and Interactions between users and news) into an innovative reinforcement learning model known as reinforced adaptive learning false news identification.

**5.2.3.4** Crowdsource-deep learning approach In order to achieve certain organizational goals, crowdsourcing is described as "a production paradigm, an online and dispersed problem-solving system that utilizes the collective intellect of online communities." The majority of machine learning systems designed to automatically identify fake stories have frequently failed. This is due to the fact that not all news has the same style of writing and incorporates a variety of themes with important details. Low accuracy was found to be an issue with automatic fake news detection (Shabani and Sokhn 2018). This is because the majority of linguistic devices used to create fake news avoid the detection process. To identify fake news and satire, they suggest a technique that takes a hybrid crowdsource-machine technique. They achieved a total high accuracy of 87% by combining ML approaches. In (C. Wang and Zhu 2019) study, the author used a hybrid machine-crowdsource technique to identify social media fake news. The author then designed a hybrid CNN to combine meta-data with text and achieved better results.

## 6 Monolingual and multilingual fake news detection models

Based upon the above analysis and review, mostly the research conducted to date has centred on identifying fake news in a specific language in English only. But in this section, fake news detection models for different languages are reviewed.

(Martínez-gallego 2016) developed a fake news detection model in Spanish and English language using Deep Learning and Machine Learning architectures. For the Spanish dataset, BETO embedding enabled them to achieve the best outcome.

(Posadas-Durán et al. 2019) Proposed an approach for detecting fake news in Spanish that is focused on style. The authors tested four machine learning classifiers: boosting, logistic regression, random forest, and support vector machine to create a model that can discriminate between fake and real information.

(Amjad et al. 2020) Introduced a novel corpus for fake news detection in the Urdu language. In order to identify fake news on the corpus, authors considered a variety of machine learning classifiers.

(Hossain et al. 2020) For the development of automated fake news detection systems for the Bangla language, a 50 K news dataset with annotations has been proposed. The development of this system involved both conventional linguistic aspects and neural network techniques.

(Abonizio et al. 2020) To investigate text complexity, stylometric, and psychological aspects, collections of news writing in American English, Brazilian Portuguese, and Spanish were examined. The features that were extracted help in the identification of real, fake, and satirical news. Support Vector Machine (SVM), Extreme Gradient Boosting (XGB) k-Nearest Neighbors (KNN), and Random Forest (RF) were four machine learning methods that the authors compared in order to create the detection model, with which the RF gave the accuracy of 85.3%.

(Dementieva and Panchenko 2020) Examined a novel method for identifying fake news based on data from multiple languages English, Spanish, French, Russian, and, German. The sentence embedding from the Multilingual BERT was obtained to create the vector representation.

(Du et al. 2021) Crossfake was a framework that jointly encodes the multilingual news body texts and fully captures the news content. A Chinese true & false news dataset is first curated by the authors using information from fact-checking sources. The authors build a neural classifier using the combined BERT embeddings of an English news sub-text that has been fact-checked. A Chinese news piece is first translated into English for verification, and the final forecasts are made by adding up all of the subtext predictions.

(Majumder and Das 2021) Used two different models, one of which is language-dependent and the other of which tries to look into numerous language-independent concerns using BERT. For languages like Bengali, English, and Hindi, the authors' results in a language-independent model were better. When it comes to European languages like German, French, Italian, and Spanish, both language-dependent and independent models work admirably, and their outcomes are comparable.

(Azad et al. 2021) Built a fake news detection model in the Kurdish language. TF-IDF was used as a feature of selection before numerous classifiers were applied to the corpus. Among the other classifiers, Support Vector Machine (SVM) had the highest accuracy rating of 88.71%.

(Chu et al. 2021) Investigated how different textual characteristics in Chinese and English affect the ability to spot fake news. Additionally, the outcomes demonstrated that the bidirectional encoder representations from the transformers (BERT) model outperformed other techniques.

(Buzea et al. 2022) The suggested method is based on a corpus of Romanian news that includes 25,841 real news and 13,064 fake news. In this work, the findings from recurrent neural networks with the gated recurrent unit and long short-term memory are compared to those from convolutional neural networks. The output of Nave Bayes and Support Vector Machine, two traditional classification methods, is compared to that of deep learning systems.

(Ahuja and Kumar 2023) Mul-FaD is a system for multilingual fake news detection. A dataset of about 40,000 articles in English, German, and French has been developed by the authors. Fast text embeddings for many languages were employed and the model is capable of 93.73% accuracy.

## 7 Comparison to previous surveys

A quick summary of the existing survey publications and our research contributions is given in Table 8. By performing a thorough survey on fake news identification, the current study seeks to address the shortcomings and strengths of earlier research. From the previous surveys, it has been found that not so much work has been done on multilingual fake news detection models. The majority of previous research has been done on detecting fake news in a specific language. However, most of the content is spread not only in the native English language, but there are users of different languages also. It is vital to identify fake news in documents written in many languages. However, there are typically some semantic, syntactic, and lexical differences between various languages, and developing a system with multi-lingual comprehension is one of the important obstacles to tackling multi-lingual fake news detection.

## 8 Technical challenges

Fake news can only be detected after it has been made and distributed across the Internet, according to online fact-checking services. Fact-checking websites can alert internet users to similar claims or issues, but they can't completely prevent misinformation from spreading through social media.

*Identification of the Spreader* Identifying prominent spreaders in social networks is a crucial topic that allows researchers to better understand the role of nodes in information dissemination and epidemic propagation.

*Embedding of unsupervised news* Semantic similarity analysis, sentiment analysis, and other related tasks are crucial components for detecting online fake news due to

**Table 8** A comparison table of existing surveys utilising deep learning to identify fake news

| Citations | NLP methodologies | Datasets | Technical challenges | Multilingual models | DL algorithms | | |
|---|---|---|---|---|---|---|---|
| | | | | | Supervised | Unsupervised | Hybrid |
| (Shu et al. 2017) | √ | | | | √ | | |
| (Parikh and Atrey 2018) | √ | √ | √ | | √ | | |
| (Oshikawa et al. 2018) | √ | √ | | | √ | | |
| (Habib et al. 2019) | | | | | √ | | |
| (Bondielli and Marcelloni 2019) | √ | √ | | | √ | | |
| (X. Zhang and Ghorbani 2019) | √ | √ | √ | | √ | √ | |
| (Meel and Vishwakarma 2020) | √ | √ | | | √ | | |
| (Zhou and Zafarani 2020) | √ | | | | √ | | |
| (Kumar et al. 2021) | | √ | | | √ | √ | |
| (Mridha et al. 2021) | √ | √ | √ | | √ | √ | √ |
| (Ali et al. 2022) | √ | √ | √ | | √ | | |
| (Rohera et al. 2022) | | | √ | | √ | √ | √ |
| (Hu et al. 2022) | | √ | | | √ | √ | √ |
| (Rastogi and Bansal 2023) | | √ | | | √ | √ | √ |
| Ours | √ | √ | √ | √ | √ | √ | √ |

the textual character of the content. The process of extracting distributed representations of raw textual data is known as embedding, and it is an important stage in natural language processing. Choosing a proper embedding approach is crucial for determining the underlying nature of the news. Unsupervised embedding models that are widely used like word2Vec, doc2Vec (Mikolov 2013) and Sent2vec (Pagliardini et al. 2018). BERT (Devlin et al. 2019) includes brand-new state-of-the-art results on eleven natural language processing tasks.

*Identification of fake information* The majority of research studies tend to focus on notifying people but provide no explanation as to why this is misinformation; instead, focus on directly engaged users for misinformation detection. However, even if the users are not affiliated, certain individuals can be quite adept at disseminating misinformation on social media sites. It's tough to tell them apart because they aren't related in any way.

## 9 Conclusion

False information and fake news can take many forms. They can also have significant consequences, as knowledge shapes our worldview. Because of the widespread use of social media platforms for the broadcast of information and news, there is a growing study trend for detecting fake news. Identifying the origin of fake news and detecting fake news has been the subject of extensive research. Approximately half of the studies looked toward the detection of fake content in textual data using deep learning. The implemented techniques of this survey were classified based on fake news, features, deep learning algorithms, hybrid deep learning algorithms, monolingual and multilingual fake news detection models. As of now, researchers have done more research only on English languages. This survey shows that very little work has been done on the Hindi language. Moreover, Hindi is also our regional language. So I think that work should be done on Hindi language also.

Using traditional methods to develop a fake news detection system is no longer sufficient. We explored fake news detection recommended methodologies in this study, including the use of machine learning for fake news classification training. It has begun to outperform classic machine learning algorithms in the detection of fake news. As evident from the results obtained, deep learning is efficient enough to detect textual and non-textual data. Deep Learning is an efficient and effective methodology for detecting fake news on online social networks. It would be motivating to code domain information in deep neural networks for this research. Examine if data features merged with neural network models for our upcoming fake news detection model using a deep learning approach. When compared to multi-layered CNN and LSTM, deep learning bidirectional transformers produce better outcomes.

Fake news detection is mainly treated as a binary classification issue in some datasets. My future plan according to the above survey is that it should be treated as a multi-class classification problem and multiple languages in fact-checking platforms as fake news is a contentious problem in many domains, and the deep learning literature has traditionally depended on studies that focused on a single platform or single language. We also go through the specifics of several datasets and their availability for future research.

**Author contributions** Both the contributors contributed equally in the manuscript.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors have declared that they have no competing interests.

## References

Abedalla A, Al-Sadi A, Abdullah M (2019) A closer look at fake news detection: a deep learning perspective. ACM Int Conf Proceeding Ser 24–28. https://doi.org/10.1145/3369114.3369149

Abonizio HQ, de Morais JI, Tavares GM, Junior SB (2020) Language-independent fake news detection: English, Portuguese, and Spanish mutual features. Future Internet 12(5):1–18. https://doi.org/10.3390/FI12050087

Abu Salem FK, Al Feel R, Elbassuoni S, Jaber M, Farah M (2019) FA-KES: a fake news dataset around the Syrian war. Proc 13th Int Conf Web Social Media ICWSM 2019 Icwsm 573–582. https://doi.org/10.1609/icwsm.v13i01.3254

Agarwal A, Mittal M, Pathak A, Goyal LM (2020) Fake news detection using a blend of neural networks: an application of Deep Learning. SN Comput Sci 1(3):1–9. https://doi.org/10.1007/s42979-020-00165-4

Aghakhani H, MacHiry A, Nilizadeh S, Kruegel C, Vigna G (2018) Detecting deceptive reviews using generative adversarial networks. Proc – 2018 IEEE Symp Secur Priv Workshops SPW 2018 89–95. https://doi.org/10.1109/SPW.2018.00022

Ahmed H, Traore I, Saad S (2017) Detection of Online Fake News Using N-Gram Analysis and machine learning techniques. Lect Notes Comput Sci 10618 LNCS:127–138. https://doi.org/10.1007/978-3-319-69155-8_9

Ahuja N, Kumar S (2023) Mul – FaD: attention based detection of multiLingual fake news. J Ambient Intell Humaniz Comput 0123456789(201414):2481–2491. https://doi.org/10.1007/s12652-022-04499-0

Alexandre L, Augusto K, Ribeiro PB (2016) Learning Spam features using restricted Boltzmann machines. IADIS Int J Comput Sci Inform Syst 11(1):99–114

Ali I, Nizam M, Ayub B, Shivakumara P (2022) Review article fake news detection techniques on Social Media: a Survey. 2022:1–17

Ali AM, Ghaleb FA, Mohammed MS, Alsolami FJ, Khan AI (2023) Web-informed-augmented fake news detection model using stacked layers of Convolutional Neural Network and Deep Autoencoder. Mathematics 11(9):1–21. https://doi.org/10.3390/math11091992

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31(2):211–236. https://doi.org/10.1257/jep.31.2.211

Alrubaian M, Al-qurishi M, Omar S, Mohamed A (2021) DeepTrust: a Deep Learning Approach for measuring social media users trustworthiness. ArXiv Comput Sci 2018(January):1–18

Alsaeedi A, Al-Sarem M (2020) Detecting rumors on Social Media based on a CNN Deep learning technique. Arab J Sci Eng 45(12):10813–10844 Springer Berlin Heidelberg. https://doi.org/10.1007/s13369-020-04839-2

Alyoubi S, Kalkatawi M, Abukhodair F (2023) The detection of fake news in arabic tweets using deep learning. Appl Sci (Switzerland) 13(14):1–21. https://doi.org/10.3390/app13148209

Alzanin SM, Azmi AM (2018) Detecting rumors in social media: a survey. Procedia Comput Sci 142:294–300. https://doi.org/10.1016/j.procs.2018.10.495

Amjad M, Sidorov G, Zhila A, Gómez-Adorno H, Voronkov I, Gelbukh A (2020) Bend the truth: Benchmark dataset for fake news detection in Urdu language and its evaluation. J Intell Fuzzy Syst 39(2):2457–2469. https://doi.org/10.3233/JIFS-179905

Asghar MZ, Habib A, Habib A, Khan A, Ali R, Khattak A (2019) Exploring deep neural networks for rumor detection. J Ambient Intell Humaniz Comput 0123456789:1–19. https://doi.org/10.1007/s12652-019-01527-4

Aslam N, Ullah Khan I, Alotaibi FS, Aldaej LA, Aldubaikil AK (2021) Fake detect: a deep learning ensemble model for fake news detection. Complexity 2021(1–8). https://doi.org/10.1155/2021/5557784

Azad R, Mohammed B, Mahmud R, Zrar L, Sdiq S (2021) Fake news detection in low-resourced languages kurdish language using machine learning algorithms. Turkish J Comput Math Educ 4219 Res Artic 4219 Res Article 12(6):4219–4225. https://www.rudaw.net/

Bahad P, Saxena P, Kamal R (2019) Fake news detection using bi-directional LSTM-Recurrent neural network. Procedia Comput Sci 165(2019):74–82. https://doi.org/10.1016/j.procs.2020.01.072

BalaAnand M, Karthikeyan N, Karthik S, Varatharajan R, Manogaran G, Sivaparthipan CB (2019) An enhanced graph-based semi-supervised learning algorithm to detect fake users on Twitter. J Supercomput 75(9):6085–6105. https://doi.org/10.1007/s11227-019-02948-w

Bhutani B, Rastogi N, Sehgal P, Purwar A (2019) Fake News Detection Using Sentiment Analysis. 2019 Twelfth International Conference on Contemporary Computing (IC3), pp 1–5

Boididou C, Papadopoulos S, Dang-Nguyen DT, Boato G, Riegler M, Middleton SE, Petlund A, Kompatsiaris Y (2016) Verifying Multimedia Use at MediaEval 2016. CEUR Workshop Proc 1739:1–3

Bondielli A, Marcelloni F (2019) A survey on fake news and rumour detection techniques. Inf Sci 497:38–55. https://doi.org/10.1016/j.ins.2019.05.035

Bovet A, Makse HA (2019) Influence of fake news in Twitter during the 2016 US presidential election. Nat Commun 10(1):1–14. https://doi.org/10.1038/s41467-018-07761-2

Braşoveanu AMP, Andonie R (2019) Semantic fake news detection: a machine learning perspective. Lecture Notes Comput Sci (Including Subser Lecture Notes Artif Intell Lecture Notes Bioinformatics) 11506 LNCS:656–667. https://doi.org/10.1007/978-3-030-20521-8_54

Burfoot C, Baldwin T (2009) Automatic Satire Detection: Are You Having a Laugh ? University of Melbourne University of Melbourne. Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, August, Singapore, pp 161–164

Buzea MC, Trausan-Matu S, Rebedea T (2022) Automatic fake news detection for Romanian Online News. Information 13(3):151, 1–13. https://doi.org/10.3390/info13030151

Camacho D, Panizo-LLedot Á, Bello-Orgaz G, Gonzalez-Pardo A, Cambria E (2020) The four dimensions of social network analysis: an overview of research methods, applications, and software tools. Inform Fusion 63:88–120. https://doi.org/10.1016/j.inffus.2020.05.009

Caropreso MF, Matwin S (2006) Beyond the bag of words: a text representation for sentence selection. Lecture Notes Comput Sci (Including Subser Lecture Notes Artif Intell Lecture Notes Bioinformatics) 4013 LNAI:324–335. https://doi.org/10.1007/11766247_28

Castillo C, Mendoza M, Poblete B (2011) Information credibility on Twitter. Proc 20th Int Conf Companion World Wide Web WWW 2011 India 675–684. https://doi.org/10.1145/1963405.1963500

Cha M, Mislove A, Gummadi KP (2009) A measurement-driven analysis of information propagation in the Flickr social network. WWW'09 - Proceedings of the 18th International

World Wide Web Conference, ACM, pp 721–730. https://doi.org/10.1145/1526709.1526806
Chakraborty A, Paranjape B, Kakarla S, Ganguly N (2016) Stop Clickbait: detecting and preventing clickbaits in online news media. Proc 2016 IEEE/ACM Int Conf Adv Social Networks Anal Min ASONAM 2016 USA 9–16. https://doi.org/10.1109/ASONAM.2016.7752207
Chen C, Lin K, Shae Z A Trusting News Ecosystem against Fake News from Humanity and Technology Perspectives. 9th International Conference on Computational, Science, Applications I (2019) (ICCSA), St. Petersburg, Russia, pp 132–137. https://doi.org/10.1109/ICCSA.2019.00011
Chi Y, Zhu S, Hino K, Gong Y, Zhang Y (2009) iOLAP: a framework for analyzing the internet, social networks, and other networked data. IEEE Trans Multimedia 11(3):372–382. https://doi.org/10.1109/TMM.2009.2012912
Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. EMNLP 2014–2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, Doha, Qatar, pp 1724–1734. https://doi.org/10.3115/v1/d14-1179
Choraś M, Demestichas K, Giełczyk A, Herrero Á, Ksieniewicz P, Remoundou K, Urda D, Woźniak M (2021) Advanced Machine Learning techniques for fake news (online disinformation) detection: a systematic mapping study. Appl Soft Comput 101:1–22. https://doi.org/10.1016/j.asoc.2020.107050
Chu SKW, Xie R, Wang Y (2021) Data Inf Manag 5(1):100–109. https://doi.org/10.2478/dim-2020-0025. Cross-Language Fake News Detection
Cruz JCB, Tan JA, Cheng C (2020) Localization of fake news detection via multitask transfer learning. LREC 2020–12th International Conference on Language Resources and Evaluation, Conference Proceedings, France, pp 2596–2604. https://doi.org/10.13140/RG.2.2.23028.40322
Dementieva D, Panchenko A (2020) Fake news detection using multilingual evidence. Proc – 2020 IEEE 7th Int Conf Data Sci Adv Analytics DSAA 2020 Australia pp775–776. https://doi.org/10.1109/DSAA49011.2020.00111
Deng L (2014) A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Trans Signal Inform Process 3(1):1–29. https://doi.org/10.1017/atsip.2013.9
Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. NAACL HLT 2019–2019 Conf North Am Chapter Association Comput Linguistics: Hum Lang Technol - Proc Conf 1(Mlm):4171–4186
Drif A, Hamida ZF, Giordano S (2019) Fake news detection Method Based on text-features fake news detection Method Based on text-features. IMMM 2019: Ninth Int Conf Adv Inform Min Manage Fake Switzerland:26–31
Du J, Dou Y, Xia C, Cui L, Ma J, Yu PS (2021) Cross-lingual COVID-19 Fake News Detection. arXiv:2110.06495v2, pp 5–8
Elsaeed E, Ouda O, Elmogy MM, Atwan A, El-Daydamony E (2021) Detecting fake news in Social Media using Voting Classifier. IEEE Access 9:161909–161925. https://doi.org/10.1109/ACCESS.2021.3132022
Fouad KM, Sabbeh SF, Medhat W (2022) Arabic fake news detection using deep learning. Comput Mater Contin 71(2):3647–3665. https://doi.org/10.32604/cmc.2022.021449
Giachanou A, Zhang G, Rosso P (2020) Multimodal multi-image fake news detection. Proceedings – 2020 IEEE 7th International Conference on Data Science and Advanced Analytics, DSAA 2020, 647–654. https://doi.org/10.1109/DSAA49011.2020.00091
Giachanou A, Ghanem B, Ríssola EA, Rosso P, Crestani F, Oberski D (2022) The impact of psycholinguistic patterns in discriminating between fake news spreaders and fact checkers. Data Knowl Eng 138(November 2021):1019601–1019615. https://doi.org/10.1016/j.datak.2021.101960
Goldani MH, Safabakhsh R, Momtazi S (2021) Convolutional neural network with margin loss for fake news detection. Inf Process Manag 58(1):1–12. https://doi.org/10.1016/j.ipm.2020.102418
Grinberg N, Joseph K, Friedland L, Swire-thompson B, Lazer D (2019) Fake news on Twitter during the 2016 U. S Presidential Election Sci 363(6425):374–378
Guibon G, Ermakova L, Seffih H, Firsov A, Le G, Guibon G, Ermakova L, Seffih H, Firsov A, Multilingual GLN, Guibon G, Ermakova L, Seffih H, Firsov A (2019) Multilingual Fake News Detection with Satire. Conference on Intelligent Text Processing and Computational Linguistics, pp 1–12
Habib A, Asghar MZ, Khan A, Habib A, Khan A (2019) False information detection in online content and its role in decision making: a systematic literature review. Soc Netw Anal Min 9(1):1–20. https://doi.org/10.1007/s13278-019-0595-5
Hakak S, Alazab M, Khan S, Gadekallu TR, Maddikunta PKR, Khan WZ (2021) An ensemble machine learning approach through effective feature extraction to classify fake news. Futur Gener Comput Syst 117:47–58. https://doi.org/10.1016/j.future.2020.11.022
Halpern D, Valenzuela S, Katz JE (2017) We face, I tweet: how different social media influence political participation through collective and internal efficacy. J Comput Commun 22(6):320–336. https://doi.org/10.1111/jcc4.12198
Han W, Mehta V (2019) Fake news detection in social networks using machine learning and deep learning: performance evaluation. Proc - IEEE Int Conf Industrial Internet Cloud ICII 2019 Icii USA 375–380. https://doi.org/10.1109/ICII.2019.00070
Henrique P, Faustini A, Covões TF (2020) Expert systems with Applications fake news detection in multiple platforms and languages. Expert Syst Appl 158:1–9
Horne B, Adali S, More Similar To Satire Than Real News (2017) This Just In: Fake News Packs A Lot In Title, Uses Simpler, Repetitive Content in Text Body,. Proceedings of the International AAAI Conference on Web and Social Media, 11(1):759–766. https://doi.org/10.1609/icwsm.v11i1.14976
Hossain MZ, Rahman MA, Islam MS, Kar S (2020) BanFakeNews: A dataset for detecting fake news in Bangla. LREC 2020–12th International Conference on Language Resources and Evaluation, Conference Proceedings, France, pp 2862–2871
Hu L, Wei S, Zhao Z, Wu B (2022) Deep learning for fake news detection: A comprehensive survey. AI Open, 3(2021):133–155. https://doi.org/10.1016/j.aiopen.2022.09.001
Huang Q, Zhou C, Wu J, Liu L, Wang B (2020) Deep spatial–temporal structure learning for rumor detection on Twitter. Neural Comput Appl 35(3):12995–13005. https://doi.org/10.1007/s00521-020-05236-4
Indu V, Thampi SM (2019) A nature - inspired approach based on Forest Fire model for modeling rumor propagation in social networks. J. Netw. Comput. Appl., 125(2018):28–41. https://doi.org/10.1016/j.jnca.2018.10.003
Iwendi C, Mohan S, khan S, Ibeke E, Ahmadian A, Ciano T (2022) Covid-19 fake news sentiment analysis. Comput Electr Eng 101(March):107967. https://doi.org/10.1016/j.compeleceng.2022.107967
Jarrahi A, Safari L (2021) FR-Detect: A Multi-Modal Framework for Early Fake News Detection on Social Media Using Publishers Features. August 2018, 1–21. http://arxiv.org/abs/2109.04835
Jehad R, Yousif S (2021) Classification of fake news using multi-layer percepti on. AIP Conference Proceedings, 2334, 070004(2021). https://doi.org/10.1063/5.0042264
Jin Z, Cao J, Zhang Y, Luo J (2016) News verification by exploiting conflicting social viewpoints in microblogs. 30th AAAI Conference on Artificial Intelligence, AAAI 2016,USA, pp 2972–2978

Kaliyar RK, Goswami A, Narang P, Sinha S (2020) FNDNet – a deep convolutional neural network for fake news detection. Cogn Syst Res 61:32–44. https://doi.org/10.1016/j.cogsys.2019.12.005

Kaliyar RK, Goswami A, Narang P (2021a) A hybrid model for effective fake news detection with a novel COVID-19 dataset. ICAART 2021 - Proceedings of the 13th International Conference on Agents and Artificial Intelligence, 2(Icaart), pp 1066–1072. https://doi.org/10.5220/0010316010661072

Kaliyar RK, Goswami A, Narang P (2021b) FakeBERT: fake news detection in social media with a BERT-based deep learning approach. Multimed Tools Appl 80:1765–11788. https://doi.org/10.1007/s11042-020-10183-2

Kar D, Bhardwaj M, Samanta S, Azad AP (2020) No Rumours Please! A Multi-Indic-Lingual Approach for COVID Fake-Tweet Detection. http://arxiv.org/abs/2010.06906

Karnyoto AS, Sun C, Liu B, Wang X (2022) Transfer Learning and GRU-CRF Augmentation for Covid-19 fake news detection. Comput Sci Inf Syst 19(2):639–658. https://doi.org/10.2298/CSIS210501053K

Kaur S, Kumar P, Kumaraguru P (2020) Automating fake news detection system using multi-level voting model. Soft Comput 24(12):9049–9069. https://doi.org/10.1007/s00500-019-04436-y

Kausar N, AliKhan A, Sattar M (2022) Towards better representation learning using hybrid deep learning model for fake news detection. Social Network Anal Min, 12(1), 1–11. https://doi.org/10.1007/s13278-022-00986-6

Kishwar A, Zafar A (2023) Fake news detection on Pakistani news using machine learning and deep learning. Expert Syst. Appl., 211(April 2022):1–10 https://doi.org/10.1016/j.eswa.2022.118558

Koloski B, Pollak S, Škrlj B (2020) Multilingual Detection of Fake News Spreaders via Sparse Matrix Factorization Notebook for PAN at CLEF 2020. CEUR Workshop Proceedings, 2696(September), pp 22–25

Kumar S, Kumar S, Yadav P, Bagri M (2021) A Survey on Analysis of Fake News Detection Techniques. Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021, pp 894–899. https://doi.org/10.1109/ICAIS50930.2021.9395978

Kwon S, Cha M, Jung K, Chen W, Wang Y (2013) Prominent features of rumor propagation in online social media. Proceedings - IEEE International Conference on Data Mining, ICDM, Dallas, pp 1103–1108. https://doi.org/10.1109/ICDM.2013.61

Lee JW, Kim JH (2022) Fake Sentence Detection Based on Transfer Learning: Applying to Korean COVID-19 Fake News. Appl. Sci. 2022, 12(13):1–15. https://doi.org/10.3390/app12136402

Li Y, Jiang B, Shu K, Liu H (2020) Toward A Multilingual and Multimodal Data Repository for COVID-19 Disinformation. Proceedings –2020 IEEE International Conference on Big Data, Big Data 2020, pp 4325–4330. https://doi.org/10.1109/BigData50022.2020.9378472

Li Q, Hu Q, Lu Y, Yang Y, Cheng J (2020a) Multi-level word features based on CNN for fake news detection in cultural communication. Pers Ubiquitous Comput 24(2):259–272. https://doi.org/10.1007/s00779-019-01289-y

Li D, Guo H, Wang Z, Zheng Z (2021) Unsupervised fake news detection based on Autoencoder. IEEE Access 9:29356–29365. https://doi.org/10.1109/ACCESS.2021.3058809

Long Y, Lu Q, Xiang R, Li M, Huang C-R (2017) Fake News Detection Through Multi-Perspective Speaker Profiles. Proceedings of the Eighth International Joint Conference on Natural Language Processing, Volume 2:(8), pp 252–256. http://www.aclweb.org/anthology/I17-2043

Ma J, Gao W, Mitra P, Kwon S, Jansen BJ, Wong KF, Cha M (2016) Detecting rumors from microblogs with recurrent neural networks. IJCAI International Joint Conference on Artificial Intelligence, 2016, pp 3818–3824

Madhubala M, Yadav AK, Sucharitha G, Praveen Kumar D, P (2021) Ann Rom Soc Cell Biol 25(6):4182–4192. https://annalsofrscb.ro/index.php/journal/article/view/6212 A Deep Learning Based Algorithm Design for Fake News Detection Framework

Mahyoob M, Algaraady J, Alrahaili M (2020) Linguistic-based detection of fake news in Social Media. Int J Engl Linguist 11(1):99–109. https://doi.org/10.5539/ijel.v11n1p99

Majumder S, Das D (2021) Studies Towards Language Independent Fake News Detection. Proceedings of the 18th International Conference on Natural Language Processing (ICON), India, pp 439–446. https://aclanthology.org/2021.icon-main.53

Mallik A, Kumar S (2024) Word2Vec and LSTM based deep learning technique for context-free fake news detection. Multimed Tools Appl 83:919–940

Martínez-gallego K (2016) Fake News Detection In Spanish Using Deep Learning arXiv: 2110. 06461v1 [ cs. CL ] 13 Oct 2021. 1–10

Meel P, Vishwakarma DK (2020) Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst Appl 153:112986. https://doi.org/10.1016/j.eswa.2019.112986

Meesad P (2021) Thai Fake News Detection Based on Information Retrieval, Natural Language Processing and Machine Learning. SN Comput Sci 2(6):1–17. https://doi.org/10.1007/s42979-021-00775-6

Mikolov T (2013) Learning Representations of Text using Neural Networks (Slides). Adv. Neural Inf. Process. Syst., ACM, pp 3111–3119

Mitra T, Gilbert E (2015) CREDBANK: A large-scale social media corpus with associated credibility annotations. Proceedings of the 9th International Conference on Web and Social Media, ICWSM 2015, 258–267. https://doi.org/10.1609/icwsm.v9i1.14625

Mohapatra A, Thota N, Prakasam P (2022) Fake news detection and classification using hybrid BiLSTM and self-attention model. Multimed Tools Appl 81(2):18503–18519

Mohawesh R, Maqsood S, Althebyan Q (2023) Multilingual deep learning framework for fake news detection using capsule neural network. J Intell Inf Syst 60(3):655–671. https://doi.org/10.1007/s10844-023-00788-y

Monteiro RA, Santos RLS, Pardo TAS, de Almeida TA, Ruiz EES, Vale OA (2018) Contributions to the Study of Fake News in Portuguese: New Corpus and Automatic Detection results. Lecture notes in Computer Science (including subseries lecture notes in Artificial Intelligence and Lecture notes in Bioinformatics): Vol. 11122 LNAI. Springer International Publishing. https://doi.org/10.1007/978-3-319-99722-3_33

Mosallanezhad A, Karami M, Shu K, Mancenido MV, Liu H (2022) Domain Adaptive Fake News Detection via Reinforcement Learning. WWW 2022 - Proceedings of the ACM Web Conference 2022, pp 3632–3640. https://doi.org/10.1145/3485447.3512258

Mridha MF, Keya AJ, Hamid MA, Monowar MM, Rahman MS (2021) A comprehensive review on fake news detection with Deep Learning. IEEE Access 9:156151–156170. https://doi.org/10.1109/ACCESS.2021.3129329

Nasir JA, Khan OS, Varlamis I (2021) Fake news detection: a hybrid CNN-RNN based deep learning approach. Int Int J Inf Manag Data Insights 1(1):1–13. https://doi.org/10.1016/j.jjimei.2020.100007

Okoro EM, Abara BA, Umagba AO, Ajonye AA, Isa ZS (2018) A hybrid approach to fake news detection on social media. Nigerian Niger J Technol 37(2):454–462. https://doi.org/10.4314/njt.v37i2.22

Oshikawa R, Qian J, Wang WY (2018) A Survey on Natural Language Processing for Fake News Detection. arXiv:1811.00770, pp 1–8

Ott M, Choi Y, Cardie C, Hancock JT (2011) Finding deceptive opinion spam by any stretch of the imagination. ACL-HLT 2011 - Proceedings of the 49th Annual Meeting of the Association for

Computational Linguistics: Human Language Technologies, 1, pp 309–319
Pagliardini M, Gupta P, Jaggi M (2018) Unsupervised learning of sentence embeddings using compositional n-gram features. NAACL HLT 2018–2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1, Louisiana, pp 528–540. https://doi.org/10.18653/v1/n18-1049
Palani B, Elango S (2023) BBC-FND: an ensemble of deep learning framework for textual fake news detection. Comput Electr Eng 110(July):1–22
Pan JZ, Pavlova S, Li C, Li N, Li Y, Liu J (2018) Content Based Fake News Detection Using Knowledge Graphs. In: Vrandečić, D., et al. The Semantic Web – ISWC 2018. ISWC 2018. Lecture notes in computer science, 11136. Springer, Cham. https://doi.org/10.1007/978-3-030-00671-6_39
Parikh SB, Atrey PK (2018) Media-Rich Fake News Detection: A Survey. Proceedings - IEEE 1st Conference on Multimedia Information Processing and Retrieval, MIPR 2018, USA, pp 436–441. https://doi.org/10.1109/MIPR.2018.00093
Pérez-Rosas V, Kleinberg B, Lefevre A, Mihalcea R (2018) Automatic detection of fake news. COLING 2018–27th International Conference on Computational Linguistics, Proceedings, Coling, pp 3391–3401
Posadas-Durán JP, Gomez-Adorno H, Sidorov G, Escobar JJM (2019) Detection of fake news in a new corpus for the Spanish language. J Intell Fuzzy Syst 36(5):4868–4876. https://doi.org/10.3233/JIFS-179034
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, 2019-Novem(Icdm), pp 518–527. https://doi.org/10.1109/ICDM.2019.00062
Rashkin H, Choi E, Jang JY, Volkova S, Choi Y (2017) Truth of varying shades: Analyzing language in fake news and political fact-checking. EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings, pp 2931–2937. https://doi.org/10.18653/v1/d17-1317
Rastogi S, Bansal D (2023) A review on fake news detection 3T's: typology, time of detection, taxonomies. Int J Inf Secur 22(1):177–212. https://doi.org/10.1007/s10207-022-00625-3
Reis JCS, Correia A, Murai F, Veloso A, Benevenuto F, Cambria E (2019) Supervised learning for fake news detection. IEEE Intell Syst 34(2):76–81. https://doi.org/10.1109/MIS.2019.2899143
Rohera D, Shethna H, Patel K, Thakker U, Tanwar S, Gupta R, Hong WC, Sharma R (2022) A taxonomy of fake news classification techniques: survey and implementation aspects. IEEE Access 10:30367–30394. https://doi.org/10.1109/ACCESS.2022.3159651
Rubin VL, Chen Y, Conroy NJ (2015) Deception detection for news: Three types of fakes. Proceedings of the Association for Information Science and Technology, 52(1):1–4. https://doi.org/10.1002/pra2.2015.145052010083
Ruchansky N (2017) CSI: A Hybrid Deep Model for Fake News Detection. Conference on Information and Knowledge Management, New York, pp 797–806
Sahoo SR, Gupta BB (2019) Classification of various attacks and their defence mechanism in online social networks: a survey. Enterp Inf Syst 13(6):832–864. https://doi.org/10.1080/17517575.2019.1605542
Sahoo SR, Gupta BB (2021) Multiple features based approach for automatic fake news detection on social networks using deep learning. Appl Soft Comput 100:1–16. https://doi.org/10.1016/j.asoc.2020.106983
Saleh H, Alharbi A, Alsamhi SH (2021) OPCNN-FAKE: optimized convolutional neural network for fake news detection. IEEE Access 9:129471–129489. https://doi.org/10.1109/ACCESS.2021.3112806
Samadi M, Mousavian M, Momtazi S (2021) Deep contextualized text representation and learning for fake news detection. Inf Process Manage 58(6):1–13. https://doi.org/10.1016/j.ipm.2021.102723
Santos RLS, Wick-Pedro G, Leal S, Vale OA, Pardo TAS, Bontcheva K, Scarton C (2020) Measuring the impact of readability features in fake news detection. LREC 2020–12th International Conference on Language Resources and Evaluation, Conference Proceedings, France, pp 1404–1413
Sarker IH (2021) Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci 2(6):1–21. https://doi.org/10.1007/s42979-021-00815-1
Selvaganapathy SG, Nivaashini M, Natarajan HP (2018) Deep belief network based detection and categorization of malicious URLs. Inf Secur J 1 27(3):145–161. https://doi.org/10.1080/19393555.2018.1456577
Shabani S, Sokhn M (2018) Hybrid machine-crowd approach for fake news detection. Proceedings – 4th IEEE International Conference on Collaboration and Internet Computing, CIC 2018, USA. pp 299–306. https://doi.org/10.1109/CIC.2018.00048
Shahi GK, Nandini D (2020) FakeCovid -- A Multilingual Cross-domain Fact Check News Dataset for COVID-19. 1–9. arXiv:2006.11343 https://doi.org/10.36190/2020.14.
Shahid W, Li Y, Staples D, Amin G, Hakak S, Ghorbani A (2022) Are you a Cyborg, Bot or Human?-A survey on detecting fake news spreaders. IEEE Access 10:27069–27083. https://doi.org/10.1109/ACCESS.2022.3157724
Sharma DK, Garg S (2021) IFND: a benchmark dataset for fake news detection. Complex & Intelligent Systems. (2021):1–21.https://doi.org/10.1007/s40747-021-00552-1
Shu K, Sliva A, Wang S, Tang J, Liu H (2017) Fake news detection on social media: a data mining perspective. Special Interest Group on Knowledge Discovery in Data SIGKDD explorations. Newsletter 19(1):22–36
Shu K, Mahudeswaran D, Wang S, Lee D, Liu H (2018) FakeNewsNet: A Data Repository with News Content, Social Context and Spatiotemporal Information for Studying Fake News on Social Media. Big Data 8(3):171–188
Singhal S, Shah RR, Chakraborty T, Kumaraguru P, Satoh S (2019) SpotFake: A multi-modal framework for fake news detection. Proceedings – 2019 IEEE 5th International Conference on Multimedia Big Data, BigMM 2019,pp 39–47. https://doi.org/10.1109/BigMM.2019.00-44
Stringhini G, Kruegel C, Vigna G (2010) Detecting spammers on social networks. Proceedings - Annual Computer Security Applications Conference, ACSAC, December 2010, USA, pp 1–9. https://doi.org/10.1145/1920261.1920263
Tacchini E, Ballarin G, della Vedova ML, Moret S, de Alfaro L (2017) Some like it hoax: Automated fake news detection in social networks. ArXiv, pp 1–12
Thorne J, Vlachos A, Christodoulopoulos C, Mittal A (2018) FEVER: A large-scale dataset for fact extraction and verification. NAACL HLT 2018–2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1, Louisiana, pp 809–819. https://doi.org/10.18653/v1/n18-1074
Thota A (2018) Fake news detection: a Deep Learning Approach. SMU Data Sci Rev, 1(3)
Tzortzis G, Likas A (2007) Deep belief networks for spam filtering. Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, 2, pp 306–309. https://doi.org/10.1109/ICTAI.2007.65
Varlamis I, Michail D, Glykou F, Tsantilas P (2022) A Survey on the Use of Graph Convolutional Networks for combating fake news. Future Internet 14(3):1–19

Varol O, Ferrara E, Menczer F, Flammini A (2017) Early detection of promoted campaigns on social media. EPJ Data Sci 6(1):1–29. https://doi.org/10.1140/epjds/s13688-017-0111-y

Verdoliva L (2020) Media forensics and DeepFakes: an overview. IEEE J Sel Top Sign Proces 14(5):910–932. https://doi.org/10.1109/JSTSP.2020.3002101

Vereshchaka A, Cosimini S, Dong W (2020) Analyzing and distinguishing fake and real news to mitigate the problem of disinformation. Comput Math Organ Theory 26(3):350–364. https://doi.org/10.1007/s10588-020-09307-8

Verma PK, Agrawal P, Amorim I, Prodan R (2021) WELFake: Word Embedding over linguistic features for fake news detection. IEEE Trans Comput Social Syst 8(4):881–893. https://doi.org/10.1109/TCSS.2021.3068519

Vijayaraghavan S, Wang Y, Guo Z, Voong J, Xu W, Nasseri A, Cai J, Li L, Vuong K, Wadhwa E (2020) Fake News Detection with Different Models. pp 1–15.http://arxiv.org/abs/2003.04978

Vlachos A, Riedel S (2014) Fact Checking: Task definition and dataset construction. Proceedings ofthe ACL 2014 Workshop on Language Technologies and Computational Social Science, USA, pp 18–22

Vogel I, Jiang P (2019) Fake News Detection with the New German Dataset GermanFakeNC. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 11799:288–295 LNCS. https://doi.org/10.1007/978-3-030-30760-8_25

Vosoughi S, Mohsenvand M, 'Neo, Roy D (2017) Rumor gauge: Predicting the veracity of rumors on twitter. ACM Trans Knowl Discovery Data 11(4):1–36. https://doi.org/10.1145/3070644

Wagner W (2010) Steven Bird, Ewan Klein and Edward Loper: Natural Language Processing with Python, analyzing text with the Natural Language Toolkit. Lang Resour Eval 44(4):421–424. https://doi.org/10.1007/s10579-010-9124-x

Wang WY (2017) Liar, Liar Pants on Fire: A New Benchmark Dataset for Fake News Detection. ArXiv, pp 422–426

Wang C, Zhu J (2019) Forbidden nodes aware community search. 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, pp 758–765. https://doi.org/10.1609/aaai.v33i01.3301758

Wang Y, Yang W, Ma F, Xu J, Zhong B, Deng Q, Gao J (2020) Weak supervision for fake news detection via reinforcement learning. AAAI 2020–34th AAAI Conference on Artificial Intelligence, pp 516–523. https://doi.org/10.1609/aaai.v34i01.5389

Wani A, Joshi I, Khandve S, Wagh V, Joshi R (2021) Evaluating deep learning approaches for Covid19 fake news detection. Commun Comput Inf Sci 1402 CCIS:153–163. https://doi.org/10.1007/978-3-030-73696-5_15

Wu L, Liu H (2018) Tracing fake-news footprints: Characterizing social media messages by how they propagate. WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining, 2018-Febua, pp 637–645. https://doi.org/10.1145/3159652.3159677

Wu L, Morstatter F, Carley KM, Liu H (2019) Misinformation in Social Media: definition, Manipulation, and detection. ACM SIGKDD Explorations Newsl 21(2):80–90. https://www.nytimes.com/2016/11/21/technology/fact

Xue J, Wang Y, Tian Y, Li Y, Shi L, Wei L (2021) Detecting fake news by exploring the consistency of multimodal data. Inf Process Manag 58(5):1–14. https://doi.org/10.1016/j.ipm.2021.102610

Yang Y, Zheng L, Zhang J, Cui Q, Li Z, Yu PS (2018) TI-CNN: Convolutional Neural Networks for Fake News Detection. http://arxiv.org/abs/1806.00749, pp 1–11

Yin W, Kann K, Yu M, Schütze H (2017) Comparative Study of CNN and RNN for Natural Language Processing. http://arxiv.org/abs/1702.01923, pp 1–7

Zhang X, Ghorbani AA (2019) An overview of online fake news: Characterization, detection, and discussion. Inf. Process. Manag., August 2018, 102025. https://doi.org/10.1016/j.ipm.2019.03.004

Zhang X, Lashkari AH, Ghorbani AA A lightweight online advertising classification system using lexical-based features. ICETE 2017 - Proceedings of the 14th International Joint Conference on e-Business and, Telecommunications (2017) 4(Cic), pp 486–494. https://doi.org/10.5220/0006459804860494

Zhang Q, Guo Z, Zhu Y, Vijayakumar P, Castiglione A (2023) A deep learning-based Fast Fake News Detection Model for Cyber-physical Social Services. Pattern Recognit Lett 168(2023):31–38. https://doi.org/10.1016/j.patrec.2023.02.026

Zhou X, Zafarani R (2020) A survey of fake news: fundamental theories, detection methods, and opportunities. arXiv:1812.00315. ACM Comput Surv 53(5):1–40

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.