International Journal of Information Management Data Insights 1 (2021) 100052

Contents lists available at ScienceDirect

International Journal of Information Management Data Insights

journal homepage: www.elsevier.com/locate/jjimei

Combating the menace: A survey on characterization and detection of fake news from a data science perspective

Wazib Ansar[a,a,*], Saptarsi Goswami[b]

[a] A.K. Choudhury School of IT, University of Calcutta, Kolkata, India
[b] Bangabasi Morning College, University of Calcutta, Kolkata, India

ARTICLE INFO

Keywords:
Fake news
Fake news data sets
Transformer NLP
Deep learning
Machine learning
COVID-19

ABSTRACT

Journalism has always remained a vital constituent of our society and journalists play a key role in making people aware of the happenings and developments in society. This spread of information enables shaping the ideologies, orientations and thoughts of individuals as well as the society. Contrary to this, the spread of misinformation or fake news leads to detrimental consequences. With the advent of social media, the menace of fake news has become grievous due to the unrestrained propagation of information and difficulty to track several accounts operated by humans or bots. This menace can be mitigated through data science approaches by combining artificial intelligence with statistics and domain-based knowledge. In this paper, a survey of works aimed at characterization, feature extraction and subsequent detection of fake news has been conducted from a data science perspective. Along with it, an analysis of the 8 renowned fake news detection repositories has been presented. Furthermore, through a case study on tweets related to COVID-19 pandemic, the factors behind the spread of misinformation during critical times, distinguishing between factual and emotional tweets and viable approaches to restrain fake news has been enunciated.

1. Introduction

Human beings have always been inquisitive of the environment around them and the happenings in their surroundings. The information gathered as a result of this inquisition has helped people to make informed decisions in their day-to-day lives. This gathering and conveyance of information has evolved through the years and led to the development of the term known as "news" and journalism as a profession. News is often seen as an output of journalism which ideally provides independent, reliable, accurate, and comprehensive information account of events that significantly affect people (Kovach & Rosenstiel, 2014). Today, journalism is a prominent source of information to people and is also regarded as one of the "pillars of democracy". With journalism being so influential in the development of society there come certain responsibilities which journalists must abide by such as adherence to particular standards, being objective and accurate and above all to represent the truth (Keeble, 2008). But in the present scenario, the news is not always based upon journalists' own preferences (White, 1950). Various factors such as the government, advertisers and the taste of the audience also play a vital role in the final representation of news (Shoemaker & Reese, 2013). It has been found that news is being sold to its audiences while at the same time audiences are being sold to the advertisers (McManus, 1992). All these factors have contributed towards the negative practice of creating fake news for financial or political gains, public confusion, social stigma, discrimination and glorifying or defaming a particular individual, organization or product. Fake news has gained so much relevance in the present world that it was awarded word of the year in 2016 by Macquarie dictionary. Through a study, it has been found that the ex-president of the US, Donald J Trump used the term "Fake News" approximately 900 times in his microblogs to condescend the media.[1] In the year 2019, India experienced such a surge in fake news related to events like "Article370", "NRC", "CAA" and the "Pulwama attacks" that the year 2019 was termed as the "Year of Fake News".[2]

With the advent of the era of Web 2.0, social media has become an inevitable part of our lives and as per study, it has been found that in US social media is the primary source of news for 62% of the adults compared to 49% in 2012 (Gottfried & Shearer, 2016). This makes social media the most preferred medium for sharing fake news among individuals (Duffy, Tandoc, & Ling, 2020). Also, a large share of people

[1] https://edition.cnn.com/2020/10/25/world/trump-fake-news-legacy-intl/index.html

[2] https://economictimes.indiatimes.com/news/politics-and-nation/fake-news-still-a-menace-despite-government-crackdown-fact-checkers/articleshow/72895472.cms

* Corresponding author.
E-mail addresses: wazibansar@ymail.com (W. Ansar), sgakc@caluniv.ac.in (S. Goswami).

https://doi.org/10.1016/j.jjimei.2021.100052
Received 6 June 2021; Received in revised form 19 October 2021; Accepted 28 November 2021

who view such fake news tend to believe them and share them further (Talwar, Dhir, Kaur, Zafar, & Alrasheedy, 2019). It was found that in the 2016 US presidential elections, fake news favoring Trump was shared 30 million times while those favoring Hillary Clinton were shared 7.6 million times on Facebook (Allcott & Gentzkow, 2017). One of the most reminiscent instances of fake news was during the 2016 US Presidential elections popularly known as Pizzagate which had flooded the social media. In this scandal, US presidential candidate Hillary Clinton was falsely claimed to run an illegal child trafficking racket with its headquarters in a pizza restaurant. This fake news had very serious consequences to the extent that a man aggrieved by this news went inside the restaurant with an assault rifle and fired bullets! (Bleakley, 2021). Another instance when tremendous volumes of purported mis-information spread was during the 2020 US Presidential elections in which unverified and misleading claims ranging from the "election being rigged by the democrats"[3] to "voter fraud through mail-in ballots."[4] In India too, just after Demonetization took place on 8th November, 2016 and new notes were introduced, rumors spread that the new notes were equipped with a chip that had tracking capabilities of up to 120 m below the earth. Also, there have been claims of lynching carried out in various pockets due to rumors spread on social media platforms or instant messaging applications.[5]

In order to curb the dissemination of fake news, certain standards exist among journalists which they must abide by Keeble (2008) and various news outlets and even politicians have come together plan ways to curb fake news propagation (Butler, 2018). Detecting fake news articles manually is a subjective and painstaking task which requires expertise in that domain. With the vastness of social media platforms consisting of voluminous information and no such foolproof measure to ensure the credibility of the profiles, it is nearly impossible to entirely curtail the propagation of fake news on social media (Thota, Tilak, Ahluwalia, & Lohia, 2018). The research is in progress to deal with fake news wherein various works have been conducted for characterization and detection of fake news. However, majority of the fake news detection techniques are supervised and require to be trained on large annotated data sets (Long, 2017). Such models are domain-dependent, i.e. models trained on one domain (such as political elections) cannot be applied to other domain (like medical science) without fine-tuning. Therefore, fake news detection is a complex task and requires knowledge of multiple disciplines such as artificial intelligence including deep learning and machine learning techniques, statistics, journalism, information science, psychology and social science (Shu, Sliva, Wang, Tang, & Liu, 2017; Zhou & Zafarani, 2020).

The review methodology adopted in a survey plays a crucial role in the effective portrayal of information and achieving the objectives behind the survey. A survey may have various objectives such as presenting available information related to a term or concept, tracking the history of developments, identifying relationships among associated concepts, examining the evidence in support of any proposition or justification of the worthiness of a problem for further research (Aromataris & Pearson, 2014). Methods like bibliometric analysis analyze the research conducted based on year of publishing, citations, authors affiliations, author collaborations and other publication details (Chen, Zou, Cheng, & Xie, 2020). But, bibliometric analysis primarily focuses on the quantitative aspect of research and falters to assess the qualitative aspects like scientific merit and drawbacks of a work (Kabudi, Pappas, & Olsen, 2021). Another popular review method is systematic mapping which maps the existing works based on both qualitative as well as quantitative aspects that may serve as a primer for future research or review works (Kabudi et al., 2021). This approach carries out a broad mapping of existing works based on a set of research questions due to time constraints, leading to a shallow review with a high probability of some significant articles being excluded (Grant & Booth, 2009). Apart from these methods, literature review helps to identify, consolidate and assess the existing works to help determine the research gaps. This fosters a complete, comprehensive, as well as structured analysis of research, works (Grant & Booth, 2009). A method similar to the literature review is systematic review which too summarizes the information from numerous existing works through a comprehensive study into a single article following a set of guidelines. However, it stands apart from the literature review in focusing on reporting data from existing works instead of theoretical concepts (Aromataris & Pearson, 2014). After evaluating the benefits and shortcomings of the above-mentioned methods, we adopt systematic literature review method accompanied by a case study to comprehensively and systematically present the concepts as well as data from existing research related to fake news. This fulfills our objective to methodically present the concepts associated with the term "fake news" with an extensive survey of the research works encompassing multiple disciplines to present the accomplishments and drawbacks of existing developments accompanied by future recommendations.

In this survey, a comprehensive review of the various works aimed at understanding and curtailing the generation and dissemination of fake news has been conducted. This has been accompanied by an innovative case study to comprehend how situations of distress provide a fertile ground for propagation of fake news and grasp the variation in requirements, apprehensions and psychological conditions of people over a period of time. In addition, efforts have been directed to obtain a relationship among the credibility of the profiles from the trade-off between the percentage of factual and emotional tweets by them. This paper is a systematic literature review accompanied by a case study for a better understanding of the factors behind the dissemination of fake news and measures for its mitigation. However, the objective of this work does not include collection and analysis of data from microblogs. Here, a multidisciplinary study has been conducted including fields like data science, artificial intelligence, journalism, information science, psychology and social science. This work has been aimed at researchers, professionals as well as journalists interested in the above-mentioned disciplines. The prime contribution of this paper is as follows:

1. An interpretation of fake news has been presented supported by a comparison of its categories based upon aspects like objective, intention, distinguishability and information type.
2. A survey of related works has been conducted on the characterization, feature extraction and subsequent detection of fake news using statistical, machine learning and deep learning approaches.
3. An analysis of the renowned fake news detection repositories has been performed based upon various parameters.
4. A case study on tweets related to COVID-19 and distinguishing between factual and emotional tweets. A commentary of the factors behind the proliferation of fake news during the pandemic and the efforts to impede this spread of misinformation with associated research works.

The compendium of this paper is as follows- Section 2 describes the ***survey process*** carried out for this paper. As an outcome of the survey process, the following sections have been designed. Section 3 presents the definition of ***fake news*** and a comparison of its categories based upon aspects like objective, intention, distinguishability and information type. Section 4 discusses the ***characterization of fake news*** based upon theories related to journalism, information science, psychology and social science. Section 5 enunciates feature extraction and model construction for ***detection of fake news*** based upon statistical, machine learning and deep learning approaches. It also includes an analysis of existing annotated data sets based upon various parameters. Accompanied by it, Section 6 showcases a ***case study on COVID-19 related fake news*** wherein the contribution of the researchers and journalists have been highlighted along with an analysis of tweets related to COVID-19. This paper has been wrapped up through a ***discussion*** on the implica-

[3] https://www.bbc.com/news/54562611
[4] https://www.bbc.com/news/world-us-canada-53353404
[5] https://www.thequint.com/quintlab/lynching-in-india/

tions as well as the limitations of this study in Section 7 followed by the concluding remarks and the arenas in which the work can be extended in the future being presented in the ***conclusion*** section.

2. Survey process

The survey process (Bimba et al., 2016; Kabudi et al., 2021) comprised of a study of publications extracted from digital libraries like Google Scholar[6], Science Direct[7], IEEE Xplore[8], Semantic Scholar[9] and ACM Digital Library.[10] Both, original as well as review articles written in the English language have been surveyed. The articles were obtained from the digital libraries using keywords related to fake news and its associated terms like "news", "journalism", "rumor", "hoax", "misinformation", "disinformation", "fabricated news", etc. Additionally, research works on COVID-19 related fake news have been explored for the extensive case study on COVID-19 presented in this paper. During the survey process, 984 articles were obtained after the removal of duplicates out of 1217 articles published from the year 2010 to 2021. For this, a spreadsheet application was utilized. While selecting the articles, higher priority was given to those recently published. However, some path-breaking contributions were selected irrespective of their year of publication. In the cases when two or more articles were observed having the same concept or technique, the journal publications were selected over conference papers. Apart from this, the citations and the impact factor of the journals were taken into consideration while selecting the papers. If similar works by the same authors were noticed, the redundant articles were omitted. This resulted in 153 articles being considered for full-text review after filtering the articles depending upon the above-mentioned criteria. Based on the full-text evaluation, a total of 96 papers have been shortlisted for inclusion in this paper as listed in Appendix A. The flow of the survey process has been illustrated in Fig. 1. Finally, based on relevance, the study has been divided into various coherent segments for a methodical and systematic organization of content.

3. Fake news

Fake News has been circulated since times immemorial - In the epic Mahabharata, a piece of fake news "Ashwatthama is dead" was spread which led to the beheading of Dronacharya by Dhristadyumna [16]. In the recent context, the term "Fake News" gained momentum with the US presidential elections in 2016 which saw enormous false information or fake news being spread (Albright, 2016). A formal definition of fake news may be a news article which is delusive both intentionally and verifiably and could misinform the readers (Shu et al., 2017). Most of the time fake news is created with malicious intent, but sometimes it may also be created unintentionally or for the entertainment of its readers. With so much of false information being spread, the legitimacy of journalism is being undermined (Kang, Bae, Zhang, & Sundar, 2011) and news outlets and even politicians have come together to plan ways to curb fake news propagation (Butler, 2018).

Some recent works have expressed their reservations against the term "fake news" as it undermines the inherent credibility associated with the term "news" and prefer to use terms like "misinformation", "disinformation" or "fabricated news" instead (Allcott & Gentzkow, 2017; Ireton & Posetti, 2018). In another work, the authors acknowledge this issue and distinguish fake news as a "genre" from treating fake news as a "label" to undermine the credibility of the term "news" (Egelhofer & Lecheler, 2019). Upon observing the interest trend of terms related to fake news based on web searches over a period of 5 years from July, 2016 to July, 2021[11], it can be inferred that the interest trend of the term "fake news" outnumbers the interest trend of related terms like "misinformation", "disinformation" and "fabricated news" by a large margin. Therefore, keeping in mind the familiarity of the term "fake news" along with its implications, in this paper "fake news" has been projected as an umbrella-term which encompasses all other associated terms such as "misinformation", "disinformation", "fabricated news", "satire", "rumor", "hoax" and so on. Table 1 presents a comparison of the various categories of fake news and their similarities as well as differences based upon aspects like the objective behind it, whether the intention is malicious, its distinguishability from real news and the type of information presented.

3.1. Satire

Satire generally deals with mockery or exaggeration of real-world topics with a humorous intention (Rubin, Conroy, Chen, & Cornwell, 2016). Though not always classified as fake news due to the fact that most of the satirical news content does not involve any mal-intention and the audiences who watch them are generally able to distinguish them from real news (Highet, 2015). An example of satirical news may involve short animation films broadcast on news channels based upon current affairs crafted with punch lines and sarcastic-catchy graphics. Here entertainment of the audience is the sole objective with some hidden message.

3.2. Parody

Parody bear a certain resemblance to satire but the distinguishing feature is that humour is created through the use of non-factual information (Berkowitz & Schwartz, 2016; Maronikolakis, Villegas, Preotiuc-Pietro, & Aletras, 2020). They deal with the farcicality of issues and provide them limelight through a concoction of fabricated news (Sinclair, 2020). An example of a parody Website is The Onion[12], which has times often been recognized as a true news site. It has often posted news as whimsical and vague as "detonation of a nuclear scientist" by North Korea[13]. A parody, when crafted with excellence maintains a distinguishable line between nonsensical and conceivability.

3.3. Fabricated news

Fabricated news bears no relationship with actual news content but the style adopted is that of actual news content to create an impression of legitimacy. Here, the motive of the author is to mislead the readers as compared to satire or parody where there is an implicit understanding among the readers and the authors that the content presented is vague. Fabricated news is generally hard to distinguish from actual news content and is generally propelled by partisans of an organization, ideology or belief. Such news is often spread through the social media laminated with a fictitious film of genuineness due to the resemblance of the style and presentation with actual news content. This often misguides the readers who are unaware of the credibility of the source and believe it to be true (Frenda, Knowles, Saletan, & Loftus, 2013). It must be noted that the seeds of fabricated news germinate well in the environment of social tension and atmosphere of uncertainty. Under such circumstances, the readers are more prone to accepting fabricated news as legitimate and spread them among their acquaintances which creates a vicious web of deceptive content circulation. Another factor which leads to the propagation of fabricated news is the creation of social bots (Ferrara, Varol, Davis, Menczer, & Flammini, 2016) using which false

[6] https://scholar.google.co.in
[7] https://www.sciencedirect.com
[8] https://ieeexplore.ieee.org
[9] https://www.semanticscholar.org
[10] https://dl.acm.org

[11] https://trends.google.com/trends/explore?date=2016-07-17 2021-07-17&q=fake news, Fabricated News, Misinformation, Disinformation.
[12] http://www.theonion.com/about/
[13] http://www.theonion.com/article/north-korea-successfully-detonates-nuclear-scienti-55783

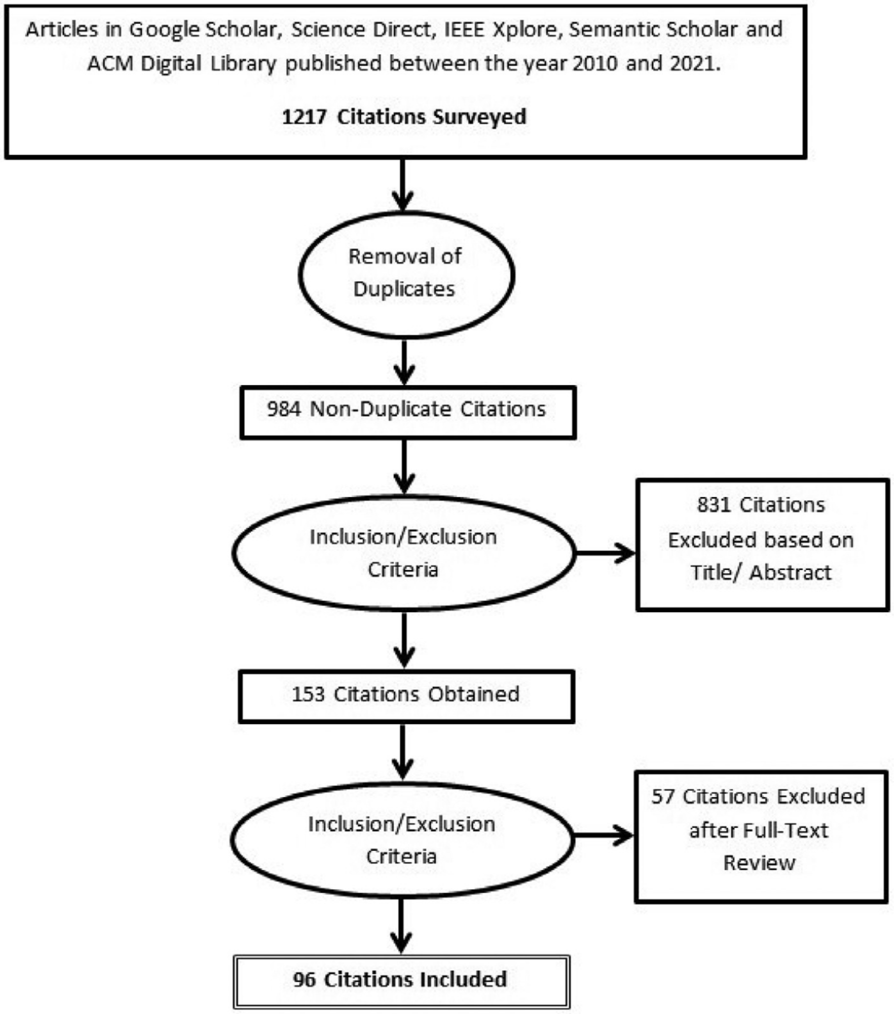


Fig. 1. Flow of the Survey Process.

Table 1
A Comparison of the Fake News Categories.

Category	Objective	Intention	Distinguishable from Real News	Information Type
Satire	Humor and entertainment	Usually benevolent	Yes	Factual
Parody	Humor and entertainment	Benevolent	Yes	Non-factual
Fabricated News	Public confusion, social stigma, propaganda, financial or political gains	Malicious	No	Non-factual
Manipulated Photos or Videos	Public confusion, social stigma, propaganda, financial or political gains	Malicious	No	Non-factual
Using Fake News as Advertisement	Financial gains	Malicious (if indistinguishable from real news)	Yes (Disclaimers present)	Factual/ Non-factual
Rumor	Propagate fear, anxiety, hatred or hope	Malicious/ Benevolent	No	Factual
Hoax	Prank or practical joke	Benevolent	No	Non-factual

impression can be created among users that a large number of readers are reading that news leading to real-time propaganda based ecosystem (Albright, 2016). A piece of fabricated news claimed Pope Francis endorsed the presidential candidate Donald J. Trump that during the US presidential elections. Approximately half of the readers considered this news to be true (Allport, 1937).

There may be various motives behind the creation and spread of fabricated news such as political benefits, financial gains, public confusion, social stigma, discrimination, glorifying or defaming a particular individual, organization or product and so on. Among such motives, the financial motive is assumed to be the most prominent one. Attracting more readers may result in attracting more advertisers and in turn lead to increased income to the authors of the stories. In the 2016 US presidential election, some authors of fabricated news in Macedonia have in turn used services like Google AdSense to make money (Subramanian, 2017). Also, in one study people who were paid to post comments on social media platforms were investigated (Chen, Wu, Srinivasan, & Zhang, 2013). It was found that they were paid to promote positive views regarding a certain product while at the same time they were asked to spread negative comments regarding the products of rival companies.

3.4. Manipulation of photos /videos

With the advent of social networking sites and instant messaging apps, a large number of photos and videos have been circulated. Apart from personal photos or videos, a major chunk of such media is related to topics such as politics, sports, entertainment, recent happenings and so on. Therefore, manipulation of photos or videos has attracted people both with and without malicious intent. The manipulations in an image may involve altering the image properties such as hue, saturation, resolution and orientation, cropping or merging with another image in order to deceive the viewers. For videos, the manipulations may deal with altering the frame rate, resolution or tampering with frames. Zubiaga and Ji in their study (Zubiaga & Ji, 2014) analyzed the morphed images spread during Hurricane Sandy via microblogs. There are various regulations which credible news agencies abide by for correct representation of truth related to images and video [30]. But in social media, there are no codes behind image manipulations which lead to the misappropriation of images. For instance, a video of Delhi chief minister Arvind Kejriwal before the 2017 assembly elections in Punjab was morphed to portray him drunk while giving a speech and spread through the social media.[14]. In another instance the picture of a sunken boat was morphed and a “pro-Trump” flag was added on top of it to give it an impression that it had sunk during the boat-parade in Texas in the run-up to the 2020 US Presidential elections.[15]

3.5. Using fake news as advertisement

Sometimes for the purpose of advertisement, some articles may be published which resemble a news article but are created solely for advertisement purpose. Such advertisements generally carry disclaimers from the news agencies that it has been created for advertisement purposes and the editor of the newspaper bear no responsibility for the claims in the advertisement (Nelson & Park, 2015). But, such type of advertisements still fall in the fake news category because there have been many instances where a major part of the audience believed such claims to be true and had unfavorable repercussions (Deziel, 2014).

[14] https://www.altnews.in/arvind-kejriwal-speaking-in-drunken-state-no-old-video-slowed-down-to-create-illusion/

[15] https://www.bbc.com/news/election-us-2020-54070406

3.6. Rumor

Rumor denotes a piece of information which has been circulated without ascertaining its truthfulness. The information circulated may eventually be true, partly true, false or remain unverified. Initially, a rumor is unverified due to lack of evidence in its favor or being officially verified by credible sources (Zubiaga, Aker, Bontcheva, Liakata, & Procter, 2018). The intention behind a rumor may be to propagate fear, anxiety, distrust or hatred in society (Yang, Li, & Giua, 2020b). For instance, the COVID-19 vaccines induced magnetic fields in the human body and the vaccine contains microchip that can be used to track the human population.[16] Contrastingly, it may even be used to generate hope and wishful thinking such as the rumor that the football player Cristiano Ronaldo donated 1.2 million dollars for flood victims in Kerala.[17]

3.7. Hoax

Hoaxes denote misinformation which are intentionally circulated to deceive people (Prasetijo et al., 2017; Tacchini, Ballarin, Della Vedova, Moret, & de Alfaro, 2017). It consists of concocted facts related to people, organizations or entities which may not even exist in the real world. It is usually so descriptive and complex that it may be difficult to distinguish them from the truth (Santoso, Yohansen, Warnars, Hashimoto et al., 2017). Although, it may be observed to be used interchangeably with fabricated news, it can be distinguished from the former by the intention behind it. Fabricated news is created with malicious intention having social, political, financial or psychological implications. On the other hand, hoaxes have comparatively benevolent intentions underlying its creation and propagation and are meant to be treated as a prank or practical joke (Kumar, West, & Leskovec, 2016). An example hoax may be the article on fictitious Australian aboriginal god Jar’Edo Wens which existed in Wikipedia for a more than nine years[18].

4. Characterization of fake news

The characterization of fake news is the first step towards research on curbing the dissemination of fake news. The characterization of fake news can be divided into the following categories as depicted in Fig. 2.

4.1. Traditional media

Traditional media is referred to as the form of media which existed before the advent of social media. The ecosystem of fake news generation and propagation has its foundations in the psychology of people and the society in which they live. This is enunciated as follows:-

- **Psychology:** Human beings have certain inbuilt psychological vulnerabilities which make them more prone to believing in fake news. “Naive-realism” is one such vulnerability wherein people believe that only the way in which they perceive the world is the reality and all others who differ in their views are misinformed or partial (Robinson, Keltner, Ward, & Ross, 1995). In “desirability-bias” people assimilate only the information which they find entertaining (Fisher, 1993). While in “confirmation-bias”, people believe only in information which is in accordance with their perception (Nickerson, 1998). Another such vulnerability is the “illusion of asymmetric-insight” in which people considerable their knowledge

[16] https://www.news18.com/news/lifestyle/no-magnetic-superpowers-or-dna-alteration-dont-pay-heed-to-these-popular-myths-about-covid-19-vaccine-3945962.html

[17] https://www.timesnownews.com/sports/football/video/is-it-true-rumour-of-cristiano-ronaldos-rs-77-crore-donation-to-kerala-flood-victims-goes-viral/272748

[18] https://en.wikipedia.org/wiki/Jar%27Edo_Wens_hoax

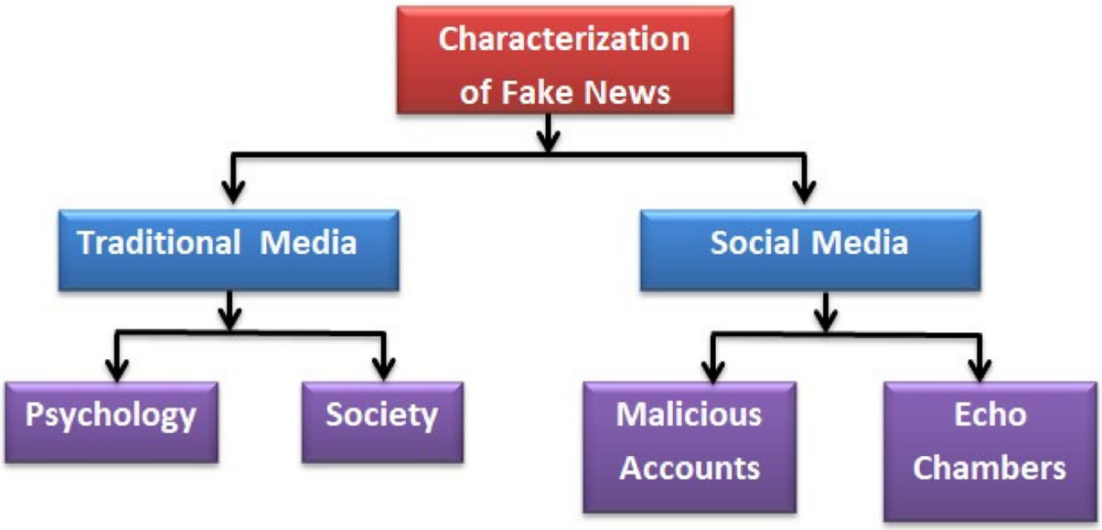


Fig. 2. Characterization of Fake News.

to be superior to the rest (Pronin, Kruger, Savtisky, & Ross, 2001). It has been observed that the presence of these psychological factors provide a very fertile ground for fake news to thrive.

- **Society:** In the world of information generation and consumption, there are two major players - publishers and consumers. News generation is in fact a process in which a certain event is taken as input and processed to form a news article in which some amount of distortion takes place due to biases involved in the presentation or representation of the article. There may also be certain expectations or interests of the publishers involved in the publishing process like financial gains and the readership extent which may be considered as short-term utilities. While the reputation and authenticity of the publisher may be considered as long-term utilities. It can be concluded that fake news is generated when the short-term utilities over-power the long-term utilities of the publishers (Shu et al., 2017).

4.2. Social media

The uncontrolled spread of fake news on social media platforms can be primarily attributed to malicious accounts and echo-chamber effect which are mentioned as follows:-

- **Malicious Accounts:** Malicious accounts are created using various means such as cyborgs, which are accounts created by humans but run certain programs and perform activities to spread fake news (Chu, Gianvecchio, Wang, & Jajodia, 2012), bots which generally do not need human intervention to interact with others and spread fake news (Ferrara et al., 2016; Polignano, de Pinto, Lops, & Semeraro, 2019) and trolls who are propaganda-driven humans who use social media as a platform to either promote to defame a particular individual or organization (Cheng, Bernstein, Danescu-Niculescu-Mizil, & Leskovec, 2017).
- **Echo-Chambers:** It can be easily understood by the old proverb "Birds of a feather flock together" which means that people with similar tastes and ideologies gel well together. In the present era, the abundance of social media has given rise to virtual connections and friendships such that people with similar tastes and views are now able to interact and form a network among themselves inspite of being physically distant from each other (Choi, Chun, Oh, Han et al., 2020). For example, on Facebook people generally associate with people who have similar perceptions and orientations (Quattrociocchi, Scala, & Sunstein, 2016). It is observed people generally trust a news article as true if others also express their trust in it due to "echo-chamber effect" (Jamieson & Cappella, 2008). Also, if a news article is circulated frequently then people start believing in it even if the source remains unverifiable (Paul & Matthews, 2016).

5. Detection of fake news

The detection of fake news is the second step towards research on curbing the dissemination of fake news. The process of fake news detection is illustrated in Fig. 3.

5.1. Feature extraction

The first step towards the detection of fake news deals with extracting the features from the news article. It can be broadly divided into two categories as follows:-

- **Based upon News Content:** A news article is generally composed of the following attributes - source consisting of the details of the author or publisher, headline descriptive of the content and topic, body which contains the actual article along with associated graphics or video.
  - *Text-Based:* For news articles primarily composed of text, the feature extraction process deals with extracting lexical features such as word-frequency, unique-words, word-count, word-size, etc. and syntactic features such as parts-of-speech, inter-word relationships and inter-sentence relationships and so on. Wynne & Wint (2019) studied the effect of n-grams and concluded that character n-grams contribute more towards improving performance compared to word n-grams for fake news detection. Gravanis, Vakali, Diamantaras, & Karadais (2019) observed that various linguistic features when combined with word embedding (Mikolov, Chen, Corrado, & Dean, 2013) for input representation, provides the most accurate results. Verma, Agrawal, Amorim, & Prodan (2021) proposed a feature extraction technique comprising of Term Frequency - Inverse Document Frequency (TF-IDF) and Count Vectorizer (CV). They showcased their approach to be more accurate than state-of-the-art approaches like Bidirectional Encoder Representations from Transformers (BERT). However, it has been observed that most of the works focus on a single language only. Faustini & Covões (2020) alleviated this impediment and devised an approach for which was applicable to varied languages with different syntax and semantics for each language.
  - *Media-Based:* Although text-based information plays a crucial part in the verification of news content, the role of media-based content cannot be ignored. Jin, Cao, Zhang, Zhou, &

Fig. 3. Detection of Fake News.

Tian (2016b) proposed a media-based approach for verification of news content wherein they extract the statistical and visual features such as histogram, image-ratio, coherence-score, clarity-score, clustering-score, hot-image ratio and so on. Agrawal & Sharma (2021) presented a study of existing techniques using neural networks and spatiotemporal information to detect the forgery in videos for creating fake news.

- **Based upon Social Context:** A news article consists of the following features related to social context-
  - *User Details:* From the user details features like age, gender, affiliations, location, follower-count, tweet-count, etc. can be extracted which can assist in detecting fake news (Shu, Wang, & Liu, 2018b). Yang, Liu, Yu, & Yang (2012) used a combination of user-based features accompanied by content-based, propagation-based, application program-based and location-based features. Furthermore, based upon these features, a tweet-retweet graph may be constructed and different users may be clustered into a group for enhanced comprehension of the characteristics of the group of users related to a particular news article (Castillo, Mendoza, & Poblete, 2013).
  - *Post Analysis:* After reading a piece of news on social media, users generally express their reactions, emotions or thoughts in the form of comments or posts. Therefore, extraction of features from the posts proves to be a major factor in the fake news detection process. (Jin, Cao, Zhang, & Luo, 2016a) observed stance to be an indication of acceptance or refusal of a particular news article. Rubin et al. (2016) inferred that a combination of features like absurdity, punctutations and semantics when combined together, may aid in predicting satirical content. Castillo et al. (2013) analyzed topic-level features, Ma, Gao, Wei, Lu, & Wong (2015) used Latent Dirichlet Allocation (LDA), while Verma et al. (2021) extracted features like TF-IDF and CV from the posts for assessing credibility. However, such approaches did not capture the variation in the characteristics of posts during propagation. To alleviate this issue, Ma et al. (2015) formulated a time-series based approach and noticed a difference in the characteristics of typical features present in rumors compared to non-rumors during propagation.
  - *Network Analysis:* In the methods based upon network analysis, the objective is to map the flow of news and identify the cluster of related users. It may be stance-based in which the nodes may represent the posts related to a piece of news and edges having weights depending upon the stance-similarity. Kwon, Cha, Jung, Chen, & Wang (2013) devised a friendship-based approach wherein the network is formed upon the followee-follower relationships accompanied by a diffusion-based in which a network is established based upon the news dissemination trajectories. For instance, Jin et al. (2016a); Tacchini et al. (2017) deploy a crowd-sourcing approach to determine the credibility of the network. Ruchansky, Seo, & Liu (2017) adopted an accompaniment-based technique in which the users are connected via the network if their posts are on similar articles. From these works, it can be inferred that a conglomeration of the content-based features, user details and analysis of the propagation of information through the network can provide an all-round and robust solution to fake news detection.
- **Based upon Spatiotemporal Content:** The spatiotemporal content comprises two parts - spatial and temporal. The spatial or location information could be extracted from the user profile details as observed from the work of Yang et al. (2012) wherein user-based features along with content-based, propagation-based, application program-based and location-based features have been considered. The temporal information provides an insight into the real-time propagation of fake news among social networks. It shows how fake news flows from one user to another based upon inter-user relationships and the variation in the topics over time (Shu, Mahudeswaran, Wang, Lee, & Liu, 2018a). For instance, Ma et al. (2015) captured the temporal information using a time-series based approach. Therefore, the inclusion of spatiotemporal information can aid in discovering vital information for distinguishing between real and fake content.

5.2. Model construction

The model for the detection of fake news can be divided into the following categories as depicted in Fig. 4.

- **Models based upon News Content:** These models focus upon the content of the news to detect fake news among them. Such models can be of the following types:-
  - *Knowledge Oriented:* The models in this category use external information to verify whether given news content is authentic or not. The methods for verification may be using humans who are experts in that domain for investigating the content. Examples for this may include PolitiFact[19] and Snopes.[20] Although, these methods have high accuracy due to the fact that experts themselves verify the content but suffer from constraints like scalability and low throughput. Another method for verification is to use the opinions of a large number of people regarding news content and then summarizing those in order to obtain the final verdict regarding the genuineness of the content. Examples for this may include Fiskkit[21] which considers user annotations in a news article to verify the authenticity. It is observed that such methods rely upon the assumption that the users are unbiased and there are sufficient numbers of users commenting upon a news article. The last and most promising method is to use computation-based software to verify news content. Such methods at first find the claims which need to be verified and then verify the claims to determine their authenticity. To extract the claims in the news article the statements are examined and the outlook is extracted

[19] http://www.politifact.com/
[20] http://www.snopes.com/
[21] http://fiskkit.com

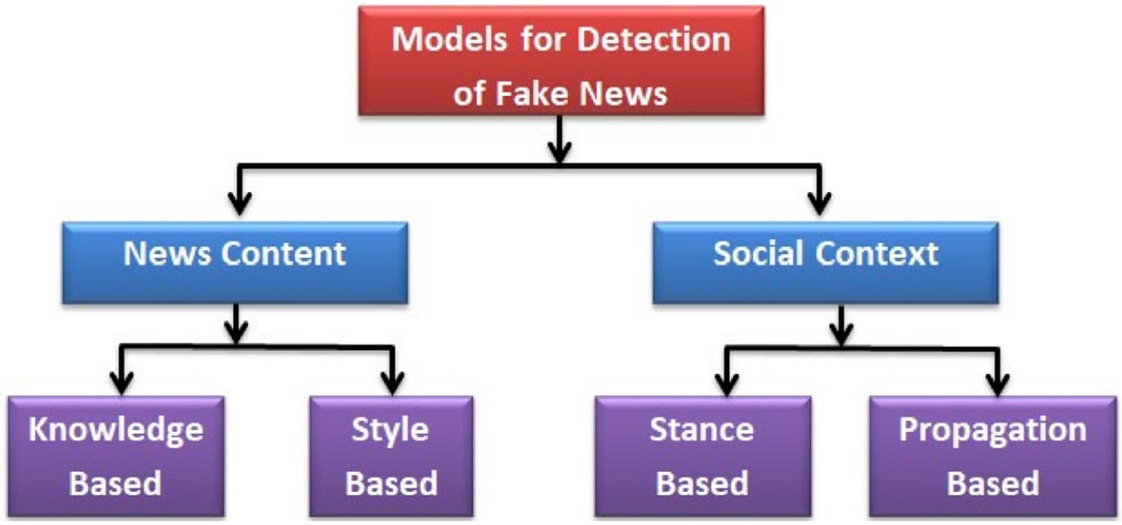


Fig. 4. Models for Detection of Fake News.

(Hassan, Li, & Tremayne, 2015). For claim verification external sources such as open-web (Fader, Soderland, & Etzioni, 2011; Magdy & Wanas, 2010) and knowledge-graphs like DB-pedia and Google-Relation-Extraction Corpus are checked (Ciampaglia et al., 2015; Shi & Weninger, 2016; Wu, Agarwal, Li, Yang, & Yu, 2014).

- *Style Oriented:* It is observed that fake news content is generally crafted in a manner to make the content captivating to the audience and represent the false claims in such a way that they seem genuine. Often, the writing style is common among various fake news articles due to the fact that the same author may write numerous fake news stories. Style based approach tries to identify the writing style in the news content in order to identify if it has been authored by a fake news author. Earlier works include Vrij (2005), Lesce (1990) in which the authors performed a study on various scientific and forensic tools to distinguish the truth. Other works such as Feng, Banerjee, & Choi (2012), Rubin & Lukoianova (2015) applied context-free-grammar to decompose the content into a rule-based structure which is further compared with the structures of other news articles for verification. Mukherjee, Venkataraman, Liu, & Glance (2013) present a style-oriented approach using Support Vector Machine (SVM) while, Granik & Mesyura (2017) propose an approach using Naive Bayes classifier. This has been taken forward by deep learning approaches such as Convolutional Neural Networks (CNN) by Wang (2017) or a combination of CNN with Recurrent Neural Networks (RNN) (Nasir, Khan, & Varlamis, 2021). However, RNN-based models failed to capture word-relationships beyond a specified length (Deepak & Chitturi, 2020). This led to the formulation of Long Short Term Memory (LSTM) model by Long (2017). This work was further improved by Roy, Basak, Ekbal, & Bhattacharyya (2018) through a Bidirectional Long Short Term Memory (Bi-LSTM) based approach which enabled the bidirectional representation of context. Recent advancements like models with attention mechanism (Paka, Bansal, Kaushik, Sengupta, & Chakraborty, 2021; Song, Ning, Zhang, & Wu, 2021) have further improved the performance of classification. However, the majority of deep-learning techniques are supervised and require to be trained on large annotated data sets (Deepak & Chitturi, 2020). This poses a limitation on application to tasks with limited data. To provide a solution to this issue, Baruah, Das, Barbhuiya, & Dey (2020); Jwa, Oh, Park, Kang, & Lim (2019) applied pre-trained models like BERT to facilitate transfer learning for fake news detection.

- **Models based upon Social Context:** The models based upon social context use the relationships among users and posts in a social network to capture relevant information to be used for the detection of fake news. Such models can be of the following types:-
  - *Based upon Stance:* Such models extract the opinions of different users related to a news article in order to verify it as fake or genuine. The opinions may consist of the number of likes a particular post has got, to the remarks provided related to an article. The opinion is further mined to find the explicit or implicit stances of users. For instance, Jin et al. (2016a) devised a method to verify news content based on analyzing the conflicting stances. In another study, Mohammad, Sobhani, & Kiritchenko (2017) proposed a more refined stance detection technique based on SVM and input features like n-grams and word embedding.
  - *Based upon Propagation:* Such models extract the relationships among social media posts which contain some degree of similarity among them. It tracks the flow of news articles among various networks and uses the credibility of users present in those networks in order to verify it as fake or genuine (Jin et al., 2016a). Yang et al. (2012) considered propagation-based information as a feature to determine credibility. Jin, Cao, Jiang, & Zhang (2014) proposed a three-tier hierarchical model comprising of event, sub-event and message-based analysis of propagation. Apart from analyzing the spatial propagation, a few works have considered temporal propagation too as a vital indicator of credibility. For instance, Ma et al. (2015) devised a time-series based approach to capture the temporal information during propagation. Apart from this, some works applied statistical measures to determine the propagation of fake news in a network. For instance, Shrivastava et al. (2020) obtained basic reproduction-number as a parameter to determine the prominence of message propagation in a network. Table 2 presents a study on the related works and models constructed for the detection of fake news using Statistical and Machine Learning Approaches. Whereas, Table 3 presents a study on the related works and models for the detection of fake news using Deep Learning Approaches.

5.3. Existing annotated data sets

In this section, a list of the available data sets employed for fake news detection task have been presented. Furthermore, a comparison of these repositories has been drawn in Table 4 based upon various metrics.

Table 2
A Comparative Study of the Fake News Detection using Statistical and Machine Learning Approaches.

Model	Description	Methods Used	Data Set(s)	Performance
Mukherjee et al. (2013)	Style oriented approach to Fake News Detection	SVM	Mechanical Turk and Yelp	Accuracy=0.861, F1-Score=0.857
Kwon et al. (2013)	Detecting Fake News based upon temporal, structural and linguistic characteristics	SVM, Random Forest and Decision Tree	Microblogs	Accuracy=0.897, F1-Score=0.878 (Random Forest)
Castillo et al. (2013)	Determination of newsworthiness and credibility of microblog propagation	LR and Random Forest	Microblogs	Mean F-Score=0.824, Mean ROC area=0.816
Ma et al. (2015)	Detecting rumors based upon social context	SVM	Microblogs	Mean Accuracy =0.871, Mean F1=0.87
Rubin & Lukoianova (2015)	Fake News Detection through discourse structure analysis	RST-VSM	Mechanical Turk	NA
Jin et al. (2016a)	Detecting relations among tweets and constructing a credibility-network	Topic Modeling	Microblogs	Accuracy=0.84
Rubin et al. (2016)	Using satirical cues for Fake news Detection	SVM and TF-IDF	News Articles	F-Score=0.87
Granik & Mesyura (2017)	Fake News Detection using NB	NB	Posts on Facebook	Accuracy=0.74
Tacchini et al. (2017)	Automated detection of hoaxes	LR and HBLC	Posts on Facebook	Mean Accuracy =0.992 (HBLC)
Gravanis et al. (2019)	Fake News Detection using content-based features	SVM, AdaBoost and Bagging	BuzzFeedNews, BS Detector and PolitiFact	Mean Accuracy=0.787
Shrivastava et al. (2020)	A mathematical model to analyze the spread of fake news in social media	Basic Reproduction Number	Online Social Networks	NA
Verma et al. (2021)	A Fake News Detection Approach through Word Embedding over Linguistic Features	TF-IDF and CV	WELFake	Accuracy=0.967

Note: NB: Naive Bayes Theorem, LR: Logistic Regression, HBLC: Harmonic Boolean-Label-Crowdsourcing, RST-VSM: Rhetorical-Structure-Theory - Vector-Space-Model, SVM: Support Vector Machine, TF-IDF: Term Frequency - Inverse Document Frequency, CV: Count Vectorizer

Table 3
A Comparative Study of the Fake News Detection using Deep Learning Approaches.

Model	Description	Methods Used	Data Set(s)	Performance
Long (2017)	Attention-Based approach to Fake News Detection	LSTM and Attention Mechanism	LIAR	Accuracy=0.415
Wang (2017)	Presented the LIAR Data Set	SVM, LR, CNN and Bi-LSTM	LIAR	Accuracy=0.274
Ruchansky et al. (2017)	Detecting Fake News based on network analysis	RNN	Microblogs	Mean Accuracy= 0.922, Mean F-Score=0.924
Roy et al. (2018)	Fake News Detection using Deep Learning Ensemble	CNN and Bi-LSTM	LIAR	F1-Score=0.43
Thota et al. (2018)	Stance Detection using TF-IDF with DNN	Neural Networks	FNC	Accuracy=0.942
Jwa et al. (2019)	Fake News Detection using BERT	BERT	FNC	F1-Score=0.746
Polignano et al. (2019)	Detection of bots and determining gender of social network users	CNN	CLEF 2019	Accuracy=0.918 (Rangel & Rosso, 2019)
Baruah et al. (2020)	Detection of Fake News Spreaders using BERT	BERT	Microblogs	Accuracy=0.690
Song et al. (2021)	Multimodal Fake News Detection	Attention Mechanism and CNN	Microblogs	Accuracy=0.922
Paka et al. (2021)	Cross-SEAN for Detecting COVID-19 based Fake News	BERT, RoBERT, Attention and Cross-Stitch Mechanism	CTF	Accuracy=0.954, F1-Score=0.953
Nasir et al. (2021)	Fake News Detection through Hybrid RNN-CNN	CNN, RNN	ISO, FA-KES	Mean Accuracy=0.795, Mean F1-Score=0.790

Note: LSTM: Long Short Term Memory, Bi-LSTM: Bidirectional Long Short Term Memory, RNN: Recurrent Neural Networks, CNN: Convolutional Neural Networks, DNN: Dense Neural Networks, BERT: Bidirectional Encoder Representations from Transformers, Cross-SEAN: Cross-Stitch Mechanism in Semi-Supervised Attention-Based Neural Model

Table 4
A Comparison of the Fake News Detection Repositories.

Data Set	News-Content		Social-Context			Spatiotemporal	
	Linguistic	Visual	User	Post	Network	Spatial	Temporal
BuzzFeedNews	Yes	-	-	-	-	-	-
LIAR	Yes	-	-	-	-	-	-
BS Detector	Yes	-	-	-	-	-	-
Fake News Challenge	Yes	-	-	Yes	-	-	-
CREDBANK	Yes	-	Yes	Yes	-	Yes	Yes
BuzzFace	Yes	-	-	Yes	-	-	Yes
FacebookHoax	Yes	-	Yes	Yes	-	-	-
FakeNewsNet	Yes	Yes	Yes	Yes	Yes	Yes	Yes

- **BuzzFeedNews**[22]: The data set is based upon the news posted by 9 news agencies during the US presidential elections collected from Facebook. It comprises of 1627 news articles whose authenticity was verified by five journalists from BuzzFeed. The data set is a mix of 356 left wing, 545 right wing and 826 mainstream articles.
- **LIAR**[23]: The data set is obtained from PolitiFact[24] using an API (Wang, 2017) and consists of 12,836 articles from various fields. The classification has been done upon categories such as “pants-fire”, “false”, “barely-true”, “half-true”, “mostly true” and “true”.
- **BS Detector**[25]: The data set is generated by self-annotating feature of the toolkit which crawls all the links in a given web-page and ascertains unreliable sources by comparing them against a domain-list provided beforehand.
- **Fake News Challenge (FNC)**[26]: The data set consists of headlines or content of news articles and aims to label the stance associated with it into categories such as “agree”, “disagree”, “discuss” and “unrelated”. While the train set consists of 49,972 records, the test set consists of 25,413 records. The domain of the data set includes various topics like politics, health, environment, lifestyle and so on.
- **CREDBANK**[27]: The data set comprises of 60 million tweets from October, 2015 over a period of 96 days. All tweets have been segregated and attributed to over 1000 events and the credibility of each event verified by associates from Amazon-Mechanical-Turk (Mitra & Gilbert, 2015).
- **BuzzFace**[28]: The data set is an extension of BuzzFeed with 2263 articles along with 1.6 million comments (Santia & Williams, 2018).
- **FacebookHoax**[29]: The data set comprises of 15,500 posts spread over 32 pages related to scientific and conspiracy theories with over 2 million likes (Tacchini et al., 2017).
- **FakeNewsNet**[30]: It is a multi-dimensional data set containing various categories of information like news-context, social-context, spatial and temporal information. The data has been taken from repositories like PolitiFact and GossipCop.[31]

6. Case study: COVID-19 related fake news

In the preceding sections, through an extensive survey, an insight into the research advancements towards understanding and curtailing the generation and dissemination of fake news has been provided. This has been accompanied by a list of annotated data sets which may serve as ground truth for devising a fake news classifier. However, an unresolved point to ponder is to determine the factors behind the proliferation of fake news and associated psychological implications. Therefore, to strengthen the understanding of the readers, a real-life case study upon the proliferation and measures adopted to combat fake news related to COVID-19 has been presented in this section. Here, we study the contribution of researchers and journalists followed by an analysis of tweets related to the pandemic wherein the variation in the significant and predominant terms attributed to the pandemic have been analyzed and efforts have been made to assess the credibility of users based on the share of factual and emotional tweets by them.

The emergence of the COVID-19 pandemic has affected the lifestyle of people with over 149 million cases and over 3 million deaths all over the globe (WHO, 2021). With most of the people confined in their homes, social media has experienced a hike in the activity of its users by about 25% (Yang, Torres-Lugo, & Menczer, 2020a). The rising concerns related to the pandemic have led to people seeking information related to COVID-19 on social media and sharing it among their peers (Huynh et al., 2020). All these factors have led to a fertile ground for the creation and propagation of fake news (Frenkel, Alba, & Zhong, 2020). Some fake news instances include false remedies such as salt-water gargle, drinking bleach and eating oregano (Apuke & Omar, 2021), conspiracy theories that COVID-19 has been bio-engineered by the Chinese (Cohen, 2020), purported claims that the 5G technology is aggravating the pandemic (van der Linden, Roozenbeek, & Compton, 2020). With the second wave of COVID-19 in India, fake news related to the pandemic has been widely circulated such as fake posts that the police in Delhi, India is restraining people from arranging help for COVID-19 patients[32], misleading claims regarding the effectiveness of COVID-19 vaccines stated with CT scan of lungs[33], a viral video claiming that surgical masks contain worms[34] and so on.

6.1. Contribution of researchers and journalists

To counter this menace of fake news in such a grave situation, various works have been carried out to analyze the dissemination of fake news and devise approaches for its detection. Moscadelli et al. (2020) reviewed news related to COVID-19 circulated in Italy and found that fake news was circulated over 2 million times which was over 23% of the total number of articles included in their study. Apuke & Omar (2021) performed a study on the motivation behind the dissemination of fake news related to COVID-19 in Nigeria using Partial-Least-Squares (PLS) method. They found that altruism was the principal motivation behind sharing fake news apart from motivations like sharing information, socializing, seeking information and utilizing spare time. van der Linden

[22] https://github.com/BuzzFeedNews/2016-10-facebook-factcheck/tree/master/data
[23] https://www.cs.ucsb.edu/william/software.html
[24] https://www.politifact.com/
[25] https://github.com/bs-detector/bs-detector
[26] https://github.com/FakeNewsChallenge/fnc-1
[27] http://compsocial.github.io/CREDBANK-data/
[28] https://github.com/gsantia/BuzzFace
[29] https://github.com/gabll/some-like-it-hoax
[30] https://github.com/KaiDMML/FakeNewsNet
[31] https://www.gossipcop.com/
[32] https://www.oneindia.com/fact-check/delhi-fake-warning-to-online-covid-volunteers-causes-panic-3250102.html
[33] https://www.altnews.in/ct-scan-of-lungs-of-covid-patient-viral-with-misleading-vaccine-claim/
[34] https://www.altnews.in/no-your-surgical-masks-dont-have-worms/

et al. (2020) proposed inoculation-theory in the context of resisting oneself from persuasion. The authors proposed that just like a virus is incapacitated through inoculation, people can be psychologically inoculated against fake news related to COVID-19 by making them aware that fake news related to COVID-19 is being spread with malicious intentions and simultaneously exposing them to facts and information refuting false-claims. Paka et al. (2021) introduced a data set with approximately 45,000 annotated real and fake microblogs and about 21 million unannotated microblogs related to COVID-19. They have even proposed a cross-stitch mechanism in a semi-supervised attention-based neural model through which they claim to address the problem of paucity of training data and improving the detection ability when dealing with instances differing from the training data. Karami, Bookstaver, Nolan, & Bozorgi (2021) applied topic modeling and frequency-wise analysis of frequently referred diseases and chemicals in papers related to COVID-19. Shahi, Dirkson, & Majchrzak (2021) perform an exploratory analysis based on the content, profiles and propagation of misinformation related to COVID-19. In their study, they highlight the discrepancy between the pace and volume of misinformation being spread and the rate of performing data analysis. Choudrie, Patil, Kotecha, Matta, & Pappas (2021) performed month-wise as well as country-wise emotion analysis on tweets related to COVID-19 using Robustly Optimized BERT Pretraining Approach (RoBERTa). From these works, one can assess the contribution of research works to comprehend the proliferation of fake news, the motivation behind dissemination and efforts to safeguard people from this menace. However, a major hurdle towards combating fake news using supervised approaches is the unavailability of annotated data sets for unforeseen catastrophes to train the models. This paves the way to the need to devise unsupervised or semi-supervised approaches to deal with such situations. Therefore, an unsupervised approach has been adopted by us as elucidated in Section 6.2 which is domain-independent and suitable for tasks with unannotated data to gauge the credibility of profiles.

Accompanied by the researchers, journalists play a vital role in combating fake-news by fact-checking the news in circulation and identifying fake news from them. They inoculate their audiences against fake news by refuting the misinformation and exposing the truth behind it (van der Linden et al., 2020). Some of the fact-checking websites involved in the task of enlightening people with the truth are Snopes, PolitiFact,[35], Alt News[36], India Today Fact Check[37] which is a part of the India Today News Group[38], WebQoof[39] is a fact-checking website certified by IFCN provided by The Quint[40] and so on. Figure 5 presents a pictorial representation of instances of misinformation related to COVID-19 being fact-checked through the fact-checking websites. Figure 5a shows the screenshot of a purported video on WHO's reversal of COVID-19 guidelines which was tagged as fake by Alt News[41]. Figure 5b is from a widely circulated video having its origin in Mexico being shared as "fake vaccination-drive" in India[42] fact-checked by India Today Fact Check. Figure 5c depicts the false-claims being refuted by India Today Fact Check regarding viral breathing tests which claim that if a person can hold their breath for 20 seconds then they have not been infected with COVID-19.[43] The contribution by the journalists and the ongoing research to detect and mitigate the spread of fake news show a ray of hope that although misinformation related to the pandemic is on the rise, the researchers and professionals in the fields of data science, artificial intelligence, journalism, information science, psychology and social science are on their front foot to save people distraught due to the pandemic from further anxiety due to fake news.

6.2. Analysis of tweets related to COVID-19

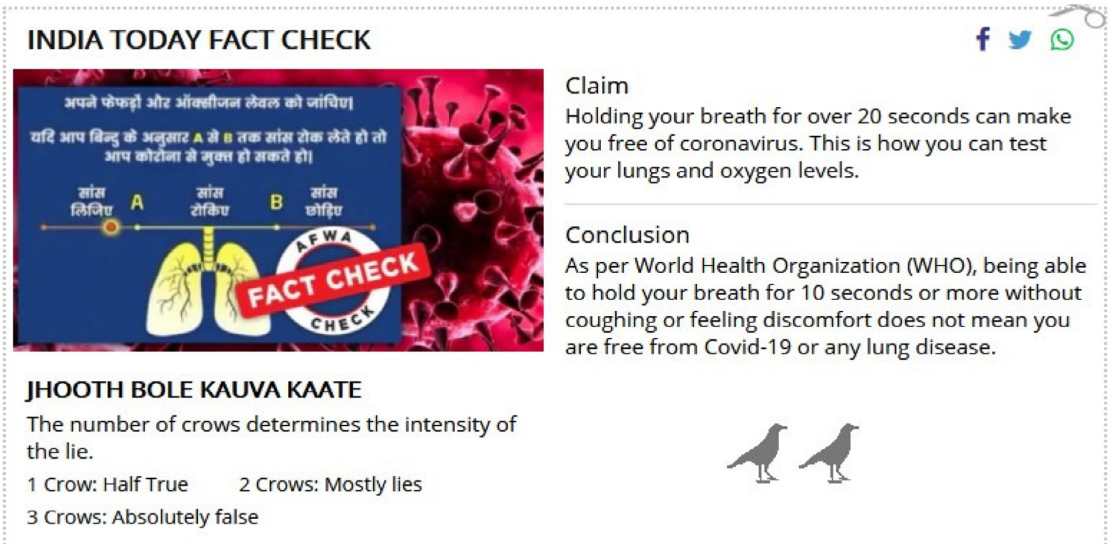
After understanding the steps taken by the researchers and journalists to counter fake news during the pandemic, in this section we direct our efforts towards analyzing the tweets related to the pandemic. Through this, an insight into the psychology of the people has been obtained and the relationship between the credibility of users and the share of factual versus emotional tweets by them has been figured out. This methodology is unsupervised, domain-independent and can be applied to tasks with unannotated data.

- **Task Definition:** The approach adopted for analysis of the tweets has been outlined in Fig. 6. The first step consisted of preparing the tweets corpus by crawling tweets. The tweets were then subjected to cleaning operations such as removal of URLs, punctuations, non-ASCII characters, line breaks, carriage returns and standardizing additional white spaces. After that, tokenization, lemmatization and stop word removal have been performed. Followed by preprocessing, the analysis process has been divided into three phases. Firstly, the predominant terms attributed to the pandemic have been visualized through a word cloud representation. Secondly, the weekly variation in the frequency of significant terms related to the pandemic have been plotted and the observations noted. Lastly, the relationship among the shares of the factual and emotional tweets has been examined to obtain an indicator of the credibility of the profiles and the tweets.
- **Description of the Data Set:** The data set was prepared by crawling tweets from Twitter[44] containing keywords like "#Covid_19", "#coronavirusindia", "#CoronaChainScare", "#StayHomeStaySafe", "#CoronavirusPandemic", "#CoronaVirusUpdate". The corpus consisted of 41,402 tweets which were crawled from 27th February, 2020 to 30th March, 2020. The data set consisted of the following fields of information- user ID, tweet content, creation date, screen name, retweeted or not and retweet count. The week-wise distribution of tweets have been depicted in Fig. 7.
- **Word Cloud Representation:** Fig. 8 illustrates the word cloud representation of all the tweets in the corpus. From Fig. 8, it can be deduced that the most frequent words apart from "covid19" and "coronavirus" are "jantacurfewmarch22", "lockdown", "pandemic", "home", "death", "hospital", "quarantine", "distancing", "mask", "government" and so on. It is worthwhile to mention that a new term "covidiot" has also surfaced through social media platforms. The definition of this term can be conceived as someone who ignores the warnings regarding public health or safety. Another definition can be a person who hoards goods, denying them from their neighbors. For example- "Did you see that covidiot with 300 rolls of toilet paper in his basket?", "That covidiot is hugging everyone he sees.".
- **Frequency variation of Terms:** In order to grasp the needs and apprehensions of people an analysis has been performed upon the frequency variation of certain relevant terms attributed to COVID-19. The week-wise frequency variation of some relevant terms has been presented in Fig. 9 where the week number corresponds to the week of the year. From Fig. 9, it can be inferred that the most frequent terms are "lockdown", "quarantine"., "mask", "doctor" and "isolation". The frequency of the term "covidiot" was highest during week 10 but declined drastically over the following weeks while for all other terms, the overall trend is almost similar.

[35] https://www.factcheck.org/
[36] https://www.altnews.in/
[37] https://www.indiatoday.in/fact-check
[38] https://www.indiatoday.in/
[39] https://www.thequint.com/news/webqoof
[40] https://www.thequint.com/
[41] https://www.altnews.in/world-doctors-alliance-video-promoting-covid-conspiracies-viral-again/
[42] https://www.indiatoday.in/fact-check/video/video-from-mexico-shared-fake-vaccination-drive-india-1796249-2021-04-29
[43] https://www.indiatoday.in/fact-check/story/fact-check-this-viral-breathing-test-to-diagnose-covid-19-is-a-hoax-1795248-2021-04-26
[44] https://twitter.com/

(a) Fake news on WHO's reversal of COVID-19 guidelines

(b) Misleading video on fake vaccination drive in India



(c) Viral breathing test for COVID-19 diagnosis

Fig. 5. Pictorial representation of instances of misinformation related to COVID-19 being fact-checked.

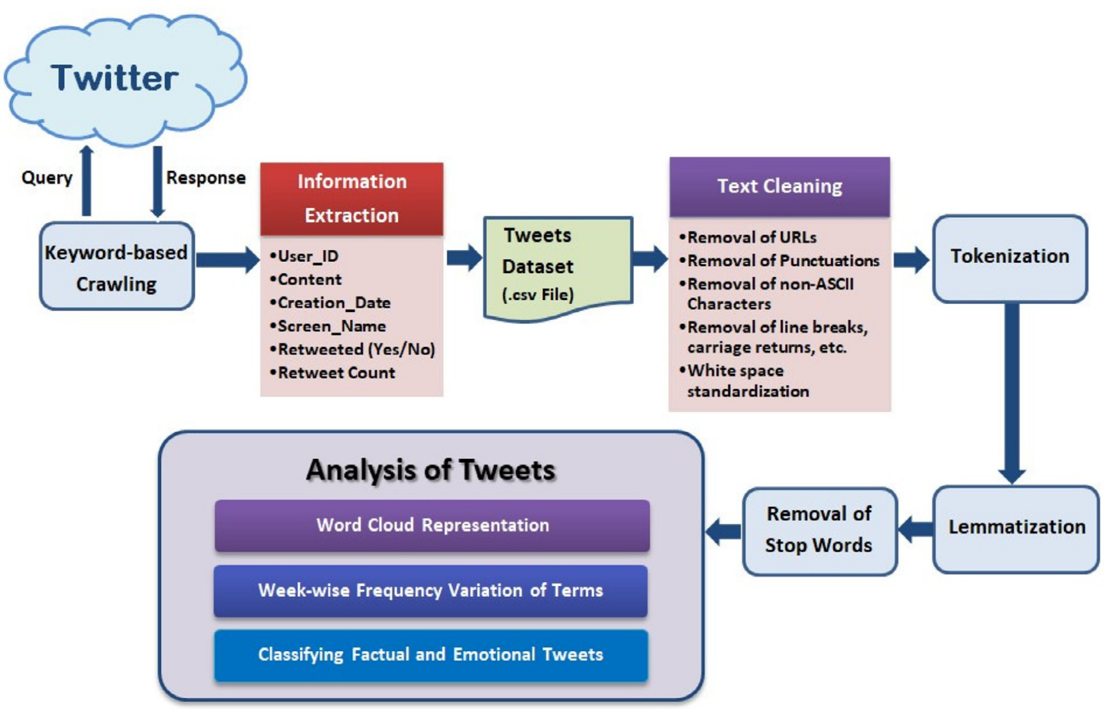


Fig. 6. Flow diagram for analysis of tweets related to COVID-19.

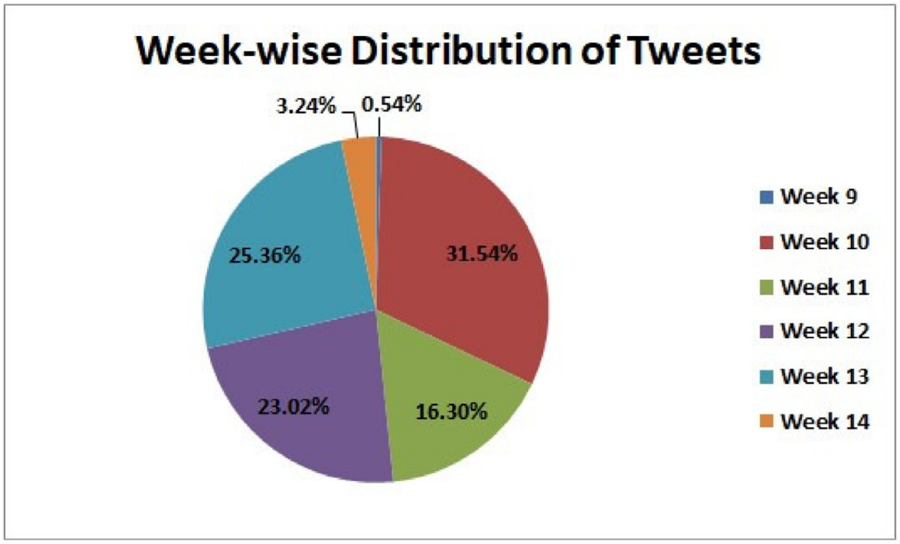
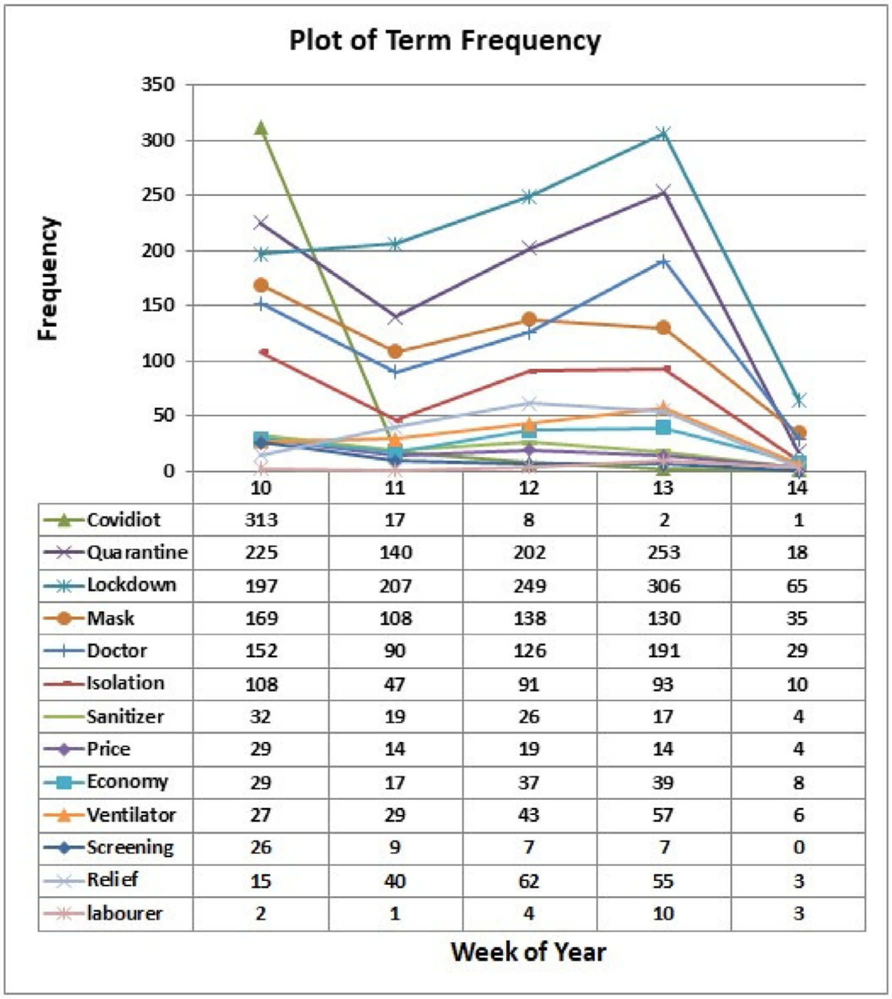


Fig. 7. Week-wise Distribution of Tweets.

Fig. 8. Word Cloud Representation of Tweets.

- **Factual versus Emotional Tweets:** In our study, we aimed to distinguish between the factual tweets and the tweets driven by emotions. For this, we deployed a lexicon-based classifier built upon a lexicon of terms annotated depending upon their orientations. The week-wise variation in the percentage of factual and emotional tweets have been presented in Fig. 10. In these plots, credible profiles such as "@WHO", "@PMOIndia", "@BorisJohnson", "@realDonaldTrump", "@CDCgov" and "@RBI" have been taken so that the authenticity can be ascertained. To aid in comprehension, call-outs along the time-line have been added to denote the significant events during the period. From our study, it can be inferred that most of the factual tweets can be attributed to credible profiles. While majority of the emotional tweets belonged to profiles whose credibility could not be ascertained. Therefore, the trade-off between the factual and emotional tweets may provide an indication of the credibility of content.



	10	11	12	13	14
Covidiot	313	17	8	2	1
Quarantine	225	140	202	253	18
Lockdown	197	207	249	306	65
Mask	169	108	138	130	35
Doctor	152	90	126	191	29
Isolation	108	47	91	93	10
Sanitizer	32	19	26	17	4
Price	29	14	19	14	4
Economy	29	17	37	39	8
Ventilator	27	29	43	57	6
Screening	26	9	7	7	0
Relief	15	40	62	55	3
labourer	2	1	4	10	3

Fig. 9. Week-wise Frequency Variation of Terms Attributed to COVID-19.

7. Discussion

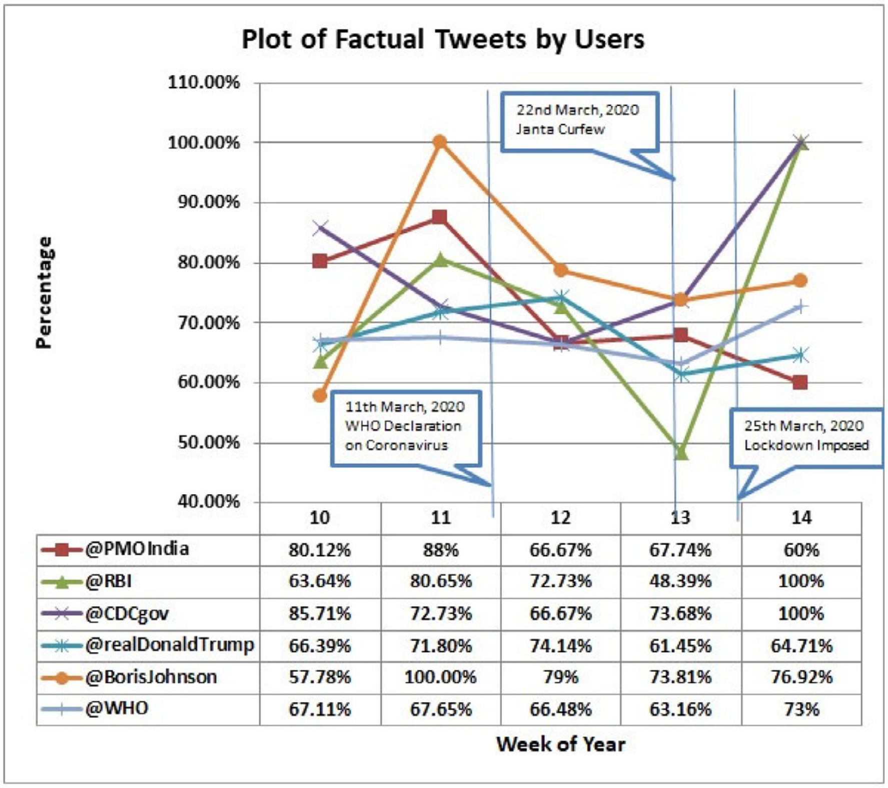
7.1. Implications of this study

The extent to which fake news adversely affects our society is a matter of concern and measures need to be taken to curb this phenomenon. This paper presents an insight into the research advancements towards understanding and curtailing the generation and dissemination of fake news through an extensive survey. Specifically, the theoretical and the practical implications of our study are as follows:

- **Theoretical Implications:** This paper provides an insight into fake news and all the terms associated with it based on the objective, intention and type of information conveyed. It sheds light upon the psychological vulnerabilities, societal factors along with the role of social media behind the dissemination of fake news. Furthermore, this paper explores various aspects like news content, social context and spatiotemporal information based on which fake news can be distinguished. This has been accompanied by an extensive study on the previous works for the detection of fake news exploiting these inherent aspects. Through this survey, it can be observed that most of the techniques to mitigate fake news follow a supervised approach which restricts their application in unforeseen domains. In this paper, the concepts have been presented through a detailed survey of existing works supported with appropriate citations to help develop a sound understanding of fake news among its readers.
- **Practical Implications:** The data sets, accompanied by the machine learning and deep learning approaches surveyed in this paper, may serve as a stepping stone to the readers who desire to develop their classifiers and take this study forward. The research works on COVID-19 help to discover the factors behind the propagation of fake news, the psychological impacts of fake news and the approaches that have been devised to detect fake news and mitigate its spread. Furthermore, the analysis of tweets helps to comprehend the needs of the people and the predominant terms associated with the pandemic. Also, from analyzing the share of factual and emotional tweets of the users, a trade-off can be determined to assess the credibility of the profiles. As the methods applied to perform the analysis of tweets are unsupervised and domain-independent, they can be applied to detect misinformation and the factors behind it during the initial stages of any natural disaster, epidemic, political or social event when the availability of annotated data sets related to it is scarce.

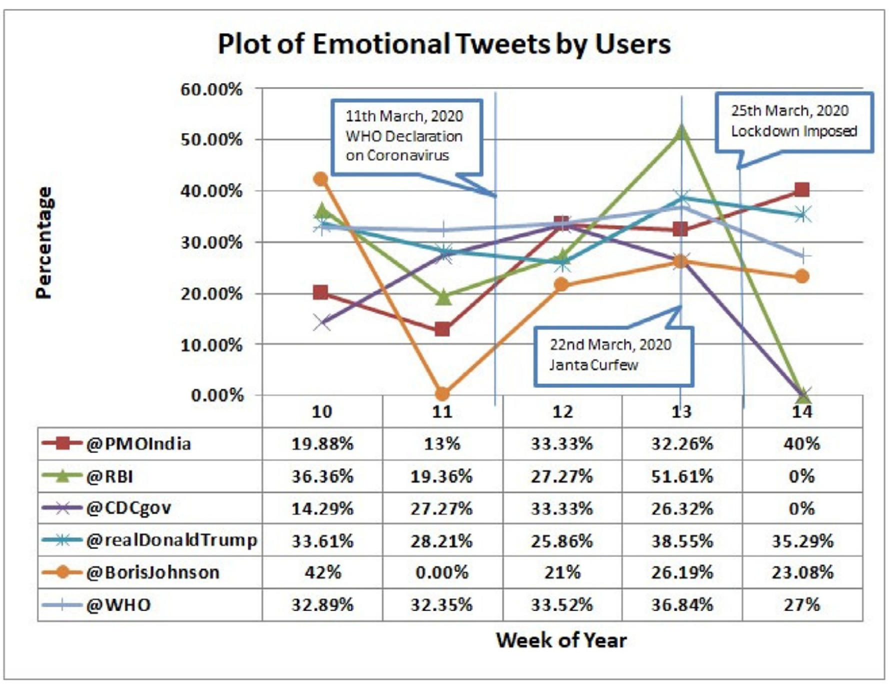
7.2. Limitations of this study

Apart from the accomplishments as stated above, this survey has some limitations too. There might be a selection bias as only articles written in English have been surveyed since it is one of the most popular languages across the globe and the authors of this paper are well-conversant in it. As the volume and the pace at which content is generated and propagated in the social media largely outnumber the rate of technological advancements and measures to detect and mitigate fake news, a single foolproof solution to all fake news in circulation could not be observed. However, we cover the multiple facets of fake news char-



	10	11	12	13	14
@PMOIndia	80.12%	88%	66.67%	67.74%	60%
@RBI	63.64%	80.65%	72.73%	48.39%	100%
@CDCgov	85.71%	72.73%	66.67%	73.68%	100%
@realDonaldTrump	66.39%	71.80%	74.14%	61.45%	64.71%
@BorisJohnson	57.78%	100.00%	79%	73.81%	76.92%
@WHO	67.11%	67.65%	66.48%	63.16%	73%

(a) Week-wise Variation of Factual Tweets



	10	11	12	13	14
@PMOIndia	19.88%	13%	33.33%	32.26%	40%
@RBI	36.36%	19.36%	27.27%	51.61%	0%
@CDCgov	14.29%	27.27%	33.33%	26.32%	0%
@realDonaldTrump	33.61%	28.21%	25.86%	38.55%	35.29%
@BorisJohnson	42%	0.00%	21%	26.19%	23.08%
@WHO	32.89%	32.35%	33.52%	36.84%	27%

(b) Week-wise Variation of Emotional Tweets

Fig. 10. Week-wise variation in the percentage of factual and emotional tweets.

acterization and detection spanning across multiple disciplines which may serve as a reference to devise more robust and accurate approaches in the future. Majority of the existing annotated data sets in our study focus primarily on text-based information with very little or no information about attributes related to visual information, network information or spatiotemporal information. Future works may be directed to build a comprehensive data set including all attributes related to news-content, social-context and spatiotemporal information. In Section 6.2, while analyzing the tweets, we acknowledge the limitation on the number of tweets extracted at a time imposed by Twitter and also the restriction on obtaining tweets older than seven days.[45] Also, the location information of the tweets has been excluded from our study due to the inability to track precise location information of all the tweets due to privacy issues.[46] It can be perceived that none of the limitations attenuate the novelty and the significance of this paper. Besides, the limitations usher new avenues for extending this work in the future.

8. Conclusion and future scope

In the present scenario, with an increasing number of people becoming consumers of social media, the spread of information is increasing with every passing day. This makes people all the more vulnerable to fake news. To envision a society where the truth will prevail, it is essential to understand the impact of fake news on society and undertake measures to combat the menace of fake news. In this paper, a survey of works aimed at restraining the creation and propagation of fake news has been performed from a data science perspective using a combination of artificial intelligence approaches with statistics and domain-based knowledge. In order to provide a coherent understanding, various aspects of fake news such as the definition of fake news, categories of fake news, characterization of fake news, extracting features from news and detection of fake news statistical, machine learning and deep learning approaches have been discussed. From this survey, it can be inferred that fake news detection is a multi-disciplinary task that required knowledge of multiple disciplines such as data science, journalism, information science, psychology and social science. Followed by it, a study of the existing fake news detection repositories has been conducted based upon various parameters. However, the application of such data sets is found to be limited to supervised approaches based on a particular domain. Also, capturing additional information such as profile details, network information or spatiotemporal information may aid in improving the accuracy of detection of fake content. To aid in understanding fake news and its implications, a real-life case study on fake news related to COVID-19 has been presented. The research works on COVID-19 enunciate how the situation provided a fertile ground for the proliferation of fake news and propose measures to impede the spread of misinformation. From the analysis of tweets on COVID-19, the word cloud representation and frequency variation of significant terms attributed to the pandemic provide an understanding of the gravity and vulnerability of the situation while the trade-off between the factual and emotional tweets imparts an indication to the credibility.

This survey is aimed at researchers and professionals in the fields of data science, artificial intelligence, journalism, information science, psychology and social science who desire to develop an in-depth understanding of fake news and devise measures to combat the generation and propagation of fake news. It can also be useful to students who use social media in their day-to-day lives to understand how fake information can be spread online and not become victims or further propagators of fake news.

In the future, this survey can be extended to various related domains. For instance, detecting “clickbaits”, i.e. using catchy text to attract users to click upon links to fraudulent or irrelevant websites. This discrepancy between news content and headlines can be applied to fake news verification. Another field of interest is to detect spam in social networks propagated through individuals, group of users or social bots.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[45] https://developer.twitter.com/en/docs/twitter-api/rate-limits
[46] https://help.twitter.com/en/safety-and-security/tweet-location-settings

Acknowledgements

None. The author(s) received no financial or any other kinds of support for the research, authorship, and/or publication of this article.

Appendix A. Shortlisted Papers for the Survey

1. Agrawal, R. and Sharma, D.K., 2021, March. A Survey on Video-Based Fake News Detection Techniques. In 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 663–669). IEEE.
2. Allcott, H. and Gentzkow, M., 2017. Social media and fake news in the 2016 election. Journal of economic perspectives, 31(2), pp. 211–236.
3. Allport, F.H., 1937. Toward a science of public opinion. Public opinion quarterly, 1(1), pp. 7–23.
4. Apuke, O.D. and Omar, B., 2021. Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. Telematics and Informatics, 56, p.101475.
5. Baruah, A., Das, K.A., Barbhuiya, F.A. and Dey, K., 2020. Automatic Detection of Fake News Spreaders Using BERT. In CLEF (Working Notes).
6. Berkowitz, D. and Schwartz, D.A., 2016. Miley, CNN and The Onion: When fake news becomes realer than real. Journalism practice, 10(1), pp. 1–17.
7. Bleakley, P., 2021. Panic, pizza and mainstreaming the alt-right: A social media analysis of Pizzagate and the rise of the QAnon conspiracy. Current Sociology, p.00113921211034896.
8. Butler, A., 2018. Protecting the democratic role of the press: A legal solution to fake news. Wash. UL Rev., 96, p.419.
9. Castillo, C., Mendoza, M. and Poblete, B., 2013. Predicting information credibility in time-sensitive social media. Internet Research.
10. Chen, C., Wu, K., Srinivasan, V. and Zhang, X., 2013, August. Battling the internet water army: Detection of hidden paid posters. In 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013) (pp. 116–120). IEEE.
11. Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C. and Leskovec, J., 2017, February. Anyone can become a troll: Causes of trolling behavior in online discussions. In Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing (pp. 1217–1230).
12. Choi, D., Chun, S., Oh, H. and Han, J., 2020. Rumor propagation is amplified by echo chambers in social media. Scientific reports, 10(1), pp. 1–10.
13. Choudrie, J., Patil, S., Kotecha, K., Matta, N. and Pappas, I., 2021. Applying and understanding an advanced, novel deep learning approach: A Covid 19, text based, emotions analysis study. Information Systems Frontiers, pp. 1–35.
14. Ciampaglia, G.L., Shiralkar, P., Rocha, L.M., Bollen, J., Menczer, F. and Flammini, A., 2015. Computational fact checking from knowledge networks. PloS one, 10(6), p.e0128193.
15. Chu, Z., Gianvecchio, S., Wang, H. and Jajodia, S., 2012. Detecting automation of twitter accounts: Are you a human, bot, or cyborg?. IEEE Transactions on dependable and secure computing, 9(6), pp. 811–824.
16. Cohen, J., 2020. Scientists ‘strongly condemn’rumors and conspiracy theories about origin of coronavirus outbreak; Science; AAAS.
17. Deepak, S. and Chitturi, B., 2020. Deep neural approach to Fake-News identification. Procedia Computer Science, 167, pp. 2236–2243.
18. Duffy, A., Tandoc, E. and Ling, R., 2020. Too good to be true, too good not to share: the social utility of fake news. Information, Communication & Society, 23(13), pp. 1965–1979.
19. Egelhofer, J.L. and Lecheler, S., 2019. Fake news as a two-dimensional phenomenon: A framework and research agenda. An-

nals of the International Communication Association, 43(2), pp. 97–116.
20. Fader, A., Soderland, S. and Etzioni, O., 2011, July. Identifying relations for open information extraction. In Proceedings of the 2011 conference on empirical methods in natural language processing (pp. 1535–1545).
21. Faustini, P.H.A. and Covoes, T.F., 2020. Fake news detection in multiple platforms and languages. Expert Systems with Applications, 158, p.113503.
22. Feng, S., Banerjee, R. and Choi, Y., 2012, July. Syntactic stylometry for deception detection. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) (pp. 171–175).
23. Ferrara, E., Varol, O., Davis, C., Menczer, F. and Flammini, A., 2016. The rise of social bots. Communications of the ACM, 59(7), pp. 96–104.
24. Fisher, R.J., 1993. Social desirability bias and the validity of indirect questioning. Journal of consumer research, 20(2), pp. 303–315.
25. Frenda, S.J., Knowles, E.D., Saletan, W. and Loftus, E.F., 2013. False memories of fabricated political events. Journal of Experimental Social Psychology, 49(2), pp. 280–286.
26. Granik, M. and Mesyura, V., 2017, May. Fake news detection using naive Bayes classifier. In 2017 IEEE first Ukraine conference on electrical and computer engineering (UKRCON) (pp. 900–903). IEEE.
27. Gravanis, G., Vakali, A., Diamantaras, K. and Karadais, P., 2019. Behind the cues: A benchmarking study for fake news detection. Expert Systems with Applications, 128, pp. 201–213.
28. Hassan, N., Li, C. and Tremayne, M., 2015, October. Detecting check-worthy factual claims in presidential debates. In Proceedings of the 24th acm international on conference on information and knowledge management (pp. 1835–1838).
29. Highet, G., 2015. Anatomy of satire. Princeton University Press.
30. Huynh, T.L., 2020. The COVID-19 risk perception: A survey on socioeconomics and media attention. Econ. Bull, 40(1), pp. 758–764.
31. Ireton, C. and Posetti, J., 2018. Journalism, fake news & disinformation: handbook for journalism education and training. Unesco Publishing.
32. Jamieson, K.H. and Cappella, J.N., 2008. Echo chamber: Rush Limbaugh and the conservative media establishment. Oxford University Press.
33. Jin, Z., Cao, J., Jiang, Y.G. and Zhang, Y., 2014, December. News credibility evaluation on microblog with a hierarchical propagation model. In 2014 IEEE International Conference on Data Mining (pp. 230–239). IEEE.
34. Jin, Z., Cao, J., Zhang, Y. and Luo, J., 2016, March. News verification by exploiting conflicting social viewpoints in microblogs. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 30, No. 1).
35. Jin, Z., Cao, J., Zhang, Y., Zhou, J. and Tian, Q., 2016. Novel visual and statistical image features for microblogs news verification. IEEE transactions on multimedia, 19(3), pp. 598–608.
36. Jwa, H., Oh, D., Park, K., Kang, J.M. and Lim, H., 2019. exbake: Automatic fake news detection model based on bidirectional encoder representations from transformers (bert). Applied Sciences, 9(19), p.4062.
37. Kang, H., Bae, K., Zhang, S. and Sundar, S.S., 2011. Source cues in online news: Is the proximate source more powerful than distal sources?. Journalism & Mass Communication Quarterly, 88(4), pp. 719–736.
38. Karami, A., Bookstaver, B., Nolan, M. and Bozorgi, P., 2021. Investigating diseases and chemicals in COVID-19 literature with text mining. International Journal of Information Management Data Insights, 1(2), p.100016.
39. Keeble, R., 2001. Ethics for journalists (p. 49). London: Routledge.
40. Kumar, S., West, R. and Leskovec, J., 2016, April. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In Proceedings of the 25th international conference on World Wide Web (pp. 591–602).
41. Kwon, S., Cha, M., Jung, K., Chen, W. and Wang, Y., 2013, December. Prominent features of rumor propagation in online social media. In 2013 IEEE 13th international conference on data mining (pp. 1103–1108). IEEE.
42. Lesce, T., 1990. Scan: Deception detection by scientific content analysis. Law and Order, 38(8), pp. 3–6.
43. van Der Linden, S., Roozenbeek, J. and Compton, J., 2020. Inoculating against fake news about COVID-19. Frontiers in psychology, 11, p.2928.
44. Long, Y., 2017. Fake news detection through multi-perspective speaker profiles. Association for Computational Linguistics.
45. Ma, J., Gao, W., Wei, Z., Lu, Y. and Wong, K.F., 2015, October. Detect rumors using time series of social context information on microblogging websites. In Proceedings of the 24th ACM international on conference on information and knowledge management (pp. 1751–1754).
46. Magdy, A. and Wanas, N., 2010, October. Web-based statistical fact checking of textual documents. In Proceedings of the 2nd international workshop on Search and mining user-generated contents (pp. 103–110).
47. Maronikolakis, A., Villegas, D.S., Preotiuc-Pietro, D. and Aletras, N., 2020. Analyzing political parody in social media. arXiv preprint arXiv:2004.13878.
48. McManus, J.H., 1992. What kind of commodity is news. Communication research, 19(6), pp.787–805.
49. Mitra, T. and Gilbert, E., 2015, April. Credbank: A large-scale social media corpus with associated credibility annotations. In Ninth international AAAI conference on web and social media.
50. Mohammad, S.M., Sobhani, P. and Kiritchenko, S., 2017. Stance and sentiment in tweets. ACM Transactions on Internet Technology (TOIT), 17(3), pp. 1–23.
51. Moscadelli, A., Albora, G., Biamonte, M.A., Giorgetti, D., Innocenzio, M., Paoli, S., Lorini, C., Bonanni, P. and Bonaccorsi, G., 2020. Fake news and covid-19 in Italy: results of a quantitative observational study. International journal of environmental research and public health, 17(16), p.5850.
52. Mukherjee, A., Venkataraman, V., Liu, B. and Glance, N., 2013, June. What yelp fake review filter might be doing?. In Seventh international AAAI conference on weblogs and social media.
53. Nasir, J.A., Khan, O.S. and Varlamis, I., 2021. Fake news detection: A hybrid CNN-RNN based deep learning approach. International Journal of Information Management Data Insights, 1(1), p.100007.
54. Nelson, M.R. and Park, J., 2015. Publicity as covert marketing? The role of persuasion knowledge and ethical perceptions on beliefs and credibility in a video news release story. Journal of Business Ethics, 130(2), pp. 327–341.
55. Nickerson, R.S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2(2), pp. 175–220.
56. Paka, W.S., Bansal, R., Kaushik, A., Sengupta, S. and Chakraborty, T., 2021. Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection. Applied Soft Computing, 107, p.107393.
57. Paul, C. and Matthews, M., 2016. The Russian “firehose of falsehood” propaganda model. Rand Corporation, pp. 2–7.
58. Polignano, M., de Pinto, M.G., Lops, P. and Semeraro, G., 2019, September. Identification Of Bot Accounts In Twitter Using 2D CNNs On User-generated Contents. In CLEF (Working Notes).
59. Prasetijo, A.B., Isnanto, R.R., Eridani, D., Soetrisno, Y.A.A., Arfan, M. and Sofwan, A., 2017, October. Hoax detection system on Indonesian news sites based on text classification using SVM and SGD. In 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) (pp. 45–49). IEEE.

60. Pronin, E., Kruger, J., Savtisky, K. and Ross, L., 2001. You don't know me, but I know you: The illusion of asymmetric insight. Journal of Personality and Social Psychology, 81(4), p.639.
61. Quattrociocchi, W., Scala, A. and Sunstein, C.R., 2016. Echo chambers on Facebook. Available at SSRN 2795110.
62. Rangel, F. and Rosso, P., 2019. Overview of the 7th author profiling task at PAN 2019: bots and gender profiling in twitter. In Proceedings of the CEUR Workshop, Lugano, Switzerland (pp. 1–36).
63. Rashkin, H., Choi, E., Jang, J.Y., Volkova, S. and Choi, Y., 2017, September. Truth of varying shades: Analyzing language in fake news and political fact-checking. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 2931–2937).
64. Reaves, S., 1995. The vulnerable image: Categories of photos as predictor of digital manipulation. Journalism & Mass Communication Quarterly, 72(3), pp.706–715.
65. Robinson, R.J., Keltner, D., Ward, A. and Ross, L., 1995. Actual versus assumed differences in construal: "Naive realism" in intergroup perception and conflict. Journal of personality and social psychology, 68(3), p. 404.
66. Rosenstiel, T., 2014. The elements of journalism: What newspeople should know and the public should expect. Three Rivers Press.
67. Roy, A., Basak, K., Ekbal, A. and Bhattacharyya, P., 2018. A deep ensemble framework for fake news detection and classification. arXiv preprint arXiv:1811.04670.
68. Rubin, V.L., Conroy, N., Chen, Y. and Cornwell, S., 2016, June. Fake news or truth? using satirical cues to detect potentially misleading news. In Proceedings of the second workshop on computational approaches to deception detection (pp. 7–17).
69. Rubin, V.L. and Lukoianova, T., 2015. Truth and deception at the rhetorical structure level. Journal of the Association for Information Science and Technology, 66(5), pp. 905–917.
70. Ruchansky, N., Seo, S. and Liu, Y., 2017, November. Csi: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (pp. 797–806).
71. Santia, G.C. and Williams, J.R., 2018, June. Buzzface: A news veracity dataset with facebook user commentary and egos. In Twelfth International AAAI Conference on Web and Social Media.
72. Santoso, I., Yohansen, I., Warnars, H.L.H.S. and Hashimoto, K., 2017, November. Early investigation of proposed hoax detection for decreasing hoax in social media. In 2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom) (pp. 175–179). IEEE.
73. Shahi, G.K., Dirkson, A. and Majchrzak, T.A., 2021. An exploratory study of covid-19 misinformation on twitter. Online social networks and media, 22, p.100104.
74. Shi, B. and Weninger, T., 2016, April. Fact checking in heterogeneous information networks. In Proceedings of the 25th International Conference Companion on World Wide Web (pp. 101–102).
75. Shoemaker, P.J. and Reese, S.D., 2013. Mediating the message in the 21st century: A media sociology perspective. Routledge.
76. Shrivastava, G., Kumar, P., Ojha, R.P., Srivastava, P.K., Mohan, S. and Srivastava, G., 2020. Defensive modeling of fake news through online social networks. IEEE Transactions on Computational Social Systems, 7(5), pp. 1159–1167.
77. Shu, K., Mahudeswaran, D., Wang, S., Lee, D. and Liu, H., 2020. Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big data, 8(3), pp. 171–188.
78. Shu, K., Sliva, A., Wang, S., Tang, J. and Liu, H., 2017. Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter, 19(1), pp. 22–36.
79. Shu, K., Wang, S. and Liu, H., 2018, April. Understanding user profiles on social media for fake news detection. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 430–435). IEEE.
80. Sinclair, C., 2020. Parody: fake news, regeneration and education. Postdigital Science and Education, 2(1), pp.61–77.
81. Song, C., Ning, N., Zhang, Y. and Wu, B., 2021. A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. Information Processing & Management, 58(1), p.102437.
82. Tacchini, E., Ballarin, G., Della Vedova, M.L., Moret, S. and de Alfaro, L., 2017. Some like it hoax: Automated fake news detection in social networks. arXiv preprint arXiv:1704.07506.
83. Talwar, S., Dhir, A., Kaur, P., Zafar, N. and Alrasheedy, M., 2019. Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. Journal of Retailing and Consumer Services, 51, pp. 72–82.
84. Thota, A., Tilak, P., Ahluwalia, S. and Lohia, N., 2018. Fake news detection: a deep learning approach. SMU Data Science Review, 1(3), p.10.
85. Verma, P.K., Agrawal, P., Amorim, I. and Prodan, R., 2021. WELFake: Word Embedding Over Linguistic Features for Fake News Detection. IEEE Transactions on Computational Social Systems.
86. Vrij, A., 2005. Criteria-Based Content Analysis: A Qualitative Review of the First 37 Studies. Psychology, Public Policy, and Law, 11(1), p.3.
87. Wang, W.Y., 2017. " liar, liar pants on fire": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648.
88. White, D.M., 1950. The "gate keeper": A case study in the selection of news. Journalism quarterly, 27(4), pp. 383–390.
89. Wu, Y., Agarwal, P.K., Li, C., Yang, J. and Yu, C., 2014. Toward computational fact-checking. Proceedings of the VLDB Endowment, 7(7), pp. 589–600.
90. Wynne, H.E. and Wint, Z.Z., 2019, December. Content based fake news detection using n-gram models. In Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services (pp. 669–673).
91. Yang, F., Liu, Y., Yu, X. and Yang, M., 2012, August. Automatic detection of rumor on sina weibo. In Proceedings of the ACM SIGKDD workshop on mining data semantics (pp. 1–7).
92. Yang, K.C., Torres-Lugo, C. and Menczer, F., 2020. Prevalence of low-credibility information on twitter during the covid-19 outbreak. arXiv preprint arXiv:2004.14484.
93. Yang, L., Li, Z. and Giua, A., 2020. Containment of rumor spread in complex social networks. Information Sciences, 506, pp. 113–130.
94. Zhou, X. and Zafarani, R., 2020. A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys (CSUR), 53(5), pp. 1–40.
95. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M. and Procter, R., 2018. Detection and resolution of rumours in social media: A survey. ACM Computing Surveys (CSUR), 51(2), pp.1–36.
96. Zubiaga, A. and Ji, H., 2014. Tweet, but verify: epistemic study of information verification on twitter. Social Network Analysis and Mining, 4(1), p.163.

References

Agrawal, R., & Sharma, D. K. (2021). A survey on video-based fake news detection techniques. In *2021 8th International conference on computing for sustainable global development (indiacom)* (pp. 663–669). IEEE.

Albright, J. (2016). The# election2016 micro-propaganda machine. *Retrieved, 20*(June), 2019.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives, 31*(2), 211–236.

Allport, F. H. (1937). Toward a science of public opinion. *Public Opinion Quarterly, 1*(1), 7–23.

Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: Modelling the predictors of fake news sharing among social media users. *Telematics and Informatics, 56*, 101475.

Aromataris, E., & Pearson, A. (2014). The systematic review: An overview. *AJN The American Journal of Nursing, 114*(3), 53–58.

Baruah, A., Das, K., Barbhuiya, F., & Dey, K. (2020). Automatic detection of fake news spreaders using bert. *Clef*.

Berkowitz, D., & Schwartz, D. A. (2016). Miley, CNN and the onion: When fake news becomes realer than real. *Journalism Practice, 10*(1), 1–17.
Bimba, A. T., Idris, N., Al-Hunaiyyan, A., Mahmud, R. B., Abdelaziz, A., Khan, S., & Chang, V. (2016). Towards knowledge modeling and manipulation technologies: A survey. *International Journal of Information Management, 36*(6), 857–871.
Bleakley, P. (2021). Panic, pizza and mainstreaming the alt-right: A social media analysis of Pizzagate and the rise of the QAnon conspiracy. *Current Sociology*. 00113921211034896
Butler, A. (2018). Protecting the democratic role of the press: A legal solution to fake news. *Washington University Law Review, 96*, 419.
Castillo, C., Mendoza, M., & Poblete, B. (2013). Predicting information credibility in time-sensitive social media. *Internet Research*.
Chen, C., Wu, K., Srinivasan, V., & Zhang, X. (2013). Battling the internet water army: Detection of hidden paid posters. In *2013 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM 2013)* (pp. 116–120). IEEE.
Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of computers & education. *Computers & Education, 151*, 103855.
Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: Causes of trolling behavior in online discussions. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing* (pp. 1217–1230).
Choi, D., Chun, S., Oh, H., Han, J., et al. (2020). Rumor propagation is amplified by echo chambers in social media. *Scientific Reports, 10*(1), 1–10.
Choudrie, J., Patil, S., Kotecha, K., Matta, N., & Pappas, I. (2021). Applying and understanding an advanced, novel deep learning approach: A COVID19, text based, emotions analysis study. *Information Systems Frontiers*, 1–35.
Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2012). Detecting automation of twitter accounts: Are you a human, bot, or cyborg? *IEEE Transactions on Dependable and Secure Computing, 9*(6), 811–824.
Ciampaglia, G. L., Shiralkar, P., Rocha, L. M., Bollen, J., Menczer, F., & Flammini, A. (2015). Computational fact checking from knowledge networks. *PloS one, 10*(6), e0128193.
Cohen, J. (2020). Scientists 'strongly condemn' rumors and conspiracy theories about origin of coronavirus outbreak. *Science, 2*.
Deepak, S., & Chitturi, B. (2020). Deep neural approach to fake-news identification. *Procedia Computer Science, 167*, 2236–2243.
Deziel, M. (2014). *Women inmates: Why the male model doesn't work*. The New York Times.
Duffy, A., Tandoc, E., & Ling, R. (2020). Too good to be true, too good not to share: The social utility of fake news. *Information, Communication & Society, 23*(13), 1965–1979.
Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association, 43*(2), 97–116.
Fader, A., Soderland, S., & Etzioni, O. (2011). Identifying relations for open information extraction. In *Proceedings of the 2011 conference on empirical methods in natural language processing* (pp. 1535–1545).
Faustini, P. H. A., & Covões, T. F. (2020). Fake news detection in multiple platforms and languages. *Expert Systems with Applications, 158*, 113503.
Feng, S., Banerjee, R., & Choi, Y. (2012). Syntactic stylometry for deception detection. In *Proceedings of the 50th annual meeting of the association for computational linguistics (volume 2: Short papers)* (pp. 171–175).
Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104.
Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research, 20*(2), 303–315.
Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology, 49*(2), 280–286.
Frenkel, S., Alba, D., & Zhong, R. (2020). *Surge of virus misinformation stumps facebook and twitter*: vol. 8. The New York Times.
Gottfried, J., & Shearer, E. (2016). News use across social media platforms 2016,.
Granik, M., & Mesyura, V. (2017). Fake news detection using naive bayes classifier. In *2017 IEEE first ukraine conference on electrical and computer engineering (UKRCON)* (pp. 900–903). IEEE.
Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91–108.
Gravanis, G., Vakali, A., Diamantaras, K., & Karadais, P. (2019). Behind the cues: A benchmarking study for fake news detection. *Expert Systems with Applications, 128*, 201–213.
Hassan, N., Li, C., & Tremayne, M. (2015). Detecting check-worthy factual claims in presidential debates. In *Proceedings of the 24th ACM international on conference on information and knowledge management* (pp. 1835–1838).
Highet, G. (2015). *Anatomy of satire*. Princeton University Press.
Huynh, T. L., et al. (2020). The COVID-19 risk perception: A survey on socioeconomics and media attention. *Economics Bulletin, 40*(1), 758–764.
Ireton, C., & Posetti, J. (2018). *Journalism, fake news & disinformation: Handbook for journalism education and training*. Unesco Publishing.
Jamieson, K. H., & Cappella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford University Press.
Jin, Z., Cao, J., Jiang, Y.-G., & Zhang, Y. (2014). News credibility evaluation on microblog with a hierarchical propagation model. In *2014 IEEE international conference on data mining* (pp. 230–239). IEEE.
Jin, Z., Cao, J., Zhang, Y., & Luo, J. (2016a). News verification by exploiting conflicting social viewpoints in microblogs. In *Proceedings of the AAAI conference on artificial intelligence: vol. 30*.
Jin, Z., Cao, J., Zhang, Y., Zhou, J., & Tian, Q. (2016b). Novel visual and statistical image features for microblogs news verification. *IEEE Transactions on Multimedia, 19*(3), 598–608.
Jwa, H., Oh, D., Park, K., Kang, J. M., & Lim, H. (2019). Exbake: Automatic fake news detection model based on bidirectional encoder representations from transformers (bert). *Applied Sciences, 9*(19), 4062.
Kabudi, T., Pappas, I., & Olsen, D. H. (2021). Ai-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence, 2*, 100017.
Kang, H., Bae, K., Zhang, S., & Sundar, S. S. (2011). Source cues in online news: Is the proximate source more powerful than distal sources? *Journalism & Mass Communication Quarterly, 88*(4), 719–736.
Karami, A., Bookstaver, B., Nolan, M., & Bozorgi, P. (2021). Investigating diseases and chemicals in COVID-19 literature with text mining. *International Journal of Information Management Data Insights*, 100016.
Keeble, R. (2008). *Ethics for journalists*. Routledge.
Kovach, B., & Rosenstiel, T. (2014). *The elements of journalism: What newspeople should know and the public should expect*. Three Rivers Press (CA).
Kumar, S., West, R., & Leskovec, J. (2016). Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of the 25th international conference on world wide web* (pp. 591–602).
Kwon, S., Cha, M., Jung, K., Chen, W., & Wang, Y. (2013). Prominent features of rumor propagation in online social media. In *2013 IEEE 13th international conference on data mining* (pp. 1103–1108). IEEE.
Lesce, T. (1990). Scan: Deception detection by scientific content analysis. *Law and Order, 38*(8), 3–6.
van der Linden, S., Roozenbeek, J., & Compton, J. (2020). Inoculating against fake news about COVID-19. *Frontiers in Psychology, 11*, 2928.
Long, Y. (2017). *Fake news detection through multi-perspective speaker profiles*. Association for Computational Linguistics.
Ma, J., Gao, W., Wei, Z., Lu, Y., & Wong, K.-F. (2015). Detect rumors using time series of social context information on microblogging websites. In *Proceedings of the 24th ACM international on conference on information and knowledge management* (pp. 1751–1754).
Magdy, A., & Wanas, N. (2010). Web-based statistical fact checking of textual documents. In *Proceedings of the 2nd international workshop on search and mining user-generated contents* (pp. 103–110).
Maronikolakis, A., Villegas, D. S., Preotiuc-Pietro, D., & Aletras, N. (2020). Analyzing political parody in social media. arXiv preprint arXiv:2004.13878.
McManus, J. H. (1992). What kind of commodity is news. *Communication Research, 19*(6), 787–805.
Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
Mitra, T., & Gilbert, E. (2015). Credbank: A large-scale social media corpus with associated credibility annotations. In *Proceedings of the international AAAI conference on web and social media: vol. 9*.
Mohammad, S. M., Sobhani, P., & Kiritchenko, S. (2017). Stance and sentiment in tweets. *ACM Transactions on Internet Technology (TOIT), 17*(3), 1–23.
Moscadelli, A., Albora, G., Biamonte, M. A., Giorgetti, D., Innocenzio, M., Paoli, S., … Bonaccorsi, G. (2020). Fake news and COVID-19 in italy: Results of a quantitative observational study. *International journal of environmental research and public health, 17*(16), 5850.
Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013). What yelp fake review filter might be doing?. In *Proceedings of the international AAAI conference on web and social media: vol. 7*.
Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights, 1*(1), 100007.
Nelson, M. R., & Park, J. (2015). Publicity as covert marketing? The role of persuasion knowledge and ethical perceptions on beliefs and credibility in a video news release story. *Journal of Business Ethics, 130*(2), 327–341.
Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology, 2*(2), 175–220.
Paka, W. S., Bansal, R., Kaushik, A., Sengupta, S., & Chakraborty, T. (2021). Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection. *Applied Soft Computing*, 107393.
Paul, C., & Matthews, M. (2016). *The russian "firehose of falsehood" propaganda model* (pp. 2–7). Rand Corporation.
Polignano, M., de Pinto, M. G., Lops, P., & Semeraro, G. (2019). Identification of bot accounts in twitter using 2D CNNs on user-generated contents. *Clef (working notes)*.
Prasetijo, A. B., Isnanto, R. R., Eridani, D., Soetrisno, Y. A. A., Arfan, M., & Sofwan, A. (2017). Hoax detection system on indonesian news sites based on text classification using SVM and SGD. In *2017 4th International conference on information technology, computer, and electrical engineering (ICITACEE)* (pp. 45–49). IEEE.
Pronin, E., Kruger, J., Savtisky, K., & Ross, L. (2001). You don't know me, but i know you: The illusion of asymmetric insight. *Journal of Personality and Social Psychology, 81*(4), 639.
Quattrociocchi, W., Scala, A., & Sunstein, C. R. (2016). Echo chambers on facebook. *Available at SSRN 2795110*.
Rangel, F., & Rosso, P. (2019). Overview of the 7th author profiling task at pan 2019: bots and gender profiling in twitter. *Working notes papers of the CLEF 2019 evaluation labs volume 2380 of CEUR workshop*.
Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68*(3), 404.
Roy, A., Basak, K., Ekbal, A., & Bhattacharyya, P. (2018). A deep ensemble framework for fake news detection and classification. arXiv preprint arXiv:1811.04670.
Rubin, V. L., Conroy, N., Chen, Y., & Cornwell, S. (2016). Fake news or truth? Using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection* (pp. 7–17).

Rubin, V. L., & Lukoianova, T. (2015). Truth and deception at the rhetorical structure level. *Journal of the Association for Information Science and Technology, 66*(5), 905–917.
Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on conference on information and knowledge management* (pp. 797–806).
Santia, G., & Williams, J. (2018). BuzzFace: A news veracity dataset with facebook user commentary and egos. In *Proceedings of the international AAAI conference on web and social media: vol. 12*.
Santoso, I., Yohansen, I., Warnars, H. L. H. S., Hashimoto, K., et al. (2017). Early investigation of proposed hoax detection for decreasing hoax in social media. In *2017 IEEE international conference on cybernetics and computational intelligence (cyberneticscom)* (pp. 175–179). IEEE.
Shahi, G. K., Dirkson, A., & Majchrzak, T. A. (2021). An exploratory study of COVID-19 misinformation on twitter. *Online Social Networks and Media, 22*, 100104.
Shi, B., & Weninger, T. (2016). Fact checking in heterogeneous information networks. In *Proceedings of the 25th international conference companion on world wide web* (pp. 101–102).
Shoemaker, P. J., & Reese, S. D. (2013). *Mediating the message in the 21st century: A media sociology perspective*. Routledge.
Shrivastava, G., Kumar, P., Ojha, R. P., Srivastava, P. K., Mohan, S., & Srivastava, G. (2020). Defensive modeling of fake news through online social networks. *IEEE Transactions on Computational Social Systems, 7*(5), 1159–1167.
Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2018a). FakeNewsNet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media. arXiv preprint arXiv:1809.01286.
Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter, 19*(1), 22–36.
Shu, K., Wang, S., & Liu, H. (2018b). Understanding user profiles on social media for fake news detection. In *2018 IEEE conference on multimedia information processing and retrieval (MIPR)* (pp. 430–435). IEEE.
Sinclair, C. (2020). Parody: Fake news, regeneration and education. *Postdigital Science and Education, 2*(1), 61–77.
Song, C., Ning, N., Zhang, Y., & Wu, B. (2021). A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks. *Information Processing & Management, 58*(1), 102437.
Subramanian, S. (2017). Inside the macedonian fake-news complex, wired, february 15.
Tacchini, E., Ballarin, G., Della Vedova, M. L., Moret, S., & de Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. arXiv preprint arXiv:1704.07506.
Talwar, S., Dhir, A., Kaur, P., Zafar, N., & Alrasheedy, M. (2019). Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services, 51*, 72–82.
Thota, A., Tilak, P., Ahluwalia, S., & Lohia, N. (2018). Fake news detection: A deep learning approach. *SMU Data Science Review, 1*(3), 10.
Verma, P. K., Agrawal, P., Amorim, I., & Prodan, R. (2021). Welfake: Word embedding over linguistic features for fake news detection. *IEEE Transactions on Computational Social Systems*.
Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law, 11*(1), 3.
Wang, W. Y. (2017). "Liar, liar pants on fire": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648.
White, D. M. (1950). The "gate keeper": A case study in the selection of news. *Journalism Quarterly, 27*(4), 383–390.
WHO (2021). International COVID-19 dashboard. https://covid19.who.int/, Accessed: April 29, 2021.
Wu, Y., Agarwal, P. K., Li, C., Yang, J., & Yu, C. (2014). Toward computational fact-checking. *Proceedings of the VLDB Endowment, 7*(7), 589–600.
Wynne, H. E., & Wint, Z. Z. (2019). Content based fake news detection using n-gram models. In *Proceedings of the 21st international conference on information integration and web-based applications & services* (pp. 669–673).
Yang, F., Liu, Y., Yu, X., & Yang, M. (2012). Automatic detection of rumor on Sina Weibo. In *Proceedings of the ACM SIGKDD workshop on mining data semantics* (pp. 1–7).
Yang, K.-C., Torres-Lugo, C., & Menczer, F. (2020a). Prevalence of low-credibility information on twitter during the COVID-19 outbreak. arXiv preprint arXiv:2004.14484.
Yang, L., Li, Z., & Giua, A. (2020b). Containment of rumor spread in complex social networks. *Information Sciences, 506*, 113–130.
Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR), 53*(5), 1–40.
Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys (CSUR), 51*(2), 1–36.
Zubiaga, A., & Ji, H. (2014). Tweet, but verify: Epistemic study of information verification on twitter. *Social Network Analysis and Mining, 4*(1), 163.