A Review of Methodologies for Fake News Analysis

Article *in* IEEE Access · January 2023

DOI: 10.1109/ACCESS.2023.3294989

CITATIONS
16

READS
1,146

4 authors, including:

Tajrian Mehedi
Charles Sturt University
5 PUBLICATIONS **18** CITATIONS
SEE PROFILE

Ashad Kabir
Charles Sturt University
170 PUBLICATIONS **2,433** CITATIONS
SEE PROFILE

Md Rafiqul Islam
Charles Sturt University
198 PUBLICATIONS **3,295** CITATIONS
SEE PROFILE

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 June 2023, accepted 10 July 2023, date of publication 13 July 2023, date of current version 24 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294989

RESEARCH ARTICLE

A Review of Methodologies for Fake News Analysis

MEHEDI TAJRIAN, (Member, IEEE), AZIZUR RAHMAN, MUHAMMAD ASHAD KABIR, (Member, IEEE), AND MD. RAFIQUL ISLAM, (Senior Member, IEEE)
School of Computing, Mathematics and Engineering, Charles Sturt University, Wagga Wagga, NSW 2678, Australia
Corresponding author: Mehedi Tajrian (mtajrian@csu.edu.au)

ABSTRACT Nowadays, with the proliferation of different news sources, fake news detection is becoming a crucial topic to research. Millions of articles are published daily in the press, on social media, and in electronic media, and many of them may be fake. It is common for scammers to spread fake news to mislead people for malicious purposes. For researchers to be able to evaluate fake news, it is necessary to understand its diversity, how to study it, how to detect it, and its limitations. A descriptive literature review has been conducted in this paper to identify more appropriate methodologies for analysing fake news. The review found two broad classifications in the fake news research methodologies: fake news study perspectives and fake news detection techniques. Based on our literature review, we suggest four perspectives to study fake news and two major approaches to detecting it. Fake news can be studied in terms of knowledge, style, propagation and source. In order to detect fake news, there are two major approaches: manually and automatically. There are two types of manual fact-checks: expert-based and crowd-sourced. Automatic techniques are based mainly on data science methods, specifically deep learning and machine learning. A machine learning-based method was found to be more appealing when we evaluated all the automatic methods. Further research will focus on investigating the efficacy of using Bayesian methods for detecting fake news statistically because it is a flexible approach that allows for rapid updating of models in response to new data and has been successfully applied to a wide range of problems across different domains.

INDEX TERMS Bayesian modelling, classification, deep learning, fake news detection, machine learning, NLP.

I. INTRODUCTION

In today's world, we are inundated with information every second by a variety of media. Since data is easily accessible and loosely authenticated, much of it cannot be verified and can be fake. By posting, sharing, and commenting on fake news, people who have little time to verify it are contributing to propagating this false information. Consumers of this information are thus susceptible to fake details, which negatively affect society at all levels: local, national, and global. Fake information can be hazardous to the people who are the subject of the information. Especially when it comes to key world events like central elections, fake news propagation can have horrible results on the public, such as transforming their opinions and actions. As an example, fake news motivated more than 50% of voters during the 2016 U.S. Presidential elections [4], [5]. The impact of it goes beyond politics. Fake news propagation makes People demotivated to vaccinate [6].

From different perspectives, fake news has several definitions, including maliciously false news, misleading information, satire news, misinformation, conspiracy theories, and disinformation. Fake news has not been universally defined [7], [8], [9]. Research communities often define fake news as "Fake news is a news article that is intentionally and verifiably false" [4], [10]. The main intention behind writing fake news is to mislead readers [11], [12], [13]. Fake news has diverse aspects, such as rich imagination, extensive scams, and clever trickery. Social media usually goes viral with malicious intentions when weighty imagination is shared. A comprehensive scam, on the other hand, frequently targets

The associate editor coordinating the review of this manuscript and approving it for publication was Wei-Yen Hsu.

public figures or ideas and asserts false information. It is common for clever fakes to be written with a humorous intention [12], [14], [15], [16]. On TikTok, there is an account dedicated to Tom Cruise's deep fakes, in which videos show Cruise doing everything from golfing to performing magic tricks.

Social media content is used in the news in the US media more than twice [17]. News from social media becomes more famous after being shared on Facebook, and people believe news from social media more than mainstream media [18], [19], [20]. It is common for news to contain declarations, posts, statements, and speeches. Users' ignorance, awareness, or even mental preferences can lead to users spreading fake news unintentionally [21], [22]. These users can unconsciously spread fake news through blogs, articles, comments, and tweets. According to a study [23], some fake images of Hurricane Sandy posted by a few users caused massive retweets. Two groups of people are involved in fake news movements (i) malicious users and (ii) naive users. The malicious users reproduce fake news to mislead people, and the naive users spread the fake news without concern. Bots and sockpuppets are used by those who create and spread false information [24], [25], [26], [27]. Unlike humans, bots are mindless. They automatically create and spread false content. A sockpuppet is a pseudonym that is given by an author so that they may increase sales or spread false information regarding their own product. For instance, Russian sock puppets are spreading misinformation about Ukraine on social media. This type of fake account is maintained by malicious users [28], [29]. Social media bots are implicated in the spreading of fake news, according to a study [30]. Despite its critical importance for our individual and social contexts, finding an appropriate mechanism for detecting fake news remains challenging.

Several studies have been conducted on fake news. We demonstrated a flowchart of methodologies for analysing fake news in our study. Based on our proposed methodology, a fake news analysis can be approached from four perspectives. Our methodology enables the detection of fake news manually and automatically. Expert-based or crowd-sourced methods can be used to detect fake news manually, while deep learning and traditional machine learning methods can be used to detect fake news automatically.

Following is a breakdown of the paper's structure. The methodology for analysing fake news has been briefly reviewed in Section II. We discussed different perspectives of fake news in Section III. A number of techniques have been discussed in Section IV for detecting fake news. In Section V, we discussed some limitations of the methods for detecting fake news. We summarised the main findings and concluded with a few remarks in Section VI.

II. METHODOLOGIES FOR FAKE NEWS ANALYSIS

Research on fake news uses a wide range of methodologies. Figure 1 illustrates a flowchart of those methodologies for fake news analysis in this descriptive review. Various methodologies for fake news detection are available based on our findings. The overall methodologies can be classified in the following manner.

Fake news can be studied from four main perspectives, which have been discussed in Section III. Furthermore, fake news can be detected manually or automatically. Manual fake news detections are done by either expert-based techniques or crowed sourced based techniques. Manual-based methods were mainly used before automatic methods were invented. This method was popular when technology was not so good and computation power was not very high.

With technological developments and increasing computation power, contemporary methods such as automatic techniques were invented. Automatic fake news detections are done using deep learning and traditional machine learning techniques. A variety of deep learning approaches, including RNNs, LSTMs, Bi-LSTMs, CNNs, and others, have been proposed for detecting fake news. Moreover, researchers have proposed methods for detecting fake news using traditional machine learning techniques, including SVM, NBC, RF, DT, LR, and BM. Based on the comparison, we found that machine learning-based methods are more appealing and used by many researchers. Further detailed explanations of these methods are described in the next sections.

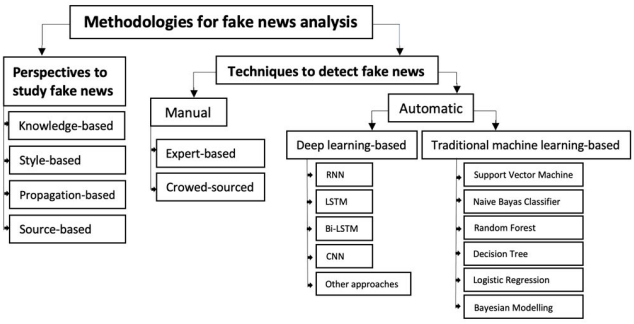

FIGURE 1. A classification of the overall methodologies in fake news analysis.

The following two sections will describe the details of the types of study perspectives and detection techniques of fake news. Firstly, we have given an idea about the perspectives on how fake news can be studied. Secondly, we provided many detection ways based on data science techniques that we have found from our literature study.

III. PERSPECTIVES TO STUDY FAKE NEWS

Before beginning to understand how to identify fake news, a researcher must develop a comprehensive understanding of its variety. Additionally, it is important to study the causes and motivations behind the spread of fake news, such as malicious actors, financial incentives, and social media algorithms. Moreover, it is essential to be aware of the effects of fake news, such as the erosion of trust in the media, the difficulty of political discourse, and the potential to incite violence. Once a researcher has a thorough understanding of

fake news, they can then begin to explore ways to detect it. Different researchers have different perspectives on studying fake news. The study of fake news can take many forms. Many researchers investigate the percentages of clarity in the news, while others explore how fake news is created. Contrarily, many studies directly investigate the propagation paths of fake news while others examine the source credibility. There are four main study perspectives of fake news detection [7], [8], [31], as depicted in Figure 1. In two perspectives, fake news is studied during its creation, knowledge and style-based studies, and in two perspectives, fake news is studied after its spread, propagation and source-based studies.

A. KNOWLEDGE-BASED STUDY

Knowledge-based approaches, also known as fact-checking, involve collecting raw facts from the open web to construct knowledge. However, this data must be further processed to address redundancy, invalidity, conflicts, unreliabilities, and incompleteness issues. Five tests are necessary to ensure the accuracy of the facts: entity resolution, time recording, consistency evaluation, completeness evaluation, and accuracy evaluation. Entity resolution is used to link records of similar facts, while time recording is used to measure the temporal validity of facts. Consistency evaluation ensures that facts are not contradictory, completeness evaluation ensures that all facts are accounted for, and accuracy evaluation ensures that the facts are correct [7], [8]. The data can be effectively cleansed by running these five tests, and knowledge-based approaches can be successfully implemented.

B. STYLE-BASED STUDY

The purpose of style-based studies is to assess whether the news is intended to mislead the public or not [7], [8]. Style-based approach attempt to capture the writing style of news content. It is possible to distinguish fake news from the truth by defining styles based on some quantifiable characteristics. Fake news is identified using theories, patterns, and strategies. Text and images distinguish fake news from real news. The text patterns concern informality, diversity, subjectivity, and inspirational language. In fake images, clarity and coherence are often high, but diversity is often poor due to a lack of clustering scores.

C. PROPAGATION-BASED STUDY

Propagation-based studies try to understand how true information propagates in networks where the news created by anomalies is predicted as false information. This kind of study is mainly about how a user is involved in spreading fake news. The input of this kind of study may be either a news cascade or a self-defined graph [8]. In a news cascade, news propagation is represented directly, whereas self-defined graphs are indirect representations that capture more information about propagation. A cascade is a structure that is similar to a tree and is defined by the initiator of fake news. It comprises multiple nodes, each determined by the users who posted or forwarded the fake news. Conversely, a self-defined graph is a network in which fake news spreads and can be of various types. A homogeneous network is one type of network that is composed of a single type of node and edge. A heterogeneous network is another type comprising multiple nodes and edges and is usually more complicated than a homogeneous network. Finally, a hierarchical network is composed of nodes arranged systematically in layers or tiers, where the root node initiates the fake news.

D. SOURCE-BASED STUDY

Source-based study investigates the credibility of its creators and spreaders. The study can be done based on news and social information. Source credibility can be assessed by observing the role of the news author and publisher. Malicious social media users can initiate and spread news stories inside and outside their networks. Normal users forward or share manufactured news without any verification. Source-based studies are about how users engage with fake news and their roles in creating it, publishing it, and spreading it on media.

There are several types of research that examine the life cycle of fake news, its spreading nature, reason of propagation, and detection challenges [2], [11], [12], [17], [31], [45], [56], [70], [71] and investigate the different perspectives of fake news studies [7], [8], as described in Table 1. For example, according to Bondielli et al. [11], fake news can be defined in many ways, collected in many ways, and witnessed using different techniques. They informed that there is a limited dataset that can be addressed as standard in every required way. They stated that deep learning techniques are very efficient in identifying fake news.

IV. TECHNIQUES TO DETECT FAKE NEWS

While there are numerous techniques available for detecting fake news, none of them can distinguish fake news completely because of their limitations, including the lack of datasets, the difficulty in preprocessing large amounts of data, the variety and vibrant characteristics of data, the lack of investigation of multimodal data, and many more. Thus, various kinds of fake news detection methods have been categorised in further research. Fake news can be witnessed manually and automatically [32], as described in Figure 1.

A. MANUAL DETECTION TECHNIQUES

Manual Fact-Checking is done by obtaining fact checks which allow a reader to critically assess writing and consider its relevancy and integrity [7], [33]. There are more than 100 engaged fact-checking websites worldwide [34]. However, these fact-checking sites often need help to link with all the pages carrying the suspicious claims because of the mutation and fast-spreading manner of fake news, which reduces the possible influence of fact checks [33]. The manual process of fact-checking is further classified into two groups, as described in Figure 1.

TABLE 1. Different articles about fake news analysis approaches.

Citations	Research Aim	Methodology	Significant Findings	Strategy/Limitation/Recommendation
2	They studied the Web's fraudulent data life cycle, different kinds of fraudulent data, performers, and reasons.	They studied users' perception of false information, how it propagates, the community efforts to witness and manage the spread of fraudulent data on the Web and the impact of false information on politics.	Automated solutions can reduce the propagation of fraudulent data on the Web without human input.	They recommended that researchers should concentrate on planning and inventing actual outlets.
7	This research aims to identify and specify fundamental theories across various disciplines to facilitate and enhance the interdisciplinary research of fake news.	They studied fake news using four views: the fraudulent knowledge it holds, its writing manner, its spreading patterns, and the investigation of the user behaviours.	They suggested attribute-based or relational-based methods for fake news analysis, classification, comparison, and determination of existing fake news.	Fake news should be detected early, as once it gains people's confidence, it is challenging to correct. Further investigation can be done on different domains, languages, websites, and topics.
8	They reviewed and evaluated methods that can detect fake news from four perspectives: the false knowledge it carries, its writing style, its propagation patterns, and the credibility of its source.	They studied various definitions of fake news, its effects on the public, and some fundamental theories. They categorised automatic fake news investigation techniques in four ways: knowledge, style, propagation, and source.	Particularly, they identified the basic approaches across many domains in detail to boost multi-order investigation of fake news.	Non-traditional and partially fake news can be detected along with early detection, identifying check-worthy content, and cross-domain detection.
11	They surveyed various techniques for intuitive verification of fake news and rumours.	They studied different types of fake news and rumours, data collection approaches, features for fake news and rumour detection systems, and methods used to witness rumours and fake news.	Deep learning techniques have obtained higher accuracy in many instances.	There is a shortage of largely acknowledged standard datasets that have to be addressed.
12	They provided a typology of integrity estimation strategies from two specific types-network analysis approaches and linguistic cue approaches and designed a fake news detector.	They aimed to propose a hybrid approach to system design, utilising the disparate technique. They categorised two major methods: linguistic and network approaches. Linguistic approaches extract the content of deceptive messages and analyse them to associate language patterns with deception, whereas network approaches aggregate deception measures.	Linguistic and network techniques in classification tasks within finite environments had observed the highest accuracy.	For maximum performance, linguistic depuration should be made from lexical examination to the highest discourse-level observation to get the highest outcome.
17	This survey spans	They categorised the	A significant challenge is	Capable automatic

TABLE 1. *(Continued.)* **Different articles about fake news analysis approaches.**

	diverse aspects of false information about the spreader, mechanisms, rationale, impact, characteristics, and detection.	false information identification algorithms among feature engineering-based, graph-based, and modelling-based.	to differentiate between false and true information. In most cases, fake data is pretended as fact, making it more difficult to recognise.	information-matching algorithms are required for fact-checking. Crowdsourced signals can be helpful in the premature verification of false information and resource allotment of fact-checkers.
31	In this study, researchers, give an idea of the standard methods, structures, datasets, and testing arrangements for scope and behaviour estimation of fake online data.	They studied various works reported in fake news and rumour detection. They talked about the origination of falsified content by focusing on the technical aspects along with different models and diffusion patterns, different stylometric and feature-oriented machine learning methods, deep learning and other methods of credibility analysis by which fraudulent content can be segregated.	They gave information about the data corruption lifecycle, classification of fake data, various social and digital transmission media ecosystems, creators and spreaders of fake data, propagation and several source estimation outlets.	Future research can be done on cross-platform detection, real-time learning, unsupervised models, multilingual platform, complex and dynamic network structure, cross-domain analysis, multimedia false information detection and so on.
45	They described the challenges in fake news detection, the task formulations, datasets, NLP solutions, their potential, and limitations.	They mainly focused on artificial news investigation of text content. Due to different labelling or scoring strategies, manufactured information investigation can be cultivated as a classification or regression problem.	They investigated whether the manufactured attributes can be merged with neural network standards, proper usage of non-textual information, and opening the mode of confirmation with scopes.	In a phony information investigation corpus, truthful and deceptive instances should be available.
56	This research systematically analyses, examine and incorporate recent studies in which several ML and DL methods were utilised to verify COVID-19 deceptions.	They figured out three databases and, using the PRISMA method, illustrated a gradual, systematic process for choosing the papers.	They recognised diverse COVID-19 deception datasets and examined various attribute pulling, data processing and classification methods to witness COVID-19 deception.	A dataset with all relevant features collected from major reliable sources that are balanced, complete, sufficiently explained, and extensively standard on COVID-19 deception is still missing.
70	They studied the various uses of sentiment investigation in detecting fake news and described two methods for identifying fake news.	They discussed different approaches to incorporating sentiment in detecting fake news, the meaning of false data, its detecting process, application of sentiment analysis in catching false data as an element in a machine-learning approach.	Clear differentiation between methods and systems is challenging due to the limitations of data sets.	Future research can be done on multilingualism, multimedia content, and detecting slightly modified authentic news stories.

TABLE 1. *(Continued.)* **Different articles about fake news analysis approaches.**

71	They mainly had done sophisticated categorisations of many recent articles that illustrated the current research movement in sentiment investigation and its corresponding dimensions.	They collected fifty-four articles that presented important enhancements of sentiment analysis, feature selection, and sentiment classification techniques by illustrating their work's algorithms and data.	This survey uniquely gives a refined categorisation to various SA techniques with a discussion on new related fields. These fields include Emotion Detection, Building Resources and Transfer Learning.	This is qualitative research.

1) EXPERT-BASED MANUAL FACT-CHECKING

Fact-checking is usually done by requesting experts such as reporters to check assertions against proof established on earlier expressed or reported validities [75]. It is the traditional way to complete this task which is also time-consuming and costly. However, many websites allow expert-based fact-checking, for example, PolitiFact, The Washington Post Fact Checker, FactCheck, Snopes, TruthOrFiction, FullFact, HoaxSlayer and many more [75].

2) CROWD-SOURCED MANUAL FACT-CHECKING

Crowd-source refers to many folks of normal people functioning as fact-checkers, for example, collective intelligence. This technique repeatedly requires purifying doubtable users and fixing clashing fact-checking outcomes [7]. It is moderately challenging to maintain and less reasonable than expert-based fact-checking. Both prerequisites evolve more crucial with the increasing number of fact-checkers. Crowd-sourced fact-checking websites are still need evolution, such as Fiskkit.

However, according to several research studies, humans are not good judges of false information [1], [35], [36], [37], [75]. Specifically, the accuracies are between 53% and 78% of humans' ability to identify false information with various kinds of untrue manners, including hoaxes, fake reviews, and fake news. It is easy to fool both trained and casual readers with well-written, long, and well-referenced false information [31].

B. AUTOMATIC DETECTION TECHNIQUES

Automatic fake news detection helps to minimise the consequences, including human time and contribution to detecting

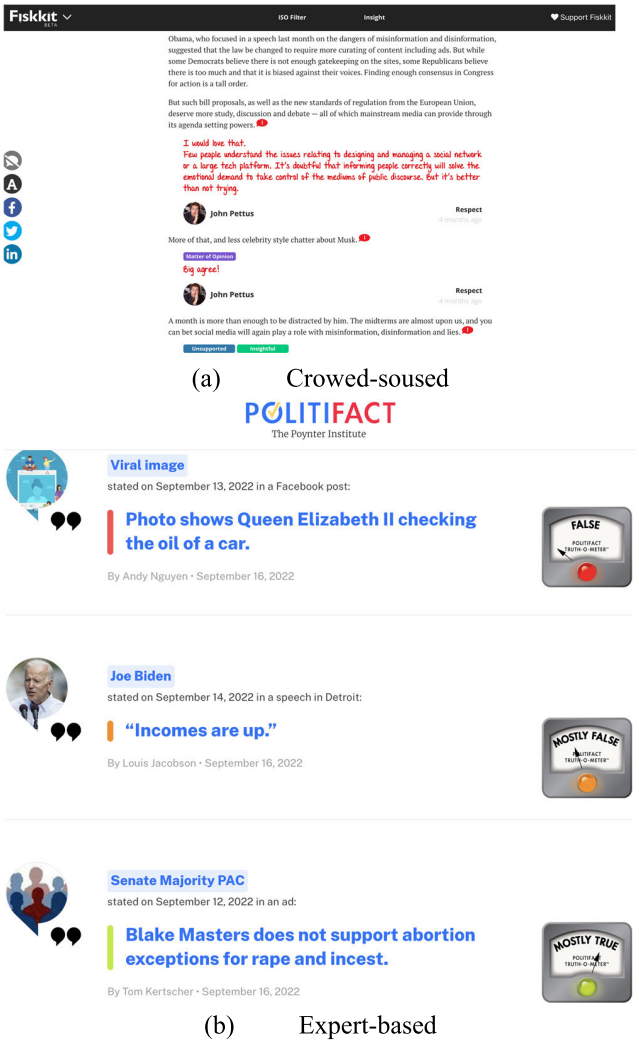

FIGURE 2. Two illustrations of manual fact-checking websites.

fake news and preventing distribution [32], [75]. Several labelling or scoring strategies do exist to detect fake news. Deep learning and machine learning are two main branches of data science used by different researchers in different ways to detect fake news.

1) DEEP LEARNING-BASED TECHNIQUES

Deep learning has been investigated extensively among many research topics in machine learning. Deep Learning classifiers have witnessed exceptional advancement because of promising outcomes in many research fields, including text mining and NLP. Deep learning models can learn hidden illustrations from easier inputs in context and content interpretations [38]. Neural networks (NNs) are the most popular strategies among deep learning techniques. Many research articles used deep learning to detect fake news, as following descriptions.

RNN: Recurrent Neural Networks (RNNs) are becoming increasingly popular for solving fake news detection

TABLE 2. Different articles about fake news detection techniques.

Citations	Research Aim	Methodology	Significant Findings	Strategy/Limitation/Recommendation
3	They aimed to solve rumour detection tasks beneath the representation learning the framework by assembling the propagation graph, following the propagation form of posts on Twitter.	They suggested a gated graph neural network-based algorithm named PGNN. It frequently renovates node depiction by swapping facts between the neighbour nodes through association routes within a fixed time activity.	Their suggested models perform considerably finer than state-of-the-art approaches to rumour detection and early detection tasks.	Future work can profoundly integrate users' social network information by enriching the propagation graph. Transfer learning methods can be used to work with unlabeled data.
33	They presented an advanced Related Fact Checks service that supports a person in censorious assessing writing and judging its integrity by obtaining fact checks that apply to the report.	The main contribution of the RFC service is to make it effortless to access via a browser extension, which can be accomplished by adding fake news with fact checks that examine it.	They examined the individual writing truthiness depending on three dimensions. The number of 'Irrelevant' results was relatively high. Only using topics does not do well on 'On Claim'.	Computational tools are needed to classify articles as fake or true. The automatic identification of the topics needs to explore more. Articles should be fact-checked before the user interacts with the browser considering privacy.
39	They presented a mixed deep learning representative incorporating CNN and RNN for fake news classification.	They used CNN to take out local attributes and used LSTM to learn long-term dependencies in their model.	Compared with other non-mixed basic models, it showed considerably better results for fake news detection after applying it to datasets like ISO and FA-KES.	Deep learning models have difficulty uncovering the optimal hyperparameters for individual problems and datasets, which can be solved using hybrid approaches.
40	They developed different deep learning standards that detect phony information by organising the information among pre-determined fine-grained types.	They first created models based on CNN and Bi-directional Long Short-Term Memory (Bi-LSTM) networks. Then the acquired models are provided into a Multi-layer Perceptron Model (MLP) for the conclusive classification.	They achieved 44.87% accuracy on a benchmark dataset which outperforms the existing works.	Using some semi-supervised or active learning models for labelling data and the actual statements can help train models better. Applying sarcasm detection methods in the fake news domain can be effective for short statement detection.
42	They utilised TI-CNN (Text and Image information-based Convolutional Neural Networks) to merge the explicit and latent features of image and text data to suitable attribute lengths, which employed intellectual elements for recognising the fake news.	Excluding explicit features, they utilised two parallel CNNs to extract latent components after which latent and explicit components are launched to corresponding attribute dimensions for constructing further models of images and texts.	They investigated a dataset collected before the presidential election. It demonstrates their proposed model could strongly recognise fake news.	Identifying social network information, for example, users' behaviours, and social network establishments and the pertinence between headlines and news texts, will be instrumental in identifying fake news.
43	They presented the multi-level convolutional neural network (MCNN) that captures semantic information from article texts that can be utilised to	They used a technique of estimating the consequence of susceptible words (TFW). MCNN pulls article illustrations, and WS estimates the weight of susceptible	The evaluation outcomes indicate that MCNN-TFW surpasses recent strategies by accuracy and efficiency, which is 88.82% and 90.10% on Weibo and	Reviewing the execution of their technique in a broader spectrum of applications can be done in the future.

TABLE 2. *(Continued.)* **Different articles about fake news detection techniques.**

	categorise fake news and real news.	terms for unique news.	NewsFN datasets, respectively.	
44	They suggested a multimodal fake news recognition framework founded on Crossmodal Attention Residual and Multichannel convolutional neural Networks (CARMN).	The CARMN pulls the relevant data connected to a prey paradigm with another reference paradigm and preserves the prey paradigm's details.	The experiment result on four real-world datasets demonstrated that the suggested standard showed better results than other recent techniques and retained better additional variant feature presentations.	This work can be extended by exploiting the visual information, more modalities, multimodal fusion modules and event-level multimodality to improve fake news detection.
53	They proposed a fully automated end-to-end neural network standard, DeClarE, for a proof-aware integrity check of natural language.	DeClarE captures signals from exterior proof articles and standards of mutual exchanges between aspects like the context of an assertion, the wording of documenting reports, and the dependability of their origins.	They tested using Snopes, PolitiFact, News Trust, and SemEval datasets, which showed that DeClarE outperforms the existing works.	This is quantitative research.
55	They showed that Facebook posts could be picked as scams or founded on the counting of likes.	The writers employed two classification methods founded on logistic regression and the unexplored transformation of boolean crowdsourcing algorithms.	By experimenting on a dataset, they acquired 99% accuracies for classification where the training set includes smaller than 1%.	This is quantitative research.
56	They proposed a costume category model for fake news detection	They extracted essential features from the datasets of fake news and then comprised three mainly used machine learning strategies, i.e., decision tree, random forest, and extra tree classifier.	Their classification model has performed more reasonably accurately than the recent studies for detecting fake news.	Hyper-parameter tuning prototypes established on meta-heuristic algorithms can be used for future works.
57	They proposed an n-gram fake news detection model with two distinct feature extraction methods and six machine-learning classification approaches.	They presented a recognition strategy that merges text analysis using n-gram features, terms frequency metrics and machine learning classification.	Their standard outperforms existing work. For example, 90% and 87%, whereas previous works had 89% and 71% for the same dataset using 50,000 and 10,000 features.	Accurately labelled real-world datasets are unavailable. Adding statistical and text features can improve dimensionality. So, more feature selection techniques should be explored along with unsupervised and semi-supervised methods.
58	The proposed C-LSTM is a sentence depiction and text classification model.	C-LSTM utilises CNN to drag an order of higher-level word illustrations to obtain the sentence representation that detain both local features of words and global and temporal sentence semantics.	According to their experiment result, C-LSTM beats both CNN and LSTM.	The standard convolution can be replaced with tensor-based operations or tree-structured convolutions, in which LSTM will profit from better-structured higher-level models.

TABLE 2. *(Continued.)* **Different articles about fake news detection techniques.**

59	They proposed a new F1-based metric delivering a modified design order which directs to a fiction component-rich piled LSTM representative that functions on par with the most suitable methods.	They assessed the execution of three top-scoring strategies, critically evaluated the experimental design, and completed a detailed feature investigation.	They identified better functioning features, performed error analysis, and discovered that the models primarily lean on lexical overlap for classification .	More advanced machine learning approaches are required with more in-depth semantic knowledge, which can decide the perspective established on the content.
60	They presented the multi-task learning strategy that permits the joint exercise of the primary and supplementar y assignments, enhancing the execution of rumour confirmation.	The rumour solution approach has been described as a pipeline concerning four sub-assignments: rumour investigation, rumour hunt, perspective category, and rumour validation.	Multi-task techniques improved over single-task studying in accuracy and macro F-score for integrity category by using job corresponde nce with supplementa ry tasks, particularly rumour detection and stance classification .	Further improvement s are possible on primary and supplementar y assignments with multi-task learning by adjusting various dataset measures, including the order between assignments into the standard and adding additional features (such as a user).
61	They suggested a new standard, DUAL, for catching fake news by connecting two dimensions of features with suitable attention and comprehendin g the undercover illustration of these two features.	They used an attention based bi-directional GRU to drag attributes from the information scope and a deep model to pull inconspicuous expressions of flank data by combining two invisible vectors.	They experimente d on two real-world standard datasets and discovered that their strategy beats multiple criteria in the accuracy of witnessing fake news.	This is quantitative research.
62	They proposed a multistage intervention framework.	It embarks on fake news in social networks, integrating support knowledge with a point process network training representative and developing a policy iteration method to optimise the actions for a maximal total reward.	This approach delivers profitable execution in real-time intervention investigation s on a Twitter network to mitigate a surrogate fake news movement and surpasses choices on manufacture d datasets.	Additional complex real-world investigations that are not content-neutral can be done with better complex features, for instance, quadratic and nonlinear features, aside from subjecting real users to fake news.
63	They introduced MANDOLA, a big-data processing procedure using big-data methods.	MANDOLA consists of six separate segments shared ingest, function, stock, and imagine statistical data concerning hate speech circulated online.	It is a robust method for reporting and monitoring online hate speech by usability and functionality .	By applying it to current standard datasets and spreading this work using the multi-lingual aspect, performance can be enhanced.
64	They proposed a Deep Learning-based model for Misinformatio n Detection that assumes various text features terms, for example, general and exact word embeddings.	They trained ten Deep Learning models for multi-class variety, employing a large dataset labelled with ten classes. They also employed three word embeddings methods that maintain the word context,	They obtained the most promising outcomes, and the observed outcomes demonstrate that the standards are favourably trustworthy.	Further investigations can be done on strategies for creating textual features, for example, sentiment analysis or topic modelling. An attention mechanism can be added to the

TABLE 2. *(Continued.)* **Different articles about fake news detection techniques.**

		such as Word2Vec, FastText, and GloVe.		bidirectional LSTM.
65	They designed SpotFake: a multimodal framework for fake news investigation, which detects if a shared news report is real or fake.	The exceptional originality of SpotFake is to connect the power of vocabulary prototypes. The illustrations from both modalities are then joined concurrently to create the expected news vector.	This model outperforms the existing studies on Twitter and Weibo datasets by 3.27% and 6.83%, respectively.	More improvement strategies are needed to confound the difficulties of detecting longer-length reports, more complex fusion methods and different modalities.
66	They suggested a strategy founded on textual and uniform resource locator (URL) features that investigate and catch if the news is manufactured or faithful.	They instructed different machine learning algorithms with preprocessing techniques, for instance, bag-of-words and TF-IDF, utilising textual and URL information belongings. Additionally, they applied re-sampling techniques to handle class inequality because of real-world data.	The over-sampling technique showed adequate results, but the under-sampling technique could not improve prototype performance and exhibited abysmal effects because of the small specimen length.	Phishing methods and frequently utilised antagonistic URL verification techniques should be evaluated for fake news verification for creating classifiers on lexical and host-depended features of the URL.
67	They introduced automatic fake news verification approaches in a chrome background to identify fake news on Facebook.	In this study, researchers used numerous attributes correlated with Facebook accounts to analyse the account's behaviour via deep learning. The resultant output will be displayed in the condition of a pop-up box in the chrome background at the user's homepage.	Their planned fake news detection system had more increased accuracy with 99.4% than the current methods, which resulted in an analysis of real-world information.	More profitable decisions can be made by examining the features and utilising other deep learning algorithms such as bidirectional LSTM, bidirectional GRU and other hybrid techniques.
68	They proposed a data-driven automatic fake news detection method.	They applied the BERT standard to glimpse fake news by exploring the connection between the headline and the body text of news.	BERT has a deep-contextualizing nature, excellently fitted for this analysis and enhanced the 0.14 F-score more than existing standards.	Further experimentation can be done by using extra news information in the pre-training stage with different points of fake news detection tasks.
72	They proposed a Deep Normalized Attention-based method to extract dual emotion features and an Adaptive Genetic Weight Update-Random Forest for classification to investigate fake news.	Using deep normalized attention-based strategy integrating BiGRU, feature values were enhanced to remove gradient explosion issues by bringing out long-range context data to capture more valuable features.	They adjusted the genetic weight for the prototype to Random Forest and accomplished better hyper parameter value which enriched detection accuracy of classifier. Experiment result showed accuracy of 5%, 11% and 14% for three different datasets compared to other existing models.	This method is not suitable for early detection due to insufficient comments for reliable verification. Multimodal dual emotion and cross-correlation features from social media data can be considered for further research. Additionally single emotion and sentimental impact can be observed to check accuracy of the model

TABLE 2. *(Continued.)* **Different articles about fake news detection techniques.**

73	They innovated a latest strategy to detect fake news using sentiment analysis incorporate with Multiple Imputation Chain Equation method to manage multivariate missing variables with two datasets: ISOT and LIAR	Sentiment analysis using lexicon-based scoring algorithm was utilized in this method to increase the accuracy. Additionally, TF-IDF used to determine long term features. Utilizing Naïve Bayes, passive-aggressive and DNN they classified correlation of missing variable and important features.	Their method gained 99.8% accuracy to detect fake news by labelling the data among five classes: barely true, half true, true, mostly true and false.	This method only applied on two datasets. More datasets can be investigated with this method to examine the accuracy.
74	This study investigated the effect of attaching contextual data on hate speech detection.	They examine a subdomain of Twitter data containing replies to online newspaper posts a built a new corpus in Spanish language	According to their finding, observing transformer-based machine learning methods, their macro F1 was grown to 4.2 and 5.5 points.	For further research, other source of data can be included using knowledge graph.

problems, especially regarding text, speech, and video content. RNNs can process data sequentially, allowing them to learn from past data, which can be incredibly powerful when detecting fake news [39]. RNNs can also be used to detect patterns in the data, which can be used to detect the presence of fake news. Furthermore, RNNs can identify underlying trends that may indicate fake news, such as the use of specific words or phrases that may indicate that the news is not genuine. In addition, RNNs can also be used to detect inconsistencies in the data, such as if the same phrase is used across multiple sources, which could be a sign of fake news. RNNs are well-suited to tackling the problem of fake news detection, particularly regarding text, speech, and video content. With their ability to process data sequentially and identify patterns in the data, RNNs are potent tools for detecting fake news. Furthermore, their ability to detect inconsistencies in the data can be useful for identifying potential signs of fake news.

LSTM: Long Short-Term Memory (LSTM) networks are one of the most powerful tools available for detecting fake news. LSTM networks are a type of recurrent neural network (RNN) that can store information for long periods of time and are adept at recognising patterns. This makes them well-suited for detecting false information, as they can quickly recognise similarities between articles and detect any discrepancies. LSTM networks have been used to detect fake news by recognising the patterns in the language used in the articles. They can identify common words and phrases used in false stories as well as detect any irregularities or inconsistencies in the writing style. The LSTM network can easily distinguish between real and fake news by analysing the text for these features. In addition to detecting fake news, LSTM networks can also be used to detect the source of the news. By analysing the text for common words or phrases, the LSTM can identify which sources are most likely to be reliable and which are more likely to be unreliable. It solves the vanishing gradient problem to grab longer-term reliance [39], [67].

Bi-LSTM: Bi-LSTM (Bidirectional Long Short-Term Memory) is a deep learning algorithm that is capable of detecting fake news. Bi-LSTM is a type of recurrent neural network that considers both the past and future context of a given time step [40], [41]. This makes it ideal for detecting fake news, as it is able to identify patterns in the data that are indicative of false information. Bi-LSTM works by leveraging two separate sub-networks that each process the data in different directions. The first sub-network processes the data from left to right, while the second sub-network processes the data from right to left. This allows the network to capture information from both directions. The result is a deep network that can detect both subtle and complex patterns in the data. One advantage of using Bi-LSTM for fake news detection is its ability to capture more complex patterns than traditional machine-learning techniques. For instance, it can detect patterns in the data that may be too nuanced for a human to detect. Additionally, Bi-LSTM is able to process large amounts of data quickly and accurately. This makes it particularly useful for detecting fake news, which often involves large amounts of data.

CNN: In recent years, CNN models have gained much attention in detecting fake news. Leveraging the ability of convolutional neural networks to extract features from images, these models are able to detect subtle differences between real and fake news. This allows them to identify signs of potential falsehoods in news stories based on features such as language, visual content, and other factors. Additionally, CNN models can analyse large amounts of data quickly and accurately. This is especially useful in large-scale fake news detection tasks, where they can work through hundreds of articles in a short amount of time. Their accuracy and speed make them popular for organisations and researchers looking to detect and address the spread of false information.

Different researchers have used CNN, RNN and LSTM by mixing their abilities for better results. For example, [39]

proposed a hybrid approach combining CNN and RNN for fake news detection. Their results were significantly better than when they used individual traditional basic structures with CNN for local feature extraction and LSTM for long-term reliance. In [40], they first developed models based on CNN and Bi-LSTM. Then the illustrations conveyed from these two standards are provided in a Multi-layer Perceptron Model (MLP) for the complete classification. Another research [41] utilised many architectures for detecting fake news, named CNN, Bi-LSTM, and Residual Network, integrated with pre-experienced word embedding. According to the result, the Bi-LSTM architecture exceeded CNN and ResNet on every tested dataset.

Other Approaches: Text Image Convolutional Neural Network (TI-CNN) [42] is a methodology to detect fake news where researchers consider both text and image contents to detect fake news. No previous work considered this combination of text and image contents to detect fake news like them. They also considered explicit and latent features a nice area, after which they used intellectual features to recognise fake news. For extracting latent features, they utilised two parallel CNNs. The multi-level convolutional neural network (MCNN) presented in [43] utilised a procedure of estimating the consequence of sensitive words (TFW) and developed MCNN-TFW as a fake news detection system. They achieved 88.82% and 90.10% accuracy on Weibo and NewsFN datasets, respectively. Propagation Graph Neural Network (PGNN) [3] aims to solve rumour detection tasks and can bargain strong representations for each node in the propagation graph. It continuously revises node representations within a limited time by swapping data between the neighbour nodes via relation paths. They performed significantly better than other state-of-art methods. Cross-modal Attention Residual and Multichannel convolutional neural Networks (CARMN) [44] pull the relevant data connected to a prey paradigm with another reference paradigm and preserves the prey paradigm's details. The experiment result showed better results than other recent techniques and retained better additional variant feature presentations.

2) TRADITIONAL MACHINE LEARNING-BASED TECHNIQUES

Machine learning techniques have been used to solve fake news detection problems in both supervised and unsupervised ways. Many researchers have used basic machine learning techniques and have some modifications in them.

SVM: Support Vector Machines (SVMs) have been used to detect the presence of fake news by using features such as the text's writing style, the news source, and the article's content. The basic approach of SVM for fake news detection involves feeding in the text to be analysed and then measuring the relevant features. Once the features have been identified, the SVM can classify the text as real or fake news [45]. They can process large amounts of data quickly and accurately, and they are also able to make predictions in real time. Additionally, SVMs identify complex relationships between different features, allowing for more precise detection of fake news.

NB: Naive Bayes is a machine learning algorithm that is based on Bayes theorem, which states that given a hypothesis and some evidence, the likelihood of the hypothesis being true can be calculated. When detecting fake news, Naive Bayes can analyse a large amount of data and identify patterns that may indicate false information [46]. For example, it can identify words and phrases often associated with false news stories and identify them in each article. It can also look at the writing style of an article and compare it to other sources to see how similar it is. It can also look at the overall sentiment of the article and determine if it is likely to be factual or false.

LR: Logistic Regression is a powerful tool that can be used to detect fake news. It can be used to detect fake news by analysing the characteristics of the input. It looks at elements such as the source, the headline, the content, the author, and other features that may indicate the likelihood that the news is fake. By analysing these characteristics, the algorithm can calculate the probability that the news is fake. The advantage of using Logistic Regression to detect fake news is that it can make an accurate decision regardless of the size of the dataset. This makes it particularly useful for dealing with large datasets of news articles. In [47], they state that logistic regression has the second-best classification accuracy after SVM for fake news detection.

DT: Decision Tree is an effective machine learning algorithm that can be used to detect fake news. It has the ability to analyse large amounts of data quickly and accurately, making it a valuable tool in the fight against misinformation. One of the key features of decision tree algorithms is their ability to identify relationships between data points. For instance, it can look at the source of the news article, the language used, the content, and the context in which it is presented. By analysing these factors, the algorithm can determine whether the article will likely be authentic. The decision tree can also be used to identify particular types of fake news. For example, it can be used to detect stories that have been created to manipulate public opinion or influence elections. It can also detect stories treated to spread fear or hatred. The decision tree can also be used to identify the actors behind fake news. This can help to identify the sources of fake news and take steps to limit their influence. In [48], they state that the decision trees method shows a better result than SVM, with an adequate accuracy of 95 per cent for fake news detection.

RFC: Random Forest is a versatile classifier that can detect fake news. When using Random Forest to detect fake news, the algorithm first looks at the news article's content. It then looks at the style of writing, the tone, and the source of the article. It then evaluates all of the features and combines them to form a prediction about whether or not the news article is fake. Random Forest can also detect bias in the data by looking at the data points selected for the decision tree. Random Forest can detect fake news with high accuracy [49]. It is also a fast, efficient and easy way to identify fake news.

The Random Forest algorithm is also less prone to overfitting, which can generalise well to unseen data. This makes it an effective tool for detecting fake news.

BM: Bayesian modelling uses probability theory to identify and analyse patterns in data and draw conclusions from them. It is well suited to the task of identifying fake news because it can be used to identify the likelihood of news stories being false, based on their content. For example, if a piece of news appears to contain false information, it may be associated with certain phrases or words that are commonly used in fake news stories. With Bayesian modelling, these patterns can be identified and used to develop a model that identifies and classifies stories. Additionally, Bayesian priors are based on the probability of an event occurring, which in the context of fake news allows researchers to measure how likely a piece of news is to be true. For example, if a piece of news is reported by a well-known media outlet, researchers may assign a higher probability of that news being true than if it was reported by an unknown source. Similarly, when multiple reliable sources report a piece of news, researchers may assign a higher probability of the news being true than if it was reported by only a single source. Many modelling techniques, including deep learning algorithms, that use frequencies or standard conditional probability concepts cannot take advantage of such prior information or knowledge. Moreover, Bayesian methods can be used to estimate the optimal neural network distribution for fake news [50]. Using the Bayesian strategy, we can be certain or against our hypothesis rather than coming to a single conclusion [51].

A combination of techniques proposed for fake news detection is described in Table 2. Every research has a strategy to explain its methodology for detecting fake news, which is unique compared to the others. For instance, TI-CNN [42] is a methodology to detect fake news where researchers consider the text and image contents detected as fake news. No previous work considered this combination of text and image contents to detect fake news like them. They also considered explicit and latent features as a nice feature area, after which they used intellectual features to recognise fake news. For extracting latent features, they utilised two parallel CNNs. The experiment result shows that TI-CNN successfully detects fake news established on the latent features and the explicit features comprehended from the convolutional neurons.

V. DISCUSSION

These techniques can detect fake news but are not 100% accurate because of some limitations. To assess the usefulness of each method and compare the procedures between them, the widely accepted benchmark datasets must be opened. Information from diverse sources is rarely included in most datasets. In some information sets, class allocation was imbalanced [75], influencing the classification's general performance [52]. According to the study, most investigators focus on the process and neglect data pre-processing. The preprocessing of a large amount of data and selecting features are challenging. A multimodal approach to fake news investigation is still needed. Currently, studies are limited to a single social network. Acquiring deep knowledge and early detection of fake news can be achieved by investigating fake news among different disciplines, websites, issues, and speeches [52]. The complexity and variety of online transmission data in social media make it difficult to detect fake news accurately [53]. A variety of data can be distributed very quickly due to the superactivity of users on social media. The negative impact generated by fake news has grown over time, making it essential to recognise [3], [7], [8], [11], [31]. There may be insufficient aids for achieving new perceptions and building models that are capable of correctly operating in real-world settings [11], [54]. For the modality to be accepted independently, it must be more flexible.

Further research can be done to minimise the limitations of fake news detection, and new strategies can be invented for higher accuracy. For this concern, Bayesian modelling can be a promising approach among these methods for further study. Because Bayesian modelling is a very robust method for its extensively flexible characteristics. So, it could be suitable for use in big data contexts and any field.

VI. CONCLUSION

Reviewing relevant literature for fake news analysis methods is the main focus of this paper. In a comprehensive review, we found that the overall methodology can be separated into two categories: study perspectives and fake news detection techniques. Fake news can be studied from four perspectives, and detecting fake news can be done manually or automatically. Manual fact-checks fall into two categories: expert-based and crowd-sourced. Many researchers have used data science techniques to detect fake news automatically. The main techniques used for detecting fake news are deep learning techniques such as CNN, RNN, LSTM, and Bi-LSTM. Fake news can be detected using SVM, NB, LR, DT, RFC, and BM, among other traditional machine learning techniques. As a future study method, Bayesian modelling could be the most promising. Due to the fact that it updates the prior distribution every time new data is received, Bayesian modelling is a very robust method. This makes it suitable for use in any field, including big data contexts.

REFERENCES

[1] S. Kumar, R. West, and J. Leskovec, "Disinformation on the web: Impact, characteristics, and detection of Wikipedia hoaxes," in *Proc. 25th Int. Conf. World Wide Web*, Apr. 2016, pp. 1–12.

[2] S. Zannettou, M. Sirivianos, J. Blackburn, and N. Kourtellis, "The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans," *J. Data Inf. Quality*, vol. 11, no. 3, pp. 1–37, Sep. 2019, doi: 10.1145/3309699.

[3] Z. Wu, D. Pi, J. Chen, M. Xie, and J. Cao, "Rumor detection based on propagation graph neural network with attention mechanism," *Exp. Syst. Appl.*, vol. 158, Nov. 2020, Art. no. 113595, doi: 10.1016/j.eswa.2020.113595.

[4] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *J. Econ. Perspect.*, vol. 31, no. 2, pp. 36–211, 2017.

[5] C. Dewey, "Facebook has repeatedly trended fake news since firing its human editors," Washington, DC, USA, Tech. Rep., 2016.

[6] D. Condon. (2017). *Fake News Impacting Vaccination Rates*. [Online]. Available: https://www.irishhealth.com/article.html?id=25780
[7] X. Zhou and R. Zafarani, "Fake news: A survey of research, detection methods, and opportunities," 2018, *arXiv:1812.00315v1*.
[8] X. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Comput. Surveys*, vol. 53, no. 5, pp. 1–40, Sep. 2021.
[9] Q. Su, M. Wan, X. Liu, and C.-R. Huang, "Motivations, methods and metrics of misinformation detection: An NLP perspective," *Natural Lang. Process. Res.*, vol. 1, nos. 1–2, pp. 1–13, Jul. 2020, doi: 10.2991/nlpr.d.200522.001.
[10] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newslett.*, vol. 19, no. 1, pp. 22–36, Sep. 2017, doi: 10.1145/3137597.3137600.
[11] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Inf. Sci.*, vol. 497, pp. 38–55, Sep. 2019, doi: 10.1016/j.ins.2019.05.035.
[12] N. K. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, Jan. 2015, doi: 10.1002/pra2.2015.145052010082.
[13] L. Cui and D. Lee, "CoAID: COVID-19 healthcare misinformation dataset," 2020, *arXiv:2006.00885*.
[14] V. L. Rubin, Y. Chen, and N. K. Conroy, "Deception detection for news: Three types of fakes," *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, Jan. 2015, doi: 10.1002/pra2.2015.145052010083.
[15] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, Mar. 2018, doi: 10.1126/science.aao2998.
[16] A. Gupta, N. Kumar, P. Prabhat, R. Gupta, S. Tanwar, G. Sharma, P. N. Bokoro, and R. Sharma, "Combating fake news: Stakeholder interventions and potential solutions," *IEEE Access*, vol. 10, pp. 78268–78289, 2022, doi: 10.1109/ACCESS.2022.3193670.
[17] M. M. U. Rony, M. Yousuf, and N. Hassan, "A large-scale study of social media sources in news articles," 2018, *arXiv:1810.13078v1*.
[18] C. Silverman and J. Singer-Vine, "Most Americans who see fake news believe it, new survey says," BuzzFeed News, Tech. Rep., 2016.
[19] A. Perrin, "Social media usage," Pew Res. Center, Washington, DC, USA, Tech. Rep., 2015.
[20] E. Shearer and J. Gottfried, "News use across social media platforms," Pew Res. Center, Washington, DC, USA, Tech. Rep., 2017.
[21] D. Fallis, "A conceptual analysis of disinformation," Tech. Rep., 2009.
[22] B. Skyrms, *Signals: Evolution, Learning, and Information*. Oxford, U.K.: Oxford Univ. Press, 2010.
[23] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: Characterizing and identifying fake images on Twitter during hurricane sandy," in *Proc. 22nd Int. Conf. World Wide Web Companion*. New York, NY, USA: ACM, 2013, pp. 729–736, doi: 10.1145/2487788.2488033.
[24] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Commun. ACM*, vol. 59, no. 7, pp. 96–104, Jun. 2016, doi: 10.1145/2818717.
[25] S. Kumar, J. Cheng, J. Leskovec, and V. S. Subrahmanian, "An army of me: Sockpuppets in online discussion communities," in *Proc. 26th Int. Conf. World Wide Web*, Apr. 2017, pp. 1–7, doi: 10.1145/3038912.3052677.
[26] N. Shah, H. Lamba, A. Beutel, and C. Faloutsos, "The many faces of link fraud," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2017, pp. 1069–1074, doi: 10.1109/ICDM.2017.140.
[27] V. S. Subrahmanian, A. Azaria, S. Durst, V. Kagan, A. Galstyan, K. Lerman, L. Zhu, E. Ferrara, A. Flammini, and F. Menczer, "The DARPA Twitter bot challenge," *Computer*, vol. 49, no. 6, pp. 38–46, Jun. 2016, doi: 10.1109/MC.2016.183.
[28] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, "Everyone's an influencer: Quantifying influence on Twitter," in *Proc. 4th ACM Int. Conf. Web search data mining*, Feb. 2011, pp. 1–4, doi: 10.1145/1935826.1935845.
[29] J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec, "Can cascades be predicted?" in *Proc. 23rd Int. Conf. World wide web*, Apr. 2014, pp. 925–936, doi: 10.1145/2566486.2567997.
[30] C. Shao, V. Ciampaglia, and A. Flammini, "The spread of fake news by social bots," 2017, *arXiv:1707.07592v1*.
[31] S. Kumar and N. Shah, "False information on web and social media: A survey," 2018, *arXiv:1804.08559v1*.
[32] R. Oshikawa, J. Qian, and W. Y. Wang, "A survey on natural language processing for fake news detection," 2020, *arXiv:1811.00770v2*.
[33] S. Guha, "Related fact checks: A tool for combating fake news," 2017, *arXiv:1711.00715*.
[34] (2017). *Duke Reporter'S Lab*. [Online]. Available: https://www.reporterslab.org
[35] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Hum. Language Technol.*, 2011, pp. 1–11.
[36] V. P. Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic detection of fake news," 2017, *arXiv:1708.07104*.
[37] Y. Yao, B. Viswanath, J. Cryan, H. Zheng, and B. Y. Zhao, "Automated crowdturfing attacks and defenses in online review systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2017, pp. 1–7.
[38] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2016, pp. 3818–3824.
[39] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 1, Apr. 2021, Art. no. 100007, doi: 10.1016/j.jjimei.2020.100007.
[40] A. Roy, K. Basak, A. Ekbal, and P. Bhattacharyya, "A deep ensemble framework for fake news detection and classification," 2018, *arXiv:1811.0467*.
[41] I. K. Sastrawan, I. P. A. Bayupati, and D. M. S. Arsa, "Detection of fake news using deep learning CNN–RNN based methods," *ICT Exp.*, vol. 8, no. 3, pp. 396–408, Sep. 2022, doi: 10.1016/j.icte.2021.10.003.
[42] Y. Yang, L. Zheng, J. Zhang, Q. Cui, X. Zhang, Z. Li, and P. S. Yu, "TI-CNN: Convolutional neural networks for fake news detection," 2018, *arXiv:1806.00749*.
[43] Q. Hu, Q. Li, Y. Lu, Y. Yang, and J. Cheng, "Multi-level word features based on CNN for fake news detection in cultural communication," *Pers. Ubiquitous Comput.*, vol. 24, no. 2, pp. 259–272, Apr. 2020, doi: 10.1007/s00779-019-01289-y.
[44] C. Song, N. Ning, Y. Zhang, and B. Wu, "A multimodal fake news detection model based on crossmodal attention residual and multichannel convolutional neural networks," *Inf. Process. Manage.*, vol. 58, no. 1, Jan. 2021, Art. no. 102437.
[45] J. Chiu, A. Gokcen, W. Wang, and X. Yan, "Classification of fake and real articles based on support vector machines," in *Proc. Language Statist. Spring*, 2013, pp. 1–6.
[46] M. Granik and V. Mesyura, "Fake news detection using naive Bayes classifier," in *Proc. IEEE 1st Ukraine Conf. Electr. Comput. Eng. (UKRCON)*, May 2017, pp. 900–903, doi: 10.1109/UKRCON.2017.8100379.
[47] N. Islam, A. Shaikh, A. Qaiser, Y. Asiri, S. Almakdi, A. Sulaiman, V. Moazzam, and S. A. Babar, "Ternion: An autonomous model for fake news detection," *Appl. Sci.*, vol. 11, no. 19, p. 9292, Oct. 2021, doi: 10.3390/app11199292.
[48] S. Lyu and D. C.-T. Lo, "Fake news detection by decision tree," in *Proc. SoutheastCon*, 2020, pp. 1–2, doi: 10.1109/SoutheastCon44009.2020.9249688.
[49] C.-M. Lai, M.-H. Chen, E. Kristiani, V. K. Verma, and C.-T. Yang, "Fake news classification based on content level features," *Appl. Sci.*, vol. 12, no. 3, p. 1116, Jan. 2022, doi: 10.3390/app12031116.
[50] M. Sahan, V. Smidl, and R. Marik, "Batch active learning for text classification and sentiment analysis," in *Proc. 3rd Int. Conf. Control, Robot. Intell. Syst.*, New York, NY, USA, Aug. 2022, pp. 1–7, doi: 10.1145/3562007.3562028.
[51] [Online]. Available: https://learning.edanz.com/frequentist-bayesian-statistics/#:~:text=The%20main%20advantage%20of%20Bayesian,no%20predefined%20set%20of%20priors
[52] A. R. Sana Ullaha, A. Dasa, A. Dasb, M. A. Kabir, and K. Shu, "A survey of COVID-19 misinformation: Datasets, detection techniques and open issues," 2021, *arXiv:2110.00737v1*.
[53] K. Popat, S. Mukherjee, A. Yates, and G. Weikum, "DeClarE: Debunking fake news and false claims using evidence-aware deep learning," 2018, *arXiv:1809.06416v1*.
[54] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Inf. Process. Manage.*, vol. 57, no. 2, Mar. 2020, Art. no. 102025, doi: 10.1016/j.ipm.2019.03.004.

[55] E. Tacchini, G. Ballarin, M. L. D. Vedova, S. Moret, and L. de Alfaro, "Some like it hoax: Automated fake news detection in social networks," 2017, *arXiv:1704.07506v1*.
[56] S. Hakak, M. Alazab, S. Khan, T. R. Gadekallu, P. K. R. Maddikunta, and W. Z. Khan, "An ensemble machine learning approach through effective feature extraction to classify fake news," *Future Gener. Comput. Syst.*, vol. 117, pp. 47–58, Apr. 2021, doi: 10.1016/j.future.2020.11.022.
[57] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spams and fake news using text classification," *Secur. Privacy*, vol. 1, no. 1, p. e9, Jan. 2018, doi: 10.1002/spy2.9.
[58] C. Zhou, C. Sun, Z. Liu, and F. C. M. Lau, "A C-LSTM neural network for text classification," 2015, *arXiv:1511.08630*.
[59] A. Hanselowski, B. Schiller, F. Caspelherr, D. Chaudhuri, C. M. Meyer, and I. Gurevych, "A retrospective analysis of the fake news challenge stance detection task," 2018, *arXiv:1806.05180*.
[60] E. Kochkina, M. Liakata, and A. Zubiaga, "All-in-one: Multi-task learning for rumour verification," 2018, *arXiv:1806.03713v1*.
[61] M. Dong, L. Yao, X. Wang, B. Benatallah, Q. Z. Sheng, and H. Huang, "DUAL: A deep unified attention model with latent relation representations for fake news detection," in *Proc. Int. Conf. Web Inf. Syst. Eng.*, 2018, pp. 199–209, 2018, doi: 10.1007/978-3-030-02922-7_14.
[62] M. Farajtabar, J. Yang, X. Ye, H. Xu, R. Trivedi, E. Khalil, S. Li, L. Song, and H. Zha, "Fake news mitigation via point process based intervention," 2017, *arXiv:1703.07823v2*.
[63] D. Paschalides, D. Stephanidis, A. Andreou, K. Orphanou, G. Pallis, M. D. Dikaiakos, and E. Markatos, "MANDOLA: A big-data processing and visualization platform for monitoring and detecting online hate speech," *ACM Trans. Internet Technol.*, vol. 20, no. 2, pp. 1–21, May 2020, doi: 10.1145/3371276.
[64] V. Ilie, C. Truica, E. Apostol, and A. Paschke, "Context-aware misinformation detection: A benchmark of deep learning architectures using word embeddings," *IEEE Access*, vol. 9, pp. 162122–162146, 2021.
[65] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh, "SpotFake: A multi-modal framework for fake news detection," in *Proc. IEEE 5th Int. Conf. Multimedia Big Data (BigMM)*, Dec. 2019, pp. 1–18.
[66] V. Mazzeo, A. Rapisarda, and G. Giuffrida, "Detection of fake news on COVID-19 on web search engines," *Frontiers Phys.*, vol. 9, Jun. 2021, Art. no. 685730, doi: 10.3389/fphy.2021.685730.
[67] S. R. Sahoo and B. B. Gupta, "Multiple features based approach for automatic fake news detection on social networks using deep learning," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106983, doi: 10.1016/j.asoc.2020.106983.
[68] H. Jwa, D. Oh, K. Park, J. M. Kang, H. Lim, "exBAKE: Automatic fake news detection model based on bidirectional encoder representations from transformers (BERT)," *Appl. Sci.*, vol. 9, no. 19, p. 4062, 2019, doi: 10.3390/app9194062.
[69] M. A. Alonso, D. Vilares, C. Gómez-Rodríguez, and J. Vilares, "Sentiment analysis for fake news detection," *Electronics*, vol. 10, no. 11, p. 1348, Jun. 2021, doi: 10.3390/electronics10111348.
[70] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014.
[71] P. Meel and D. K. Vishwakarma, "Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities," *Exp. Syst. Appl.*, vol. 153, Sep. 2020, Art. no. 112986, doi: 10.1016/j.eswa.2019.112986.
[72] A. M. Luvembe, W. Li, S. Li, F. Liu, and G. Xu, "Dual emotion based fake news detection: A deep attention-weight update approach," *Inf. Process. Manage.*, vol. 60, no. 4, Jul. 2023, Art. no. 103354, doi: 10.1016/j.ipm.2023.103354.
[73] S. V. Balshetwar and A. Rs, "Fake news detection in social media based on sentiment analysis using classifier techniques," *Multimedia Tools Appl.*, Mar. 2023, doi: 10.1007/s11042-023-14883-3.
[74] J. M. Pérez, F. M. Luque, D. Zayat, M. Kondratzky, A. Moro, P. S. Serrati, J. Zajac, P. Miguel, N. Debandi, A. Gravano, and V. Cotik, "Assessing the impact of contextual information in hate speech detection," *IEEE Access*, vol. 11, pp. 30575–30590, 2023, doi: 10.1109/ACCESS.2023.3258973.
[75] N. Capuano, G. Fenza, V. Loia, and F. D. Nota, "Content-based fake news detection with machine and deep learning: A systematic review," *Neurocomputing*, vol. 530, pp. 91–103, Apr. 2023, doi: 10.1016/j.neucom.2023.02.005.

MEHEDI TAJRIAN (Member, IEEE) was born in Dhaka, Bangladesh, in 1993. She received the B.S. degree in information technology from the University of Information Technology and Sciences, Dhaka, in 2016, and the M.S. degree in computer science engineering from Ajou University, Suwon, South Korea, in 2019. She is currently pursuing the Ph.D. degree in data science with Charles Sturt University, Wagga Wagga, NSW, Australia. She was a Researcher with the Cyber Security Laboratory, Ajou University, from 2017 to 2019. She is researching with the Data Mining Research Group, CSU. Her research interests include data science, machine learning, deep learning, the IoT, scheduling algorithms, blockchain, cyber security, network security, and cloud computing.

AZIZUR RAHMAN received the Ph.D. degree. He is currently an Associate Professor of applied statistician and data scientist with expertise in both developing and applying novel methodologies, models and technologies. He is also the Leader of the "Statistics and Data Mining Research Group," Data Science Research Unit, Faculty of Business, Justice and Behavioural Sciences, Charles Sturt University. He is able to assist in understanding multi-disciplinary research issues within various fields, including how to understand the individual activities which occur within very complex scientific, behavioral, socio-economic, and ecological systems.

MUHAMMAD ASHAD KABIR (Member, IEEE) received the Ph.D. degree in computer science from the Swinburne University of Technology, Melbourne, Australia. He is currently the Deputy Leader of the Data Mining Research Group and an Associate Professor with the School of Computing Mathematics and Engineering, Charles Sturt University, Australia. He has published over 100 peer-reviewed articles. His research interests include data mining, data analytics and visualization, blockchain and security, smart mobile applications, health informatics, human–computer interactions, adaptive, and context-aware software.

RAFIQUL ISLAM (Senior Member, IEEE) is currently an Associate Professor with the School of Computing, Mathematics and Engineering, Faculty of Business, Justice and Behavioral Sciences, Charles Sturt University, Australia. He is leading the Cybersecurity Research Team and has developed a strong background in leadership, sustainability, and collaborative research. He has a strong publication record and has published more than 190 peer-reviewed research articles, book chapters, and books. He has a strong research background in cybersecurity with a specific focus on malware analysis and classification, authentication, security in the cloud, privacy in social media, and the Internet of Things (IoT). His contribution is recognized both nationally and internationally by achieving various rewards, such as professional excellence, research excellence, and leadership awards.