International Journal of Multimedia Information Retrieval (2023) 12:28
https://doi.org/10.1007/s13735-023-00296-3

REGULAR PAPER

A comprehensive survey of multimodal fake news detection techniques: advances, challenges, and opportunities

Shivani Tufchi[1] · Ashima Yadav[1] · Tanveer Ahmed[1]

Received: 17 March 2023 / Revised: 9 July 2023 / Accepted: 27 July 2023 / Published online: 23 August 2023

Abstract

The escalating prevalence of disinformation, or "fake news," on social media platforms represents a growing societal concern with far-reaching implications. Its ubiquitous presence across platforms such as Facebook, Twitter, and Instagram exacerbates the criticality of this issue. Consequently, the development of efficacious counter strategies—such as fact-checking mechanisms and media literacy initiatives—is paramount. This domain of crafting robust defensive strategies against disinformation has burgeoned into an expansive field of investigation, with primary emphasis on the identification and categorization of fake news within the digital media landscape. This study offers a comprehensive survey of the contemporary research landscape on disinformation, or "fake news," detection, and mitigation strategies. We meticulously analyze the life cycle of disinformation—from inception and propagation to detection—shedding light on both supervised and unsupervised learning techniques such as generative adversarial networks (GANs). The paper provides a comparative analysis of different classification models across a variety of text- and image-based datasets related to fake news. Furthermore, it provides an in-depth discussion of the key evaluation metrics used in assessing the accuracy of news authenticity. Challenges have been put forth concerning the sophisticated task of precisely detecting fake news, an undertaking obscured by its subtle and often deceptive aspects. By offering in-depth insights into the phenomenon of fake news and potential counterstrategies, this survey is anticipated to enrich scholarly understanding, thereby catalyzing the development of innovative solutions to tackle this persistent global issue.

Keywords Fake news · Fact-checking · Deep learning · Handmade features · Adversarial networks · Attention mechanism

1 Introduction

The sphere of technology is undergoing precipitous advancements, with an associated parallel growth in the realm of social media. Empirical data, procured by a comprehensive investigation led by Smart Insights [20], demonstrates that a significant majority—surpassing the semi-centennial percentile of the global populace—are engaging in habitual utilization of these social media interfaces, with usage patterns potentially extending into excess as indicated in Fig. 1.

While social media is a potent instrument, capable of catapulting businesses to unprecedented success levels, it is not devoid of substantial disadvantages. The digital universe encompasses a spectrum of grave and pressing issues, spanning from cybercrime and cyberbullying to digital scams and manipulations, among others. A particularly worrisome concern is the rampant propagation of inauthentic news [44, 126, 172], which has evolved into a critical societal problem impacting diverse societal sectors [37].

False or misleading news, or "fake news," is any media content that is contrived or purposely deceptive. While this form of spurious journalism has a long history, the term "fake news" has gained prominence recently, owing to its widespread incidence in the digital epoch. This phenomenon poses a significant challenge in the media landscape as it can be exploited to manipulate public sentiment, disseminate misinformation, and even meddle in political elections.

Ashima Yadav and Tanveer Ahmed have been contributed equally to this work.

✉ Shivani Tufchi
e21soep0017@bennett.edu.in

Ashima Yadav
ashima.yadav@bennett.edu.in

Tanveer Ahmed
tanveer.ahmed@bennett.edu.in

1 Computer Science, Bennett University, Greater Noida, UP 201310, India

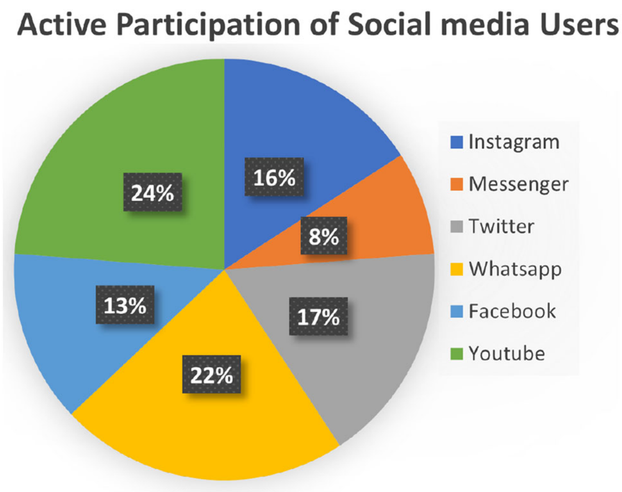


Fig. 1 The engagement of individuals on social media platforms

Its primary objective is to distort, deceive, or manipulate individuals' beliefs and opinions.

A notable manifestation of this issue was witnessed during the 2016 U.S. Presidential election [125]. During this political event, fabricated narratives and conspiracy theories were liberally disseminated via social media platforms. Among the numerous misleading narratives, the "Pizzagate" conspiracy [95] gained notoriety. This unfounded theory, alleging that Hillary Clinton and her affiliates orchestrated a child trafficking network from a pizzeria in Washington D.C., gained significant online traction, leading to a weaponized individual investigating the claim on-site. Another instance of misleading news can be seen in the "Yellow Vest" protests [149] in France, sparked by false narratives concerning immigration, crime, and economics. In this instance, inauthentic stories were intentionally disseminated via social media to manipulate public opinion and provoke violence. These instances underscore the severe repercussions of fake news, including fostering mistrust in media and public institutions, tarnishing reputations, and inciting violence. Overall, fake news is a significant concern with potentially serious implications. Individuals must stay informed of the potential risks posed by fake news and take necessary precautions to verify the authenticity of any news content they encounter. By doing so, they can play an integral role in ensuring that the truth is not overshadowed by falsehoods and misinformation [172].

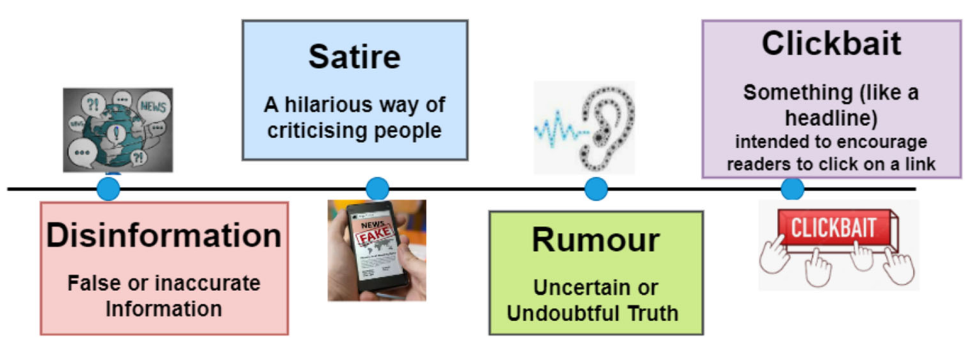
False information disseminated in the media can manifest in various forms, including satire, rumor, clickbaits, and misinformation [76, 108], as illustrated in Fig. 2. Satire, typically steeped in humor, serves to underscore salient points about specific issues and is often utilized to critique societal and political dynamics. Rumors pertain to invalidated narratives or assertions that gain considerable circulation, often in the form of gossip.

Clickbaits, on the other hand, consist of headlines or titles engineered to garner attention and provoke user interaction via "clicks." These are frequently characterized by sensationalism and lack empirical backing. Misinformation constitutes false or inaccurate data that is purposely disseminated. In light of these diverse modalities of false information, it becomes imperative for individuals to exercise due diligence regarding the information they consume, placing reliance only on credible and validated information sources. Implementing such a discerning approach can mitigate the risk of being misled by such fallacious information.

Figure 3 presents a few instances of such misleading images. For instance, Fig. 3a refutes the erroneous claim that COVID-19 vaccination can be transferred via blood transfusion, a myth debunked by a leading fact-checking authority, Poltifact.com. Implementing rigorous critical evaluation of information sources, coupled with a reliance on proven fact-checking resources, is pivotal in navigating the current information ecosystem. Figure 3b exhibits an image originally associated with an aircraft accident that transpired in Nepal in 2011, inaccurately presented as an illustration of a purported aircraft incident in Nepal in 2023. This falsification was subsequently rectified by The Quint [21], a reputable fact-checking entity based in India.

In the ongoing battle against fake news, unimodal text-based analysis has been widely adopted, demonstrating its value through its effective use of Natural Language Processing (NLP) to detect linguistic patterns that suggest false news content. This fact reinforces the utility of NLP in discerning

Fig. 2 The various shades of fake news



(a) (b)

Fig. 3 Examples showing fake news. **a** Fake image showing blood transfusion does not transfer COVID-19 vaccine. **b** Second fake image shows an old image of a plane crashed in place of an accident that happened in 2023 in Nepal

deceptive information. To further elaborate, Fig. 4 delineates earlier tried-and-tested techniques for text preprocessing and feature extraction. A precursor to this was the application of Part-of-Speech (POS) tagging for feature selection. This technique was followed by the application of Naïve Bayes (NB) and random forest (RF) models for fake news classification. Empirical results from this study illustrated the effectiveness of POS tagging and the subsequent application of NB and RF models in classifying fake news [63].

To improve model efficiency, an optimized feature selection strategy was introduced, specifically designed to decrease the number of input variables. This strategy combined chi2, univariate, and information gain as feature selection techniques with various machine learning (ML) methodologies. The resultant enhancement in classification accuracy substantiates the benefits of this optimized feature selection strategy [41]. The employment of two well-known feature selection methods, Bag of Words and TF-IDF vectorization, was observed in another study. These techniques, coupled with ML models such as passive aggressive classifier, NB, LR, and RF, led to a noticeable improvement in the classification of fake news, substantiating the efficacy of these models in this context [124].

Additional feature selection methods, such as N-grams and Simple Bag-of-Words, along with term frequency techniques like TF-IDF weighting, were employed. The observed improvements in fake news detection provide empirical evidence for the utility of these feature selection techniques [25]. Advancements in feature extraction techniques have significantly improved the performance in fake news classification. Beyond traditional methods such as Bag of N-grams, Count Vectorizer, and Tfidf Vectorizer, recent approaches leverage transformer-based models such as bidirectional encoder of representations from transformers (BERT), GPT, and RoBERTa, which generate contextualized word embeddings that capture both the semantic meaning of words and their context within a sentence. These models offer a richer feature set, leading to substantial improvements in classification accuracy. Also, techniques like ELMo and the Universal Sentence Encoder (USE) generate embeddings that carry the semantic meanings of entire sentences or paragraphs, further enhancing the accuracy and efficiency of fake news detection [2]. The use of Word2Vec [23], an efficient feature extraction method, in conjunction with other text features significantly enhanced fake news detection performance. This efficient way of extracting features has been corroborated by a comparative analysis that used a deep convolutional neural network (CNN) to extract various features across each layer. The observed improvements in classification accuracy provided empirical evidence for the efficacy of these feature extraction techniques [59, 110].

In response to the escalating issue of fake news dissemination, advancements in classification methodologies have been made across both text- and image-based modalities. To support this, a novel semi-supervised method built on the foundational principles of GANs has been introduced. This technique is primarily designed for synthetic fake image creation, demonstrating its efficacy in classifying counterfeit images, providing evidence for the importance of GANs in this realm [82].

Several deep learning (DL) algorithms, renowned for their performance in various applications, have also found their place in fake news detection. Notably, Bert has been employed for textual analysis, presenting a significant improvement in the identification of false narratives [2]. Similarly, ResNet's application in image processing has shown remarkable progress in deciphering manipulative visual content [2]. To manage the interplay between different modalities, attention mechanisms have been successfully utilized, thereby establishing their potential in dealing with inter-modality and intra-modality relationships [101].

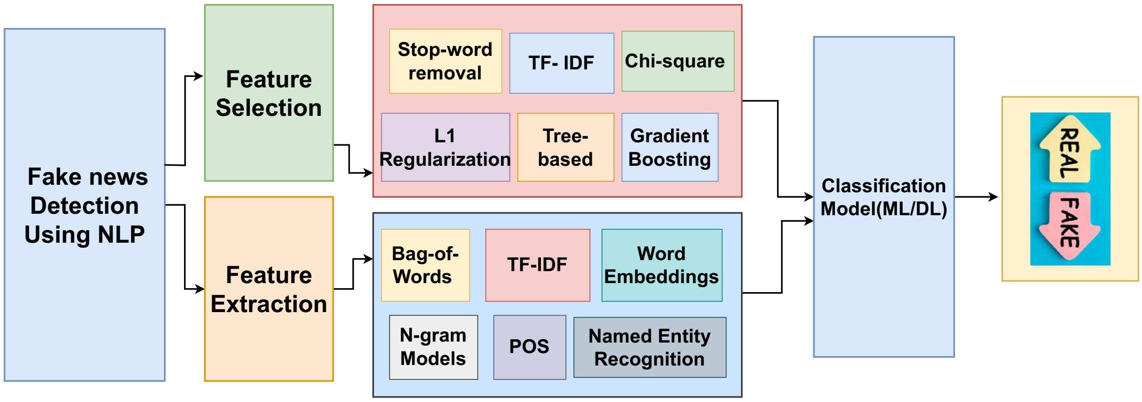


Fig. 4 Utilization of diverse feature extraction techniques for text-based classification

In the specific area of image classification for fake news detection, pre-trained CNN models, such as ResNet, ImageNet, and InceptionNet, have demonstrated their robustness. They significantly improve the accuracy of fake image detection, further solidifying their indispensability in this area [101].

A multimodal framework, a more comprehensive approach to fake news classification, effectively combines EfficientNetB0 for image analysis and RoBERTa for text processing. This hybrid model has exhibited a superior performance, outperforming the standalone counterparts in their respective domains [130]. Furthermore, the advent of a sophisticated multimodal fusion network (FMFN) signifies another leap in this field. The FMFN model employs CNNs to extract a wide array of image features and uses RoBERTa to produce deep contextualized word embeddings, offering a broader and deeper understanding of the content being analyzed [151].

Beyond these models, the supremacy of BERT, XLNET, and GPT-3 across various natural language processing (NLP) tasks has been reaffirmed by [85]. This provides empirical evidence for the utility of these algorithms in text-based fake news detection. Similarly, the attention-based recurrent neural network (att-RNN) has been acknowledged as the state-of-the-art model for multimodal fake news detection, enabling an efficient fusion of textual, visual, and social context features. Its superior performance over traditional models underscores its potential as a comprehensive solution for fake news detection [55].

Lastly, the use of the event adversarial neural networks (EANN) approach has provided a fresh perspective on multimodal fake news detection strategies. The method, combining the strengths of Text CNN and VGG-19 for image analysis, has shown a significant improvement in discerning fake news, further demonstrating the effectiveness of combining multiple techniques for fake news detection [154]. Our survey provides an up-to-date and comprehensive look at the available research on false news detection with deep learning architectures and attention mechanism.

1.1 Motivation and key contribution

The remarkable expansion of digital and social media landscapes has simultaneously led to a substantial increase in the spread of fraudulent news, making its detection and mitigation a complex issue. Given the gravity of this situation, we are motivated to write an extensive survey paper that concentrates on multimodal fake news detection strategies. Our motivation stems from the aspiration to gather and explicate a variety of methodologies that utilize multimodal techniques in the scrutiny of fake news content.

This comprehensive survey contributes to the existing body of knowledge in the following unique ways:

(1) This exploration introduces a concise discourse on the phenomenon of fake news, offering insights into its characterization and the potential societal hazards it poses. A nuanced discussion encompassing diverse facets of fake news, such as misleading information, false information, disinformation, satire, clickbaits, and rumors, is incorporated.
(2) The survey furnishes a detailed elucidation of the phases involved in the lifecycle of fake news, tracing its trajectory from inception to propagation and eventual debunking. This discussion underlines how each phase contributes to the proliferation of false information and discusses potential countermeasures to curb this spread.
(3) A comprehensive exposition of the current state-of-the-art supervised and unsupervised machine learning, alongside deep learning methodologies for the detection of fake news, is presented.
(4) The study showcases benchmark fake news datasets in both textual and visual modalities, complete with

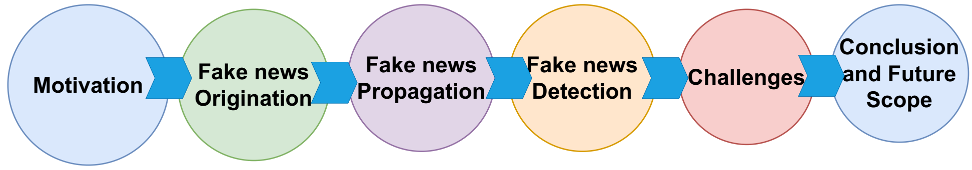


Fig. 5 Organization of the survey

their corresponding ground truths and performance metrics when applied to a variety of machine learning and deep learning models. These datasets serve as invaluable resources for ongoing and future research in the field.

(5) This survey offers an incisive explanation of the handcrafted features employed by machine learning models, simultaneously highlighting the benefits of employing deep learning models over traditional machine learning methods.
(6) The survey provides an overview of the artificial intelligence technology known as deepfake, which harnesses the power of generative adversarial networks (GANs) for the efficient detection of fake images and videos, outperforming conventional deep learning models.
(7) The transformative potential of parallel processing of words, as opposed to sequential processing, on the field of deep learning is extensively discussed.
(8) Finally, the survey identifies various challenges that remain unresolved within this field. By highlighting these obstacles, the survey aims to catalyze future research endeavors toward addressing these issues, thereby fostering societal benefits.

These distinctive contributions ensure that this survey stands as a significant addition to the scholarly literature within the field, catering to both researchers and practitioners alike.

1.2 Organization of the survey

Conducting a survey to investigate the prevalence of fake news can be a challenging undertaking. To ensure the success of our survey, we have carefully planned and organized the various sections, as depicted in Fig. 5. In Sect. 1, we delve into the concept of fake news, exploring its different types. Section 2 focuses on the origins of fake news, seeking to understand how and where it originates. The propagation of fake news is examined in Sect. 3, highlighting the various means through which it spreads. In Sect. 4, we discuss the use of handmade feature extraction for supervised and unsupervised deep learning approaches in combating fake news. Section 5 addresses the challenges posed by fake news, considering the complexities and implications associated with this phenomenon. Lastly, in Sect. 6, we present our concluding remarks and outline future directions for research in this field.

Overall, by strategically organizing our survey into these sections, we aim to comprehensively investigate the prevalence of fake news and provide valuable insights into this important issue.

2 Architecture/life cycle

This research is designed to provide a detailed examination of the life cycle of fake news, segmented into various stages that include its origination, the paths of its propagation, and finally, its detection and mitigation to limit its extensive reach. The creation, spread, and identification of fake news represent the three central elements of its detection. These elements are graphically depicted in Fig. 6.

2.1 Origination of fake news

"The proliferation of disinformation," or "fake news," has escalated alarmingly over the recent decade. It appears that no medium is immune—from social media platforms to print and broadcast media, disinformation permeates our daily lives. But where did this phenomenon originate? Interestingly, the roots of "fake news" can be traced back to the mid-nineteenth century [135]. In those times, print media was burgeoning, with the democratization of publishing enabling practically anyone to disseminate stories under the guise of factual news. This freedom to publish resulted in the propagation of sensationalized narratives and misinformation.

A case in point is the infamous 1835 "Moon Hoax" perpetuated by The New York Sun, one of the most prominent newspapers of the time [9]. It fabricated tales of exotic lunar life forms, including unicorns, bipedal beavers, and winged humanoids. These fabricated reports whipped up public frenzy, significantly boosting newspaper sales. As we entered the twentieth century, the advent of radio and television enabled the rapid broadcasting of information—true or false—to a much larger audience [144]. Fact-checking became increasingly challenging, thus providing fertile ground for the propagation of fake news.

In our contemporary era, the Internet and social media platforms have turbocharged the spread of disinformation. False narratives can now reach millions of people almost instantaneously. The deluge of information often makes it a Herculean task to sift fact from fiction. The architects of disinformation could potentially range from governments to

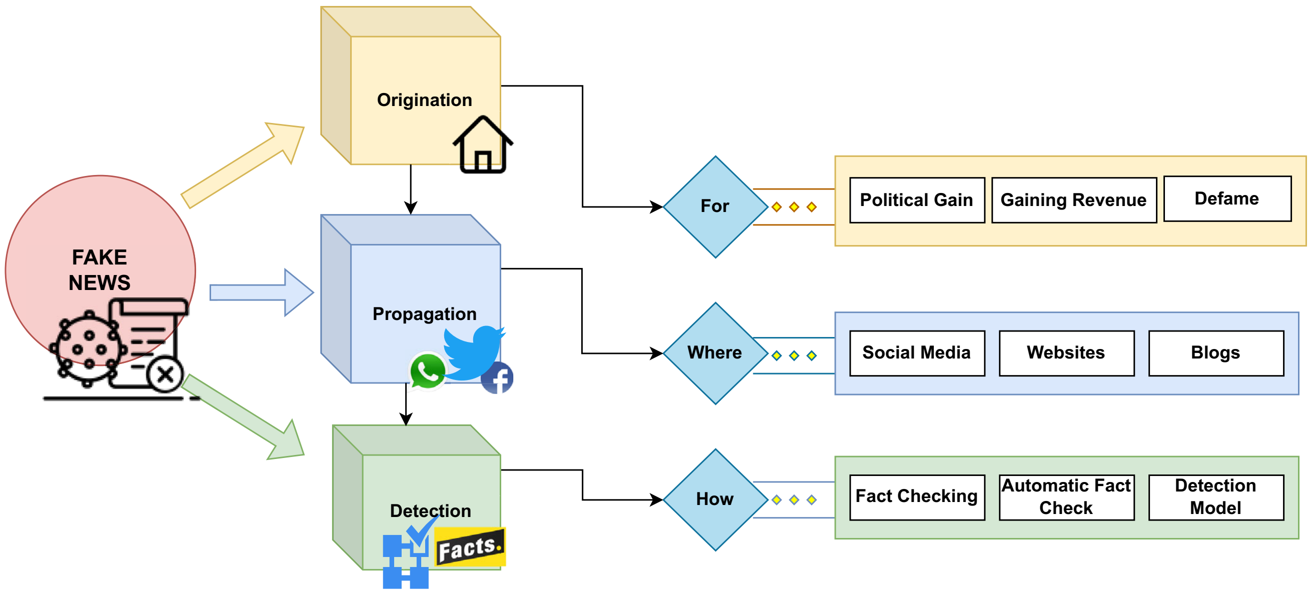


Fig. 6 Stages in the life cycle of fake news detection

corporations and other entities [60]. Hence, it is critical to remain alert and verify the information before accepting it as the gospel truth. Efforts are in place to combat this menace. One such system, named the Fuzzy Anonymous Complaint Tally System (FACT), uses a counting Bloom filter, a probabilistic data structure to track the actual origin of messages [73].

Ants are known for their search algorithms, where they leave pheromone trails to identify the shortest path to a food source. This biological model known as ant colony clustering algorithm has been adopted and transformed into a computational algorithm for detecting the source of false information [97]. It is imperative that we curtail the spread of fake news given its potential to cause significant harm. Possible countermeasures include enhancing media literacy, imposing regulations on social media platforms, and encouraging fact-checking. By adopting these strategies, we can help ensure that the public is not misled by deceptive narratives. In turn, we can contribute to a more informed, truthful societal discourse.

2.2 Current state of fact-checking

The importance of fact-checking has dramatically increased in our current society given the vast amount of accessible information. The explosion of social media has provided a platform for anyone to disseminate misinformation effortlessly, hence complicating the task of distinguishing between what is true and what is not. In light of this, fact-checking serves as a crucial weapon in the fight against misinformation. Fact-checking involves the careful examination of statements or claims to verify their truthfulness, a process thoroughly described by [85] in 2022. It entails delving into the stated claim to evaluate its veracity, a task typically undertaken by journalists, researchers, and other professionals with expertise in fact authentication.

Lately, there has been a noticeable increase in the utilization of automated systems for fact-checking tasks. Though these systems are not flawless, they have undoubtedly bolstered our capacity to counteract the propagation of false information. These automated fact-checking systems incorporate machine learning and natural language processing to scrutinize text and assess its truthfulness.

Despite the progress, an alarmingly high number of untruthful statements continue to circulate online and through various media channels. Combining the efficiency of automated systems with the precision of manual fact-checking can expedite the identification and refutation of these false claims. This dual approach aids in curtailing the propagation of false information, promoting the consumption of only reliable and factual content. Despite the room for enhancement, the current state of fact-checking is heading in the right direction. Significant strides in technology, such as Facebook's repost-verification and Google's fact-checking feature, have been developed by tech firms. These tools were created to confirm and restore the integrity of online information sources [92].

2.2.1 Manual fact-checking

Currently, more than 300 official initiatives do manual fact-checking, some of them are mentioned in Table 1, but

Table 1 Fact-checking websites

Source	Class	Type	Ref.	URL
FactCheck	Politics	Manual	[115]	https://www.factchecker.in
PolitiFact	American politics	Manual	[107]	https://www.politifact.com
Gossibcop	Entertainment stories	Manual	[128]	https://www.suggest.com
Snopes	Journalism, rumors	Manual	[71]	https://www.snopes.com
AFP	Environment, health, politics	Manual	[77]	https://factcheck.afp.com
Full Fact	Politics, Journalism	Manual	[38]	https://fullfact.org/facts
Truth or Fiction	Democracy, Journalism	Manual	[94]	https://www.truthorfiction.com
Poynter"2	Business, Tech and tools	Automatic	[147]	https://www.poynter.org/
Reporterslab	Real Time news	Automatic	[109]	https://reporterslab.org/fact-checking
ClaimBuster	Real Time news	Automatic	[174]	https://idir.uta.edu/claimbuster
Kauwa-Kaate	Real Time news	Automatic	[12]	https://download.cnet.com

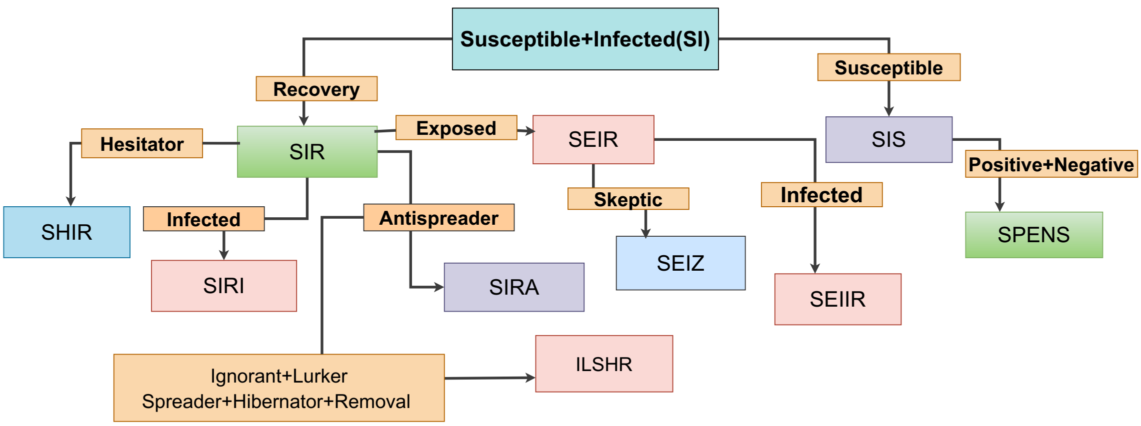


Fig. 7 Epidemiological models demonstrating spread patterns

still they are unable to keep up with the rising volume of claims that ought to be verified. The manual fact-checking process is further divided different categories as shown in Fig. 7. Crowd-sourced fact-checking relies on the collective intelligence of a large number of citizens, whereas expert-based fact-checking relies on the judgments of subject-matter experts (i.e., fact-checkers) to confirm the content of the given news item [31]. A small group of very reliable fact-checkers often conducts expert-based fact-checking, which might result in very accurate conclusions.

On the other hand, knowledge-based fact-checking [5] validates facts by researching them. It is a way of ensuring that the information presented is accurate and true. It involves verifying facts by researching various sources such as books, newspapers, journals, websites, and other reliable sources. The researcher must also take into consideration the context of the information as well as its source. Another type of fact-checking known as context-based fact-checking [42] which aims to utilize the user opinions expressed in social media posts to verify the reality and mainly focuses on (time, location) of the news created. The two groups for context based have been further separated. The first type of analysis is stance based, which seeks to ascertain if a reader is for, against, or neutral on a certain piece of news and second is propagation-based fact-checking which aims to focus on how misinformation circulates among users. In the study [5], authors have exposed content-based fact-checking which aims to focus on the content (text, images, video). Lastly, style-based fact-checking aims focuses on the style of a particular piece of writing. It is usually done by a professional editor, or by a writer themselves, and can involve the use of tools such as dictionaries, style guides, and thesauruses to ensure that all facts are accurately [16]. Several automatic tools have also been developed to differentiate the fake news automatically using algorithms based on machine learning [22], deep learning, and natural language processing aiming to automate the fact-checking process. A well-known advance technique has been presented in the study [158] that uses zero shot fact-checking by using CLAIMGEN-BART with KBIN achieving 90% accuracy on supervised learning algorithms.

2.2.2 Automatic fact-checking

The use of automated fact-checking has become increasingly popular in recent years as news organizations have sought to reduce the time and costs associated with manual fact-checking. In addition, automated fact-checking can help weed out false information and reduce the spread of misinformation. Previous research has been focused on origination, propagation, and finding the differences and similarities between the many terms associated with "fake news." Now, the present study's goal is to take things a step further by finding certain operational traits or indications that may be used to train a machine learning algorithms and a deep learning algorithms to accurately distinguish between various sorts of information that fall under the general category of fake news.

In the study, authors [148] have presented an idea to import classification models across social networking networks in the form of browser extensions or mobile apps. This approach could revolutionize how people interact with information online, introducing an additional layer of scrutiny and trust to social media content consumption. [48] uses Google's NLP API for automatically identifying fake news. Fact-checkers frequently make use of a range of tools, including news alerts, ASR, and translation software, all of which typically rely on underlying AI capabilities. The greatest level of accuracy of 99.30% has been achieved using machine learning such as k-nearest neighbor, LR, SVM, linear discriminant analysis, decision trees, and random forests classifiers. Some of the most often used deep learning models used by the authors are CNNs and LSTMs. Among these, convolutional networks have been mostly used [7].

BRENDA [14] a browser extension based on deep neural network has been used to automate the entire process of credibility assessments of false claims. Being a browser extension, BRENDA provides quick automatic fact. A standard approach has been proposed for automatic detection of fake news by incorporating two prominent tools: ClaimRank and ClaimBuster. These have been designed to identify check-worthy factual claims in text. ClaimRank uses machine learning algorithms to rank sentences in a document based on their check-worthiness, while ClaimBuster employs a supervised learning approach to detect important factual claims that need fact-checking. This method facilitates quick identification and verification of information, effectively reducing the spread of fake news. For the purpose of identifying false news, a web tool called FactFinder has been developed [91] in which the claimed sentences have been highlighted in yellow, and the words highlighted in red have been the search terms used for evidence mining. In the study, authors have [89] presented Google's Fact Check Tools API which comes with a fast AI-based fake news detector in terms of speed and accuracy for the automatic detection of fake news. A lab initiative known as CheckThat! [84] has been designed to automate the fact-checking process to help stop the spread of false and misleading information. For automatic fact-checking, in the study [87] authors have presented graph-based neural networks for classifying fake news [96]. A pre-trained model Roberta-base with an Adam optimizer has been implemented by the authors [25] for detecting fake news. In a recent study referenced by [25], a cutting-edge, fully automated fact-checking platform known as the "Holy Grail" has been introduced. This platform stands out due to its ability to efficiently detect and evaluate the veracity of a claim in real-time. The system not only identifies a potentially misleading claim as it emerges but also instantaneously provides an evaluation of its accuracy.

A mobile application has been developed by [140] which utilizes ALBERT, a language model for detecting fake news more accurately at real time as ALBERT, a derivative of BERT, is a widely recognized unsupervised language representation learning algorithm that aims to be more lightweight [161]. ALBERT incorporates parameter reduction techniques to enable the use of large-scale configurations, surpassing previous memory limitations. Additionally, it exhibits improved performance in terms of model degradation, making it a noteworthy advancement in the field.

3 Propagation

The impact of fake news can be quite severe, inducing a range of effects from mistrust in media to societal and political instability. A major medium for the propagation of fake news is social media, which is frequently utilized to circulate baseless rumors or disinformation quickly. This propagation can be deliberate, with the intent to mislead or deceive, or inadvertent, as people often share misleading information without verifying its authenticity. The structure of social media allows users to distribute fake news without assuming responsibility for its veracity, as it enables anonymous or pseudonymous sharing.

Traditional media sources also serve as conduits for the dissemination of fake news. Often, these channels circulate manufactured disinformation intended to sway public sentiment or control public perception. The main vehicles for this dissemination are newspapers, radio, and television. Identifying fake news in these media can be challenging, as they typically command more trust than digital platforms.

Furthermore, digital platforms, such as websites, blogs, and other online portals, serve as catalysts for the spread of fake news. These channels, having a broader reach than traditional media, have the capacity to rapidly disseminate disinformation to vast audiences, thereby potentially causing substantial harm. The propagation of fake news mirrors the transmission of an infectious disease, hence it can be thought of as an epidemic.

To illustrate this situation, graph-based neural networks (GNNs) have been employed to identify the patterns of propagation. In these networks, G = (N,E), where "N" signifies the nodes (individuals) and "E" symbolizes the edges connecting these nodes. The state of each node can be categorized as Ignorant, Spreader, Carrier, or Recovered, as depicted in Fig. 7.

3.1 Epidemic models unveiling propagation patterns

Epidemic models, a particular class of mathematical models, are extensively applied to delineate the dissemination of diseases or information, notably the spread of misinformation, or "fake news." Over the recent years, the phenomenon of fake news has witnessed escalating attention, attributed to its profound influence on the diminishing public trust in media. Consequently, it becomes imperative to decipher its dissemination patterns to devise effective strategies countering the proliferation of false information.

A significant proportion of researchers have employed epidemic models to demonstrate and simulate the distribution patterns of fake news. These models are founded on analogous principles that are typically used to simulate the spread of diseases, such as the SIR (Susceptible, Infectious, Removed) model [19]. This model classifies the population into three categories: Susceptible, referring to those not yet exposed to the misinformation; Infectious, implying individuals who have been exposed and might be actively disseminating it; and Removed, representing those no longer influenced by the fake news, either due to their recognition of its falsity or their cessation of its propagation. The SIR model aims to emulate the spread of fake news chronologically, by estimating the counts of Susceptible, Infectious, and Removed individuals at any given instant.

A prevalent model for analyzing the diffusion of infectious diseases is the SEIR (Susceptible–Exposed–Infected–Recovered) model [123]. Within the context of fake news, the SEIR model serves as an effective tool to comprehend the process through which individuals become cognizant of and interact with false information. This model incorporates four phases: Susceptible, signifying individuals who have yet to encounter the misinformation; Exposed, denoting those who have encountered the fake news and are conscious of it; Infected, illustrating those actively promulgating it; and Recovered, marking those who have diverted their attention to other topics. By providing insights into the mechanics of misinformation spread, this model enables researchers to determine optimal intervention points."

The Susceptible–Infected–Susceptible (SIS) model [171] shares similarities with the SEIR model, but omits the recovery phase. This model is generally applicable to incurable diseases or scenarios wherein individuals can be re-infected. In the realm of fake news, the SIS model facilitates understanding of exposure and re-exposure patterns to false information, along with its dissemination process.

The Susceptible-Infected-Recovered-Immune (SIRI) model [18], while akin to the SEIR and SIS models, incorporates an additional stage termed "immune." Within the context of fake news, this model can be employed to analyze how individuals develop immunity to misinformation, either by actively pursuing accurate information or by building an immunity over time. The Susceptible-Infected-Recovered-Aware (SIRA) model [67] resembles the SIRI model, but introduces a fifth stage: "aware." In terms of fake news, this model can elucidate how individuals become conscious of misinformation and can be educated to identify it.

The Susceptible-Exposed-Infected-Zombie (SEIZ) model [27], sharing similarities with the SEIR and SIS models, introduces a fourth stage: "zombie." In the realm of fake news, this model aids in understanding how individuals transform into "zombies," i.e., continue to propagate misinformation despite being aware of its fallacy. The Susceptible-Protected-Exposed-Non-susceptible-Susceptible (SPENS) model [62] is akin to the SEIZ model, but incorporates two additional stages: "protected" and "nonsusceptible." This model, in the context of fake news, allows us to comprehend how individuals can be shielded from misinformation and how they can become nonsusceptible to it. The Susceptible-Exposed-Infected-Inactive-Recovered(SEIIR) model [111] bears similarities to the SEIZ and SPENS models but introduces a fifth stage: "inactive." In the context of fake news, this model can facilitate understanding of how individuals become inactive in the spread of misinformation, either by diverting attention to other topics or ceasing to share it.

Lastly, the Interactive–Likely–Susceptible–Heard–Recognized (ILSHR) model [3] is akin to the SEIIR model but integrates two additional stages: "likely" and "heard." In terms of fake news, this model can elucidate how people become prone to sharing false information and how they gain awareness of it.

In conclusion, epidemiological models serve as valuable tools for comprehending the dynamics of fake news propagation and devising strategies to mitigate its effects. This article discusses seven distinct models, namely SIR, SEIR, SIS, SIRI, SIRA, SEIZ, SPENS, SEIIR, and ILSHR, each contributing to the understanding of various aspects of fake news transmission.

Although current propagation-based fake news detection techniques predominantly focus on static networks, which assume complete access to the information propagation network structure before employing learning techniques, it is important to acknowledge the frequent emergence of new nodes and edges in information diffusion. To address this issue, a novel approach is proposed, involving the utilization of a dynamic graph neural network with adversarial learning [169]. This approach enables the simultaneous capture of temporal evolution patterns in news diffusion and the nonlinear information interactions. By incorporating this novel method, researchers can better comprehend the complex dynamics of fake news dissemination and effectively combat this pressing issue.

Overall, each model presented in this article offers unique insights into various facets of fake news transmission, thereby providing researchers with invaluable tools to enhance their understanding and countermeasures against this pervasive problem. The authors [40] proposed the Factual News Graph (FANG), a graph-based neural network that incorporates social context and interactions to effectively capture the spread of false information on Twitter. By utilizing graph neural networks (GNNs) and advanced techniques for social interaction representation, FANG offers an innovative approach to modeling information diffusion and detecting misinformation, contributing to the field of graph-based neural networks in combating the challenges posed by the dissemination of false information.

Markov chains continue to be a reliable tool for modeling virus spread and predicting infection times accurately. In the context of detecting and estimating fake content, a technique proposed [65] has demonstrated improved performance by efficiently estimating fake content through a sampling strategy. This approach allows for effective simulation and analysis of the spread of false information. Additionally, [159] have developed a hybrid kernel function that combines a random walk graph and an RBF (radial basis function) kernel to simulate the propagation of false information. This hybrid kernel function enhances the ability to capture the complex dynamics and patterns of fake content dissemination. These advancements contribute to the growing body of research aimed at understanding and combating the spread of misinformation.

The distribution of fake news has been empirically shown to pose potential hazards to individuals, businesses, and various societal sectors. In response, researchers have proposed a model that utilizes differential equations to track the spread of messages, providing a valuable tool for understanding and mitigating the impact of fake news [123]. The surveillance of fake news propagation has emerged as a prominent research topic, highlighting the increasing recognition of its significance [67]. One notable model in this area, known as propagation-based fake news detection (PropFND), categorizes news as genuine or false by examining the interaction of propagation patterns and user profile characteristics, offering a promising approach to identifying and combatting the dissemination of misinformation [146].

These advancements play a crucial role in bolstering ongoing initiatives aimed at confronting the challenges brought about by fake news and its impact on society by providing solutions to mitigate its repercussions.

4 Detection

The accessibility, low cost, and rapid information sharing capabilities of the Internet have contributed to the popularity of social media as a news consumption platform. However, this ease of dissemination has also made it easier to spread fake content with the intent to deceive or harm reputations. Consequently, there has been a growing interest in research focused on detecting and combating fake content on social media platforms. Figure 8 provides an overview of state-of-the-art DL and ML models that have been applied to detect fake news, considering both text and image-based approaches.

Conventional fake news detection algorithms have demonstrated limited effectiveness in identifying fake news on social media platforms. This limitation stems from their inability to effectively utilize auxiliary information available in the vast amount of incomplete and noisy data generated on these platforms. To tackle these challenges, researchers have explored various automatic detection techniques, including browser-based applications, mobile applications, and machine learning (ML) and deep learning (DL) methods. These techniques employ diverse models such as CNNs, RNNs, multi-layer perceptron (MLP), recursive neural networks (RvNN), generative adversarial networks (GANs), and others. Furthermore, ensemble methods have gained popularity in fake news detection, combining the predictions of multiple models to improve overall accuracy and robustness. These ensemble techniques often employ a combination of different models or variations of the same model with varying parameters, leveraging their diverse strengths to enhance detection performance.

The allied recurrent and CNN (ARCNN) is a novel framework proposed for identifying bogus news by employing CNNs for image analysis and RNNs for text analysis. This multimodal approach allows for a comprehensive understanding of the content, improving the accuracy of fake news detection. In addition, BERT has emerged as a leading classification model, achieving a notable accuracy of 78% on the Fakeddit dataset [116]. BERT's success can be attributed to its utilization of transformer-based architectures, which effectively capture contextual information and dependencies in text data.

A significant contribution to fake news detection is the proposal of a multimodal variational autoencoder (MVAE) [64]. The MVAE utilizes encoders to transform data into vector representations, a decoder to reconstruct the original data, and a fake news detector that leverages a latent vector for predicting fake news. This framework integrates multiple modalities, such as text and images, to provide a comprehensive analysis of content and enhance the accuracy of fake news identification. The MVAE approach offers a promising solution for effectively detecting and combating fake news

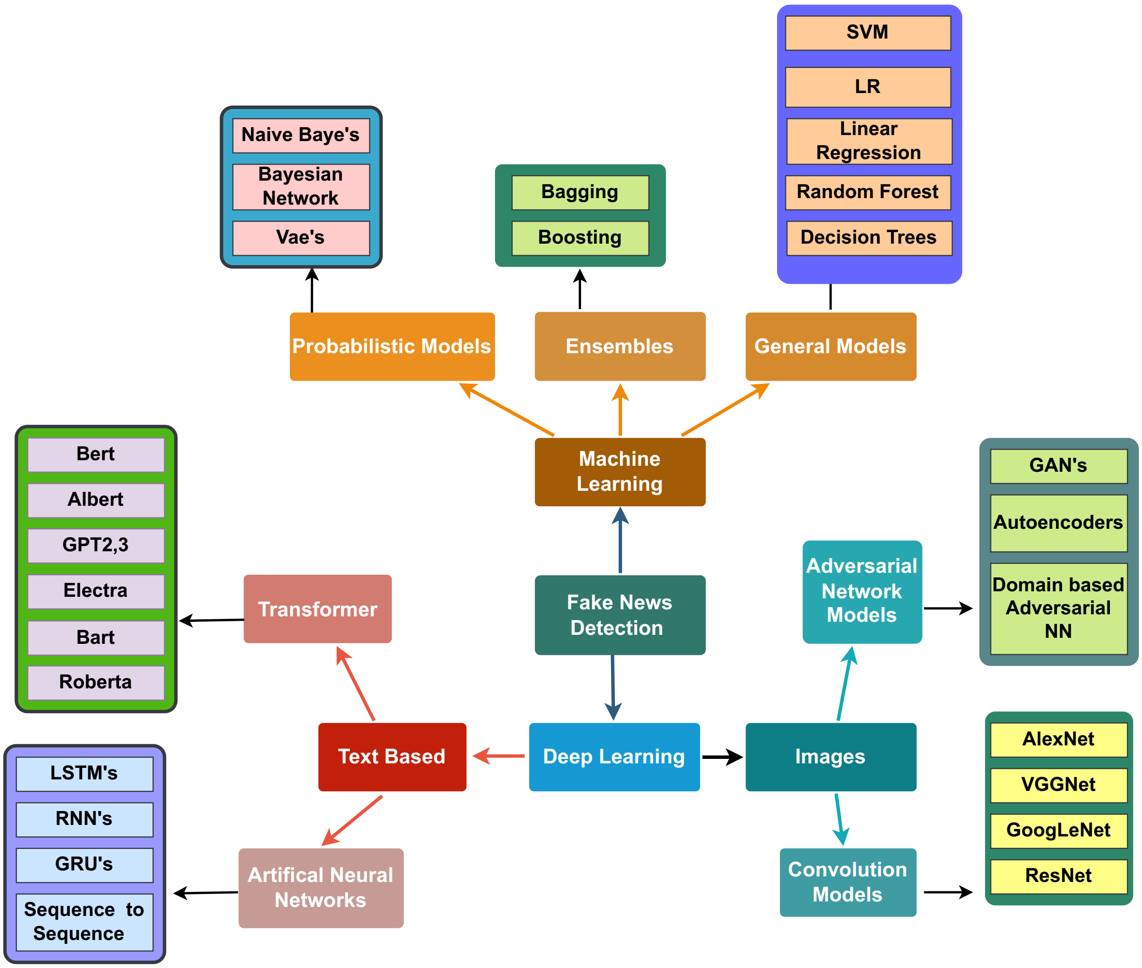


Fig. 8 Comprehensive analysis of machine and deep learning models for fake news detection using text and image data

by combining the power of variational autoencoders with the ability to process diverse data types.

The SAFE technique [173] presents a novel approach to fake news detection by leveraging neural networks to analyze both text and image content. This technique automatically obtains latent representations using a similarity metric, effectively capturing the underlying patterns and features of the data. By utilizing these joint representations, the SAFE technique enables accurate prediction of fake news. This approach offers a valuable contribution to the field of fake news detection, as it combines the strengths of neural networks in analyzing multimodal data, providing a robust framework for identifying deceptive information.

The SpotFake model has been designed that utilizes BERT [169] for classifying and VGG-19 for image [129]. In the realm of fake news detection, the use of a refined model called RoBERTa has shown significant advancements. RoBERTa, which is a variant of the BERT (bidirectional encoder representations from transformers) model, utilizes a sentence transformer for text analysis and exhibits improved performance compared to BERT in comprehending context [130]. The refined architecture and training techniques employed in RoBERTa enable it to capture richer semantic and contextual information from text data, leading to enhanced accuracy in predicting fake news.

The adoption of RoBERTa represents a notable development in the field of fake news detection, emphasizing the importance of leveraging advanced models and techniques for effectively addressing the challenges posed by misinformation.

For the purpose of identifying bogus news, pre-trained BERT and ResNet models have been used to develop accurate representations of text words and image areas, respectively [154]. In the realm of fake news detection, [51] have proposed a straightforward network that integrates contextual embedding, word embedding, and attention mechanisms to effectively analyze relevant information. This model aims to enhance the accuracy of fake news identification by leveraging these techniques. Moreover, the proposed model has been evaluated using the LIAR dataset, demonstrating a

Table 2 Handmade text-based features

S. no.	Handmade text-based features	Other features
1	Tweet length	Linguistic features
2	No. of tokens	
3	No. of URLs	
4	No. of hashtags	
5	Third person pronouns (he, she)	
6	No. of punctuation marks	Lexical features (No. of pronouns)
7	No. of adjectives, noun, verbs	
8	No. of "?!"	
9	Unigram, Bigram, N gram	Correlation-based feature selection
10	No. of "?"	
11	Frequently used words	
12	No. of "!"	
13	Unique words	Source and environment features
14	No. of emojis	
15	Type/Token	
16	Neutral emoji sentiment sum	
17	Negative emoji sentiment sum	
18	Average word length	
19	Positive emoji sentiment sum	
20	POS tagging, word2vec	Pyscholinguistic features
21	Number of periods	
22	Swear words	
23	No. of stop words	
24	Anger Words	Temporal features
25	No. of words	
26	Average no. of characters per sentence	
27	Average no. of characters per word	
28	Overall text sentiment	
29	No. of consonants	
30	No. of vowels	

Table 3 Handmade image-based features

S. no.	Handmade image-based features	Others
1	No. of faces in an image, Resolution of image, Image Intensity	
2	Image visual features	
3	Statistical properties of images	SIFT
4	Color features of an image	
5	Edges of an image	
6	Blobs in an image	SURF
7	Corner features	Shape-based
8	Texture of an image	
9	Image size, Image width, Image height	Texture-based

performance improvement of 1.49% compared to the state-of-the-art methods.

4.1 Handmade features extraction for machine learning models

Handmade feature extraction is an important part of the machine learning process, and it refers to the manual extraction of features from data. This process can be used to identify patterns and relationships between variables, which can then be used to build predictive models. The process of handmade feature extraction involves manually selecting and extracting features from a dataset. Once the features have been extracted, they can then be used to train a machine learning model. The chosen features should be relevant to the task at hand, and they should also be able to provide useful information about the data. This way, the model can be trained to accurately identify patterns and relationships in the data.

Additionally, it can help to reduce the amount of time and effort required to build a model, as the data scientist does not need to rely on algorithms to do the work for them. For object recognition and computer vision, these hand-crafted features are frequently combined with "conventional" machine learning techniques like support vector machines. For detecting fake news, previous methods rely on handmade features to employ machine learning algorithms that require a huge manual effort [49].

However, "newer" methods like CNNs often do not require the addition of such hand-crafted characteristics because they have the ability to "learn" the features directly from the text or image input. In this survey for feature extraction, handmade features from text and image has been stated. Tables 2 and 3 outline major handmade features utilized in prior state of the arts to create machine learning models.

4.2 Machine learning models

Machine learning has played a major role in classifying the fake news. In the present literature, multiple machine learning approaches have been used to identify and classify fake news. NLP for textual analysis and traditional machine learning models such as decision trees, random forest, Naïve Baye's, and K-nearest neighbor has been to classify the fake news [63]. For binary classification NB, LR, RF has been considered for classifying fake news articles [122].

In [132], authors have identified the most important features related to fake news and then applied those features to a variety of machine learning techniques, including NB, LR, SVM, RF, and K-nearest neighbors, with the latter having the highest accuracy of 93.84% on images and 81.93 % on text. Another fake news detection model has been proposed that uses the n-gram technique and ML approaches. In [4] a comparison between features extraction techniques (TF-IDF) and machine classification techniques as Linear Support Vector Machine (LSVM) has been made achieving a 92% accuracy. For Classic Machine Learning Multinomial Naive Bayes, SGD Classifier, Random forest Decision trees, LR, SVM have been implemented to classify the news as fake or real [15]. With a weighted average F1 score of 95.19% a linear SVM, machine learning method has been employed with a number of linguistic variables, including readability, n-grams, emotional tone, and punctuation [32]. To evaluate the performance for fake news identification, a thorough comparison of SVM, LR, Adaboost, Decision Trees, and NB has been conducted which shows NB among the best [11].

4.3 Multimodal ensemble methods

To increase the overall accuracy for the aim of classifying an article as fake or real, many ensemble approaches coupled with textual and image characteristics as feature input have been proposed. Ensemble models aim to train more than one model at a time thus increasing the performance and decreasing error rates. LinearSVC and Bert Model have been used as ensemble models for the classification of fake news [10].

Deep neural network (DNN) and SVM technique has been mostly used as ensembles in the binary classification of algorithms [70]. An ensemble of ML, as well as DL, approaches such as CNN, Naive Bayes, SVM, and KNN has been implemented with LSTM giving higher accuracy results of 97% [1]. For the four-class multi-class classification of false news, a gradient boosting algorithm (an ensemble machine learning framework) has been implemented, giving an accuracy of 86% [58]. An ImageFake ensemble model has been proposed that uses a variety of pre-trained CNN models. Those pre-trained models include VGG-19, Inception v3, VGG-16, Squeeze Net, and ResNet-101 for identifying and classifying fake news [26].

An ensemble classification model combining decision tree, RF, and extra tree classifier for the detection of fake news that has been implemented with an accuracy of 99.8% and 44.15% [39]. Machine learning algorithms such as SVM and LR have been implemented to accurately predict fake news [138]. An evaluative comparison has been conducted between ensembles, specifically CNN-LSTM, and other contemporary models to demonstrate the superiority of the proposed CNN-LSTM approach. The results illustrate that CNN-LSTM outperforms its counterparts by delivering the most robust performance, achieving an accuracy rate of 81.6% [137].

4.4 Handmade versus deep learning models

With the emergence of deep learning, many aspects of data analysis and machine learning have been revolutionized. Deep learning has enabled efficient and accurate solutions to problems that were previously difficult, such as computer vision, natural language processing, and speech recognition. However, handmade feature extraction is still a powerful tool for extracting relevant information from data.

These handmade techniques sometimes suffer from some limitations as extracting features consumes a lot of time and mostly extracts features without capturing a structural comprehensive depiction of the news dissemination pattern. These may help in capturing linguistic and psychological factors, but occasionally they fail in categorizing the news as real or fake, therefore limiting the performance. In order to address the problem of handmade approaches, existing DL models have been acquainted to identify fake news based on better performances.

It has been proven by [109] that the performance of the deep learning feature extractions techniques is better than the traditional handmade feature extraction because of the deep learning's ability to extract features automatically. In Figs. 9 and 10, handmade and deep learning-based feature extraction techniques are shown very clearly.

4.5 Deep learning approaches

Deep learning has gained significant popularity in recent years, primarily due to its capability to tackle complex problems with minimal human intervention. It has emerged as the gold standard in the machine learning (ML) industry and is now considered the most prominent computational method in the field. Deep learning has demonstrated remarkable performance on challenging cognitive tasks, often matching or surpassing human-level performance. One key advantage is its ability to learn from vast amounts of data, leading to breakthroughs in natural language processing (NLP), robotics, control systems, and medical information processing. Figure 10 exemplifies the superiority of deep learning in automatically extracting features compared to manual feature extraction approaches. Wang et al. [154] propose an adversarial network-based EANN model for determining the veracity of reports on unidentified incidents, utilizing event-invariant features. In contrast, [116] present a multimodal approach based on a CNN architecture that achieves an impressive accuracy of 87% on Fakeddit data. This accuracy surpasses the text-based BERT model's 78% accuracy,

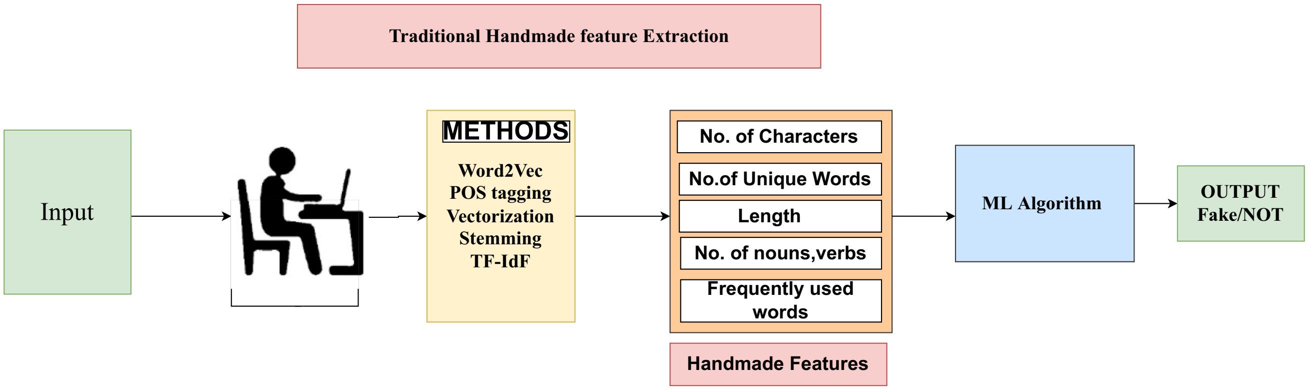


Fig. 9 Traditional feature extraction approach

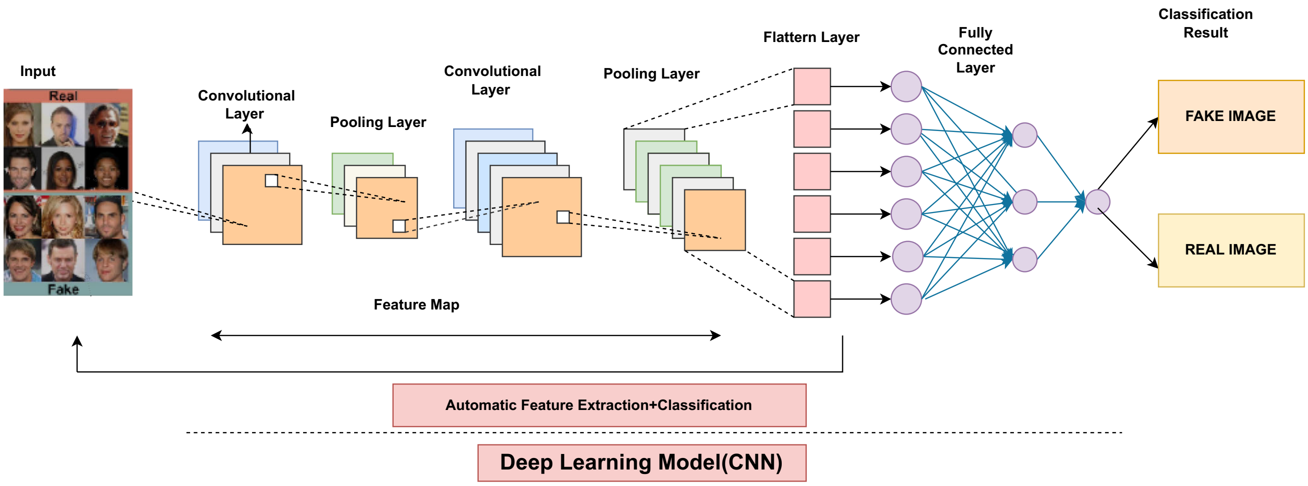


Fig. 10 Deep learning-based feature extraction approach

highlighting the benefits of incorporating multiple modalities for enhanced fake news detection. A model has been proposed that uses att-RNN for multimodal news detection integrating text, image, and social context characteristics and leverages attention mechanism to determine the relationship between modalities [55].

Khattar et al. [64] introduce a multimodal variational autoencoder (MVAE) that effectively learns shared representations for both text and visual modalities, thus enhancing the detection of fake news. By capturing common underlying features across different modalities, the MVAE improves the accuracy of fake news identification. Similarly, [169] propose the BDANN model, which incorporates the BERT preprocessing model into the EANN framework to enhance text feature extraction and, consequently, improve fake news detection. The SpotFake model, proposed by [169], combines text and picture information using the BERT and VGG-19 pre-trained models. By leveraging these models, the SpotFake model effectively captures both textual and visual features to enhance fake news detection. Additionally, the EMAF (entity multimodal alignment and fusion) model is introduced, focusing on aligning and fusing multimodal entities to capture specific properties. This approach aids in the comprehensive analysis of inter-modal entities, further enhancing the accuracy and effectiveness of fake news detection [72].

These advancements demonstrate the importance of integrating multiple modalities and leveraging pre-trained models for robust and efficient identification of fake news.

4.5.1 DeepFakes

Deepfake technology has become one of the most discussed subjects in the field of artificial intelligence (AI) in recent times [6]. Deepfake is a form of AI-based technology that uses a type of generative adversarial networks (GANs) to generate synthetic media. This technology is used to create highly realistic videos [157] or images that depict a person doing or saying something that they never did or said. Deepfake videos are increasingly being used for malicious purposes, such as spreading misinformation and false news, creating and perpetuating harmful stereotypes, and creating targeted harassment campaigns.

The legal implications of using deepfakes are also significant [117]. Depending on the jurisdiction, the creation and distribution of deepfakes may be illegal. For example, in the USA, it is illegal to create deepfakes for malicious purposes, such as to defame or harass someone. Additionally, many countries have laws that prohibit the distribution of non-consensual pornography, which could also apply to deepfakes. Fortunately, there are ways to detect deepfakes. One way is to analyze the video for inconsistencies, such as unnatural lighting, unnatural motions, or unnatural facial expressions. Additionally, AI-based deepfake detection systems can be used to detect deepfakes [106]. These systems use machine learning algorithms to identify suspicious aspects of a video that may indicate it as deepfake.

4.5.2 Generative adversarial network

The primary objective of GANs is to fabricate artificial images that appear as realistic as possible. The fundamental strategy behind GANs [36] entails the training of two distinct networks: a generator and a discriminator. The generator's purpose is to synthesize an image, while the discriminator's role is to differentiate between artificially constructed "fake" images and "real" images produced by the generator.

Recently, numerous efforts have been exerted to tackle the challenge of video generation utilizing GANs. In this context, a generator transforms a low-dimensional random vector x into a corresponding output image y. The loss function, denoted as L, defines the objective function of GANs.

$$min_G max_D L = H_x log[D(x)] + H_z log[1 - D(G(z))] \quad (1)$$

where

H_p= the expectation with reference to a distribution p.
p_j = the input noise vector distribution z.
p_i = the real data distribution r.

The discriminator has been trained to recognize real-world samples x and generate higher values close to 1. Therefore, $H_{x\sim p_i(x)} log[D(x)]$ should be increased. Likewise, the discriminator has been trained to identify fake samples $G(z)$ generating lower values closer to 0, which implies maximizing $H_{z\sim p_z(z)} log[(1 - D(G(z)))]$. The generator $G(z)$, on the other hand, generates data that are identical to real data attempting to fool the discriminator $H_{z\sim p_J(z)} log[(1 - D(G(z)))]$. Generating an image from different input descriptions such as text, audio, and video, many GAN's frameworks have been discussed so far.

Recent advancements in artificial intelligence technology have led to the emergence of the fake detection fine-tuned network (FDFtNet). This novel neural network-based classifier architecture has been specifically designed for the detection of fake images, which are typically generated from deepfake-based datasets and progressive growing GANs [53]. Despite its lightweight construction, the FDFtNet exhibits robustness and resilience, offering performance that surpasses the contemporary state-of-the-art techniques with an impressive overall accuracy rate of 90.29%. The performance underscores FDFtNet's potential as a tool for fake image detection and heralds new opportunities for future investigations in this area. Parallel to the evolution of AI technology is the increased capacity to create highly convincing false visuals. GANs have considerably simplified the creation of such counterfeit visuals, with deepfakes as prime examples. In response to this challenge, the dual attention fake detection fine-tuning network (DA-FDFtNet) has been introduced [13]. This advanced network relies on training datasets derived from FaceForensics++ and several other GAN-generated sources. Furthermore, in a separate study, a unique deep learning solution known as contrastive loss has been proposed as a method for identifying fake images [46]. This approach operates by training a deep neural network using pairs of images—one genuine and one counterfeit. The fundamental principle behind this method is to train the network to distinguish between the traits of real and fake images, using the observed disparity between the paired images as the basis for the learning signal.

With the best accuracy in recognizing fake faces produced by both humans and GANs, a technique called FakeFaceDetect is suggested to recognize various false face images [139]. Though generative adversarial network (GAN) may produce a realistic images, incorrect usage of these technologies may raise unaddressed concerns. For that, a deep forgery discriminator (DeepFD) has been proposed which efficiently and effectively detect the computer-generated images achieving an accuracy of 94.7 % which is somehow better than several state-of-the-art GANs [45]. A new method for forgery detection based on GANs has been proposed which classifies image as fake or real with an accuracy of 98%on Deeper Forensics dataset [162]. A novel framework known as two-stream CNN has been proposed to identify GAN produced fraudulent pictures, with RGB and photo-response non-uniformity (PRNU) streams, respectively [152].

In the study, authors [43] have considered two-stream CNN framework offers a new approach for identifying GAN-generated fraudulent images, leveraging both RGB and photo-response non-uniformity (PRNU) streams. By combining the two streams, the model can better detect subtle differences between real and fake images, allowing for more accurate identification of fraudulent images. This novel framework is expected to revolutionize the way GAN-generated images are identified, providing a powerful tool for combating fraud. Another study by the authors [166] has

Table 4 Comparative analysis of classification models across diverse text-based datasets

ML/DL model	Ref.	Year	Accuracy (%)	F1 score (%)	Text datasets
Roberta(Text) + CNN(Image)	[130]	[2021]	81.2	87.9	Weibo
Bert(Text) + ResNet(Image)	[165]	[2021]	87.9	87.9	Weibo
Bertbase(Text) + VGG(Image)	[100]	[2021]	84.9	83.6	Weibo
CNN(Image) + Roberta(Text)	[151]	[2022]	88.5	86.4	Weibo
RNN(Text)	[170]	[2019]	84.8	–	Weibo
Bert(Text) + MLP(Image)	[24]	[2022]	84	–	Weibo
Bert(Text) + VGG-19(Image)	[156]	[2022]	89.4	89	Weibo
BERT(T) + VGG-16,19(Image)	[131]	[2022]	86.83	–	Weibo
ML/DL model	Ref.	Year	Accuracy (%)	F1 score (%)	Datasets
BERT(T) + VGG-19 pre-trained(Image)	[133]	[2019]	77.7	–	Twitter
GRU-2(Text)	[8]	[2022]	98.5	98.5	Twitter
CNNs(Image) + Bi-LSTM(Text)	[170]	[2010]	85.8	78.5	Twitter
Electra, Bert(Text)	[131]	[2021]	85.80	–	Twitter
ML/DL model	Ref.	Year	Accuracy (%)	F1 score (%)	Datasets
Rule-based	[148]	[2019]	–	69.3	Pheme
CNN	[116]	[2022]	88.5	82.3	Pheme
Conv-1Dlayer + Max pooling(Image)	[100]	[2021]	86.4	81.3	Pheme
Naive Baye's, KNN, and Decision Trees	[57]	[2021]	88.90	–	Pheme
Bert(Text) + ResNet50(Image)	[164]	[2021]	75.8	75.5	Pheme

taken GANs and self-supervised learning that has enabled the rapid detection of fake images. This technique allows for a more efficient detection of fake images, as it can more accurately identify the differences between real and fake images. In the study, authors have presented two approaches that detect deepfake images: The first is based on a CNN deep learning network and the second is based on transfer learning.

With the CNN approach, the model is trained from scratch on a dataset of deepfake images. This approach requires a large dataset of deepfake images to train the model and is computationally intensive [68]. In this study, authors [113] have compared deep convolutional generative adversarial networks (DCGANs) and conditional generative adversarial networks (CGANs) with CNNs and GANs in detecting fake images. The results show that CGANs perform better than CNNs and GANs, achieving higher accuracy in fake image detection.

4.5.3 Common datasets and evaluation parameters

Today, the fight against fake news is more important than ever. In order to combat this issue, we have compiled a comprehensive set of datasets to support the development of machine learning applications to detect and classify fake news. These datasets cover a wide range of topics and have been manually verified to ensure accuracy. They can be used to train ML models to accurately detect and classify fake news. In this survey, we have mentioned all the state-of-the art datasets with their results allowing for better selection when designing deep learning systems. These datasets for text and images are shown in Tables 4 and 5, respectively. With these datasets, researchers and developers can create powerful tools to help reduce the spread of fake news and improve public discourse.

Text-based datasets

1) **Liar dataset**: This dataset comprises 12.8K hand-labeled brief statements gathered from diverse situations on PolitiFact. The minutely detailed articles within this dataset are organized into six groups: pants-fire, false, slightly true, half-true, largely true, and true. Machine learning models, including but not limited to Rf, LR, SVM, K-nearest neighbors, and NB classifier, are primarily employed. Furthermore, deep neural network models such as CNN, long short-term memory (LSTM), and bidirectional LSTM are predominantly utilized on this dataset [104].
2) **PHEME:** This dataset is a compendium of breaking news rumors and non-rumors from Twitter. It includes rumors about nine different occurrences, each of which has been annotated with its credibility value as true, false, or unver-

Table 5 Comparative analysis of classification models across diverse image-based datasets

ML/DL model	Ref.	Year	Accuracy (%)	F1 score (%)	Image datasets
VGG-19 Pre-trained	[34]	2020	80.8	–	ImageNet
DNNs	[130]	2021	87.13	87	Casia 2.0
VGG-16	[102]	2021	82.72	–	Ti-CNN
CNN(Image) and Bert(Text)	[116]	2022	87 (CNN),78 (Bert)	–	Fakeddit
VGG-19(Image) +Text CNN	[119]	2022	80.4	83.8	Fakeddit

ified. Bert for text, ResNet50 for image, 4 attention heads are giving fine results on this dataset [50, 101].

3) **Weibo:** The Weibo NER dataset is indeed a Chinese named entity recognition dataset derived from the Sina Weibo social networking site. RoBERTa for extracting the features of images and text has been mostly applied on the dataset with good performance results [130, 156].
4) **BuzzFeed:** This dataset is an exhaustive compilation of news articles published on Facebook between September 19 and 23, 2016, and September 26 and 27, 2016, by nine news organizations [17].
5) **FakenewsNet:** The compilation of the dataset utilized in this study draws upon information sourced from two reputable fact-checking websites, namely PolitiFact and GossipCop. These websites consist of fake news articles that have been carefully annotated and verified by experts and professional journalists, ensuring the accuracy and reliability of the dataset [127].
6) **PolitiFact:** There are 21,152 claims in the dataset that have been verified as accurate by professionals. All of the assertions have been classified as mostly true, true, half-true, false, mostly false, or pants on fire [121].
7) **Indian datasets:** There has not been enough Indian context dataset available before 2020. Now, more numbers of datasets have been made available by the authors in their work.

(a) *FactDrill:* This dataset comprises news stories from 2013 to 2020, covering 13 different languages spoken in India [134].
(b) *IFND dataset:* This dataset consists of 56,868 news real and fake from different authentic websites [120].
(c) *FakeNewsIndia:* The FakeNewsIndia dataset contains 4,803 false news incidences that occurred in India between June 2016 and December 2019. This dataset comprises fake news stories with the image URLs for supporting fake stories news [30].
(d) *Factify:* This dataset has been considered as the largest multimodal fact verification dataset which is publicly available and consists of 50K data points, covering news stories from India, US. FACTIFY includes textual claims, textual content with three broad categories, namely support, no evidence, and refute [79].

Image-based datasets

1) **MuMin**: The Mumin dataset is a collection of images used to detect fake news. The dataset consists of over 150,000 images of social media posts, news articles, and other webpages, which have been labeled as either real or fake news. The dataset is divided into two categories—real and fake, with each category containing over 75,000 images [86].
2) **Fakeddit:** The Fakeddit dataset is a collection of images used to detect manipulated images. The dataset consists of over 18,000 images which have been manipulated using various techniques such as image splicing, image scaling, image warping, and more. The dataset is divided into two categories—original and manipulated, with each category containing over 9000 images [83].
3) **ImageNet:** The ImageNet dataset is commonly used for image classification and fake news detection. It contains over 14 million images from over 20,000 categories. The images come from a variety of sources, including public domain images, personal collections, and professional photography [29].
4) **The MS-COCO dataset:** The MS-COCO dataset is commonly used for object detection. It contains over 330,000 images which have been labeled with over 80 object categories. The dataset also contains annotations for segmentation, keypoints, and captions. This dataset is particularly useful for training deep learning models that specialize in object detection [90].
5) **Flickr30k:** The Flickr30k dataset is a collection of 30,000 images, each of which is annotated for fake news relevance [141]. The images are categorized into classes such as "political," "non-political," "satire," and "hoax." Each image contains a caption that is used to help determine the relevance of the image to the topic of fake news.
6) **CASIA v2.0**: The CASIA v2. 0 database contains 7491 real and 5123 fake JPEG, BMP, and TIFF photos, with image sizes ranging from 240 by 160 to 900 by 600 pixels [143].
7) **FNID:** The fake news image dataset is an open-source dataset created by researchers at the University of Southern California. It contains over 10,000 images, many of which are labeled with whether they contain fake news

Table 6 Benchmark datasets description in brief

Dataset	Ref.	Year	Real articles	Fake articles	Images	Public
Liar	[153]	[2017]	6400	6400	No	Yes
Weibo	[156]	[2022]	4779	4749	Yes	Yes
PolitiFact	[33]	[2020]	–	–	No	Yes
Twitter	[8]	[2022]	6026	7898	Yes	Yes
FakeNewsNET	[127]	[2020]	18,000	6000	Yes	Yes
FacebookHoax	[155]	[2018]	6577	8923	No	Yes
BuzzFeedNews	[17]	[2017]	826	901	No	Yes
BuzzFace	[114]	[2018]	1656	607	No	Yes
CREDBANK	[72]	[2021]	1049	–	No	Yes
CASIA 2.0	[52]	[2019]	1701	3274	Yes	Yes
ISOT	[39]	[2021]	1000	–	No	Yes
IFND	[120]	[2021]	37809	7271	Yes	Yes
Ti-CNN	[103]	[2021]	10000	10000	Yes	Yes
MICC-F220	[136]	[2019]	110	110	Yes	Yes
Fakeddit	[83]	[2019]	1million	–	Yes	Yes
NewsBag Test	[56]	[2020]	11,000	18,000	Yes	Yes
NewsBag	[56]	[2020]	2,00,000	15,000	Yes	Yes
MS-COCO	[75]	[2019]	328K	–	Yes	Yes
Flickr30k	[80]	[2019]	31,000	–	Yes	Yes
FakeNewsIndia	[30]	[2022]	–	4803	Yes	No

or real news. The images come from a variety of sources, including social media, news websites, and blogs. The dataset is designed to be used for the training and evaluation of deep learning models for fake news detection [66].

4.6 Comparison between existing fake news datasets

The swift expansion of online media has led to a notable surge in the incidence of fake news, causing confusion, mistrust, and sometimes widespread panic globally. In an effort to tackle this issue, various datasets designed for the detection of fake news have been devised by research community. Consequently, there exists a necessity to analyze and compare these available datasets for fake news identification, allowing researchers to conveniently gauge their accessibility, supported by a concise summary as displayed in Table 6.

4.7 Attention mechanism

Attention, in cognitive processes, involves selectively focusing on a small subset of information while discarding less relevant data [99]. This concept has been integrated into neural networks to simulate human brain functionalities, giving rise to the Attention Mechanism. This mechanism has demonstrated its utility across various tasks, such as abstractive summarization, learning task-agnostic phrase representations, and textual entailment.

Particularly in the context of machine translation models, the attention mechanism enhances the encoder-decoder model's performance by assigning weights to all encoded input vectors, thereby allowing the decoder to utilize key segments of the input sequence flexibly. Notable variations of attention mechanisms include multi-head dot product attention and scaled dot-product attention (Fig. 11).

Single dot-product attention: A single attention head takes three values as input, namely Query (Q), Key (K), and Value (V). An attention function maps given Query to Key–Value pairs, and the corresponding result can be thought of as weighted values that describe which Key–Value pair is more important to query. This value is then passed through the softmax function to get normalized weights, and then, the dot product is taken with values [145].

$$Attention(Q, K, V) = Softmax(QK^T/\sqrt{d}) * V \tag{2}$$

Multi-head attention: Instead of applying Query, Key, and Values through a single dot-product attention, we divide the dimension up into h components. The computation happens in parallel, and the results are then concatenated to get the final result. This allows the model to learn multiple concepts together rather than focus on a single concept [145].

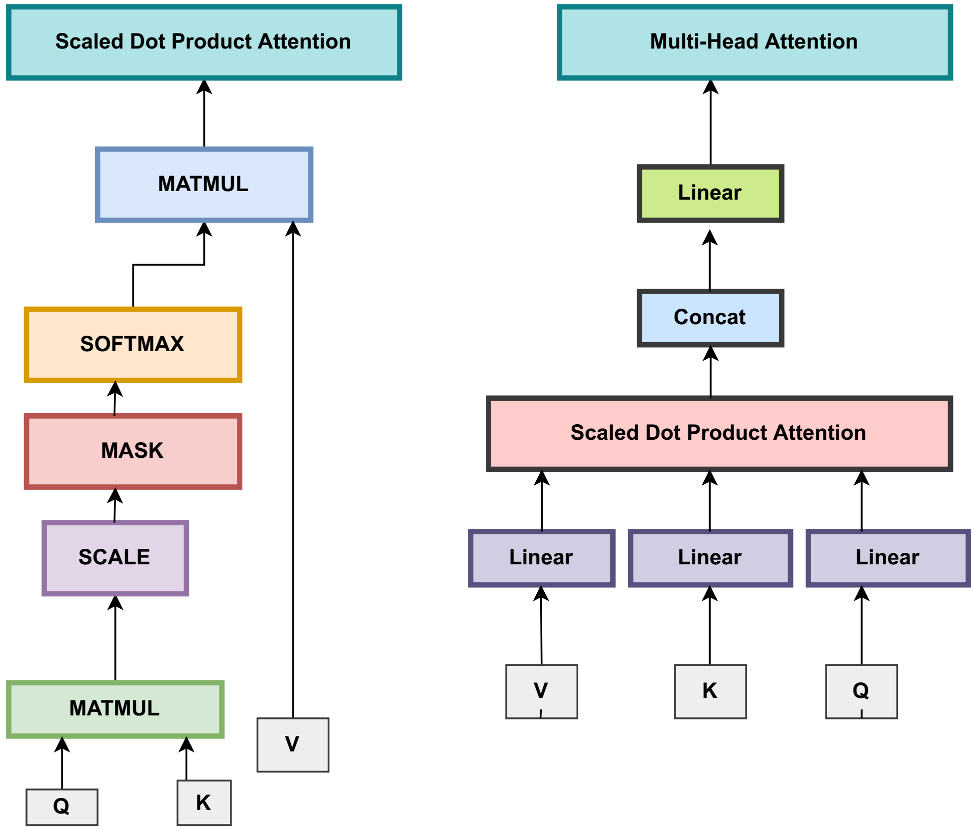


Fig. 11 Attention mechanism

$$Multi-Head(Q, K, V) = softmax(QK^T//\sqrt{d_k}) * V \tag{3}$$

where Q, K, and V are the Query, Key, and Value matrices, respectively. The concept of attention mechanism has been used extensively in the field of NLP, and it has recently been applied to the problem of fake news detection.

A study by [150] proposed a two-stage attention-based model for fake news detection. The first stage was a content-based attention model, which used word-level and sentence-level attention to identify important parts of the news article. The second stage was a user-based attention model, which used user behavior data to identify users who were likely to spread false information. The model has been evaluated on a dataset of Chinese news articles and found to be effective in identifying fake news. A study by [167] proposed a multi-head attention neural network model for fake news detection. The model used an attention mechanism to identify the most important sentences in a news article and then used the features extracted from these sentences to classify the article as real or fake.

In [105], authors have taken convolution deep learning neural network model with attention mechanism applied giving promising results on LIAR dataset. Another architecture has been proposed that combines word embeddings and contextual by leveraging attention mechanisms with accessible relevant information for detecting fake news [51]. In the study [142], authors have applied a pre-trained BERT and pre-trained VGG-19 on the ImageNet dataset. Further, a scaled dot-product attention mechanism has been used to record the relationship between text and visual features for detecting fake news articles [142].

In the context of feature extraction, a combination of Roberta for text and CNN for images has been used to great success. However, a novel technique known as the fine-grained multimodal fusion network (FMFN) has been proposed to further improve feature extraction accuracy. This novel approach uses a scaled dot-product attention mechanism to fuse word embedding of words and several feature vectors representing various aspects of the image. This approach not only takes into account the relationships between various visual aspects but more accurately reflects the interdependence between textual and visual features with an accuracy of 88.5% [151]. A novel multimodal attention adversarial fusion technique has been proposed which is based on the pre-training language model BERT, and has two essential parts: an attention mechanism and an adversarial mechanism. The attention captures the modalities of the data, while the adversarial mechanism captures the correlation between modalities. Experiments have been done on the Chinese dataset, indicating that the method achieves 5% higher in the F1-score than the state-of-the-art [93]. Ma et al.

[74] proposed a fake news detection model which is based on a multi-head attention mechanism and convolution neural networks. The model was trained on a dataset of real and fake news articles and was shown to achieve an accuracy of 95.3%.

Moreover, a recent study by [168] proposed a novel approach for fake news detection based on graph attention networks. The model was trained on a dataset of news headlines and was shown to achieve an accuracy of 97.4%. In addition, [150] proposed an attention-based RNN model for fake news detection. The model was trained on a dataset of tweets and was shown to outperform existing models in terms of accuracy. Following this, [112] proposed an attention-based RNN model for fake news detection. The model utilized two attention layers to identify the most relevant words in a news article. The results showed that the model outperformed traditional methods of fake news detection, with an F1 score of 0.85%.

In addition, [98] proposed an attention-based CNN model for fake news detection. The model is based on a multi-head self-attention mechanism to capture the semantic and syntactic features of the news articles. The results showed that the model was able to detect fake news with an F1 score of 0.89%. In particular, the multimodal multi-image system has been taken into consideration that uses a semantic representation of the text-image by calculating the similarity cosine similarity between the title and image tags embeddings, thus significantly outperforming the BERT baseline by 4.19% [34]. A new framework has been proposed that uses attention-based multilevel CNN-RNN (ABM-CNN-RNN) for image feature extraction [69] and attention-based stacked Bi-LSTM for text-based feature representation.

4.8 Evaluation parameters

In the process of assessing deep learning models, it is crucial to contemplate multiple parameters. Among these, a set of key evaluative parameters, employed frequently by scholars for the appraisal of performance in the sphere of fake news detection, are succinctly explicated in the ensuing discourse.
1. AUC: The area under the curve (AUC), specifically the receiver operating characteristic (ROC) AUC, is an essential metric for fake news classification tasks. It plots the true positive rate against the false positive rate at varying thresholds, effectively assessing a model's ability to distinguish between "fake" and "real" news. An AUC score of 1.0 denotes a perfect classifier, while a score of 0.5 suggests performance equivalent to random guessing. A high AUC indicates a model that can effectively differentiate between real and fake news, even in imbalanced datasets. However, it is crucial to consider AUC alongside other metrics, such as precision, recall, or F1 score, for a comprehensive performance assessment [81, 130].

2. Accuracy: Typically, accuracy is determined by comparing the model's output to the expected output. The model is called accurate if it is capable of appropriately classifying data. It generally describes how the model performs across all classes. It is beneficial when all classes are equally important [17]. TP—True positive, TN—true negative, FN—false negative, FP—false positive.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

3. Recall: Recall evaluates how well a model can identify Positive samples, the more positive samples detected. The higher the recall, the better the model is at identifying data [35].

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

4. Precision: Precision is one measure of model performance. The higher the precision, the better the model's ability to identify data [88].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

5. F1-score: It is defined as the weighted average of Precision and Recall, and it takes both false positives and false negatives into consideration [54].

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Overall, accuracy, precision, and recall are the three most important evaluation parameters used in deep learning for evaluating fake news detection. It is important to ensure that all three of these parameters are taken into consideration when evaluating a model's performance. By doing so, developers can ensure that their model is able to accurately and effectively identify data and solve complex problems.

5 Challenges in existing literature

The endeavor to develop robust systems for the detection of fake news presents a myriad of challenges that must be navigated effectively. These challenges stem from the unique characteristics of fake news and the complexities inherent in its detection and mitigation. To understand the obstacles and pave the way for effective solutions, we have found some of the significant challenges associated with detecting and mitigating fake news, and these are as follows:

1. **Multimodal data analysis:** In the realm of fake news detection, the task of multimodal data analysis presents a

significant challenge. Current research exhibits a strong focus on unimodal, predominantly text-based, analyses. This precludes the simultaneous examination of diverse data modalities such as images, videos, and social network metadata, all of which can convey critical, and sometimes contrasting, information. The development of robust models, capable of synthesizing insights from these disparate data modalities, is a non-trivial problem.
The inherent complexity arises from the need to integrate computational techniques specific to each data modality—such as natural language processing for text, computer vision for images, and graph algorithms for social networks—into a unified framework [160]. This integration is paramount for a comprehensive understanding of information authenticity. Although preliminary efforts have been made in this direction, the scientific community still stands at the nascent stages of research concerning multimodal fake news detection, indicating a pertinent need for further investigation [78].
2. **Lack of explainability:** The issue of explainability presents a substantial challenge in the field of fake news detection. Many of the current methodologies, particularly those leveraging complex machine learning models, function akin to "black boxes." These models, while adept at discerning between real and fake news, do not elucidate the rationale behind their classifications [163]. This lack of transparency can engender mistrust in these systems and hinder their refinement, as understanding the basis of correct or incorrect classifications is integral for improvement [118]. Consequently, there is an imperative need to evolve toward more explainable artificial intelligence models. These models should strike a balance between robust predictive performance and interpretability, fostering greater confidence and transparency in fake news detection mechanisms.
3. **Temporal dynamics**: The veracity of a news item may change over time as new information becomes available. Existing detection models struggle with this temporal aspect, as they generally evaluate veracity based on the information available at a single point in time [28].
4. **Manual fact-checking process:** As manual fact-checking is a process of verifying the accuracy of a piece of news by researching its sources and checking the veracity of the information given. However, manual fact-checking has some challenges in detecting fake news [86, 109]. First, it is labor-intensive and time-consuming, requiring a great deal of effort and resources to verify the accuracy of a single piece of news. Additionally, manual fact-checking can be unreliable due to human errors or biases, as well as the fact that the person conducting the check may not be knowledgeable about the subject matter. Furthermore, manual fact-checking is not always successful in detecting false information, as it is difficult to distinguish between accurate and inaccurate sources [61]. Lastly, manual fact-checking is unable to detect more sophisticated forms of fake news such as deepfakes and doctored images. These challenges make manual fact-checking a less reliable method of detecting fake news.
5. **Addressing the deficit of multilingual datasets:** The challenge of insufficient multilingual datasets presents a significant obstacle in the area of fake news detection. Each language possesses unique linguistic patterns and cultural contexts that are vital for effective fake news detection. However, the current bias toward English in fake news detection models and datasets leaves non-English speaking regions more vulnerable to misinformation. To mitigate this, a global effort to prioritize the collection and development of multilingual datasets is necessary. This effort could involve collaboration between technology firms, academic institutions, governments, and global organizations. Furthermore, the application of transfer learning and multilingual model training can enhance detection models' generalization across various languages, providing a more globally effective defense against the spread of fake news [174].
6. **Less attention on other auxiliary information:** In recent years, fake news detection has become a major focus of research in the field of natural language processing. As fake news can spread rapidly and cause great harm to society, it is necessary to develop effective methods for detecting fake news. However, current fake news detection methods often focus on the content of the news, while ignoring other auxiliary information. This can lead to inaccurate results and make it difficult to detect fake news. The auxiliary information here mainly includes the source and the time of the news. For example, when the source of a news is not a mainstream media, it is more likely to be fake. In addition, when the news is too exaggerated or sensational, it is also more likely to be fake. Therefore, when detecting fake news, it is necessary to pay attention to the source and time of the news.
However, recent fake news detection models do not consider these auxiliary information. They only use the content of the news as the input, ignoring the source and time information. This makes it difficult to effectively detect fake news. In addition, the accuracy of the detection is also greatly reduced. Thus, it is important to pay more attention to the source and time information of the news when detecting fake news. Existing fake news detection models need to be improved to consider both the content and auxiliary information of the news. This will help to more accurately detect fake news and reduce its negative impact on society [47, 125].
7. **Early detection of fake news**: The timely interception of fake news is instrumental in mitigating its harmful con-

sequences, as once falsified information permeates the media landscape, containing it becomes an arduous task, especially in the context of social media platforms where information can rapidly diffuse [109, 126]. As such, it is crucial to create systems capable of identifying and classifying fake news in its nascent stages, thereby thwarting its wide dissemination. Today, the swift and accurate detection of deceptive information stands as a critical challenge, necessitating the deployment of sophisticated and integrated solutions that leverage machine learning, natural language processing, network analysis, and human-centric approaches. These comprehensive strategies are vital to combat the burgeoning issue of fake news in our information-driven society.

6 Conclusion and future work

Detecting fake news is a sophisticated task, necessitating a well-rounded strategy to be fruitful. The evolution in natural language processing (NLP) and machine learning (ML) has been pivotal in enabling researchers to devise a myriad of models and algorithms to discern fake news. Nevertheless, these models present diverse degrees of precision and might harbor biases. In conclusion, this study serves as a comprehensive examination of the current state of research on disinformation, also known as “fake news,” focusing on detection and mitigation tactics. It provides an in-depth exploration of the life cycle of disinformation, from its creation to its dissemination and eventual detection, accentuating the value of both supervised and unsupervised learning techniques, including generative adversarial networks (GANs). The study not only scrutinizes existing datasets related to fake news but also clarifies the key evaluation metrics used to determine the authenticity of news. The survey further identifies and discusses the significant challenges present in the realm of fake news detection, derived from an exhaustive survey of 179 research papers as mentioned in the reference list below.

Further progress can be achieved by delving into sophisticated machine learning methodologies, integrating multimodal data, enriching the explainability of models, employing data augmentation and transfer learning, examining social network dynamics, embedding real-time fact-checking, augmenting contextual comprehension, and endorsing collaborative methods. Uninterrupted endeavors in these domains can lead to a more successful identification of fake news, thus empowering users to make educated decisions and counteract the dissemination of incorrect information. In general, the strides made in the sphere of fake news detection are commendable. Key contributors in facilitating the more effective and precise identification of fake news include AI, natural language processing, and machine learning. These technologies are now leveraged in the battle against the propagation of fake news on social media, assuring users of accurate and trustworthy information. As technology progresses, it is probable that these fake news detection methodologies will become progressively sophisticated.

Author Contributions T wrote the complete paper, AY and TA prepared the structure of the paper, All authors reviewed the paper.

Data availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval For this type of study, formal consent is not required.

Informed consent This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Agarwal A, Dixit A (2020) Fake news detection: an ensemble learning approach. In: 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp 1178–1183. https://doi.org/10.1109/ICICCS48265.2020.9121030
2. Agarwal V, Sultana HP, Malhotra S, Sarkar A (2019) Analysis of classifiers for fake news detection. Procedia Comput Sci 165:377–383
3. Ahmadi K, Khafaie T, Ganjoo M (2021) Rumor propagation detection in complex networks based on ILSR model and nodes degree. J Commun Eng 11(42):55–68
4. Ahmed H, Traore I, Saad S (2017) Detection of online fake news using n-gram analysis and machine learning techniques. In: International conference on intelligent, secure, and dependable systems in distributed and cloud environments, Springer, pp 127–138
5. Ahmed S, Hinkelmann K, Corradini F (2022) Combining machine learning with knowledge engineering to detect fake news in social networks-a survey. arXiv preprint arXiv:2201.08032
6. Ahmed SR, Sonuç E, Ahmed MR, Duru AD (2022) Analysis survey on deepfake detection and recognition with convolutional neural networks. In: 2022 International Congress on Human–Computer Interaction, Optimization and Robotic Applications (HORA), pp 1–7
7. Al-Asadi MA, Tasdemir S (2022) Using artificial intelligence against the phenomenon of fake news: a systematic literature review. In: Lahby M, Pathan ASK, Maleh Y, Yafooz WMS (eds) Combating fake news with computational intelligence techniques. Springer, Cham, pp 39–54
8. Alghamdi J, Lin Y, Luo S (2022) A comparative study of machine learning and deep learning techniques for fake news detection. Information 13(12):576
9. Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31(2):211–36
10. Althabiti S, Alsalka MA, Atwell E (2022) SCUoL at CheckThat! 2022: fake news detection using transformer-based models. Working Notes of CLEF
11. Alzaidi MS, Subbalakshmi C, Roshini T, Shukla PK, Shukla SK, Dutta P, Alhassan M (2022) 5G-telecommunication allocation

network using IoT enabled improved machine learning technique. Wirel Commun Mobile Comput 2022
12. Bagade A, Pale A, Sheth S, Agarwal M, Chakrabarti S, Chebrolu K, Sudarshan S (2020) The Kauwa-Kaate fake news detection system. In: Proceedings of the 7th ACM IKDD CoDS and 25th COMAD, pp 302–306
13. Bang YO, Woo SS (2021) Da-FDFtNet: dual attention fake detection fine-tuning network to detect various AI-generated fake images. arXiv preprint arXiv:2112.12001
14. Botnevik B, Sakariassen E, Setty V (2020) Brenda: browser extension for fake news detection. In: Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, pp 2117–2120
15. Braşoveanu AM, Andonie R (2021) Integrating machine learning techniques in semantic fake news detection. Neural Process Lett 53(5):3055–3072
16. Brookes S, Waller L (2022) Communities of practice in the production and resourcing of fact-checking. Journalism p 14648849221078465
17. Buntain C, Golbeck J (2017) Automatically identifying fake news in popular twitter threads. In: 2017 IEEE international conference on smart cloud (SmartCloud), pp 208–215
18. Buonomo B (2020) Effects of information-dependent vaccination behavior on coronavirus outbreak: insights from a SIRI model. Ricerche Mat 69(2):483–499
19. Campan A, Cuzzocrea A, Truta TM (2017) Fighting fake news spread in online social networks: actual trends and future research directions. In: 2017 IEEE International Conference on Big Data (Big Data), pp 4453–4457
20. Chaffey BD (2022) smartinsights. https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/
21. Chatterjee M, Pal S (2019) Busting fake news: need for digital media literacy. In: Rise of the Digital Human, 4th All India Media Conference, Udaipur, Rajasthan. Accessed February, vol 14, p 2022
22. Chaudhary L, Sharma S, Sajwan M (2022) Comparative analysis of supervised machine learning algorithm. Available at SSRN 4143890
23. Chen MY, Lai YW, Lian JW (2022a) Using deep learning models to detect fake news about COVID-19. ACM Trans Internet Technol
24. Chen Y, Li D, Zhang P, Sui J, Lv Q, Tun L, Shang L (2022) Cross-modal ambiguity learning for multimodal fake news detection. In: Proceedings of the ACM Web Conference 2022, pp 2897–2905
25. Choi H, Ko Y (2022) Effective fake news video detection using domain knowledge and multimodal data fusion on youtube. Pattern Recogn Lett 154:44–52
26. Choudhary A, Arora A (2021) Imagefake: an ensemble convolution models driven approach for image based fake news detection. In: 2021 7th International Conference on Signal Processing and Communication (ICSC), pp 182–187
27. Das M, Singh P, Majumdar A (2022) Investigating dynamics of polarization of youtube true and fake news channels. In: Causes and Symptoms of Socio-Cultural Polarization, pp 73–112
28. Davoudi M, Moosavi MR, Sadreddini MH (2022) DSS: a hybrid deep model for fake news detection using propagation tree and stance network. Expert Syst Appl 198:116635
29. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, pp 248–255
30. Dhawan A, Bhalla M, Arora D, Kaushal R, Kumaraguru P (2022) FakeNewsIndia: a benchmark dataset of fake news incidents in India, collection methodology and impact assessment in social media. Comput Commun 185:130–141
31. Draws T, La Barbera D, Soprano M, Roitero K, Ceolin D, Checco A, Mizzaro S (2022) The effects of crowd worker biases in fact-checking tasks. In: 2022 ACM Conference on Fairness, Accountability, and Transparency, pp 2114–2124
32. Felber T (2021) Constraint 2021: machine learning models for COVID-19 fake news detection shared task. arXiv preprint arXiv:2101.03717
33. Garg S, Sharma DK (2020) New Politifact: a dataset for counterfeit news. In: 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART), IEEE, pp 17–22
34. Giachanou A, Zhang G, Rosso P (2020) Multimodal multi-image fake news detection. In: 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), pp 647–654
35. Gilda S (2017) Notice of violation of IEEE publication principles: evaluating machine learning algorithms for fake news detection. In: 2017 IEEE 15th student conference on research and development (SCOReD), pp 110–115
36. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. Adv Neural Inf Process Syst, 27
37. Gray J, Bounegru L, Venturini T (2020) 'Fake news' as infrastructural uncanny. New Media Soc 22(2):317–341
38. Guo Z, Schlichtkrull M, Vlachos A (2022) A survey on automated fact-checking. Trans Assoc Comput Linguist 10:178–206
39. Hakak S, Alazab M, Khan S, Gadekallu TR, Maddikunta PKR, Khan WZ (2021) An ensemble machine learning approach through effective feature extraction to classify fake news. Future Gener Comput Syst 117:47–58
40. Han Y, Karunasekera S, Leckie C (2020) Graph neural networks with continual learning for fake news detection from social media. arXiv preprint arXiv:2007.03316
41. Hannah Nithya S, Sahayadhas A (2022) Automated fake news detection by LSTM enabled with optimal feature selection. J Inf Knowl Manag 21:2250036
42. Harrag F, Djahli MK (2022) Arabic fake news detection: a fact checking based deep learning approach. Trans Asian Low-Resour Lang Inf Process 21(4):1–34
43. He Y, Yu N, Keuper M, Fritz M (2021) Beyond the spectrum: detecting deepfakes via re-synthesis. arXiv preprint arXiv:2105.14376
44. Horne BD, Nørregaard J, Adali S (2019) Robust fake news detection over time and attack. ACM Trans Intell Syst Technol (TIST) 11(1):1–23
45. Hsu CC, Lee CY, Zhuang YX (2018) Learning to detect fake face images in the wild. In: 2018 international symposium on computer, consumer and control (IS3C), pp 388–3 91
46. Hsu CC, Zhuang YX, Lee CY (2020) Deep fake image detection based on pairwise learning. Appl Sci 10(1):370
47. Huang YF, Chen PH (2020) Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms. Expert Syst Appl 159:113584
48. Ibrishimova MD, Li KF (2019) A machine learning approach to fake news detection using knowledge verification and natural language processing. In: International Conference on Intelligent Networking and Collaborative Systems, pp 223–234
49. Islam MR, Liu S, Wang X, Xu G (2020) Deep learning for misinformation detection on online social networks: a survey and new perspectives. Soc Netw Anal Min 10(1):1–20
50. Jain DK, Kumar A, Shrivastava A (2022) CanarDeep: a hybrid deep neural model with mixed fusion for rumour detection in social data streams. Neural Comput Appl 34:15129–15140
51. Jain V, Kaliyar RK, Goswami A, Narang P, Sharma Y (2022) AENeT: an attention-enabled neural architecture for fake news detection using contextual features. Neural Comput Appl 34(1):771–782. https://doi.org/10.1007/s00521-021-06450-4

52. Jaiswal AK, Srivastava R (2019) Image splicing detection using deep residual network. In: Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE)
53. Jeon H, Bang Y, Woo SS (2020) Fdftnet: facing off fake images using fake detection fine-tuning network. In: IFIP international conference on ICT systems security and privacy protection, pp 416–430
54. Jiang T, Li JP, Haq AU, Saboor A, Ali A (2021) A novel stacking approach for accurate detection of fake news. IEEE Access 9:22626–22639
55. Jin Z, Cao J, Guo H, Zhang Y, Luo J (2017) Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In: Proceedings of the 25th ACM international conference on Multimedia, pp 795–816
56. Jindal S, Sood R, Singh R, Vatsa M, Chakraborty T (2020) Newsbag: a multimodal benchmark dataset for fake news detection. In: CEUR Workshop Proc, vol 2560, pp 138–145
57. Kaliyar RK, Dash P (2021) Rueval20: improving rumour detection on social media using a deep convolutional neural network. In: 8th ACM IKDD CODS and 26th COMAD, p 439
58. Kaliyar RK, Goswami A, Narang P (2019) Multiclass fake news detection using ensemble machine learning. In: 2019 IEEE 9th International Conference on Advanced Computing (IACC), pp 103–107
59. Kaliyar RK, Goswami A, Narang P, Sinha S (2020) FNDNet-a deep convolutional neural network for fake news detection. Cogn Syst Res 61:32–44
60. Kalsnes B (2018) Fake news. In: Oxford Research Encyclopedia of Communication
61. Kazemi A, Garimella K, Gaffney D, Hale SA (2021) Claim matching beyond English to scale global fact-checking. arXiv preprint arXiv:2106.00853
62. Khan T, Michalas A, Akhunzada A (2021) Fake news outbreak 2021: can we stop the viral spread? J Netw Comput Appl 190:103112
63. Khanam Z, Alwasel B, Sirafi H, Rashid M (2021) Fake news detection using machine learning approaches. In: IOP Conference Series: Materials Science and Engineering, IOP Publishing, vol 1099, p 012040
64. Khattar D, Goud JS, Gupta M, Varma V (2019) Mvae: multimodal variational autoencoder for fake news detection. In: The world wide web conference, pp 2915–2921
65. Kim S, Breen J, Dudkina E, Poloni F, Crisostomi E (2022) On the use of Markov chains for epidemic modeling on networks. arXiv preprint arXiv:2207.02737
66. Koloski B, Perdih TS, Robnik-Šikonja M, Pollak S, Škrlj B (2022) Knowledge graph informed fake news classification via heterogeneous representation ensembles. Neurocomputing 496:208–226
67. Kumar A, Aggarwal N, Kumar S (2022) SIRA: a model for propagation and rumor control with epidemic spreading and immunization for healthcare 5.0. Soft Comput 1–14
68. Kumar N, Pranav P, Nirney V, Geetha V (2021) Deepfake image detection using CNNs and transfer learning. In: 2021 International Conference on Computing, Communication and Green Engineering (CCGE), pp 1–6
69. Kumari R, Ekbal A (2021) Amfb: attention based multimodal factorized bilinear pooling for multimodal fake news detection. Expert Syst Appl 184:115412
70. Lahby M, Aqil S, Yafooz W, Abakarim Y (2022) Online fake news detection using machine learning techniques: a systematic mapping study. In: Lahby M, Pathan ASK, Maleh Y, Yafooz WMS (eds) Combating fake news with computational intelligence techniques. Springer, Cham, pp 3–37
71. Lampridis O, Karanatsiou D, Vakali A (2022) Manifesto: a human-centric explainable approach for fake news spreaders detection. Computing 104(4):717–739
72. Li P, Sun X, Yu H, Tian Y, Yao F, Xu G (2021) Entity-oriented multi-modal alignment and fusion network for fake news detection. IEEE Trans Multimedia 24:3455–3468
73. Liu L, Roche DS, Theriault A, Yerukhimovich A (2021) Fighting fake news in encrypted messaging with the fuzzy anonymous complaint tally system (facts). arXiv preprint arXiv:2109.04559
74. Ma K, Tang C, Zhang W, Cui B, Ji K, Chen Z, Abraham A (2022) DC-CNN: dual-channel convolutional neural networks with attention-pooling for fake news detection. Appl Intell 1–16
75. Mahfoudi G, Tajini B, Retraint F, Morain-Nicolier F, Dugelay JL, Marc P (2019) DEFACTO: image and face manipulation dataset. In: 2019 27Th european signal processing conference (EUSIPCO), IEEE, pp 1–5
76. Meel P, Vishwakarma DK (2020) Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst Appl 153:112986
77. Mengoni P, Yang J (2022) Empowering COVID-19 fact-checking with extended knowledge graphs. In: International Conference on Computational Science and Its Applications, pp 138–150
78. Mishra S, Shukla P, Agarwal R (2022) Analyzing machine learning enabled fake news detection techniques for diversified datasets. Wirel Commun Mobile Comput 2022:1–18
79. Mishra S, Suryavardan S, Bhaskar A, Chopra P, Reganti A, Patwa P, Das A, Chakraborty T, Sheth A, Ekbal A, et al. (2022) Factify: a multi-modal fact verification dataset. In: Proceedings of the First Workshop on Multimodal Fact-Checking and Hate Speech Detection (DE-FACTIFY)
80. Mohamad Nezami O, Dras M, Anderson P, Hamey L (2019) Face-cap: image captioning using facial expression analysis. In: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10–14, 2018, Proceedings, Part I 18, Springer, pp 226–240
81. Monti F, Frasca F, Eynard D, Mannion D, Bronstein MM (2019) Fake news detection on social media using geometric deep learning. arXiv preprint arXiv:1902.06673
82. Mridha MF, Keya AJ, Hamid MA, Monowar MM, Rahman MS (2021) A comprehensive review on fake news detection with deep learning. IEEE Access 9:156151–156170
83. Nakamura K, Levy S, Wang WY (2019) r/fakeddit: a new multimodal benchmark dataset for fine-grained fake news detection. arXiv preprint arXiv:1911.03854
84. Nakov P, Barrón-Cedeño A, Da San Martino G, Alam F, Struß JM, Mandl T, Míguez R, Caselli T, Kutlu M, Zaghouani W, et al. (2022) The clef-2022 checkthat! lab on fighting the covid-19 infodemic and fake news detection. In: European Conference on Information Retrieval, pp 416–428
85. Nakov P, Barrón-Cedeño A, da San Martino G, Alam F, Struß JM, Mandl T, Míguez R, Caselli T, Kutlu M, Zaghouani W, et al. (2022) Overview of the clef–2022 checkthat! lab on fighting the covid-19 infodemic and fake news detection. In: International Conference of the Cross-Language Evaluation Forum for European Languages, Springer, pp 495–520
86. Nielsen DS, McConville R (2022) Mumin: a large-scale multilingual multimodal fact-checked misinformation social network dataset. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 3141–3153
87. Oshikawa R, Qian J, Wang WY (2018) A survey on natural language processing for fake news detection. arXiv preprint arXiv:1811.00770

88. Ozbay FA, Alatas B (2020) Fake news detection within online social media using supervised artificial intelligence algorithms. Physica A: Stat Mech Appl 540:123174
89. Pankovska E, Schulz K, Rehm G (2022) Suspicious sentence detection and claim verification in the COVID-19 domain. In: Proceedings of the Workshop Reducing Online Misinformation through Credible Information Retrieval (ROMCIR 2022). CEUR-WS, Stavanger
90. Parekh Z, Baldridge J, Cer D, Waters A, Yang Y (2020) Crisscrossed captions: extended intramodal and intermodal semantic similarity judgments for MS-COCO. arXiv preprint arXiv:2004.15020
91. Pathak A (2022) An integrated approach towards automated fact-checking. PhD thesis, State University of New York at Buffalo
92. Pavleska T, Školkay A, Zankova B, Ribeiro N, Bechmann A (2018) Performance analysis of fact-checking organizations and initiatives in Europe: a critical overview of online platforms fighting fake news. Soc Media Converg 29:1–28
93. Peng X, Xintong B (2022) An effective strategy for multi-modal fake news detection. Multimedia Tools Appl 81(10):13799–13822
94. Pennycook G, Rand DG (2021) The psychology of fake news. Trends Cogn Sci 25(5):388–402
95. Piazza JA (2022) Fake news: the effects of social media disinformation on domestic terrorism. Dyn Asymmetric Confl 15(1):55–77
96. Pritzkau A, Blanc O, Geierhos M, Schade U (2022) NLytics at CheckThat! 2022: hierarchical multi-class fake news detection of news articles exploiting the topic structure. Working Notes of CLEF
97. Probierz B, Kozak J, Stefański P, Juszczuk P (2021) Adaptive goal function of ant colony optimization in fake news detection. In: International Conference on Computational Collective Intelligence, pp 387–400
98. Qazi M, Khan MU, Ali M (2020) Detection of fake news using transformer model. In: 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), pp 1–6
99. Qi P, Cao J, Yang T, Guo J, Li J (2019) Exploiting multi-domain visual information for fake news detection. In: 2019 IEEE international conference on data mining (ICDM), pp 518–527
100. Qian S, Hu J, Fang Q, Xu C (2021) Knowledge-aware multi-modal adaptive graph convolutional networks for fake news detection. ACM Trans Multimedia Comput, Commun, Appl (TOMM) 17(3):1–23
101. Qian S, Wang J, Hu J, Fang Q, Xu C (2021) Hierarchical multi-modal contextual attention network for fake news detection. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 153–162
102. Raj C, Meel P (2021) Convnet frameworks for multi-modal fake news detection. Appl Intell 51(11):8132–8148
103. Raj C, Meel P (2021) Convnet frameworks for multi-modal fake news detection. Appl Intell 51:8132–8148
104. Ramya S, Eswari R (2021) Attention-based deep learning models for detection of fake news in social networks. Int J Cogn Inform Nat Intel (IJCINI) 15(4):1–25
105. Ramya S, Eswari R (2022) Performance of optimization algorithms in attention-based deep learning model for fake news detection system. In: Proceedings of International Conference on Computational Intelligence, pp 113–126
106. Rana MS, Nobi MN, Murali B, Sung AH (2022) Deepfake detection: a systematic literature review. IEEE Access 10:25494–25513
107. Rashkin H, Choi E, Jang JY, Volkova S, Choi Y (2017) 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
108. Rastogi S, Bansal D (2022) A review on fake news detection 3T's: typology, time of detection, taxonomies. Int J Inf Secur 1–36
109. Raza S, Ding C (2022) Fake news detection based on news content and social contexts: a transformer-based approach. Int J Data Sci Anal 13(4):335–362
110. Riedel B, Augenstein I, Spithourakis GP, Riedel S (2017) A simple but tough-to-beat baseline for the fake news challenge stance detection task. arXiv preprint arXiv:1707.03264
111. Rohera D, Shethna H, Patel K, Thakker U, Tanwar S, Gupta R, Hong WC, Sharma R (2022) A taxonomy of fake news classification techniques: survey and implementation aspects. IEEE Access 10:30367–30394
112. Sagnika S, Mishra BSP, Meher SK (2021) An attention-based CNN-LSTM model for subjectivity detection in opinion-mining. Neural Comput Appl 33:17425–17438
113. Saji R, Anand SK, Chandavarkar B (2021) Comparing CNNs and GANs for image completion. In: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp 1–7
114. 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, pp 531–540
115. Schoenmueller V, Blanchard SJ, Johar GV (2022) Who will share fake-news on twitter? psycholinguistic cues in online post histories discriminate between actors in the misinformation ecosystem. arXiv preprint arXiv:2203.10560
116. Segura-Bedmar I, Alonso-Bartolome S (2022) Multimodal fake news detection. Information 13(6):284
117. Seow JW, Lim MK, Phan RCW, Liu JK (2022) A comprehensive overview of deepfake: generation, detection, datasets, and opportunities. Neurocomputing 513:351–371
118. Shahid W, Jamshidi B, Hakak S, Isah H, Khan WZ, Khan MK, Choo KKR (2022) Detecting and mitigating the dissemination of fake news: challenges and future research opportunities. IEEE Trans Comput Soc Syst
119. Shao Y, Sun J, Zhang T, Jiang Y, Ma J, Li J (2022) Fake news detection based on multi-modal classifier ensemble. In: Proceedings of the 1st International Workshop on Multimedia AI against Disinformation, pp 78–86
120. Sharma DK, Garg S (2021) IFND: a benchmark dataset for fake news detection. Complex Intell Syst 1–21
121. Sharma K, Qian F, Jiang H, Ruchansky N, Zhang M, Liu Y (2019) Combating fake news: a survey on identification and mitigation techniques. ACM Trans Intell Syst Technol (TIST) 10(3):1–42
122. Sharma U, Saran S, Patil SM (2020) Fake news detection using machine learning algorithms. Int J Creative Res Thoughts (IJCRT) 8(6):509–518
123. Shrivastava G, Kumar P, Ojha RP, Srivastava PK, Mohan S, Srivastava G (2020) Defensive modeling of fake news through online social networks. IEEE Trans Comput Soc Syst 7(5):1159–1167
124. Shrivastava S, Singh R, Jain C, Kaushal S (2022) A research on fake news detection using machine learning algorithm. In: Smart Systems: Innovations in Computing: Proceedings of SSIC 2021, pp 273–287
125. Shu K, Sliva A, Wang S, Tang J, Liu H (2017) Fake news detection on social media: a data mining perspective. ACM SIGKDD Explor Newsl 19(1):22–36
126. Shu K, Wang S, Liu H (2019) Beyond news contents: The role of social context for fake news detection. In: Proceedings of the twelfth ACM international conference on web search and data mining, pp 312–320
127. Shu K, Mahudeswaran D, Wang S, Lee D, 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):171–188

128. Silva A, Han Y, Luo L, Karunasekera S, Leckie C (2021) Propagation2vec: embedding partial propagation networks for explainable fake news early detection. Inf Process Manag 58(5):102618
129. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
130. Singh B, Sharma DK (2021) Predicting image credibility in fake news over social media using multi-modal approach. Neural Comput Appl 34:21503–21517
131. Singh P, Srivastava R, Rana K, Kumar V (2022) SEMI-FND: stacked ensemble based multimodal inference for faster fake news detection. arXiv preprint arXiv:2205.08159
132. Singh VK, Ghosh I, Sonagara D (2021) Detecting fake news stories via multimodal analysis. J Am Soc Inf Sci 72(1):3–17
133. Singhal S, Shah RR, Chakraborty T, Kumaraguru P, Satoh S (2019) Spotfake: a multi-modal framework for fake news detection. In: 2019 IEEE fifth international conference on multimedia big data (BigMM), pp 39–47
134. Singhal S, Shah RR, Kumaraguru P (2022) FactDrill: a data repository of fact-checked social media content to study fake news incidents in India. In: Proceedings of the International AAAI Conference on Web and Social Media, vol 16, pp 1322–1331
135. Skinnell R (2021) Teaching writing in the (new) era of fake news. Coll Compos Commun 72(4):546–569
136. Soni B, Das PK, Thounaojam DM (2019) Geometric transformation invariant block based copy-move forgery detection using fast and efficient hybrid local features. J Inf Secur Appl 45:44–51
137. Sorour SE, Abdelkader HE (2022) AFND: Arabic fake news detection with an ensemble deep CNN-LSTM model. J Theor Appl Inf Technol 100(14):5072–5086
138. Sudhakar M, Kaliyamurthie K (2023) Efficient prediction of fake news using novel ensemble technique based on machine learning algorithm. In: Information and Communication Technology for Competitive Strategies (ICTCS 2021), pp 1–8
139. Tariq S, Lee S, Kim H, Shin Y, Woo SS (2019) Gan is a friend or foe? A framework to detect various fake face images. In: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, pp 1296–1303
140. Tian L, Zhang X, Peng M (2020) FakeFinder: twitter fake news detection on mobile. In: Companion Proceedings of the Web Conference 2020, pp 79–80
141. Tiwary T, Mahapatra RP (2022) An accurate generation of image captions for blind people using extended convolutional atom neural network. Multimedia Tools Appl 1–30
142. Tuan NMD, Minh PQN (2021) Multimodal fusion with BERT and attention mechanism for fake news detection. In: 2021 RIVF International Conference on Computing and Communication Technologies (RIVF), pp 1–6
143. Tyagi S, Yadav D (2022) MiniNet: a concise CNN for image forgery detection. Evol Syst 1–12
144. Vargo CJ, Guo L, Amazeen MA (2018) The agenda-setting power of fake news: a big data analysis of the online media landscape from 2014 to 2016. New Media Soc 20(5):2028–2049
145. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I (2017) Attention is all you need. In: Guyon I, Von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R (eds) Advances in neural information processing systems, vol 30. Curran Associates Inc., New York
146. Verma PK, Agrawal P (2022) PropFND: propagation based fake news detection. In: Unhelker B, Pandey HM, Raj G (eds) Applications of artificial intelligence and machine learning. Springer, Singapore, pp 557–568
147. Vijjali R, Potluri P, Kumar S, Teki S (2020) Two stage transformer model for COVID-19 fake news detection and fact checking. arXiv preprint arXiv:2011.13253
148. Vishwakarma DK, Varshney D, Yadav A (2019) Detection and veracity analysis of fake news via scrapping and authenticating the web search. Cogn Syst Res 58:217–229
149. Wagener A (2020) Hypernarrativity, storytelling, and the relativity of truth: digital semiotics of communication and interaction. Postdigital Sci Educ 2(1):147–169
150. Wang J, Sun Z, Wang J, Wu H, Hu X (2020) A two-stage attention-based model for fake news detection. arXiv preprint arXiv:2004.14420
151. Wang J, Mao H, Li H (2022) FMFN: fine-grained multimodal fusion networks for fake news detection. Appl Sci 12(3):1093
152. Wang J, Zeng K, Ma B, Luo X, Yin Q, Liu G, Jha SK (2022) GAN-generated fake face detection via two-stream CNN with PRNU in the wild. Multimedia Tools Appl 81:42527–42545
153. Wang WY (2017) "liar, liar pants on fire": a new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648
154. Wang Y, Ma F, Jin Z, Yuan Y, Xun G, Jha K, Su L, Gao J (2018) Eann: event adversarial neural networks for multi-modal fake news detection. In: Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining, pp 849–857
155. Wang Y, Ma F, Jin Z, Yuan Y, Xun G, Jha K, Su L, Gao J (2018) Eann: event adversarial neural networks for multi-modal fake news detection. In: KDD 2018 - Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
156. Wei P, Wu F, Sun Y, Zhou H, Jing XY (2022) Modality and event adversarial networks for multi-modal fake news detection. IEEE Signal Process Lett 29:1382–1386
157. Westerlund M (2019) The emergence of deepfake technology: a review. Technol Innov Manag Rev 9(11):39–52
158. Wright D, Wadden D, Lo K, Kuehl B, Cohan A, Augenstein I, Wang LL (2022) Generating scientific claims for zero-shot scientific fact checking. arXiv preprint arXiv:2203.12990
159. Wu K, Yang S, Zhu KQ (2015) False rumors detection on sina weibo by propagation structures. In: 2015 IEEE 31st international conference on data engineering, pp 651–662
160. Xue J, Wang Y, Tian Y, Li Y, Shi L, Wei L (2021) Detecting fake news by exploring the consistency of multimodal data. Inf Process Manag 58(5):102610
161. Yang C, Xu B, Khan JY, Uddin G, Han D, Yang Z, Lo D (2022) Aspect-based api review classification: How far can pre-trained transformer model go?. In: 2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER), IEEE, pp 385–395
162. Yang J, Xiao S, Lv Z (2022) Protecting the trust and credibility of data by tracking forgery trace based on GANs. Digit Commun Netw 8:877–884
163. Yang Z, Ma J, Chen H, Lin H, Luo Z, Chang Y (2022) A coarse-to-fine cascaded evidence-distillation neural network for explainable fake news detection. arXiv preprint arXiv:2209.14642
164. Ying L, Yu H, Wang J, Ji Y, Qian S (2021) Fake news detection via multi-modal topic memory network. IEEE Access 9:132818–132829
165. Ying L, Yu H, Wang J, Ji Y, Qian S (2021) Multi-level multi-modal cross-attention network for fake news detection. IEEE Access 9:132363–132373
166. Yu C, Wang W (2022) Fast transformation of discriminators into encoders using pre-trained GANs. Pattern Recogn Lett 153:92–99
167. Zhan J, Li X, Wang J, Liu H, Huang S (2019) A multi-head attention neural network model for fake news detection. arXiv preprint arXiv:1910.09871
168. Zhang H, Qian S, Fang Q, Xu C (2021) Multi-modal meta multi-task learning for social media rumor detection. IEEE Trans Multimedia 24:1449–1459

169. Zhang T, Wang D, Chen H, Zeng Z, Guo W, Miao C, Cui L (2020) BDANN: BERT-based domain adaptation neural network for multi-modal fake news detection. In: 2020 international joint conference on neural networks (IJCNN), pp 1–8
170. Zhou K, Shu C, Li B, Lau JH (2019a) Early rumour detection. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol 1 (Long and Short Papers), pp 1614–1623, Association for Computational Linguistics, Minneapolis, Minnesota, https://doi.org/10.18653/v1/N19-1163, https://aclanthology.org/N19-1163
171. Zhou X, Zafarani R (2018) Fake news: a survey of research, detection methods, and opportunities. arXiv preprint arXiv:1812.00315 2
172. Zhou X, Zafarani R (2020) A survey of fake news: fundamental theories, detection methods, and opportunities. ACM Comput Surv (CSUR) 53(5):1–40
173. Zhou X, Wu J, Zafarani R (2020) Safe: similarity-aware multi-modal fake news detection (2020). Preprint arXiv:2003.04981
174. Zubiaga A, Aker A, Bontcheva K, Liakata M, Procter R (2018) Detection and resolution of rumours in social media: a survey. ACM Comput Surv (CSUR) 51(2):1–36