



A systematic literature review and existing challenges toward fake news detection models

Minal Nirav Shah¹ · Amit Ganatra²

Received: 19 May 2022 / Revised: 27 September 2022 / Accepted: 29 October 2022 / Published online: 14 November 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Emerging of social media creates inconsistencies in online news, which causes confusion and uncertainty for consumers while making decisions regarding purchases. On the other hand, in existing studies, there is a lack of empirical and systematic examination observed in terms of inconsistency regarding reviews. The spreading of fake news and disinformation on social media platforms has adverse effects on stability and social harmony. Fake news is often emerging and spreading on social media day by day. It results in influencing or annoying and also misleading nations or societies. Several studies aim to recognize fake news from real news on online social media platforms. Accurate and timely detection of fake news prevents the propagation of fake news. This paper aims to conduct a review on fake news detection models that is contributed by a variety of machine learning and deep learning algorithms. The fundamental and well-performing approaches that existed in the past years are reviewed and categorized and described in different datasets. Further, the dataset utilized, simulation platforms, and recorded performance metrics are evaluated as an extended review model. Finally, the survey expedites the research findings and challenges that could have significant implications for the upcoming researchers and professionals to improve the trust worthiness of automated fake news detection models.

Keywords Fake news detection models · Deep learning · Machine learning · Dataset utilized · Simulation platforms · Recorded performance metrics · Research gaps and challenges

1 Introduction

Fake news identification is one of the eminent research topics, which has been studied in recent years (Sengupta et al. 2021). Fake news is often spread by yellow journalism before digital technology with the intention of glorious news like hilarious news, accidents, rumors, and crime news (Islam et al. 2020). In the digital era, it is simpler for spreading fake news while a user may distribute fake news to neighbors, their friends, and so on due to the unique characteristics of social media (Habib et al. 2019). Thus, fake

news can be propagated in a cycle format because of the vast usage of social media by every individual (Singh and Sharma 2021). Moreover, comments on fake news can be varied every time that reducing the reliability of real news, the fake news has directly spread in a faster way while comparing to real news (Yang et al. 2021). Fake news can have several ranges of impacts by misleading or influencing governments or whole populations (Kim and Ko November 2021). Some of the techniques in detecting fake news utilize several approaches like machine learning techniques, language techniques, and knowledge-based techniques (Vereshchaka et al. 2020). Some of the strong social media platforms are Twitter, YouTube, Instagram, Facebook, and WhatsApp have offered news and entertainment with the rising employment of mobile devices and simpler WiFi connections (Ribeiro Bezerra 2021). Social media and emerging technologies have several profiles for propagating fake news (Sharma 2021). Every technology has its extremities and limitations due to the positive effects technology on society and social media (Mridha et al. 2021). Moreover,

✉ Minal Nirav Shah
minal11086@gmail.com

Amit Ganatra
provost@paruluniversity.ac.in

¹ Research Scholar, Department of Computer Science, CSPIT, CHARUSAT, Charotar University of Science and Technology, Changa, Gujarat 388421, India

² Provost, Parul University, P.O.Limda, Ta.Waghodia, Vadodara, Gujarat 391760, India

the recent literature analyzes the several advantages of fake news detection.

Nowadays, online fake news has become the main aspect of the growing interest in online media, social-networking sites, and online news portals (Bondielli and Marcelloni 2019). However, most people are generally incompetent for spending adequate time cross-checking the references and for ensuring the credibility of news (Zhou and Zafarani 2020; D'Ulizia 2021). Thus, more attention to fake news detection inspires the research community. In recent days, more research works regarding fake news detection have been implemented (Rama Krishna et al. 2021), though several studies only concentrated on news of specific categories like political or e-commerce reviews. Consequently, they have designed and developed certain features with some standard datasets with their topic of interest. These studies face poor performance in detecting news of another topic and also dataset bias (Beer and Matthee 2020). Therefore, it is necessary for studying whether these models are suitable for diverse classes of news propagated in social media through the evaluation of diverse datasets on different models and investigating their efficiency or performances (Ahmad et al. 2020). On the other hand, conventional studies on fake news detection techniques are focused on either a limited number of models or a particular category of the dataset (Dabbous et al. 2020a). Thus, there is a need of reviewing a fake news detection model.

The study on fake news detection needs a huge number of evaluations through machine learning techniques on a broad range of datasets (Kansal 2021). New methods should obtain deep knowledge regarding the nature of fake news and the way of spreading it over the world. On the other hand, the recent work contributes in this way to implementing a model through novel approaches, which verify the significance of deep learning approaches for detecting fake news (Meneses Silva et al. 2021). Among them, the "Convolutional Neural Network (CNN)" model has been utilized and shown a higher competitive performance while comparing with other existing machine learning models. In addition, "long short-term memory (LSTM)" has utilized for analyzing linguistic features and has shown noteworthy performance (Simko et al. 2021). More particularly, different variants of CNN can be suggested for detecting fake news. Although deep learning algorithms offer superior efficiency in getting classification results, they suffer from certain challenges like lack of interpretability, the necessity of large training datasets, and complexity in discovering the optimal hyperparameters for every dataset and problem (Silva et al. 2021). Thus, recent advancements in bio-inspired approaches permit the optimization of deep learning constraints, and the necessity of advanced intelligent techniques also increases for solving the problems persisted in existing works (Chauhan and Palivela 2021). Consequently, there is a need of studying the recent

research works in the field of fake news identification models to assist social media users in getting real news.

The major focus of the study on different fake news detection models is given here.

- To prepare an in-depth survey on fake news detection models by collecting noteworthy information from recent studies along with diverse algorithms utilized for achieving it.
- To present a complete study about a chronological review, their related works and contribution to fake news detection models, research designs, and general findings on fake news detection models.
- To analyze the performance metrics, applications focused, datasets used, simulation platforms utilized, and necessary research gap with the challenges present in existing fake news detection models.

The remaining sections of the paper are given here. Section 2 discusses the literature survey, research designs, and general findings on fake news detection with the chronological review. Section 3 specifies the algorithmic classification, feature extraction techniques, and dataset used in the existing fake news detection models. Section 4 gives the simulation platforms and applications focused on the existing fake news detection models. Section 5 describes the architectural view of general fake news detection models and performance measures used in state-of-the-art fake news detection models. Section 6 shows the consequences of fake news and research challenges and the future scope of fake news detection models. Section 7 concludes the survey.

2 Literature survey, research designs, and general findings on fake news detection with a chronological review

2.1 Related works

2.1.1 Existing fake news detection model approaches

In 2019, Ko et al. (2019) have used a reverse-tracking approach for defining the possibility of fake news in the articles that were taken from the cognitive system, where the designed model has been termed a Fake News Detection System (FNDS). This model has been tested in two case studies, where the first one was posted on February 9, 2017. This article was given about the "blacklisting of extreme right movie studios and Kim Jong Dae, a member of the National Assembly". The second case study was posted in September 2017. This article was about "North Korea, nuclear test, earthquake, Kim Jeong Eun". This model was faster than the conventional models. In 2019, Barbado et al.

(2019) have implemented a feature framework to detect fake reviews in the consumer electronics area, which consisted of four stages of building a dataset to classify fake reviews in four cities in the consumer electronics area through a scraping approach. Secondly, a feature scheme for proposing fake review detection was suggested with the exploitation of the social perspective. It has also focused on selecting fake reviews for organizing and characterizing the features. The results have shown that the AdaBoost classifier has performed better than other existing classifiers. In 2019, Shu et al. (2019) have implemented a new FakeNewsTracker for gathering the social context and new pieces to create valuable datasets, and then, useful features were extracted for detecting the fake news through Social Article Fusion model, and then, different machine learning models were built for detecting the fake news. In 2020, Henrique and Ferreira (2020) have implemented a new fake news detection model for detecting fake news from social media texts in Germanic, Latin, and Slavic languages. The detection was carried out through “support vector machines and random forest”. In 2020, Talwar et al. (2020) have adopted a mixed method for exploring the sharing behavior of fake news. This model has identified six behavioral manifestations correlated with fake news sharing from qualitative data. The control variables were taken as gender and age. This model had created a positive effect on fake news sharing owing to religiosity and lack of time. This model has suggested that active corrective action was engaged by social media users for sharing fake news. In 2020, Xu et al. (2020) have characterized that numerous real and fake news was shared by comments, reactions, and shares on Facebook in two ways like content understanding and domain reputation. It has revealed that the web sites of news publishers have exhibited various domain popularity, domain ranking, registration timing, and registration behaviors. Additionally, for a certain amount of time, fake news has disappeared. Further, news has been fed to “latent Dirichlet allocation (LDA) topic modeling” and TF-IDF for fake news detection when discovering the document similarity with the word and term vectors. In 2020, Oliveira et al. (2020) have implemented a “computational-stylistic analysis based on NLP”. This model has used one-class SVM to detect fake news and applied it to data for reducing the dimensionality reduction approaches like data compaction and latent semantic analysis (LSA). In 2020, Li et al. (2020) have introduced the MCNN for getting the global semantics and local convolutional features for getting the semantic information from the texts for classifying fake news. The weight of sensitive words (TFW) method was used for computing the robust significance of true or fake labels. Thus, MCNN-TFW has focused on extracting the weight of sensitive words and article representation for each news. This model has achieved higher accuracy than other existing approaches to datasets. In 2020, Vereshchaka

et al. (2020) have solved the issue of predicting fake news by getting the socio-cultural and textual characteristics of fake news features and by analyzing and detecting fake news features. Further, data analytics was investigated for constructing a concordance of phase and word frequency. They have formed binary classifiers for extracting the features through deep learning algorithms like GRU, RNN, and LSTM. In 2020, Kaur et al. (2020) have implemented a new fake news identification model through a multi-level voting ensemble model including 12 classifiers, where the features were extracted using Hashing-Vectorizer (HV), “Count-Vectorizer (CV), and Term Frequency–Inverse Document Frequency (TF–IDF)” through three datasets. It has predicted the textual or fake content from online social media websites. It has verified the performance in terms of less training time, better efficiency, and the trade-off between accuracy and efficiency. Thus, from the analysis, the best classifier was selected for both higher accuracy and efficiency. In 2021, Shahbazi and Byun (2021) have implemented an integrated model for different criteria of natural language processing and block chain for applying machine learning approaches for detecting fake news and offered a better prediction on posts and accounts on fake users. They used reinforcement learning approach that was used for this process. This scheme utilized the decentralized block chain framework for offering security, which has offered the method of outlining the digital contents. Lastly, the learning rate of the model was predicted for detection to explore the correlation among contents. In 2021, Mehta et al. (2021) have focused on the fake news classification model through “Bidirectional Encoder Representations from Transformers termed as BERT”. It has required a nominal pre-processing technique and has utilized two diverse versions of BERT, which has shown considerable improvement in terms of the fake news classification model regarding binary classification measures. The designed model has shown higher reliability in terms of multi-label classification. In 2021, Shishah (2021) has implemented a new fake news detection model through BERT with a joint learning scheme by integrating the Named Entity Recognition (NER) and Relational Features Classification (RFC). In 2021, Jiang et al. (2021) have investigated the efficiency of three deep learning models and five machine learning models. The superior performance of the designed model was verified while estimating with other existing approaches. In 2021, Kumari and Ekbal (2021) have implemented a novel multimodal fake news detection scheme with a suitable fusion of multimodal features, which leveraged the information from images and text and tried for maximizing the correlation among them for efficient multimodal distributed depiction. The performance of the designed model was improved by combining the text with images. The experiments have been conducted for validating

the efficiency of the designed model that has attained superior performance to others.

2.1.2 Fake news detection model with existing diverse classifiers

In 2018, Jang et al. (2018) have studied the problem present in the US presidential election in the year 2016. It has gathered 307,738 tweets with 30 real and 30 fake news stories. Thus, there was a need of examining the evolution patterns, producers of the source, and root content. They have utilized the evolution tree modeling method for examining misinformation, the transmission of news, and management. Finally, based on the diverse evolution patterns, fake and real news have been identified. In 2019, Altunbey and Alatas (2020) have designed a two-step method to identify the “fake news on social media”, which has several steps like pre-processing, vector conversion, and classification. The pre-processing was carried out for converting the unstructured datasets. Initially, the texts in the dataset including the news were depicted by vectors through the attained Document-term matrix and Term-Frequency (TF) weighting method. Secondly, 23 supervised algorithms, like kernel logistic regression (KLR), IBk, decision tree, bagging, sequential minimal optimization (SMO), J48, attribute selected classifier (ASC), simple cart, ordinal learning model (OLM), Ridor, multi-layer perceptron (MLP), weighted instances handler wrapper (WIHW), classification via clustering (CvC), locally weighted learning (LWL), logistic model tree (LMT), randomizable filtered classifier (RFC), CV parameter selection (CVPS), stochastic gradient descent (SGD), ZeroR, decision stump, OneR, JRip, and BayesNet, have been experimented in the dataset for transforming the structured format with the text mining algorithms. In 2019, Jadhav and Thepade (2019) have implemented a new framework for detecting and classifying fake news messages through Deep Structure Semantic Model (DSSM) and improved RNN classifiers. Initially, the twitter data was pre-processed using tokenization, and then, TF-IDF and CV were used for extracting the features. Further, semantic features and multi-layer projection were performed in DSSM, and then, the data were forwarded to the improved RNN for classifying the fake news. In 2020, Kumar et al. (2020) have developed a new fake news detection model with the help of deep CNN (FNDNet), which has learned the discriminatory features automatically to classify the fake news by several hidden layers. Further, various features were extracted at every layer for maximizing the accuracy of detection. In 2020, Singh et al. (2020) have explored Bernoulli's Naive Bayes Classifier with the help of "Multinomial Naive Bayes with predictors as Boolean variables" for detecting fake news. This model has classified the data into two classes 1 or 0, where 1 stand for unique news articles and fake new is represented by 1. In 2020, Hiriyannaiah

et al. (2020) have offered the adverse effects of fake news in society particularly advancements in the vast usage of social media. Generative adversarial networks (GANs) have offered more efficient results in getting fake news detection in terms of validation accuracy than other existing classifiers. It has learned the complex functions for getting a higher accuracy rate. This model has solved the problem of gradient descent in GANs. It has used SeqGAN, where the REINFORCE algorithm was used in the generator for updating their weights by taking the identification of the discriminator network. In 2020, Umer et al. (2020) have suggested a novel hybrid neural network architecture by LSTM and CNN for classifying the news articles with stance labels, which has also gathered data from the news articles. Initially, the features were gathered from articles using word2vectors, where the dimensionality reduction was carried out for getting the minimized feature set. Finally, the hybrid CNN-LSTM was utilized for detecting fake news for showing its effectiveness. In 2020, Agarwal et al. (2020) have suggested a new deep learning system for predicting the nature of an article while using an input. They have pre-processed the texts using “word embedding (GloVe)” for constructing a vector space of words and established a lingual relationship. They have combined CNN and RNN for getting the benchmark outcomes while predicting fake news. Moreover, this model has minimized the overfitting problem by the dropout layer and generated higher accurate values. It was shown that the designed model has attained superior outcomes while evaluated with other algorithms. In 2020, Shrivastava et al. (2020) have investigated the propagation of fake news and described the dissemination of misinformation between groups with the influence of several misinformations refuting metrics. This model has considered the fake news prediction model from online social networks during COVID-19. They have also completely analyzed the equilibrium and stability, which has also prevented the spreading of fake news. It has also been verified via examination of users in online social networks. Several conditions have been evaluated to demonstrate the performance of social network stability and verified theoretical outcomes by experimental outcomes. In 2020, Mahabub (2020) has suggested an Ensemble Voting Classifier for designing a new fake news detection model among fake and real tasks. The detection was performed by considering several classifiers. Then, better three machine learning algorithms in “Ensemble Voting Classifier” were used after cross-validation. This model has attained superior results while comparing with existing classifiers while detecting fake messages, fake profiles, etc.; finally, the results have verified the superior sufficiency scores of individual classifiers. In 2020, Choudhary and Arora (2020) have proposed a solution for detecting and classifying fake news through a linguistic model for getting the properties of content for generating the language-driven features. It has

also extracted the readability, sentimental, grammatical, and syntactic features of specific news. This model solves the problems of handcrafted features and time-consumption. Thus, for getting superior results in detecting fake news, a neural-based sequential learning model was applied in terms of accuracy and time. In 2021, Ying et al. (2021a) have designed a new "end-to-end Multi-level Multi-modal Cross-attention Network (MMCN)". The high-quality representations have been generated for image regions and text words, respectively, by pre-trained ResNet and BERT models. They have further fused the feature embeddings of the image regions and text words, respectively, for getting diverse as well as duplicate modalities. A multi-level encoding network was used for getting the higher multi-level semantics for enhancing the depictions regarding posts owing to the diverse layers of transformer architecture. In 2021, Li et al. (2021a) have suggested a new automatic model for returning and adding accurate outcomes for assisting the neural network in getting positive sample cases for enhancing the accuracy of the neural network. Initially, this model gathered data, and then, supervised and unsupervised tasks were trained simultaneously through a semi-supervised deep learning network. In 2021, Sivasankari and Vadivu (2021) have studied a new detection and identification model for learning discriminative features from Facebook posts, tweets content through social network graphs. It has often confirmed the minimization of their propagation. In 2021, Braşoveanu and Andonie (2021) have improved fake news detection through semantic features. They have suggested a new semantic fake news detection model with relational features such as facts, entities, or sentiment extracted from text. They have mostly considered short texts with several degrees of truth and shown that utilizing semantic features focused on enhancing the accuracy. In 2021, Setiawan et al. (2021) have implemented the Hybrid "Support Vector Machine (SVM)" for detecting fake news, where the data were gathered from the standard dataset that has been subjected to a feature extraction phase through TF-IDF. Then, the classification was performed by hybrid SVM. In 2021, Raj and Meel (2021) have implemented coupled ConvNet architecture with image-CNN and text-CNN modules for detecting fake news. Initially, the input data were pre-processed in both modules and then given to CNN. In addition, coupled ConvNet architecture was suggested by extending the usage of CNN. This model was also suitable for larger datasets. In 2021, Kaliyar et al. (2021a) have designed a BERT-based deep learning technique by integrating several parallel blocks of the single layer of deep CNN including filters and several kernel sizes. Thus, it can handle ambiguity and outperform the performance in terms of accuracy while being evaluated with other existing models. It was also carried out with the powerful capability of capturing long-distance and semantic dependencies in sentences. In 2021, Altunbey and

Alatas (2021) have suggested a new model for detecting fake news through "Adaptive Salp swarm optimization with oscillating strategy inertia weight(ASSO-OSIW) using an oscillating inertia weight and nonlinear decreasing coefficient and Grey Wolf Optimizer (GWO)" techniques for finding superior optimal solutions while evaluating online social media contents. It has been performed through flexible fitness functions for getting a superior performance. In 2021, Qureshi et al. (2021) have introduced a source-based approach focused on the news propagation community consisting of re-tweeters and posters for detecting fake contents from a twitter-based real-world COVID-19 dataset. It has included several features, where the complex network metrics were explored for identifying the news labels and examined the user profile features. Finally, the results have shown superior performance through CATBoost and RNN in detecting fake news. In 2021, Song et al. (2021) have implemented a new fake news detection model through a temporal propagation framework along with a graph neural network that has fused the temporal information, content semantics, and topological structures. Finally, this model has attained superior performance while evaluating other existing algorithms. In 2021, Saleh et al. (2021) have adopted an optimized CNN model for detecting fake news, where the feature extraction from input data was performed by N-gram and TF-IDF. They have used several layers for extracting low-level and high-level features. The parameters in every layer were optimized through grid search and hyperopt optimization algorithms. Finally, the high level of accuracy detected fake news efficiently while evaluating with other approaches. In 2021, Ali et al. (2021) have investigated the reliability of four diverse deep learning architectures like hybrid CNN-RNN, RNN, CNN, and multilayer perceptron (MLP). In addition, the detector complexity was explored, where the robustness of the learned model can be solved with the training loss and input sequence length. This model has also focused on solving the vulnerabilities of recent fake-news detectors. In 2021, Ni et al. (2021) have suggested a fake news detection model with multi-view attention networks (MVANs) for examining online social media. This model has included propagation structure attention and text semantic attention that has ensured the superior capturing of information. This model has ensured performance in terms of accuracy. It has also some interpretability in both ways of propagation and text structure. In 2021, Li et al. (2021b) have suggested a new fake news detection model with the help of an autoencoder, which has improved the performance. Further, the internal relationship among features and hidden information was obtained by adding the self-attention layer and bidirectional GRU layers into the autoencoder, and further, they reconstructed the remaining for detecting fake news. The experimentation was conducted on two real-world datasets and showed superior and positive results while estimated

with other approaches. In 2021, Verma et al. (2021) have a two-phase benchmark model for solving the authentication of news on social media. They have used word embedding over linguistic features, where initially data pre-processing was performed and validated the veracity of news content through linguistic features. Secondly, the linguistic features with word embedding were merged and applied to voting classification. Finally, the performance of the designed model was evaluated with other existing approaches that have specified superior efficiency in detecting fake news. In 2021, Ying et al. (2021b) have implemented a new end-to-end multi-modal topic memory network (MTMN) that incorporated the topic memory phase for an explicit characterization of final representation. For multimodal fusion, a new blended attention phase was implemented with the ability to exploit the intra-modal correlation within image regions or sentence words, which has also learned the image regions and inter-modal interrelation among sentence words for enhancing and complementing every feature for multimodal and high-quality representations. Lastly, the designed model has depicted better efficiency than others. In 2021, Han et al. (2021) have implemented a two-stream network for detecting fake videos on the Face-Forensics++ dataset, which can handle low-quality data. Further, the designed model has divided the input videos. Then, spatial-rich model filters were used for leveraging the extracted noise features in the second stream. In addition, considerable improvement was observed by a suggested model with both stream fusion and segmental fusion. It has obtained more state-of-the-art performance than others. In 2020, Dong et al. (2021b) have designed two-path deep semi-supervised learning with CNN for detecting fake news, in which one path was used for unsupervised learning, whereas another path is supervised learning. Here, the unsupervised learning path can learn a large range of unlabeled data, while the supervised learning path focused on learning the limited number of labeled data. These two paths were fed to CNN that were optimized for whole semi-supervised learning. Further, a shared CNN was constructed for getting the low-level features on both unlabeled and labeled data for feeding them into these two paths. The experimental results have verified the higher efficiency while recognizing the fake news with less labeled data. In 2021, Do et al. (2021) have implemented a generic model that considered both social context and news content for identifying fake news. Particularly, several aspects of the news content were explored through deep and shallow representations. The deep representations were created through transformer-based systems, while the shallow representations were generated with doc2vec and word2vec models. These representations can separately or jointly address the four significant tasks toxicity detection, sentiment analysis, clickbait detection, and bias detection. Additionally, graph CNN and mean-field layers were exploited for specifying the

structural information of news articles. Finally, the correlation among the articles was explored by leveraging the social context information. The efficiency of the designed model has been more verified than others. In 2021, Caravanti et al. (2021) have implemented a network-based technique through label propagation with positive and unlabeled learning, where the classification is performed by transductive and one-class semi-supervised learning techniques. They have considered languages like Portuguese and English and class balancing for specifying the superior balance among datasets. The performance of the designed model was superior to other algorithms like positive and unlabeled learning, and one-class learning. Thus, superior performance was observed even evaluating with unbalanced datasets. In 2021, Kaliyar et al. (2021b) have modeled a new deep neural network architecture for analyzing the social context and news content for getting superior results in terms of detecting fake news. Further, for getting a latent representation of news articles, a joint matrix-tensor factorization has been utilized, where the comparative analysis has been conducted on three techniques like social context-based, news content-based, and a combination of both. Thus, superior results in terms of higher accuracy were observed than existing approaches. In 2021, Kaliyar et al. (2021c) have suggested a new fake news identification model by considering the existence of echo chambers and the content of the new article in the social network. They have designed an efficient deep learning algorithm with tensor factorization. This model was implemented with several counts of filters across every dropout layer with a dense layer. Deep neural network (DNN) was implemented with optimal hyperparameters for classifying the social content and news content-based information individually. The superior efficiency of the designed model was better than the conventional models while detecting fake news. In 2021, Saad et al. (2021) have implemented an approach through three diverse models trained and created an ensemble of entire models through an aggregation approach to generate final predictions. It has extracted rich information from text reviews through parallel CNNs and bag-of-n-grams. Here, the non-textual and textual linguistic features were used for detecting fake news. In 2021, Choudhary et al. (2021) have implemented a deep learning model called BerConvoNet with the help of BERT and CNN, where news text was classified into real or fake news with the lowest error. This model has consisted of two major building blocks: multi-scale feature block and a news embedding block. Finally, the experiments with batch size, kernel size, and article embedding have been performed for ensuring the prediction quality. The experimental analysis of the designed model over existing models has presented a superior performance in terms of several performance measures. In 2021, Samadi et al. (2021) have designed three diverse classifiers like CNN, MLP, and single-layer perceptron (SLP) along

with pre-trained models like RoBERTa, GPT2, BERT, and funnel transformer for getting features from deep contextualized representation. The performance analysis was conducted on three datasets that have shown the efficiency of the designed model while estimating with conventional approaches regarding classification accuracy. In 2021, Meel et al. (2021) have implemented an intelligent CNN-based semi-supervised scheme through self-ensemble theory for considering the stylometric information and leveraging the “linguistic information of annotated news articles”, where the hidden patterns in unlabeled data were explored. This model has achieved the highest classification accuracy in terms of fake news recognition. It has also aimed to save cost, labor, and time and also solve inconsistencies derived during the data annotation procedure. In 2021, Esther et al. (2021) have implemented the “Attention-based Convolutional Bidirectional Long Short-Term Memory (AC-BiLSTM)” approach to detect fake news and classify them into six classes. The input data were gathered from the standard dataset, which was further given to the AC-BiLSTM for classifying the fake news through several layers. This model has tackled the fake news detection challenges in a multi-class environment that has improved the accuracy of fake news detection. In 2021, Scott et al. (2021) have presented a “Cross-stitch based Semi-supervised End-to-end neural Attention Network (Cross-SEAN)” model for leveraging the huge range of unlabeled data, which has generalized the fake news in COVID-19 that has learned from suitable external knowledge.

2.1.3 Fake news detection model based on Non-English existing approaches

In 2020, Silva et al. (2020) have offered a novel fake news detection model for real and fake news in Portuguese with a detailed analysis of machine learning approaches. It has manually constructed reference corpus with fake and true news. Several approaches have been used for evaluating the performance of diverse classes of features like distributed, distributive and linguistic-based features with text representations. The ensemble learning model with SVM, RF, Bagging, and AdaBoost has been utilized for evaluating the performance. In 2021, Zervopoulos et al. (2019) have detected fake news from tweets that have also focused on predicting patterns in both structures of tweets and linguistic content. A custom filtering process was used through a custom filtering procedure through hash tags co-occurrences. Through the performance analysis, the designed model has improved the performance of conventional deep learning techniques from Hong Kong protests. In 2021, Gokhan et al. (2021) have used natural language processing approaches for detecting fake news from Turkish-language posts on particular topics on Twitter. Moreover, word embeddings were used for

pre-training Turkish language structures, where word2vec and Term Frequency-Inverse Document Frequency (TF-IDF) have given superior performance on fake news detection. The social network analysis has been applied for identifying fake news from Twitter API. In 2021, Mitra et al. (2021) have presented a new neural network-based approach for detecting fake videos through CNN with a classifier network including Resnet50 and Inception V3. To classify video, the features extracted from convolutional neural network (CNN) classifier were fed to the subsequent classifier. This model has attained lower computational requirements and higher accuracy compared to conventional research works. In 2021, Meesad (2021) has suggested a new framework for reliable detection of fake news in the Thai language, which has consisted of three major modules. It has also composed of two stages like data collection stage and the building phase of the machine learning model. The web-crawler information retrieval was used for obtaining the data from Thai online news websites in the data collection phase. Then, the data were analyzed for getting suitable features from web data through natural language processing approaches. The detection was performed by LSTM, which was compared to other conventional techniques for detecting fake news.

2.1.4 Fake news detection model with existing diverse analytics approaches

In 2019, Zhang et al. (2019) have offered a new Fake News Detection (FEND) system. The designed mode was a two-layered method that consisted of identifying the fake events and fake topics. It has grouped the legitimate news into several clusters based on topic, where every cluster can have shared some common topics. Then, the events were extracted from the gathered articles through an event-extraction scheme. Further, they have proposed and implemented a credibility metric to evaluate the authenticity of news by estimating the news authenticity. Then, FEND was focused on detecting fake news by leveraging a huge database with legitimate news. In 2020, Kauffmann et al. (2020) have designed a methodology for analyzing the reviews automatically that has transformed the positive and negative user opinions into a quantitative score. This model has analyzed the online reviews on Amazon by using sentiment analysis, where the fake reviews have been removed and detected by the designed model from high-tech industries. Based on the consumer sentiments, the rating of brands was performed, which results in getting detailed and appropriate decision-making based on the scores attained.

2.1.5 Fake news detection model with existing diverse ensemble learning approaches

In 2020, Huang and Chen (2020) have suggested a new fake news detection system by deep learning algorithms. Initially, the pre-processing of news articles has been performed with diverse training models like LIWC, text analysis, grammar analysis, and tokenization words for getting bi-grams and uni-grams. Further, four diverse models like “N-gram CNN, LIWC CNN, depth LSTM, and LSTM” were combined to form an ensemble learning model for detecting fake news. Here, the “Self-Adaptive Harmony Search (SAHS) algorithm was used for optimizing the weights of the ensemble learning” model to get a higher accuracy rate. It has also solved the cross-domain intractability issue by experimenting with it on different domain-oriented datasets. In 2020, Reddy et al. (2020) have discussed several techniques for detecting fake news by utilizing the features attained from the text of the news without their metadata. They have employed an ensemble learning approach with the integration of text-based vector representations and “stylo-metric features” for the accurate prediction of fake news. The ensemble learning was designed by considering the classifiers like voting, boosting, and bagging. They have also used the media content in the news articles or no information concerning the users, which was the advantage of the designed model.

2.1.6 Fake news detection model with existing optimization algorithms

In 2021, Sheikhi (2021) has implemented a new fake news detection model through content-based features and optimized the “Extreme Gradient Boosting Tree (xgbTree) algorithm by the Whale Optimization Algorithm (WOA)”. Initially, the data were collected from the ISOT Fake News dataset and extracted the content-based features for choosing significant features. Further, the extracted features were fed to the WOA-Xgbtree algorithm for classifying the fake news from real news. Finally, the classification outcomes have revealed that the designed model has attained superior performance while evaluated with existing algorithms.

2.1.7 Detecting fake news during the COVID-19 pandemic based on existing approaches

In 2020, Wang et al. (2022) have revealed the components to determine the acceptance of fake news rebuttals on Sina Weibo. Here, the ELM has been used to analyze the central route, rebuttal acceptance, and peripheral route. The results have recommended the negative and positive effects of the given components. In 2020, Zheng et al. (2022) have analyzed the Internet users' responses to health-related online

fake news (HOFN) for the duration of the coronavirus (COVID-19) pandemic (Nistor and Zadobrischi 2022) with the help of the protective action decision model (PADM). The data were investigated using a multi-level linear model. In 2020, Gupta et al. (2022) have investigated the diverse news across the world. The datasets have been acquired from Twitter based on keywords provided by the Web crawler. Then, the investigation regarding the re-creation of the datasets has been demonstrated by the word clouds through the period of the COVID-19 pandemic (Ncube and Mare 2022).

2.2 Chronological review

The chronological review generally shows that the information concerned with the count of contributions till now carried out in the field of detecting fake news using deep learning and machine learning-based algorithms. The chronological review of the fake news detection model is depicted in Fig. 1 by considering the total contribution over the years. From the graphical representation, 1.5% of the contribution is taken from the year 2018, 9.2% of the research works are taken from 2019, 30.7% of the papers are reviewed from the year 2020 and finally, 58.4% of the works are taken from 2021, respectively. It inspires other researchers for increasing innovative techniques in the next subsequent years.

2.3 Research designs and general findings

There is no common definition for describing fake news. Fake news can be spread over the world, which can be propagated in any field like COVID-19, politics, e-commerce, marketing, and so on. So, there is a need of analyzing fake news to understand the real news in any particular field. However, some of the writers, publishers, and vendors, posting non-authentic online comments or any third-party

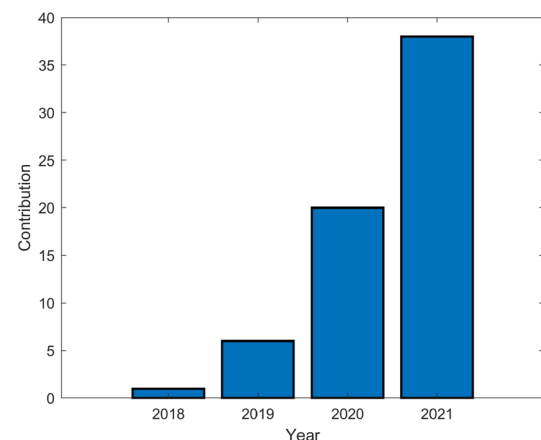


Fig. 1 A chronological review of the existing fake news detection models

monitoring online comments will act as real customers and spreads fake news on online social media for increasing product sales. Similarly, a vast number of users on social media can broadcast fake news based on their opinions. In the case of the tourism field, tweets on social media may propagate fake news based on their imagination without spending at a destination (Das et al. 2021). It may lead to the loss of genuine consumers due to fake news on online platforms. In general, fake news is considered one of the huge threats to freedom of expression, journalism, and democracy. It also influences the political impacts, from which the fake news generation can be derived due to the comments, reactions, and shares posted on Facebook, WhatsApp, Instagram, and common websites (Brenes Peralta et al. 2021; Chang 2021). More specifically, fake news detection can be divided into four perspectives like source-based approaches, propagation-based approaches, style-based approaches, and knowledge-based approaches. Finally, recent advancements in “deep learning have been utilized for detecting” fake news from online social media platforms. Deep learning has several features over machine learning approaches, which are superior accuracy, the capability of extracting high-dimensional features, and “lightly dependent on data pre-processing”. Moreover, the recent broader “availability of data and programming schemes has increased the robustness and utilization of deep learning-based algorithms”. Thus, in the past years, various research articles on fake news detection models have been implemented based on deep learning techniques.

3 Algorithmic classification, feature extraction techniques, and dataset used in the existing fake news detection models

3.1 Dataset

In this section, the datasets utilized in existing studies for evaluating the performance of their model are listed in Table 1. They have utilized benchmark datasets for both training and testing. The major problem in detecting fake news is the lack of a massive dataset and a labeled benchmark dataset with ground-truth labels. For example, some of the datasets are constructed only with political statements like PolitiFact, LIAR, Weibo, etc. The Twitter dataset consists of social media posts, whereas the FNC-1 dataset is built based on news articles. Moreover, datasets can be varied through size, labels, and modalities. Similarly, most of the studies use self-collected data from either news articles or any social media platforms.

3.2 NLP techniques used in fake news detection

Natural language processing is an innovative field in machine learning that has the ability of a computer for learning, analyzing, manipulating, and possibly generate human languages. This process has included several tasks like pre-processing, word embedding, and feature extraction techniques. Several fake news detection models utilize data pre-processing as the initial step, which is used for representing obscure attributes, managing lost words, binarization of attributes, and complicated structures with attributes. In the data pre-processing process, various visualization processes are useful. Data pre-processing helps save space and computational time, which solves the noisy data. Secondly, word vectorizing is involved in the mapping of text or words to a list of vectors. Further, a bag of words and TF-IDF are often utilized in several machine learning frameworks to detect fake news. In recent times, fake news identification models have employed pre-trained word-embedding models like word2vec and GloVe due to their ability to train larger datasets. Some of the NLP approaches and word vector models employed in deep learning-based fake news detection models are reviewed in Table 2.

While analyzing the huge number of variables, a high range of memory and computational power is necessary. Furthermore, classification techniques induce poor samples and overfitting samples. Stylometric features (Reddy et al. 2020) have also been utilized for analyzing social media content. Feature extraction is a procedure of constructing combinations of variables for overcoming the above-mentioned complications when describing the data with accurate precision. Few fake news detection models use social context features (Shu et al. 2019) for getting suitable features from the news content. N-gram (Vereshchaka et al. 2020; Agarwal et al. 2020; Saleh et al. 2021; Kaliyar et al. 2021b) generates words and characters from contents with several n-gram orders. Finally, the N-gram vectors are grouped for getting one feature vector for each information. Linguistic feature extraction (Verma et al. Aug. 2021) is used for analyzing the performance of fake news, which has several feature classes like quantity features, user credibility, stylistic features, psycho-linguistic features, and readability index. Word embedding (Kaliyar et al. 2021a; Choudhary et al. 2021; Trueman et al. 2021; Kumari and Ekbal December 2021) generates the word vectors for the downstream tasks. Though, it is complex for constructing the word vectors from scratch with several words on a large-scale dataset. Thus, this review has depicted several tasks with their features and challenges for helping future research works. The contribution of NLP tasks is depicted in Fig. 2.

Table 1 The description of publically available datasets used in conventional fake news detection models

Author name	Dataset name	Available URL like	Dataset type
Jang et al. (2018)	Crimson Hexagon database	–	Open-ended essay
Ko et al. (2019)	News taken on the web in 2017 from North Korea	–	News articles
Zhang et al. (2019)	New York Times news articles	–	News articles
Barbado et al. (2019)	“Reviews from four of the most important USA cities”	–	Social media reviews
Shu et al. (2019)	Data gathered by PolitiFact and BuzzFeed	–	Politics
Huang and Chen (2020)	“1) Fake_or_Real_news (FOR) dataset, 2) Snopes_Fake_Legit_news (SFL) dataset, and 3) Fake News Detection (FND) dataset”	–	News articles
Kaliyar et al. (2020)	U.S. Presidential Election data collected from Kaggle website	https://www.kaggle.com	News articles
Henrique and Ferreira (2020)	FakeBrCorpus, TwitterBR, Fake_or_real_news, Fakenewsdata1 and btvlifestyle	–	News articles
Singh et al. (2020b)	kaggle website	https://www.kaggle.com/c/fake-news/data , February, 2018	News articles
Altunbey and Alatas (2020)	“BuzzFeed Political News Data set, Random Political News Data set, ISOT Fake News Data set”	–	Political statements
Silva et al. (2020)	Fake.Br Corpus	–	COVID-19
Talwar et al. (2020)	Data gathered from 58 WhatsApp users,	–	Whatsapp comments on COVID-19
Hiriyannaiah et al. (2020)	105,104 news headlines	–	News articles on COVID-19
Kauffmann et al. (2020)	“Amazon Product Data by Julian McAuley”	http://jmcauley.ucsd.edu/data/amazon/ , visited on 14th of January 2019 (He & McAuley, 2016)	Amazon reviews on e-commerce
Xu et al. (2020)	Major news “outlets including New York Times, Washington Post, NBC News, USA Today, and Wall Street Journal.”	–	News articles
Umer et al. (2020)	Fake News Challenges (FNC) website	–	News articles on Culture and Tourism, football, and airport messages
Ying et al. (2021a)	public datasets (WEIBO and PHEME)	–	Political news
Oliveira et al. (2020)	33,000 tweets	–	Twitter articles
Shahbazi and Byun (2021)	900.000 records of news from 2015–2019 “from online sources and social media networks such as Twitter, Facebook, BBC news,” etc	–	Social media comments
Li et al. (2020)	LIAR, Weibo and Twitter15, NewsFN (news_dataset) and KaggleFN	Dataset 4: https://github.com/GeorgeMcIntire/fake_real Dataset 5: https://www.kaggle.com/mrisdal/fake-news	Social media data
Reddy et al. (2020)	FakeNewsNet	–	Social media data
Agarwal et al. (2020)	Kaggle	–	Social media data
Shrivastava et al. (2020)	Online social media	–	Social media data
Vereshchaka et al. (2020)	FakeNewsNet	–	Social media data

Table 1 (continued)

Author name	Dataset name	Available URL like	Dataset type
Mahabub (2020)	“6500 news from which about 3252 news is fake and 3259 news is real”	–	News articles
Kaur et al. (2020)	“News Trends, Kaggle and Reuters dataset”	https://bit.ly/2zVRLxK https://bit.ly/2Ex5VsX https://bit.ly/2BmqBQE	Social media articles
Choudhary and Arora (2020)	Horne2017_FakeNewsData	–	Social media articles
Li et al. (2021a)	FakeNewsNet	–	Social media articles
Sivasankari and Vadivu (2021)	LIAR and PolitiFact	–	Social media data
Mehta et al. (2021)	LIAR and LIAR PLUS dataset	–	Social media data
Braşoveanu and Andonie (2021)	Liar data set	–	Social media data
Jain et al. (2021)	LIAR dataset,	–	Social media data
Shishah (2021)	Politifact and Pymedia	–	Social media data
Zervopoulos et al. (2019)	Twitter’s Election Integrity Hub	–	Twitter data
Setiawan et al. (2021)	“KAGGLE3.1 dataset of 40.5 k entries”	–	Social media data
Raj and Meel (2021)	MICC-F220 dataset	–	Social media data
Kaliyar et al. (2021a)	U.S. General Presidential Election-2016	–	News article
Altunbey and Alatas (2021)	BuzzFeed Political News data set, LIAR data set, ISOT data set,	–	Social media data
Qureshi et al. (2021)	A twitter-based “real-world COVID-19 dataset from the Poynter (Meesad 2021), Politifact, and Lead Stories websites”	–	Social media data
Song et al. (2021)	Weibo dataset, FakeNewsNet dataset, Twitter dataset	–	Social media data
Saleh et al. (2021)	Kaggle, FakeNewsNet,FA-KES5, THE ISOT	–	Social media data
Ali et al. (2021)	“Kaggle fake-news dataset, ISOT dataset, and LIAR dataset”	–	Social media data
Jiang et al. (2021)	ISOT dataset and KDnugget dataset	–	Social media data
Ni et al. (2021)	Twitter15 and Twitter16	–	Social media data
Li et al. (2021b)	MediaEval 2016, Weibo dataset,	–	Social media data
Verma et al. (2021)	“four popular news data sets (i.e., Kaggle, McIntire, Reuters, and BuzzFeed Political)”	–	Social media data
Ying et al. (2021b)	WEIBO and PHEME datasets	–	Social media data
Han et al. (2021)	Face-Forensics + + dataset	–	Social media data
Dong et al. (2021b)	LIAR and PHEME data sets	–	Social media data
Do et al. (2021)	Twitter, Weibo, and PHEME	–	Social media data
Caravanti et al. (2021)	FakeNewsNet repository and GossipCop	–	Social media data
Kaliyar et al. (2021b)	BuzzFeed & PolitiFact	–	Political comments
Meesad (2021)	Gathered information using web robot crawlers	https://thaidimachine.org	politics
Gokhan et al. (2021)	Twitter API	–	Tweets on Politics, Religion, Science & technology, Society & daily news, TV & celebrities, Economy,
Kaliyar et al. (2021c)	BuzzFeed and PolitiFact	–	Social media data
Jadhav and Thepade (2019)	LIAR Dataset	–	Social media data

Table 1 (continued)

Author name	Dataset name	Available URL like	Dataset type
Mitra et al. (2021)	FaceForensics++ and Deepfake Detection Challenge datasets	–	Social media data
Saad et al. (2021)	Yelp Filtered Dataset	–	Social media data
Choudhary et al. (2021)	George McIntire Dataset and Kaggle.FakeNewsNet Repository	–	Social media data
Samadi et al. (2021)	LIAR, ISOT, COVID-19	–	Social media data
Meel et al. (2021)	Fake News Data Kaggle dataset	–	Social media data
Esther et al. (2021)	LIAR dataset	–	Social media data
Kumari and Ekbal (2021)	Twitter and Weibo dataset	–	Social media data
Scott et al. (2021)	Large-scale COVID-19 Twitter dataset	–	Social media data with labeled genuine and fake tweets
Sheikhi (2021)	ISOT Fake News dataset	–	Social media data

Table 2 Benefits and limitations of the word vector models

Approach with citation	Advantages	Challenges
BERT (Mehta et al. 2021; Shishah 2021; Ying et al. 2021b)	It captures and identifies the contextual meaning of text or words	It is computationally intensive at inference time
GloVe (Brašoveanu and Andonie 2021)	It does not depend on local statistics	However, it uses word co-occurrences or global statistics
Doc2Vec (Souza et al. 2021)	It is faster than Word2vec and has less length representation	It is not suitable for shorter documents
Word2Vec (Faustini and Covões 2020; Silva et al. 2020; Umer et al. 2020; Reddy et al. 2020; Braşoveanu and Andonie 2021; Huu Do et al. 2021; Taskin et al. 2021)	It manages the semantic meaning of several words in a text. It preserves the context information and smaller embedding vector	It does not have common representations in the shorter documents. The inability of dealing with unknown words
Bag of words (Faustini and Covões 2020; Silva et al. 2020; Souza et al. 2021)	It has a simpler implementation	It avoids semantic information among words, and also the ordering of words in a specific document
TF-IDF (Zhang et al. 2019; Barbado et al. 2019; Ozbay and Alatas 2020, 2021; Jadhav and Thepade 2019; Kaur et al. 2020; Setiawan et al. 2021; Saleh et al. 2021; Meesad 2021; Taskin et al. 2021; Dong et al. 2020)	It consists of information regarding both less significant words and more useful words	It performs slower for larger vocabularies

3.3 Algorithmic classification

Machine learning is efficiently used for reviewing the fake news detection model, which is divided into two categories, namely supervised and unsupervised learning. Here, unsupervised learning gets useful feature information from unlabeled data which makes it much easy for getting the training data. Conversely, the detection efficiency of unsupervised learning approaches is often inferior to supervised learning approaches. Supervised learning is dependent on the significant information in labeled data, where classification is the most general process, though labeling of data is often “time consuming and expensive”. Similarly, the lack of sufficient labeled data creates a major challenge to supervised learning. Deep learning is a recent research paradigm often utilized for several identification models because of

“recent achievements of these techniques in complex natural language processing tasks”. Similarly, fake news detection can be performed by deep learning algorithms. The common algorithms used for fake news identification models are categorized in Fig. 3.

The shallow models are also known as traditional machine learning models that have included several algorithms like supervised learning and unsupervised learning, where unsupervised learning includes k-means (Zhang et al. 2019) and supervised learning consists of techniques like evolution tree analysis (Jang et al. July 2018), SVM (Faustini and Covões November 2020; Kauffmann et al. October 2020; Oliveira et al. 2020), Hybrid SVM (Setiawan et al. 2021), Bernoulli’s naive Bayes (Singh et al. 2020), LDA (Reddy et al. 2020), and voting (Verma et al. 2021) classifier.

Table 3 Benefits and limitations of some machine learning-based fake news detection models

Approach with citation	Advantages	Challenges
k-means (Zhang et al. 2019)	It fits large-scale data and has strong scalability, rapid training, and simplicity	Sensitive to parameter K, and initialization, and does not perform well on non-convex data
Naive Bayes (Singh et al. 2020b),	It can learn incrementally and is robust to noise	It is not suitable for processing attribute-related data
SVM (Faustini and Covões 2020; Kauffmann et al. 2020; Oliveira et al. 2020)	It has reliable generation ability and learns significant information from the small-scale dataset	It is sensitive to kernel function parameters and not applicable for processing multiple classification tasks and big data
RNN (Shu et al. 2019; Agarwal et al. 2020)	It extracts more significant latent representations	It requires high computational costs
CNN-LSTM (Umer et al. 2020)	It performs better when compared to single networks	It may lose some significant information based on the feature extraction technique
GRU (Vereshchaka et al. 2020)	It offers a new way of exploiting both user comments and news content for predicting misleading fake news	It takes higher training time
Graph neural network (Song et al. 2021)	It accurately retrieves user information	However, it creates some inference to the detection of fake news
LSTM (Umer et al. 2020; Braşoveanu and Andonie 2021; Meesad 2021)	It retrieves a new set of features from fake news	It consumes more time for both training and testing
CNN (Kaliyar et al. June 2020; Umer et al. 2020; Agarwal et al. 2020; Kaliyar et al. 2021a; Saleh et al. 2021; Huu Do et al. 2021; Mitra et al. 2021; Samadi et al. 2021; Meel and Vishwakarma September 2021; Dong et al. Dec. 2020)	It is less prone to overfitting	It takes more time for training
BERT (Mehta et al. 2021; Shishah 2021; Kaliyar et al. 2021a)	It incorporates the additional knowledge from huge-scale data	It does not use data pre-processing, and thus, noise may affect the performance
ASSO-OSIW and GWO (Ozbay and Alatas 2021)	It overcomes complex problems, gets promising results, and is statistically important	It is not applicable for parallel and distributed datasets

Deep learning model has several deep networks, in which unsupervised learning algorithms are GAN (Srinidhi Hiriyannaiah et al. 2020), Autoencoder (Li et al. 2021b), whereas supervised learning (Li et al. 2021a; Souza et al. 2021) algorithms are AdaBoost (Barbado et al. July 2019), WOA-Xgbtree (Sheikhi 2021), RNN (Shu et al. 2019; Agarwal et al. 2020), LSTM (Umer et al. 2020; Braşoveanu and

Andonie 2021; Meesad 2021), GRU (Vereshchaka et al. 2020), DSSM-RNN (Jadhav and Thepade 2019), AC-BiLSTM (Trueman et al. 2021), Ensemble (Ozbay and Alatas February 2020; Huang and Chen November 2020; Silva et al. May 2020; Reddy et al. 2020; Jiang et al. 2021; Javed et al. 2021; Kumari and Ekbal December 2021), ensemble voting (Mahabub 2020; Qureshi et al. 2021), multi-level voting ensemble (Kaur et al. 2020), CNN (Kaliyar et al. June 2020; Umer et al. 2020; Agarwal et al. 2020; Kaliyar et al. 2021a; Saleh et al. 2021; Huu Do et al. 2021; Mitra et al. 2021; Samadi et al. 2021; Meel and Vishwakarma September 2021; Dong et al. Dec. 2020), MCNN (Li et al. 2020), C-LSTM (Zervopoulos et al. 2019), Coupled ConvNet (Raj and Meel 2021), BerConvoNet (Choudhary et al. 2021), MMCN (Ying et al. 2021a), MVAN (Ni et al. 2021), NN (Choudhary and Arora 2020; Jain et al. 2021), Graph neural network (Song et al. 2021), MTMN (Ying et al. 2021b), DNN (Ali et al. 2021; Kaliyar et al. 2021c), Lyapunov function (Shrivastava et al. Oct. 2020), social network graph (Sivasankari and Vadivu 2021; Taskin et al. 2021), Xception (Han et al. July 2021), XGBoost (Kaliyar et al. 2021b), Cross-SEAN (Paka et al. 2021), while reinforcement learning (Shahbazi and Byun 2021) includes BERT

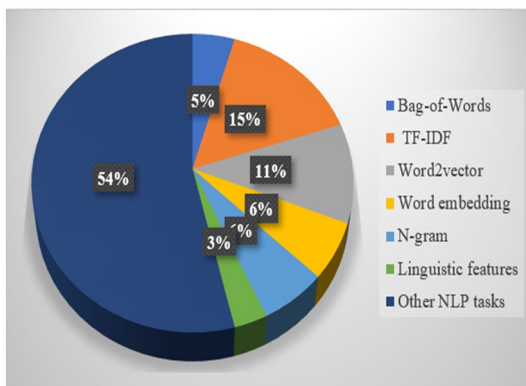
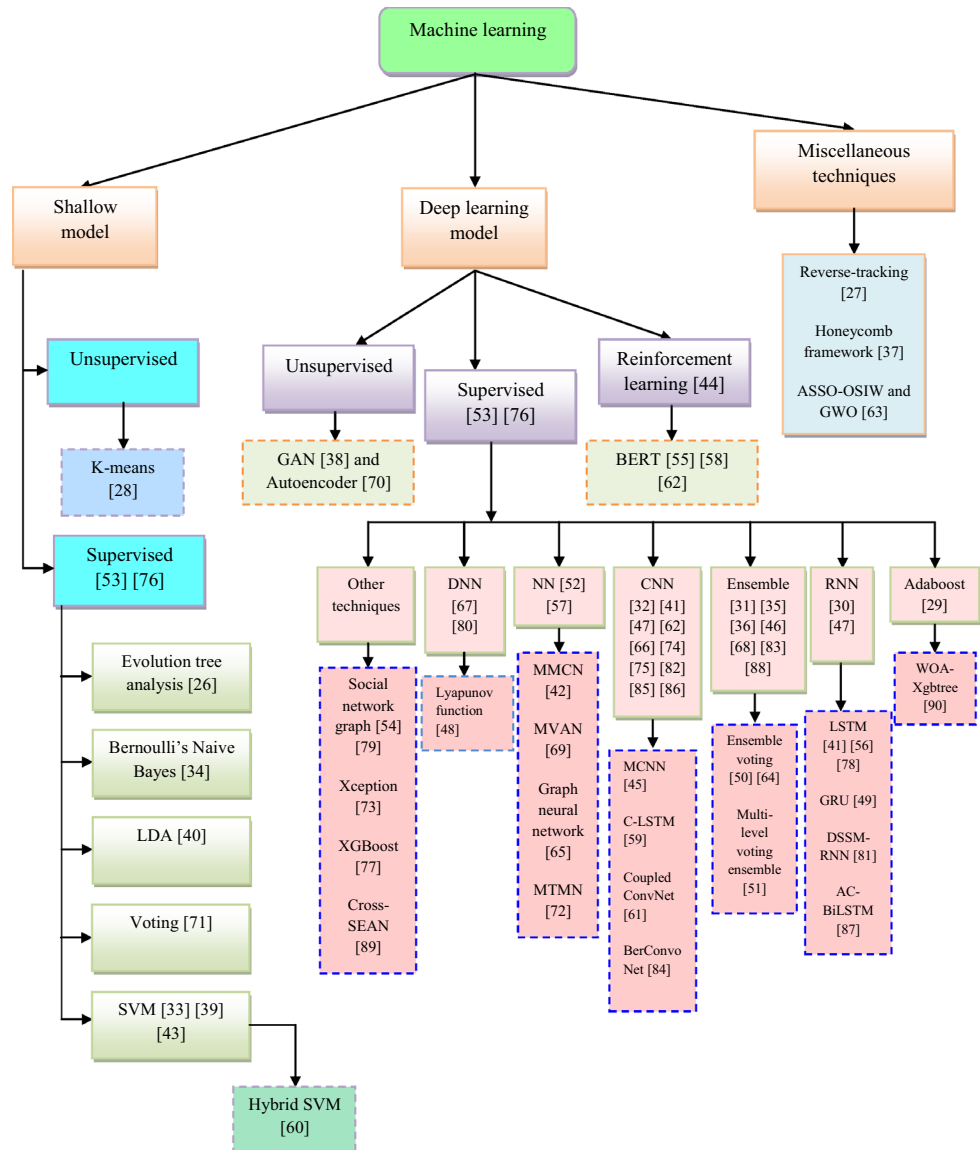


Fig. 2 Natural language processing tasks utilized in existing fake news detection models

Fig. 3 Algorithmic classification of the existing fake news detection model



(Mehta et al. 2021; Shishah 2021; Kaliyar et al. 2021a). These algorithms are focused on directly learning the features from the original data like texts and images without the need for manual feature engineering. It can be executed in an end-to-end manner. While compared with the shallow models, deep learning methods have considerable features for large datasets in terms of interpretability, learning capacity, feature representation, number of parameters, and running time. Similarly, some miscellaneous techniques are used for detecting fake news, which is the reverse-tracking approach (Ko et al. June 2019), honeycomb framework (Talwar et al. 2020), and ASSO-OSIW and GWO (Ozbay and Alatas 2021). The most used techniques with some advantages and challenges are listed in Table 3.

4 Architectural view of fake news detection models and performance measures used in traditional fake news detection models

4.1 Architectural view of the fake news detection model

The overall procedure of the fake news detection model is given in Fig. 4.

The major goal of the fake news detection model is to identify fake news to ensure the authenticity of the news by differentiating the fake news from real news. Thus, initially, the standard data can be gathered from any social-networking sites like Hike, WhatsApp, Instagram, Twitter, Facebook, and online news articles. As the gathered data requires special attention, pre-processing is required for adopting

machine learning or deep learning algorithms. Various methods have been suggested in recent years to pre-process the text and make it ready for further processing. Secondly, NLP tasks like word embedding and feature extraction will be conducted for getting the most suitable information from the data, which reduces the time and computational complexities. Finally, the features are fed to the classification model, where the classified outcomes in terms of fake vs real news are generated with the help of neural network-based algorithms. At last, the predicted outcomes demonstrate the real news that guarantees higher accuracy in detection.

4.2 Performance measures

To evaluate the efficiency of the existing fake news identification models, several performance metrics have been suggested, which are depicted in Table 4. The most commonly used techniques are discussed here.

F1 score: “harmonic mean between precision and recall. It is used as a statistical measure to rate performance”.

$$F1score = \frac{2TP}{2TP + FP + FN} \quad (1)$$

Precision: It is “the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted”.

$$Pes = \frac{TP}{TP + FP} \quad (2)$$

Accuracy: It is a “ratio of the observation of exactly predicted to the whole observations”.

$$Ac = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)$$

The recall is referred to as “the number of true positive results”

$$Re = \frac{TP}{TP + FN} \quad (4)$$

Here, “ TP, TN, FP, FN refer to the true positives, true negatives, false positives, and false negatives”, respectively.

A confusion matrix is “a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class”.

4.3 Applications considered in identifying the fake news

The fake news can be propagated in any kind of field like health, education, democracy, politics, COVID-19, etc., that could negatively affect individuals and society. Hence, the applications focused on the recent fake news detection model are diagrammatically represented in Fig. 5. From the analysis, most of the articles consider politics as their major focus. The remaining studies focus on news articles or social media comments regarding tourism, culture, COVID-19, marketing, and e-commerce topics. This helps future researchers to focus on the new field to gather information regarding the new field that gives the innovative way of research in the fake news detection model.

5 Consequences of fake news and Research challenges, future scope on fake news detection models

5.1 Consequences of fake news

Since the start of human civilization, fake news has often emerged. Though, the propagation of fake news can emerge through the utilization of the global media landscape and modern technologies. Fake news affects several fields including economic, political, and social environments (Aslam et al. 2021). On the other hand, fake news and fake information have several faces. Fake news poses tremendous impacts as information molds the view of humans around the world, though critical decisions can be made through fake information, which also leads to wrong decision-making. Similarly, good decisions cannot be made by this fabricated, distorted, false, or fake information on the Internet. The major impacts of fake news affect health, innocent people, democratic impacts, and financial impacts.

Democratic Impact: Fake news has been discussed in media due to its major role in the election. It is also considered an essential democratic problem. Hence, there is a need of predicting fake news and stop spreading fake news.

Financial Impact: Fake news is recently a complicated issue in the business world and industries. For increasing individuals' profits, dishonest businessmen may propagate fake reviews or news. Consequently, fake information can ruin the fame of a business.

Impact on Health: On the internet, health-related news is more vastly searched. People's lives can be affected due to the emergence of fake news in health. Thus, this is one of the noteworthy problems in recent times. Consequently, social media environments have created some policy changes for banning or limiting the issuing the

Table 4 Performance measures considered for evaluating the efficiency of existing fake news identification models

Citations	Accuracy	Recall	Precision	F1-measure	Confusion matrix	Friedman statistical test	ROC	AUC	Miscellaneous techniques
Barbado et al. (2019)		–	–	–	–	–	–	–	t-test
Shu et al. (2019)		–	–	–	–	–	–	–	–
Ozbay and Alatas (2020)			–		–	–	–	–	–
Jadhav and Thepade (2019)	–	–	–		–	–	–	–	Nemenyi test, Holm test
Huang and Chen (2020)					–	–	–	–	–
Kaliyar et al. (2020)					–	–	–	–	–
Faustini and Covões (2020)						–	–	–	Wilcoxon, false positive, true negative
Singh et al. (2020)		–	–		–		–	–	Nemenyi post hoc test
Silva et al. (2020)					–	–	–	–	–
Talwar et al. (2020)					–	–	–	–	–
Srinidhi Hiriyannaiah et al. (2020)	–				–	–	–	–	–
Kauffmann et al. (2020)	–	–	–	–	–	–	–	–	Reliability, Validity, average variance extracted (AVE), and composite reliability (CR) score,
Xu et al. (2020)		–	–	–		–	–	–	–
Umer et al. (2020)			–			–	–	–	–
Oliveira et al. (2020)	–	–	–	–	–	–	–	–	Jaccard similarity
Li et al. (2020)					–	–	–	–	–
Reddy et al. (2020)					–	–	–	–	–
Agarwal et al. (2020)		–		–	–	–		–	Quality of detection, sensitivity, minimum overhead, high confidence index,
Shrivastava et al. (2020)	–	–	–	–	–	–	–	–	Transaction latency, R2 score, mean absolute percentage error (MAPE)
Vereshchaka et al. ((2020)					–	–			False Positive Rate (FPR), True positive rate (TPR), false negative, true negative, True positive, false positive
Mahabub (2020)						–	–	–	–
Kaur et al. ((2020)		–	–	–	–	–	–	–	Mean Squared Error (MSE), data loss,
Choudhary and Arora (2020)		–	–	–	–	–	–	–	–
Shahbazi and Byun (2021)		–	–	–	–	–	–	–	–
Ying et al. (2021a)					–	–	–	–	Cross-validation
Li et al. (2021a)	–				–	–	–	–	specificity, FPR, and TPR
Sivasankari and Vadivu (2021)		–	–	–	–	–	–	–	–
Mehta et al. (2021)	–				–	–	–	–	–
Braşoveanu and Andonie (2021)					–	–	–	–	–
Shishah (2021)	–						–		Binary classification metrics, f2, hamming loss
Zervopoulos et al. (2019)		–	–	–	–	–	–	–	–
Setiawan et al. (2021)							–		Cross-entropy loss
Raj and Meel (2021)					–	–	–	–	–
Kaliyar et al. (2021a)					–	–	–	–	Loss
Ozbay and Alatas (2021)		–		–	–	–	–	–	–

Table 4 (continued)

Citations	Accuracy	Recall	Precision	F1-measure	Confusion matrix	Friedman statistical test	ROC	AUC	Miscellaneous techniques
Qureshi et al. (2021)					–	–	–	–	TPR and FPR
Song et al. (2021)		–	–			–	–	–	Cross-entropy loss, FPR, FNR
Saleh et al. (2021)					–		–	–	–
Ali et al. (2021)					–	–			–
Jiang et al. (2021)							–	–	–
Ni et al. (2021)					–	–	–	–	Cross-validation
Li et al. (2021b)		–	–	–	–	–	–	–	Attack Success Rate
Verma et al. (2021)					–	–	–	–	McNemar's test
Ying et al. (2021b)					–	–	–	–	–
Han et al. (2021)	–	–	–		–	–	–	–	Macro-F1, Micro-F1
Huu Do et al. (2021)					–	–	–	–	–
Souza et al. (2021)					–	–	–	–	–
Kaliyar et al. (2021b)		–	–	–	–	–	–	–	–
Meesad (2021)					–	–	–	–	Macro F, Macro P, Macro R
Taskin et al. (2021)					–	–	–	–	–
Kaliyar et al. (2021c)		–	–	–	–	–	–	–	–
Mitra et al. (2021)					–	–	–	–	–
Javed et al. (2021)					–	–	–	–	–
Choudhary et al. (2021)					–	–	–	–	Purity
Samadi et al. (2021)					–	–	–	–	–
Meel and Vishwakarma September (2021)		–	–	–	–	–	–	–	Specificity and sensitivity
Trueman et al. (2021)					–	–	–	–	–
Kumari and Ekbal (2021)					–	–	–	–	–
Paka et al. (2021)					–	–	–	–	MCC, specificity, G-mean,
Sheikhi (2021)					–	–	–	–	–
Das et al. (2021)					–	–	–	–	–
Brenes Peralta et al. (2021)					–	–	–	–	Macro average, micro average, and weighted average
Chang (2021)					–	–	–	–	–
Jain et al. (2021)		–	–		–	–	–	–	Loss function
Dong et al. (2020)		–	–		–	–	–	–	Sensitivity

misinformation on health-related articles as they affect health advocates, lawmakers, and doctors.

Impact on Innocent People: Some specific people can be affected by rumors. Social media harasses these kinds of people. People may also face threats and insults that result in real-life consequences.

5.1.1 Research gaps and future works

This review helps society from the propagation of fake news, which creates awareness of people and their impacts on social media nowadays (Alsaedi and Al-Sarem 2020). The major aim of detecting fake news is to increase the betterment of society. The existing models use several

deep learning approaches like LSTM and NNs. This NN-based training model has improved the identification of misleading news (Savyan and Bhanu 2020). The identification of fake news has posed several limitations in the existing studies. For eradicating a huge range of fake news from social media platforms, the recognition of fake news among real news can be performed by detecting the fake news subjects and creators (Kapusta and Obonya 2020). However, there is a complexity in addressing the fake news detection problems.

The major issue of fake news is inherently a multimodal and multilingual one, which consists of information in an array of languages, auditory, visual or textual forms, and generally involved in communication in a language that may be unfamiliar to users (Hakak et al. 2021). Although

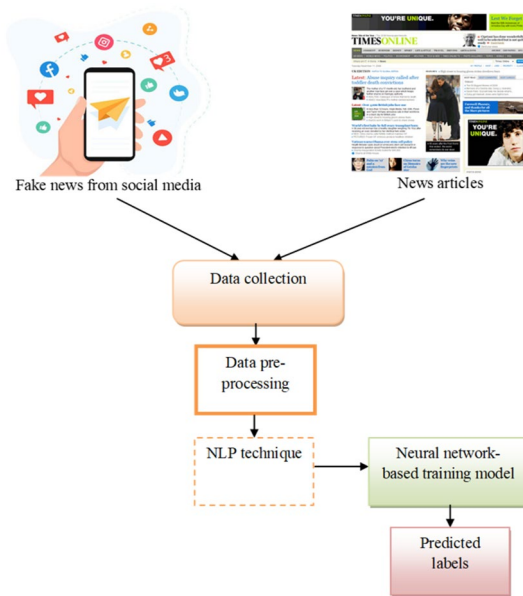


Fig. 4 The architecture of the general fake news detection model with deep learning-based models

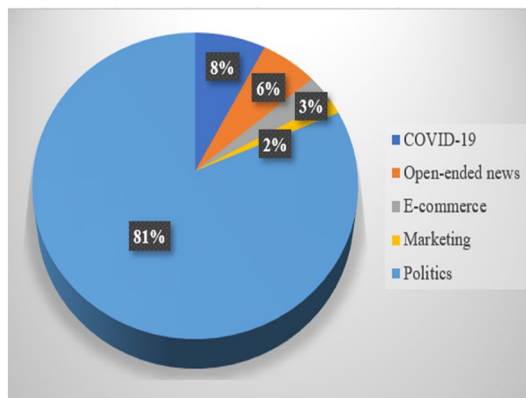


Fig. 5 Applications focused on the existing fake news detection model

deep learning-based approaches offer a superior accuracy rate while analyzing with other algorithms, thus, there is a new future perspective in making it more acceptable (Gravanis et al. 2019). Fake news detection can also be affected by selecting suitable feature extraction and classifier algorithms. The research studies must consider which classification technique is more applicable for specific features (Al-Ahmad et al. 2021). Moreover, the utilization of sequence models requires processing the long textual features. Hence, there is a need for more concentration in choosing the features and classifiers for enhancing performance. Though there are fewer probabilities of inaccurate outcomes through

deep learning models, and thus, there is a need of adopting intelligent approaches to detecting (Ambati 2021) fake news.

The future direction lies to outstretch and improvise the conventional research studies to implement the conventional works toward designing an automated system for e-commerce websites, where identification of fake news has become considerably significant (Faustini and Covões 2020). Future research works concerning fake news detection are only using supervised models, where the texts are not enough for all the cases. This problem can be solved by adding additional information like information regarding authors. The most eminent technique will be designing a knowledge-based automatic fake news detection model (Jwa et al. 2019). The results of the designed model will be extracting the information for the text and checking the information related to the dataset, which will alert the clients that the news will be considered fake. Based on this framework, the consumers can obtain awareness to solve the untrusted information (Mouratidis et al. 2021). The major problem of the existing technique is to identify the misinformation and health-related fake news, and thus, there is a new future scope for getting fake news regarding health-related fake news.

6 Conclusion

This paper has offered a detailed review of the fake news detection model with the help of a set of contributions from last year. This survey has presented information concerned with different machine learning and deep learning techniques. Additionally, this model has also given the details regarding datasets, simulation environments, algorithms, and their features and challenges. Furthermore, it has offered performance metrics used for evaluating the performance, and the research gaps and challenges for developing a new fake news detection model. Therefore, this survey can motivate future researchers to focus on novel fake news identification models with intelligent approaches. This will help the researchers in detecting fake news for gaining a concise, better perspective of conventional problems, solutions, and future directions.

Author contributions Minal Shah and Amit Ganatra designed the model, computational framework and carried out the implementation. Minal Shah performed the calculations and wrote the manuscript with all the inputs. Minal Shah, Amit Ganatra discussed the results and contributed to the final manuscript.

Declarations

Competing interests The authors declare no competing interests.

References

- Agarwal A, Mittal M, Pathak A, Goyal LM (2020) Fake news detection using a blend of neural networks: an application of deep Learning. *SN Comput Sci*. <https://doi.org/10.1007/s42979-020-00165-4>
- Ahmad I, Yousaf M, Yousaf S, Ahmad MO (2020) Fake news detection using machine learning ensemble methods, *Complexity*
- Al-Ahmad B, Al-Zoubi AM, Khurma RA, Aljarah I (2021) An evolutionary fake news detection method for COVID-19 pandemic information. *Symmetry* 13(6):1091
- Ali H, Khan MS, AlGhadhban A, Alazmi M, Alzamil A, Al-Utaibi K, Qadir J (2021) All your fake detector are belong to us: evaluating adversarial robustness of fake-news detectors under black-box settings. *IEEE Access* 9:81678–81692
- Alsaeedi A, Al-Sarem M (2020) Detecting rumors on social media based on a CNN deep learning technique. *Arab J Sci Eng* 45(12):1–32
- Ambati LS, El-Gayar O (2021) Human activity recognition: a comparison of machine learning approaches, *J Midwest Assoc Inf Syst*, 1
- Aslam N, Khan IU, Alotaibi FS, Aldaej LA, Aldubaikil AK (2021) Fake detect: a deep learning ensemble model for fake news detection. *Complexity* 2021:1–8
- Barbado R, Araque O, Iglesias CA (2019) A framework for fake review detection in online consumer electronics retailers. *Inf Process Manage* 56(4):1234–1244
- Beer DD, Matthee M (2020) Approaches to identify fake news: a systematic literature review. *Integr Sci Digit Age* 136:13–22
- Bondielli A, Marcelloni F (2019) A survey on fake news and rumour detection techniques. *Inf Sci* 497:38–55
- Braşoveanu AMP, Andonie R (2021) Integrating machine learning techniques in semantic fake news detection. *Neural Process Lett* 53:3055–3072
- Brenes Peralta CM, Sánchez RP, González IS (2021) Individual evaluation vs fact-checking in the recognition and willingness to share fake news about Covid-19 via whatsapp, *Journal Stud*
- Chang C (2021) "Fake news: audience perceptions and concerted coping strategies," *Digital Journal*, 9(5)
- Chauhan T, Palivela H (2021) Optimization and improvement of fake news detection using deep learning approaches for societal benefit. *Int J Inf Manag Data Insights*. <https://doi.org/10.1016/j.jjimei.2021.100051>
- Choudhary A, Arora A (2020) Linguistic feature based learning model for fake news detection and classification. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2020.114171>
- Choudhary M, Chouhan SS, Pilli ES, Vipparthi SK (2021) BerConvoNet: a deep learning framework for fake news classification, *Appl Soft Comput*, 110
- D'Ulizia A, Caschera MC, Ferri F, Grifoni P (2021b) Fake news detection: a survey of evaluation datasets, *Peer J Comput Sci*, 7
- Dabbous A, Aoun Barakat K, de Quero Navarro B (2020a) Fake news detection and social media trust: a cross-cultural perspective, *Behav Inf Technol*
- Das SD, Basak A, Dutta S (2021) "A heuristic-driven uncertainty based ensemble framework for fake news detection in tweets and news articles, *Neurocomputing*
- de Oliveira NR, Medeiros DSV, Mattos DMF (2020) A sensitive stylistic approach to identify fake news on social networking. *IEEE Signal Process Lett* 27:1250–1254
- de Souza MC, Nogueira BM, Rossi RG, Marcacini RM, Dos Santos BN, Rezende SO (2021) A network-based positive and unlabeled learning approach for fake news detection, *Mach Learn*
- Dong X, Victor U, Qian L (2020) Two-path deep semisupervised learning for timely fake news detection. *IEEE Trans Comput Soc Syst* 7(6):1386–1398
- Faustini PHA, Covões TF (2020) Fake news detection in multiple platforms and languages. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2020.113503>
- Faustini PH, Covões TF (2020) Fake news detection in multiple platforms and languages. *Expert Syst Appl* 158:15
- Gravanis G, Vakali A, Diamantaras K, Karadais P (2019) Behind the cues: a benchmarking study for fake news detection. *Expert Syst Appl* 128:201–213
- Gupta A, Bansal A, Mamgain K, Gupta A (2022) An exploratory analysis on the unfold of fake news during COVID-19 pandemic. *Smart Syst: Innov Comput* 235:259–272
- Habib A, Asghar MZ, Khan A, Habib A, Khan A (2019) False information detection in online content and its role in decision making: a systematic literature review, *Soc Netw Anal Mining*, 9(50)
- Hakak S, Alazab M, Khan S, Gadekallu TR, Maddikunta PKR, Khan WZ (2021) An ensemble machine learning approach through effective feature extraction to classify fake news. *Futur Gener Comput Syst* 117:47–58
- Han B, Han X, Zhang H, Li J, Cao X (2021) Fighting fake news: two stream network for deepfake detection via learnable SRM. *IEEE Trans Biom Behav Identity Sci* 3(3):320–331
- Huang YF, Chen PH (2020) Fake news detection using an ensemble learning model based on Self-Adaptive Harmony Search algorithms. *Expert Syst Appl* 159:30
- Huu Do T, Berneman M, Patro J, Bekoulis G, Deligiannis N (2021) Context-aware deep markov random fields for fake news detection. *IEEE Access* 9:130042–130054
- 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 Mining*, 10(82)
- Jadhav SS, Thepade SD (2019) Fake news identification and classification using dssm and improved recurrent neural network classifier. *Appl Artif Intell Int J*. <https://doi.org/10.1080/08839514.2019.1661579>
- Jain V, Kaliyar RK, Goswami A, Narang P, Sharma Y (2021) "AENeT: an attention-enabled neural architecture for fake news detection using contextual features, *Neural Comput Appl*
- Jang SM, Geng T, Li JY, Xia R, Huang CT, Kim H, Tang J (2018) A computational approach for examining the roots and spreading patterns of fake news: evolution tree analysis. *Comput Hum Behav* 84:103–113
- Javed MS, Majeed H, Mujtaba H, Beg MO (2021) Fake reviews classification using deep learning ensemble of shallow convolutions. *J Comput Soc Sci* 4:883–902
- 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
- Jwa H, Oh D, Park K, Kang J, Lim H (2019) ExBAKE: automatic fake news detection model based on bidirectional encoder representations from transformers (BERT). *Appl Sci* 9(19):4062
- 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
- Kaliyar RK, Goswami A, Narang P (2021a) FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimed Tools Appl* 80:11765–11788
- Kaliyar RK, Goswami A, Narang P (2021b) DeepFakeE: improving fake news detection using tensor decomposition-based deep neural network. *J Supercomput* 77:1015–1037
- Kaliyar RK, Goswami A, Narang P (2021c) EchoFakeD: improving fake news detection in social media with an efficient deep neural network. *Neural Comput Appl* 33:8597–8613
- Kansal A (2021) Fake news detection using pos tagging and machine learning. *J Appl Secur Res*
- Kapusta J, Obonya J (2020) Improvement of misleading and fake news classification for effective languages by morphological

- group analysis. *Informatics*. <https://doi.org/10.3390/informatics7010004>
- Kauffmann E, Peral J, Gil D, Ferrández A, Sellers R, Mora H (2020) A framework for big data analytics in commercial social networks: a case study on sentiment analysis and fake review detection for marketing decision-making. *Ind Mark Manage* 90:523–537
- Kaur S, Kumar P, Kumaraguru P (2020) Automating fake news detection system using multi-level voting model. *Soft Comput* 24:9049–9069
- Kim G, Ko Y (2021) Effective fake news detection using graph and summarization techniques. *Pattern Recogn Lett* 151:135–139
- Ko H, Hong JY, Kim S, Mesicek L, Na IS (2019) Human-machine interaction: a case study on fake news detection using a backtracking based on a cognitive system. *Cogn Syst Res* 55:77–81
- Kumari R, Ekbal A (2021) AMFB: attention based multimodal factorized bilinear pooling for multimodal fake news detection. *Expert Syst Appl* 184:1
- Li Q, Hu Q, Lu Y, Yang Y, Cheng J (2020) Multi-level word features based on CNN for fake news detection in cultural communication. *Pers Ubiquit Comput* 24:259–272
- Li D, Guo H, Wang Z, Zheng Z (2021b) Unsupervised fake news detection based on autoencoder. *IEEE Access* 9:29356–29365
- Li X, Lu P, Hu L, Wang X, Lu L (2021a) A novel self-learning semi-supervised deep learning network to detect fake news on social media. *Multimedia Tools Appl*
- Mahabub A (2020) A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers. *SN Appl Sci*. <https://doi.org/10.1007/s42452-020-2326-y>
- Meel P, Vishwakarma DK (2021) A temporal ensembling based semi-supervised ConvNet for the detection of fake news article. *Expert Syst Appl* 177:1
- Meesad P (2021) Thai fake news detection based on information retrieval, natural language processing and machine learning. *SN Comput Sci*, 2(425)
- Mehta D, Dwivedi A, Patra A, Anand Kumar M (2021) A transformer-based architecture for fake news classification. *Social Netw Anal Mining*, 11(39)
- Meneses Silva CV, Silva Fontes R, Colaço Júnior M (2021) Intelligent fake news detection: a systematic mapping. *J Appl Secur Res*, 16(2)
- Mitra A, Mohanty SP, Corcoran P, Kougianos E (2021) A machine learning based approach for deepfake detection in social media through key video frame extraction. *SN Comput Sci*, 2(98)
- Mouratidis D, Nikiiforos MN, Kermanidis KL (2021) Deep learning for fake news detection in a pairwise textual input schema. *Computation*. <https://doi.org/10.3390/computation9020020>
- 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
- Ncube L, Mare A (2022) Fake news” and multiple regimes of “truth” during the COVID-19 pandemic in Zimbabwe. *Afr Journal Stud* 43(2):71–89
- Ni S, Li J, Kao H-Y (2021) MVAN: multi-view attention networks for fake news detection on social media. *IEEE Access* 9:106907–106917
- Nistor A, Zadobrischi E (2022) The influence of fake news on social media: analysis and verification of web content during the COVID-19 pandemic by advanced machine learning methods and natural language processing. *Sustainability* 14(17):10466
- Ozbay FA, Alatas B (2020) Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A* 540:15
- Ozbay FA, Alatas B (2021) Adaptive Salp swarm optimization algorithms with inertia weights for novel fake news detection model in online social media. *Multimed Tools Appl* 80:34333–34357
- Paka WS, Bansal R, Kaushik A, Sengupta S, Chakraborty T (2021) Cross-SEAN: a cross-stitch semi-supervised neural attention model for COVID-19 fake news detection. *Appl Soft Comput*
- Qureshi KA, Malick RAS, Sabih M, Cherifi H (2021) Complex network and source inspired COVID-19 fake news classification on twitter. *IEEE Access* 9:139636–139656
- Raj C, Meel P (2021) ConvNet frameworks for multi-modal fake news detection. *Appl Intell* 51:8132–8148
- S. Rama Krishna, S. V. Vasantha, K. Mani Deep (2021) Survey on fake news detection using machine learning algorithms. *ICACT*, 09(08)
- Reddy H, Raj N, Gala M, Basava A (2020) Text-mining-based fake news detection using ensemble methods. *Int J Autom Comput* 17:210–221
- Ribeiro Bezerra JF (2021) Content-based fake news classification through modified voting ensemble. *J Inf Telecommun*. <https://doi.org/10.1080/24751839.2021.1963912>
- Saleh H, Alharbi A, Alsamhi SH (2021) OPCNN-FAKE: optimized convolutional neural network for fake news detection. *IEEE Access* 9:129471–129489
- Samadi M, Mousavian M, Momtazi S (2021) Deep contextualized text representation and learning for fake news detection. *Inf Process Manag*, 58(6)
- Savyan P, Bhanu SMS (2020) UbCadet: Detection of compromised accounts in Twitter based on user behavioural profiling. *Multimed Tools Appl* 79:1–37
- Sengupta E, Nagpal R, Mehrotra D, Srivastava G (2021) ProBlock: a novel approach for fake news detection. *Clust Comput* 24:3779–3795
- Setiawan R, Ponnamp VS, Sengan S, Anam M, Subbiah C, Phasinam K, Vairaven M, Ponnusamy S (2021) Certain investigation of fake news detection from facebook and twitter using artificial intelligence approach. *Wirel Pers Commun*. <https://doi.org/10.1007/s11277-021-08720-9>
- Shahbazi Z, Byun Y-C (2021) Fake media detection based on natural language processing and blockchain approaches. *IEEE Access* 9:128442–128453
- Sharma DK, Garg S (2021) IFND: a benchmark dataset for fake news detection. *Compl Intell Syst*
- Sheikhi S (2021) An effective fake news detection method using WOA-xgbTree algorithm and content-based features. *Appl Soft Comput*, 109
- Shishah W (2021) Fake news detection using BERT model with joint learning. *Arab J Sci Eng* 46:9115–9127
- 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
- Shu K, Mahudeswaran D, Liu H (2019) FakeNewsTracker: a tool for fake news collection, detection, and visualization. *Comput Math Organ Theory* 25:60–71
- Silva RM, Santos RL, Almeida TA, Pardo TA (2020) Towards automatically filtering fake news in Portuguese. *Expert Syst Appl* 146:15
- 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*. <https://doi.org/10.1016/j.ipm.2021.102618>
- Simko J, Racsco P, Tomlein M, Hanakova M, Moro R, Bielikova M (2021) A study of fake news reading and annotating in social media context. *New Rev Hypermedia Multimed*
- Singh M, Bhatt MW, Bedi HS, Mishra U (2020) Performance of bernoulli’s naive bayes classifier in the detection of fake news. *Mater Today Proc*. <https://doi.org/10.1016/j.matpr.2020.10.896>
- Singh B, Sharma DK (2021) Predicting image credibility in fake news over social media using multi-modal approach. *Neural Comput Appl*

- Sivasankari S, Vadivu G (2021) Tracing the fake news propagation path using social network analysis, *Soft Comput*
- Song C, Shu K, Wu B (2021) Temporally evolving graph neural network for fake news detection, *Inf Process Manag*, 58(6)
- Hiriyanaiyah S, Srinivas AM, Shetty GK, Siddesh GM, Srinivasa KG (2020) A computationally intelligent agent for detecting fake news using generative adversarial networks. *Hybrid computational intelligence for pattern analysis and understanding*, pp. 69–96
- Talwar S, Dhir A, Singh D, Virk GS, Salo J (2020) Sharing of fake news on social media: application of the honeycomb framework and the third-person effect hypothesis. *J Retail Consum Serv*. <https://doi.org/10.1016/j.jretconser.2020.102197>
- Taskin SG, Kucuksille EU, Topal K (2021) Detection of Turkish Fake News in Twitter with Machine Learning Algorithms, *Arab J Sci Eng*
- Trueman TE, Kumar A, Narayanasamy P, Vidya J (2021) Attention-based C-BiLSTM for fake news detection, *Appl Soft Comput*, 110
- Umer M, Imtiaz Z, Ullah S, Mehmood A, Choi GS, On B-W (2020) Fake news stance detection using deep learning architecture (CNN-LSTM). *IEEE Access* 8:156695–156706
- Vereshchaka A, Cosimini S, Dong W (2020) Analyzing and distinguishing fake and real news to mitigate the problem of disinformation. *Comput Math Organ Theory* 26:350–364
- Verma PK, Agrawal P, Amorim I, Prodan R (2021) WELFake: word embedding over linguistic features for fake news detection. *IEEE Trans Comput Soc Syst* 8(4):881–893
- Wang X, Chao F, Yu G, Zhang K (2022) Factors influencing fake news rebuttal acceptance during the COVID-19 pandemic and the moderating effect of cognitive ability. *Comput Hum Behav* 130:107174
- Xu K, Wang F, Wang H, Yang B (2020) Detecting fake news over online social media via domain reputations and content understanding. *Tsinghua Sci Technol* 25(1):20–27
- Yang C, Zhou X, Zafarani R (2021) CHECKED: Chinese COVID-19 fake news dataset, *Soc Netw Anal Mining*, 11(58)
- Ying L, Yu H, Wang J, Ji Y, Qian S (2021a) Multi-level multi-modal cross-attention network for fake news detection. *IEEE Access* 9:132363–132373
- Ying L, Yu H, Wang J, Ji Y, Qian S (2021b) Fake news detection via multi-modal topic memory network. *IEEE Access* 9:132818–132829
- Zervopoulos A, Alvanou AG, Bezas K, Papamichail A, Maragoudakis M, Kermanidis K (2021) "Deep learning for fake news detection on Twitter regarding the 2019 Hong Kong protests. *Neur Comput Appl*. <https://doi.org/10.1007/s00521-021-06230-0>
- Zhang C, Gupta A, Kauten C, Deokar AV, Qin X (2019) Detecting fake news for reducing misinformation risks using analytics approaches. *Eur J Oper Res* 279(3):1036–1052
- Zheng L, Elhai JD, Miao M, Wang Y, Wang Y, Gan Y (2022) Health-related fake news during the COVID-19 pandemic: perceived trust and information search. *Internet Res* 32:768–789
- Zhou X, Zafarani R (2020) A survey of fake news: fundamental theories, detection methods, and opportunities. *ACM Comput Surv* 53(5):1–40

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

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