TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 22, No. 4, August 2024, pp. 869~886
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v22i4.25929

Rumor detection based on deep learning techniques: a systematic review

Lifan Zhang[1,2], Shafaf Ibrahim[1], Ahmad Firdaus Ahmad Fadzil[1]
[1]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Shah Alam, Malaysia
[2]Department of Computer Engineering, Taiyuan Institute of Technology, Taiyuan, China

Article Info

Article history:

Received Dec 18, 2023
Revised Jan 19, 2024
Accepted Feb 15, 2024

Keywords:

Deep learning
Feature selection
Rumor detection
Social media
Systematic review

ABSTRACT

The rise of social media platforms has led to an increase in the flow and dissemination of information, but it has also made generating and spreading rumors easier. Rumor detection requires understanding the context and semantics of text, dealing with the evolving nature of rumors, and processing vast amounts of data in real-time. Deep learning (DL)-based techniques exhibit a higher accuracy in detecting rumors on social media compared to many traditional machine learning approaches. This study presents a systematic review of DL approaches in rumor detection, analyzing datasets, pre-processing methods, feature taxonomy, and frequently used DL methods. In the context of feature selection, we categorize features into three areas: text-based, user-based, and propagation-based. Besides, we surveyed the trends in DL models for rumor detection and classified them into convolutional neural networks (CNN), recurrent neural networks (RNN), graph neural networks (GNN), and other methods based on the model structure. It offers insights into effective algorithms and strategies, aiming to guide researchers, developers, social media users, and governments in detecting and preventing the spread of false information. The study contributes to enhancing research in this field and identifies potential areas for future exploration.

Corresponding Author:

Shafaf Ibrahim
College of Computing, Informatics and Mathematics, Universiti Teknologi MARA
40450 Shah Alam, Selangor Darul Ehsan, Malaysia
Email: shafaf2429@uitm.edu.my

1. INTRODUCTION

The prosperity and development of social media platforms like Weibo and Twitter have allowed people to freely share information and create direct connections and communication. This has sped up the flow and dissemination of public information and made it easier and faster to create and spread rumors. According to the Edelman trust barometer report 2022 and 2021 [1], [2], fake news concerns at all-time highs. In addition, trust in media declines in 15 of 27 countries. Trust in all information sources at record lows from the past decades. It has become increasingly challenging to distinguish between factual and deceptive content [3]–[9].

Rumor detection revolves around identifying and classifying rumors or false information within a given set of news articles, social media posts, or online content. The main objective is to distinguish between accurate and deceptive claims by examining the credibility and veracity of the information presented [10]–[13]. Rumor detection involves several key challenges. Firstly, it requires combining the context with the semantics of the text contents to differentiate between factual statements and misleading information [14], [15]. This often calls for deep linguistic analysis and reasoning abilities. Additionally, the brevity and informality of social media messages pose difficulties because of the lack of explicit cues and limited context [16]. Another challenge lies in dealing with the evolving nature of rumors [17]. They can manifest in different forms such as conspiracy theories, hoaxes, misinformation, or partially true claims. Rumors can also mutate over time, making it necessary to account for temporal dynamics when designing effective detection systems. Furthermore, the mass of data produced on social media platforms necessitates scalable and efficient algorithms for rumor detection [18]. To combat the spread of false information quickly and mitigate its negative effects, real-time processing is essential.

To detect and prevent the spread of false information on various online platforms, researchers have explored various effective techniques. Traditional approaches often rely on linguistic features, such as syntactic patterns, lexical cues, and sentiment analysis [9]. These methods typically involve rule-based systems or supervised classification algorithms. With the wider application of deep learning (DL) techniques, researchers have also turned to neural network models for rumor detection. These models can understand complicated contexts and represent them automatically from the mass-labeled data. Recurrent neural networks (RNNs) [19]–[24] and convolutional neural networks (CNNs) [25]–[28] have been widely employed in rumor detection tasks, often combined with word embeddings like Word2Vec or global vectors for word representation (GloVe) for capturing semantic information. Besides, the development of graph neural network (GNN) [29]–[37] models has demonstrated obvious advantages in the field of rumor detection. GNNs are constructed to work effectively with graph-structured data, making them well-suited for examining the spread of rumors within social networks or other interconnected systems. Rumor detection involves identifying and verifying the veracity of information spreading through a network. GNN models excel in this task by capturing both the network’s structural characteristics and the content of individual nodes.

This study employs a combination of well-researched papers on rumor detection based on DL methods to provide a comprehensive analysis of datasets, pre-processing methods, feature taxonomy, and frequently used DL methods. The study offers several significant contributions concerning the practical and the body of knowledge. By referring to the study, developers may be able to refer to the detection architectural model to plan for the integration of verification into rumor detection. Social media users may be able to automatically get an early prediction and reduce loss before rumors harm them. Social media platforms may be able to provide a security mode to avoid attacking by rumors and guarantee a safe environment for users. Governments may be able to effectively supervise and orient the posts and comments from social media platforms.

For the field of study, the following succinctly describes our work’s primary contribution:

- We have conducted a systematic review of different DL approaches applied specifically in rumor detection, encompassing relevant literature published between 2017 and 2023 June. This review is based on predefined resources and follows predefined inclusion/exclusion criteria. Out of the 168 publications obtained using the search term, 71 were ultimately selected for this review. As far as we know, this is the first systematic study that has been done specifically on this subject.
- We conducted a classification study of DL models in rumor detection and analyzed the application trends of mainstream models.
- We encompassed the datasets, pre-processing methods, feature taxonomy, and methodology for feature extraction that are often utilized in DL approaches in this field.
- We outlined future research directions by highlighting the significant potential of these approaches. This information aims to guide researchers towards the most effective algorithms, related features, and pre-processing methods, while also identifying gaps and potential areas for future exploration in this research field.

These contributions serve to enhance DL-based research in rumor detection by providing researchers with valuable insights into the most effective techniques, associated features, pre-processing strategies, and areas of untapped potential.

2. METHOD

2.1. Development of the review protocol

We performed a systematic review of rumor detection based on DL utilizing the recommended reporting items for systematic reviews and meta-analysis (PRISMA) technique to methodically examine and classify the research status of rumor detection. The creation and organization of SRs and other meta-analyses are guided by PRISMA, a minimal set of elements based on evidence.

Searching numerous digital libraries and databases for pertinent research is the initial stage in the review process. Subsequently, the search criteria are employed to minimize the quantity of chosen studies to enhance the caliber of the papers and incorporate the greatest variety of DL techniques. Following that, a series of research questions was developed in order to fully address the research on the state of rumor detection at the time.

2.2. Related surveys

Before conducting our review, we inspected existing survey papers on rumor detection techniques published within the last six years. We sought survey papers that investigated rumor detection based on DL and discussed research trends, techniques, and future directions. In this study, we searched the available literature on rumor detection comparison to find the most recent method. Our primary goal is to investigate the most recent rumor detection methodologies that are relevant to these research projects. From 2018 to 2023, we found 14 linked surveys for the rumor detection approach. Table 1 contains the specifics. The surveys cover topics such as fake news detection, rumor detection, fact-checking, and misinformation detection, as well as related models. The current work conducts a systematic evaluation on rumor detection based on DL approaches in an effort to thoroughly examine the efforts made in earlier studies.

Table 1. Relates surveys

Ref.	Author	Year	Topics
[38]	Mishima and Yamana	2022	They explored the models, datasets, evaluation methods, visualization procedures, and potential advancements in fake news detection
[39]	Mridha *et al.*	2021	They tried to examine advanced fake news detection mechanisms pensively. They highlighted the DL-based techniques, the prominent evaluation metrics, and further recommendations
[40]	Kotonya and Toni	2020	They focused on the explanation functionality and summarized existing strategies to explain the results of fact-checking systems.
[41]	Guo *et al.*	2022	This article reviews automated fact-checking covering both claim detection and claim validation.
[42]	Pathak *et al.*	2020	They clarified supervised and unsupervised techniques as well as DL approaches for rumor detection.
[18]	Islam *et al.*	2020	Provide an effective and scalable technique for misinformation detection based on DL
[15]	Varshney and Vishwakarma	2021	Summary rumor definition, generalized model, data collection, features, and models.
[43]	Ahsan and Kumari	2019	Analyze rumor diffusion, features, rumor detection, and rumor veracity approaches.
[44]	Bondielli and Marcelloni	2019	Elaborate different definitions of fake news and rumors, collection data methods, features, and rumor detection approaches.
[45]	Al-Sarem *et al.*	2019	Compare performance evaluation, dataset, and the DL model used per each work.
[46]	Reis *et al.*	2019	Analyze the variables influencing the model's decisions in fake news detection.
[47]	Zubiaga *et al.*	2018	Summary of the efforts and achievements so far toward the development of rumor classification systems.
[48]	Cao *et al.*	2018	Introduce a formal definition of rumor, and summarize hand-crafted features, propagation features, and deep neural networks.
[49]	Alzanin and Azmi	2018	Investigating rumor detection methods from three perspectives: hybrid methods, supervised-based methods, and unsupervised-based methods.

2.3. Definition of DL literature probe

This systematic review (SR) will categorize and review the existing relevant methods for studying the scientific and technical documentation produced by those searches to cover the area of the literature probed to locate relevant publications in our context. There are two main steps in the proposed procedure:
- Determination of the term of a search to obtain a set of keywords from previous research questions;
- Determination of queries to be used by the Boolean operators AND/OR to find and gather all relevant results.

The rumor detection literature probes used in this paper are shown in Table 2 with the motivations. We collated and analyzed eight literature probes in recent research work for rumor detection based on DL techniques and the corresponding answers are given in subsequent chapters.

2.4. Search strategy

Identifying information sources from digital libraries, search engines, and social networks is the initial stage in our search strategy. Table 3 shows the publisher's website, the type of website, and the website names that are being used to find the LR. We searched papers from these five well-known websites in Table 3.

Table 2. Rumor detection literature probe

Literature probes	Motivation
LP1. What is the distribution by year, publisher, country, and language of datasets?	The answer to this question provides an understanding of the background of earlier work done on rumor detection.
LP2. What is the source of the datasets employed?	The answer to this question identifies the primary contributors of datasets used for rumor detection.
LP3. What is the domain that the research papers studied?	The answer to this query identifies the domain most studied by recent rumor detection papers.
LP4. What pre-possessing techniques were used?	The answer to this question discerns pre-possessing methods used in rumor detection works.
LP5. What kind of features are used for rumor detection and how to use?	The answer to this question would mine the features of rumor in massive data from recent research papers.
LP6. What is the most recent rumor detection method devised and which produced salient performance?	The answers to these questions reveal the most notable rumor detection methods explored.
LP7. What evaluation techniques were formulated for rumor detection?	The answer to this question recognizes evaluation metrics widely used for rumor detection.
LP8. What are the potential future research directions and perspectives on rumor detection?	The answer to this question assists in finding potential avenues in rumor detection.

Table 3. Publisher website

Source	Type	URL
ScienceDirect	Digital library	http://www.sciencedirect.com/
IEEE Xplore	Digital library	http://ieeexplore.ieee.org/Xplore/home.jsp
Wiley online library	Digital library	https://onlinelibrary.wiley.com/
MDPI	Digital library	https://www.mdpi.com/
ACL anthology	Conference	https://aclanthology.org/

2.5. Search term

The next step is to create one or more search queries that will offer the coverage that our review objectives require. Boolean operators and AND are used to do this. Our composed search terms are listed in Table 4. Within the chosen year range, S1 seeks to collect all literature related to rumor detection. S2 searches for literature on social media rumor detection. S3 focuses on any literature on rumor detection that is based on comparisons. The search keywords searching relevant literature are “Rumor detection” OR “Rumor detection” AND “social media” OR” rumor detection” AND “comparison”.

Table 4. Search term

TITLE-ABS-KEY
S1 rumor detection
S2 rumor detection AND social media
S3 rumor detection AND comparison

2.6. Inclusion and exclusion criteria

We employ a set of inclusion criteria (IC) and exclusion criteria (EC) to identify associated items in Table 5. Papers that do not react to EC are excluded, and a screening technique is employed to choose publications that are relevant to our setting.

Table 5. Inclusion and exclusion table

Inclusion criteria
Papers should be included in the research databases.
Papers for method should be published between 2009 and 2023.
Papers for the domain should be published between 2017 and 2023.
Papers should meet at least one of the search terms.
Papers should be published at a journal or conference.
Papers should provide answers to the research questions.
The search is conducted based on the title, abstract, and full text.
Exclusion criteria
Papers that are not written in English.
Repeated papers.
Papers that are missing full text.
Papers that do not have DOI.
Papers that are not relevant to rumor detection.

The screening procedure consists of three IC steps:

– Abstract-based step: we use data and keywords searched in paper abstracts to weed out irrelevant results. Papers were held for further consideration if their abstracts met the requirements of at least 40% of the IC.
– Full-text-based step. we exclude results from papers that only contain a tiny portion of the search terms given in their abstracts, i.e., articles that do not discuss the research words contained in Table 5.
– Step based on quality analysis: Based on quality analysis, we exclude results that don’t fit any of the following criteria:
  C1: the paper discusses DL.
  C2: related works are included in the publication.
  C3: the results achieved are discussed in the study.

3. RESULTS AND DISCUSSION

The only electronic databases reviewed for the literature survey comprise the most trustworthy publications, conference proceedings, and research. During the first search, 168 papers were found; however, only 100 articles were taken into consideration once the inclusion-exclusion criteria were applied. Figures 1 to 3 are the distribution of the publications we selected in terms of year, website, and country research. In the recent seven years, the years 2021 and 2022 have the most research papers, and most related articles published in IEEE, China has the largest number of researchers.

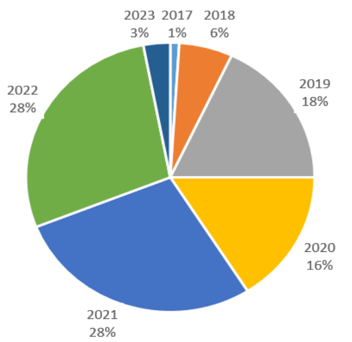

Figure 1. Distribution of year research published

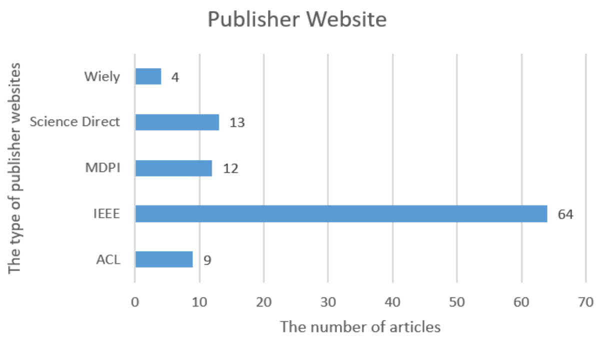

Figure 2. Distribution of year research publisher website

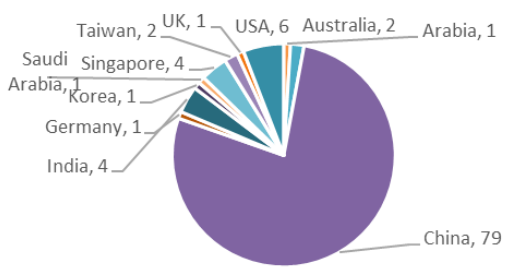

Figure 3. Distribution of country research published

3.1. Datasets, domain, and pre-processing

Social media platforms have recently become one of the most prevalent communication platforms in the world. The users on social media platforms can post and comment as much as they want without confirmation. However, this may provide a great hotbed for the spread of rumors. Therefore, it is essential for rumor detection research to gather a good amount of data from platforms like Twitter, Sina Weibo, YouTube, and Facebook or rumor debunking websites (Snopes, Politifact, and FactCheck) [47]. For experiment and evaluation, most authors have collected data via application programming interfaces (APIs), scrapping the web or Selenium web driver [15]. Some researchers created datasets by themselves while others used publicly available benchmark datasets.

A bar chart of the used datasets is given in Figure 4, where the content “crawled” in brackets indicates that authors constructed datasets from the corresponding social media platforms by crawling technology, and the remaining datasets are benchmarked datasets. From 60 papers, we have 25 kinds of datasets related to rumor detection, such as CED, Kaggle, LIAR, Snopes, RumourEval 2019, SemEval 2017, CR-Dataset, DataFoundation, Science, Tencent (crawled), Mixed media (Toutiao), COVID19, ArCOV-19, opinion spam dataset, WeChat (crawled), Zubiagaset, PHEME, Twitter 15, Twitter 16, Weibo (Ma), Twitter (crawled), Twitter (Ma), and Weibo (crawled). The more frequently used datasets are Twitter 15, Twitter 16, Weibo (Ma), PHEME, Weibo(crawled), Twitter (Ma), and SemEval 2017. The most important English datasets are Twitter 15, and Twitter 16, while the most used Chinese dataset is Weibo (Ma). Table 6 shows the details of Twitter 15, Twitter 16, and Weibo (Ma). In Table 6, we present detailed information about the moeality, size, labels and URLs of the more frequently used datasets.

There is 1 paper each related to health and science. The rest of the papers are not related to any obvious domain. Text pre-processing is the initial stage of cleaning up text data before feeding it to the model. Text data includes, among other things, noise in the form of web linking, punctuation, and text in different cases. Unstructured data is turned into structured data based on the need for summarization. Filtering “#” and lowercase were used the most frequently among 12 recent studies since 2017. Generally, filtering “# means filtering “#+topic+#”. Each current post’s topic is shown in this section of the information. On the other hand, event microblog texts should be promptly filtered out and discarded since they are gathered and arranged based on the same subject. These characters will be converted to simplified Chinese and English in lowercase, respectively, due to the mix of traditional and simplified Chinese or English in lowercase and uppercase. Stop words remove low-level information from a corpus, making room for crucial information such as “and”, “or”, and “so”. The next method is to remove web links, which were used 3 times and normalized by using special symbols to replace them. Tokenize is tied for the second used pre-processing method. Tokenizing is a technique for breaking down phrases, paragraphs, or texts into discrete tokens or components. Punctuation is used two times, along with special character removal. This method, for example, will remove full stops “.”, “commas”, question marks “?”, plus “+” or equal “=“ from the text. For English datasets, the words are segmented by white space. For the Chinese datasets, the words are segmented by word segmentation tools, such as Jieba2 library, LTP tools, and so on. Other pre-processing techniques include deleting foreign texts, extracting timestamps, lemmatization, filtering multimedia, handling repeat words, handling emoticons, filtering multimedia, normalizing, parser, removing diacritics, and tagging.

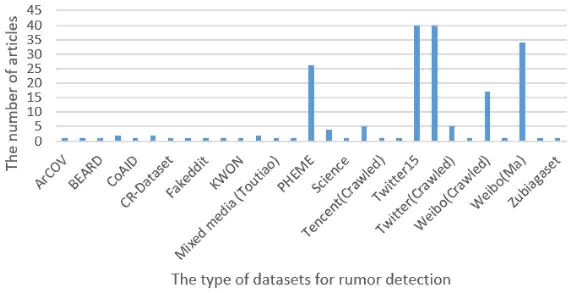

Figure 4. A bar chart of the datasets used in the studies of rumor detection

Table 6. The details of the datasets

Dataset	Modality	Size	Labels	URL
Twitter 15	Propagation trees	1,381 propagation trees. 276,663 users	Unverified, true, false, non-rumor	http://www.dropbox.com/s/7ewzdrbelpmrnxu/rumdetect2017.zip?dl=0
Twitter 16	Propagation trees	1,381 propagation trees. 276,663 users	1,381 propagation trees.276,663 users	http://www.dropbox.com/s/7ewzdrbelpmrnxu/rumdetect2017.zip?dl=0
Weibo (Ma)	Text and image	40 K tweets	Rumor, Non-rumor	https://drive.google.com/file/d/14VQ7EWPiFeGzxp3XC2DeEHi-BEisDINn/view

3.2. Features

Most researchers consider rumor detection to be a binary classification problem. Accurate classification relies heavily on feature extraction. To dig deeper into the details, we considered three aspects that are text features, user features, and propagation features to explore feature extraction. In total, there were

52 papers using text features only, 12 text and user features, 12 propagation features, 17 propagation and text features, four user and text and propagation features, two user and propagation features, and one user feature.

3.2.1. Text-based features

The conventional text features for rumor detection can be grouped into three categories: lexical features, syntactic features, and topic features. Language qualities that are taken from a word or single word level are referred to as lexical features. The "number of words," "number of @," "number of #," the message's length, and other indicators were determined using manual rumor detection features based on the words in a claim/comment. Syntactic features are those that originate at the sentence level, like word frequency and part-of-speech labeling. The goal of the topic features is to comprehend the text information and its potential semantics. It is the text feature that has been taken from the complete message collection.

The recent studies mostly derived semantic features from texts, as opposed to manually extracting features from the text in the past. Out of 100 studies, 85 of them use text features for rumor detection. It has been proved that text-based features are crucial for rumor detection.

Studies extracted text features based on word-level by exploiting pre-trained models, DL models, and statistical methods. Most of the studies used pre-trained models to represent text features. For example, study by [16], [23], [50]–[54] calculated the word embedding vector of each word from claims by word2vec. Researchers [55]–[63] employed bidirectional encoder representations from transformers (BERT) to extract textual information. The authors of [3], [64]–[66] reported that features extracted from the tweets are represented in vector form using the GloVe embedding technique [67]. Learned representations of rumors by adopting the Embedding from language from models (ELMo). Zou *et al.* [68] encoded entities and external knowledge by using pre-trained enhanced representation through knowledge integration (ERNIE) [69]. Used generalized autoregressive pretraining for language understanding (XLNet) to generate word vectors to represent Semantic and structural information from words. The authors of [70]–[73] Took different advantages of multiple pre-training techniques to construct their embedding layers, while [26] used multiple pre-training techniques for different language datasets. Some researchers passed the single tweet text to the DL model for capturing word embedding. Ma *et al.* [74] represented word embeddings by using the Doc2vec model. Researchers [75], [76] used CNN to capture word embeddings. Lan *et al.* [77] utilized a bidirectional Gated recurrent neural network (GRU) to gain representations of each word. Li and Qian [78] encoded post representations by using a two-layer graph convolutional networks (GCN). Zhang *et al.* [79] employed the variational autoencoder (VAE) as the foundational model for textual representation. Some studies employed statistical features from tweets, replies tweets, or comments to represent text features. Chen *et al.* [80] extracted statistical features from tweets and comments, while Zhang *et al.* [81] represented text statistical vectors from users and texts. Luo *et al.* [82] combined statistical features and word vectors. Xu *et al*. [83] used the one-hot method to encode tweets while reserving the sentence structure and the semantic information. Song *et al.* [84] adopted tf*idf and CNN to represent each repost.

Some researchers viewed rumor detection as event-level instead of identifying each single post. They represent events based on different strategies while representing text features by employing pre-trained models, DL models, machine learning models, or statistical methods. Peng and Wang [54] grouped the rumor data by events, and considered each event data as a paragraph, employed Doc2Vec to generate paragraph vectors. Tarnpradab and Hua [85], the main tweet and the corresponding reply tweets constituted a social network. They converted each word in the texts (including the posts and replies) into pre-trained modeling to form embeddings. Li *et al.* [86], a tweet tree consists of the source tweet and reply tweet. They used word embeddings (including source tweets and reply tweets) for rumor detection. Luo *et al.* [87] utilized Doc2Vec to convert microblogs to sentence vectors. Bai *et al*. [32], the source tweet and its replies composed a conversation. They trained word vectors by using the Word2Vec model. Zhong *et al*. [88] used BERT to extract text content features. Zeng and Gao [89], an event consists of a set of relevant posts in chronological order. They used pre-trained words and the TF-IDF method to represent the texts from different datasets. Wang *et al*. [90], each event includes all related posts at timestamps, and texts were represented by Word2vec. The authors [19], [91], each event also includes all relevant posts at timestamps. They used RNN to extract semantics information in events. Wang *et al*. [92] considered the posts based on the same topic as an event, and used a dynamic time series algorithm based on a fuzzy clustering algorithm to represent events. Kim and Yoon [60] used the CNN model to obtain the word sequence embeddings' semantic representation of the text. Ma [93] employed tf*idf to represent the original post sequence.

3.2.2. User-based features

The user's social network is the source of their features. A few users create rumors, and many users disseminate them because of financial incentives. Critical hints for rumor detection can be obtained from the examination of user features. Both individual and group characteristics are considered user features. Among them, specific attributes like "age," "identity authentication," and "registration time" are taken from a particular user. In the early studies, user traits like "user location" and "user credibility" were employed to assess users' dependability from the viewpoint of reporters. The traits taken out of the user group, like "verified user ratio," and so forth, are known as features of the user group.

Since a rumor tweet (whether true or false) can trigger a wide variety of user responses. Therefore, many studies have shown that both texts and user profiles could provide more useful features to detect rumors. Jiang *et al*. [35] reported that the content that each account's user has registered as well as some statistical details related to the account are the primary components of the user information characteristics. To extract the text content's representation vectors, they applied the pre-trained BERT Chinese model. Huang *et al.* [94] fused the nodes' information to represent source tweets. By using the transformation matrix, user features were learned from the user behaviors or user profile data. Akhtar *et al*. [72] used GloVe for word representations. They believed that rumors are rarely transmitted by trustworthy sources, whereas rumor mongers typically use anonymity to propagate inaccurate information and do not want to be identified. They defined five statistical features to represent a user. In particular, Tarnpradab and Hua [85] only employed usr2vec to initialize a representation of each unique user. Xu *et al*. [21] adopted the bidirectional long short-term memory network (Bi-LSTM) model to extract text features from the original posts and retweets, and encoded user features by statistical features. Islam *et al*. [95] used TF*IDF to represent text vectors, and then passed them through long short-term memory (LSTM) models to obtain text embeddings. They applied VAE to obtain user feature representation. Islam *et al*. [95], the BERT model was used to represent word embeddings, while the CNN model learned user vectors. Bing *et al*. [96] used the BERT model to capture post representations, and used a graph attention network (GAT) to model the network of users. Tian *et al*. [97] utilized a CNN to encode the source post, used a GCN-based component to represent user publishing, and used a GAT to based component to represent user interaction. Huang *et al*. [98] adopted the self-mechanism to represent user and content interactions. Malhotra and Vishwakarma *et al*. [34] Extracted text features from source tweets by using the RobustlyoptimizedBERT approach (RoBERTa) model, and used statistical features from users.

3.2.3. Propagation-based features

Rumors will spread through being liked, commented on, and forwarded by the majority of people. Features that are retrieved from data that arise during communication, such as "user comments," "number of reposts," "number of likes," "number of clicks," and so forth, are known as propagation features. Table 7 (in Appendix) [99]–[118] shows details of using propagation features for rumor detection tasks in the latest research.

From Table 7 (in Appendix) and Figure 5, we can observe that there are 32 papers to exploit propagation features. Additionally, it should be noted that, in comparison to other propagation features for rumor detection, the majority of works have used based-text and based-propagation features. Their use is still ongoing and is followed by the combination of text-based, user-based, and propagation-based features as well as only propagation-based features, indicating that these features are demonstrated to be effective for the rumor detection task.

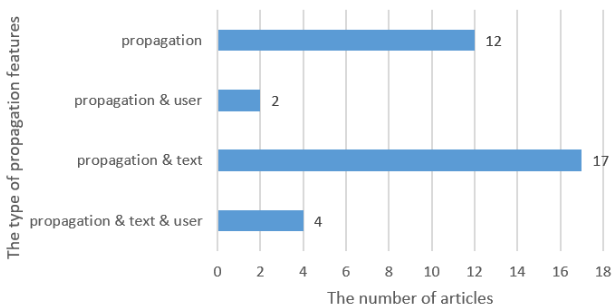

Figure 5. The distribution based on propagation features

3.3. DL techniques for rumor detection

DL models have shown extraordinarily progressed in many fields, including computer vision, speech recognition, as well as NLP. In contrast to machine learning methods, DL significantly outperforms. Particularly, Ma *et al.* [24] started to detect rumors in microblogs via DL. Many researchers explored applying different DL models in rumor detection tasks. Among these models, there are affiliation relationships between models. For example, GRU is a simplified structure of LSTM, while LSTM is an optimized structure of RNN. Bi-LSTM is a model that combines a forward LSTM with a backward LSTM. The logic of Bi-GRU and Bi-LSTM is the same, both do not change the internal structure but apply the model twice and in different directions. Based on the architecture of the models, we categorized the models used in the research literature. DL models in rumor detection were classified into eight types, which are CNN, RNN, ordinary differential equation network (ODE-net), generative adversarial network (GAN), autoencoder (AE), attention mechanism, GNN, and hybrid. To further filter the analysis, we divided it into four types.

– Rumor detection based on CNN.
– Rumor detection based on RNN.
– Rumor detection based on GNN.
– Rumor detection based on other methods.

According to the taxonomy of the DL models used for rumor detection, we analyzed the frequency of use of different models from 2017 to 2023. We can observe that the DL models that are used more frequently are RNN, CNN, and GNN. With the emergence of the GNN model in 2019, the usage of RNN and CNN has been declining, while the GNN model has been rising. The details are shown in Figure 6.

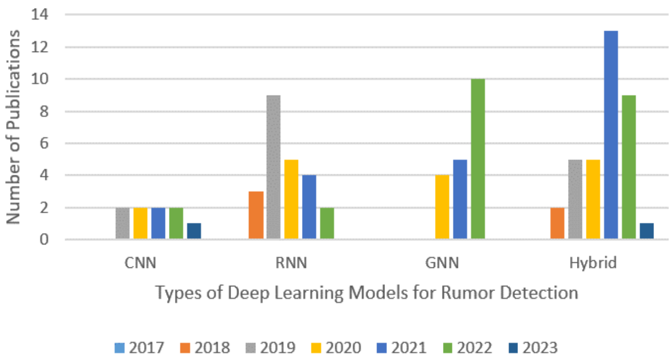

Figure 6. The trend of DL models in rumor detection (2017-2023)

3.3.1. CNN-based techniques

The CNN is a deep-structured feedforward neural network with convolutional computations. It is among the DL algorithmic exemplars. Its performance in NLP and computer vision, among other areas, has been demonstrated, as have its enhanced models.

Based on different data types, researchers use CNN models to extract different features. Yuan *et al.* [75] employed the CNN-based model to learn the semantics features from microblogs. Chen *et al.* [52] exploited text-CNN to extract textual features. Bharti *et al.* [27] used the CNN model to transmit the context of word embedding. Han *et al.* [62] used a CNN model to obtain visual information. Tu *et al.* [25] extracted high-order propagation features and the source tweet by CNN. Xu *et al.* [26] captured the dependency between the word embeddings and topic vectors by using the CNN model. Liu *et al.* [119] applied CNN-based neural networks over the generated meta-tree paths to learn the global structural representations. Moreover, some researchers utilized CNN as a classification model for rumor detection. The CNN model and the attention mechanism are integrated for rumor detection [94]. Yang *et al.* [28] constructed an accurate rumor detection model by combining the features with a CNN model.

3.3.2. RNN-based techniques

When given a data sequence as input, an RNN is recursive in the direction that the sequence is evolving, and each node is connected to the others in a chain. Its ability to preserve temporal information and the last state during recursion is the main distinction between it and CNN. As a result, recurrent neural networks are being used extensively in current research for NLP. Ma *et al.* [24] first cited RNN for the rumor detection task, employing TF-IDF to model words, automatically learning Twitter content based on time series, and RNN to learn possible rumors' content. Since then, rumor detection has made greater use of recurrent neural networks and variations. Of the papers that we investigated, 28 papers are based on RNN models. The most commonly used RNN models are LSTM, Bi-LSTM, GRU, Bi-GRU, and RNN.

It is noted that many studies used RNN-based models to encode high-level and rich features from text or learn the representation of propagation features. The authors [19], [23], [70], [72], [95] employed the LSTM model to learn the rich semantics features behind texts, and [16] has the highest accuracy 98%, which is based on self-built Chinese health datasets. Researchers [16], [51], [57], [85] utilized Bi-LSTM to learn the context features of the tweet. Xu *et al*. [21] adopted the Bi-LSTM model to learn word representation and the dynamic characteristics of retweets. Chen *et al*. [69] used Bi-GRU to capture contextual information, while [20] adopted Bi-GRU to learn representations for dynamic structures. Liu *et al*. [22] utilized the user's forwarding to get propagation features on time series via LSTM. Wang *et al*. [90] used a two-layer GRU model to learn continuous representations of microblog events, whereas Wang and Guo [66] used a two-layer cascaded gated recurrent unit (CGRU) model to detect rumor events. Wang *et al*. [92] used a two-layer GRU model to capture the hidden feature representations. Luo *et al*. [82] integrated GRU to represent post content, topology network of posts, and metadata extracted from post datasets. Zhang *et al*. [56] utilized two GRUs as meta and task networks. Xu *et al*. [83] used multiple RNN layers to learn temporal features. Zeng and Gao [89] utilized an RNN with continuous-time LSTM to capture the complex effects. Liu *et al*. [22] learned the propagation structures by using the LSTM-based models. Ni *et al.* [104] obtained word-level and event-level feature representations by using the LSTM model. Xu *et al.* [21] Employed the TF-IDF and Bi-LSTM models to encode propagation and weighted summation of the states to represent the words in the source data. Lan *et al*. [77] learned high-level semantic representations with Bi-GRU and hierarchical attention mechanism, and employed a single-layer GRU to capture latent representations of semantic information. Kotteti *et al*. [5] ensembled BiGRU, BiLSTM, GRU, LSTM, LG, and RNN to determine the prediction results by using majority voting. Han *et al*. [67] used a deep bidirectional language model to learn representations of rumors.

3.3.3. GNN-based techniques

Recently, GNN has been rapidly developed in rumor detection. One popular kind of graph data that shows the social relationships between different people or organizations is the social network. In addition, unlike CNN and RNN, the GNN preserves the structure of rumor propagation because it accepts non-euclidean graphs as input. According to recent research, there are differences in the structures used by rumors and accurate information to spread.

There are 18 papers based on GNN. These papers show that the primary applications of GNN are in the extraction of propagation structural features and user interaction structural features. Li *et al*. [31] used GNN to obtain word-level features, and post-level features via TF-IDF. Yu *et al*. [35] utilized a GCN model to get the vector representation from rumors. Zhang *et al*. [36] used GNN to learn the relations among tweets optimally. Wu *et al*. [14] employed gated GNN to encode node attributes. Li *et al*. [86] constructed an out-in-degree graph, and employed GCN to capture the semantics from the rumor propagation network. In GCN [33] considered retweet or responsive nodes as neighbors and updated all the nodes' representations simultaneously. Huang *et al*. [94] employed an attention mechanism to generate graph representation based on constructed datasets. Ke *et al*. [115] used a GCN to describe user propagation representations. Zhong *et al*. [88] used Bi-GCN to extract the aggregation and propagation features from rumors. Lin *et al*. [105] represent the propagation of each claim with GCN. Chen *et al*. [106] decomposed the propagation tree as two unsymmetric adjacency matrices to employ different GRN layers for interaction direction control. Nanjiang *et al*. [107] used a Bi-GCN to learn the propagation structure features. Zhang *et al*. [108] represented the event propagation graph structure by using two GNN encoders. Luo *et al*. [109] used GCN network to extract propagating features. Wei *et al*. [111] used GCNs to capture structural features in the graph which is constructed by the claim's information cascade. GCNs to capture the structural propagation features from rumors [112]. Liu *et al*. [113] combined GCNs to obtain propagation structure features with reduced interference. Bai *et al*. [100], each post-event is transformed into two two-directional graphs. They used a Bi-directional graph attention network (Bi-GAT) for rumor detection.

3.3.4. Other techniques

Studies have been done to develop their algorithms using alternative model structures in addition to the aforementioned techniques. CNN is not appropriate for capturing dependencies inside the field sequence because it is simple to parallelize. Long-distance sequence dependence can be captured using RNN. Nevertheless, implementing parallel processing sequences is difficult. To fuse the superiority of CNN, RNN, GNN, and other DL models, some researchers implemented a variety of hybrid methods to detect rumors. For example, (a hybrid DL model based on CNN-BiLSTM for rumor detection) combined GloVe, CNN, and BiLSTM to represent text features. Wei *et al*. [8] utilized LSTM to extract high-level representations, and adopted SENet to apply the attention and gating mechanisms. Song *et al*. [84] combined CNN and LSTM to get the representations of original microblogs and repost sequences. Chen *et al*. [29] proposed a multi-hop graph convolution layer (MHGCN) to extract user influence and susceptibility, used Bi-GRU to learn the temporal information and employed VAE to capture the uncertainty in the learned user features. Luo *et al*. [87] employed CNN to extract features from the propagation tree of a source tweet, and used a transformer-based model to extract source tweet representations. Li *et al*. [86] used BERT and BiLSTM to learn textual features, learned propagation structures by the GNN model, and obtained the temporal characteristics by self-attention mechanism. Zhang *et al*. [79] employed textual features by using the variational autoencoder and GRU. Li *et al.* [101] generated tweet representations by a deep BiLSTM, and used GraphSAGE to represent conversation structure features. Malhotra *et al*. [34] used the RoBERTa model to encode textual information, and adopted the GCN model to represent user features and relationships among related users. Chen *et al*. [30] used Bi-GRU to extract temporal patterns from the user interaction time series and vanilla GCN to encode the macroscopic diffusion of a tweet. Lin and Chen *et al*. [50] extracted semantic features by combining multihead self-attention mechanism with a transformer encoding block, and used LSTM to fuse features. Al-Sarem *et al.* [71] used word2vec, GloVe, and fast text model pre-trained embedding models to learn word vectors, and integrated the LSTM model and CNN model to represent semantic information of tweets. Kotteti *et al*. [5] combined LSTM and GRU layers (LG) to construct a neural network, and ensembled BiGRU, BiLSTM, GRU, LSTM, and RNN to determine the final prediction after cleaned data. Wu *et al*. [53] represented emotion characteristics using the ConvNet model after obtaining semantic representations using a BiLSTM-based sequence encoder. Lingyu *et al*. [102] employed the CNN model to learn grammatical features and subevent's stance features, and utilized the RNN model to capture the event-related features. Kotteti *et al*. [116] used LSTM, GRU, Bi-RNN, and CNN to learn the propagation pattern of the tweets. Poddar *et al*. [117] encoded tweet text features by using the GloVe model and a self-attention mechanism, encoded conversation sequence features by using a bi-directional RNN. Liu *et al*. [19] used the LSTM model to learn semantic representations, and used text visual geometry group (text-VGG) to learn event representations. Zhou *et al.* [120] employed GRU to mine temporal information across related microblogs under rumor events and CNN automatically generated the characteristics of the rumor microblogs. Wang *et al.* [20] used RNN to represent each sub-structure, adopted BiGRU to learn representations for dynamic structures, used a paragraph vector to represent each post, and used CNN to represent all the posts in an event. Huang *et al.* [37] learned propagation tree structure using a RvNN encoder, a GCN model to obtain the high-level user representation. Ben *et al.* [73] represented word features by a pre-trained model, and learned the semantic relations between the tweets based on the self-attention mechanism. Chen *et al*. [80] used RNN and AE to learn users' behaviors based on the statistical features from tweets and comments. Zhang *et al*. [58] used BERT, Bi-LSTM+attention, and CNN+attention to represent tweets. Bing *et al*. [96] used the BERT model to represent text contents, extracted user features from user descriptions by the text-CNN model, used Bi-GRU to learn semantic information from microblogs, and utilized dual co-attention to capture the mutual attention between the user profiles and the original tweet. Huang *et al.* [98] represented user publishing by using the GCN-based model, encoded user interaction by using the GAT-based model, and used a CNN model to extract semantic information from the post. Tian *et al.* [97] used BERT and GAT to model the comment tree, modeled the comment chain based on the transformer, and employed GAT to model the network of users. Kim and Yoon [60] only used text information to detect rumor veracity based on the BERT model by by double-channel structure. Almars *et al*. [65] employed the GloVe model to learn the embeddings of words, used CNN to extract semantic features, and combined Bi-LSTM and attention mechanism to predict the labels. Luo *et al.* [114] extracted temporal features based on the transformer model, and learned the propagation features by GCN. Bao *et al*. [86] employed self attention mechanism to learn temporal dynamics information, and encoded propagation information by GCN. Yan *et al*. [110] used a CNN-based model to extract semantics information, employed the attention mechanism to understand context, and utilized GAT to learn the propagation representations. Zou *et al*. [68] used ERNIE to encode entities and external knowledge, and used three-level co-attention network to represent the interaction between the entity information and the image, external knowledge and text representation, image and text.

Moreover, some researchers adopted different learning schemes which are adversarial learning, multi-task learning, and reinforcement learning to identify rumors. For the adversarial learning scheme, Li and Qian [78] used BiGRU to represent words from tweets, and combined a fast gradient method algorithm to detect rumors. Guo *et al.* [76] Utilized CNN to learn semantic representations from events, and combined adaptive learning to predict rumor labels. Dong *et al.* [103] used BiGRU to encode source posts, utilized BiGRU to get the temporal features, and got the propagation structures by using GCN. Ma *et al*. [93] Encoded posts by an RNN encoder, utilizing the conventional transformer to represent positional information. For the multi-task learning scheme, Zhang *et al.* [56] employed BERT to learn text features, used VGG19 to learn visual features, and utilized the attention mechanism to learn stance features. Chen *et al.* [118] modeled the process of rumor generation by examining it from the knowledge perspective. They adopted GAN and reinforcement learning to generate high-quality rumor text and used GCN to extract graph structure information. For the reinforcement learning scheme, Yuan *et al.* [55] used a stance-aware reinforcement learning method to detect rumors. For the environment part, they used CNN to encode tweets and comments, while they used

LSTM to encode tweets for the agent part. However, other researchers utilized strategies to detect rumors. For example, Zhang *et al.* [91] used the WAE model to learn the stance topic features from response propagation trees based on rumors. Ma *et al.* [74] used ODE-net to classify the rumors. Zhang *et al*. [81] considered rumors as an abnormal tweet, and tried to determine the appropriate cutoff point for the reconstructed mistakes to separate rumors from non-rumors. They learned the recent posting behavior habits of a user through an autoencoder. Zuo *et al*. [59] utilized the pre-trained model to represent claim and comment text, and used a bidirectional knowledge transfer strategy to continuously detect unseen rumors.

From the perspective of explainability, Wu *et al.* [63] engaged in deep semantic interaction with tweets to obtain false parts within them by using co-attention self-attention networks. Lu and Li [64] represented user characteristics by statistical features, utilized GRU to learn the word sequence representation, employed GRU and CNN to learn propagation representations, used two-layer GCN to learn graph-aware representations, and adopted dual co-attention mechanisms to capture the correlation between the source tweet and users' interactions/propagation. They used co-attention weights for the explainability.

3.4. Evaluation metrics for rumor detection

Among surveyed papers, 64 research papers used F1 as evaluation metrics to determine the quality of their model on DL, 55 papers with Accuracy, 48 papers with precision, 48 papers with recall, and 2 papers with false positive rate (FRP). Since F1, accuracy, precision, recall, and FRP are commonly used in NLP, we won't go into details. In addition to the above evaluation criteria, interpretability is receiving increasing attention. For example, Lu and Li [64] showed a sample by visualizing comments and semantics manually. In the work of Wu *et al.* [63], the model can be explained by the co-attention weights given to the words in the source tweet and the users that rebroadcast it. It is possible to identify evidence terms and users in identifying fake news by displaying the distribution of attention weights.

4. CONCLUSION

From past research utilizing PRISMA and eight literature probes, we can discover major criteria utilized for rumor detection. We were also able to categorize the origin of the dataset, the domain of the research, and the most recent pre-processing method. The literature probe was also able to identify the most recent rumor detection method with the performance evaluation.

As technology advances and rumors continue to present new challenges, scholars have made a lot of effort and achieved significant advancements to increase the reliability of the information on the network. Not all of the core issues have been resolved. The following future directions provide information regarding the response to research question LP8.

Fuse more information. There is growing interest in leveraging various features of information to improve the accuracy and efficiency of rumor detection systems. Based on a supervised learning model, features are mainly extracted from users, microblog text contents, and propagation structures associated with rumors, allowing for better detection and classification. Besides, knowledge in the knowledge graph, sentiment analysis, thematic analysis, and other types of data, such as videos, audio, images, and so on, can also be considered complementary to features.

Integrate NLP models. DL techniques have started to be applied more often in social media rumor identification with better generalization performance in the last several years. Based on different types of features, suitable DL models are utilized to represent the features. Especially, the emergence of large language models represented by GPT4 presents an opportunity to explore a new way to verify rumors. Expand current corpus. It is crucial to create an adequate corpus to identify and debunk rumors. Only a few datasets are currently available for research. In the future, the corpus should be expanded both in terms of data collection and data labeling. Data should be collected from different social platforms and different languages. Moreover, investigating an automatically labeled corpus method should be the focus of future research.

Joint additional tasks. By jointly training models on related natural language processing tasks, e.g., stance detection, they can benefit from shared representations and learn more robust features that capture both the context and semantics of rumors. Additionally, there is a promising direction to apply rumor detection models to downstream tasks or other relative tasks, such as cyberbullying, hate speech, and so on. Boost the explainability of models. The explainability in rumor detection is an emerging field of research, driven by the need to understand the reasoning behind model decisions and improve transparency. By making the decision-making process more transparent, users can gain trust in the model and assess its reliability. However, despite progress in this area, there are still challenges in achieving full explainability in rumor detection. DL models, such as neural networks, are often considered black boxes because they are complex and difficult to interpret. Explaining their behavior comprehensively remains a challenging task.

APPENDIX

Table 7. The details of rumor detection based on propagation features

Ref.	Author	Year	Features: Propagation	Text	User	Extraction method description
[91]	Zhang *et al.*	2021	√			Using the bag-of-word (BoW) model to learn propagation features, and learning the stance topics features based on the wavelet-like auto-encoder (WAE).
[29]	Chen *et al.*	2021	√		√	Employing a GNN layer to understand user influence and susceptibility, a randomized truncated singular value decomposition (tSVD)-based sparse matrix factorization (SMF) to learn social features, a bidirectional gated recurrent unit (Bi-GRU) model to extract the user's temporal information, and a VAE to represent user features.
[14]	Wu *et al.*	2020	√	√		Using the Doc2Vec model to capture content representations from the tweet, and using the GNN to capture propagation features.
[8]	Wei *et al*	2021	√	√	√	Representing user features, content features, and propagation features based on the kernel subtree of the event.
[87]	Luo *et al.*	2021	√	√		Using the CNN to represent propagation features from tweets, and generating source tweet embeddings by BERT and RoBERTa.
[86]	Li *et al.*	2022	√	√		Learning the temporal features by a timestamp encoding function, structural features of propagation graph by the GNN model, and extracting the textual features of source tweets by the BERT model and Bi-LSTM model.
[99]	Song *et al.*	2021	√			Applying GCN to capture high-level node characteristics from user relationship graphs.
[100]	Bai *et al.*	2021	√			Transforming each post-event into two-directional graphs to capture propagation representations based on GAT.
[101]	Li *et al.*	2020	√	√	√	Capturing the tweet representation by a deep BiLSTM, using graph sample and aggregated (GraphSAGE) to generate representations of the propagation, and utilizing user profile-based features and user credibility features to represent users' features.
[30]	Chen *et al.*	2021	√			Encoding the macroscopic diffusion of a tweet based on vanilla GCN, and capturing temporal patterns from the user engagement time series by employing a Bi-GRU.
[36]	Zhang *et al.*	2021	√			Using simplified aggregation GNN to capture propagation features.
[102]	Lingyu *et al.*	2019	√	√		Encoding root tweets and retweets by the GRU model, and capturing grammatical features from contexts by the CNN model,
[103]	Dong *et al.*	2022	√	√		Using Bi-GRU to represent textual information, and encoding propagation by two layers GCN.
[104]	Ni *et al.*	2022	√	√		Obtaining the embedding of each word from texts by GloVe, and representing the propagation by Bi-LSTM.
[105]	Lin *et al.*	2022	√	√		Utilizing cross-lingual language model-RoBERTa (XLM-RoBERTa) to encode posts, and represent the propagation of each claim with GCN.
[106]	Chen *et al.*	2022	√	√		Encoding textual features by the BERT model, and employing graph recurrent network (GRN) to represent propagation structures.
[107]	Nanjiang *et al.*	2022	√	√		Extracting the text by using the BERT model, and the propagation structure features by bidirectional graph convolutional network (Bi-GCN).
[108]	Zhang *et al.*	2022	√	√		Embedding the textual data by leveraging TF-IDF, and encoding propagation graph structure by GNN.
[109]	Luo *et al.*	2022	√			Utilizing the GCN network to obtain propagation features.
[110]	Yan *et al.*	2022	√	√		Using a CNN-based model to understand the semantics information of tweets, and representing the propagation features by GAT.
[111]	Wei *et al.*	2024	√	√		Encoding textual contents by the text embedding layer, and capturing propagation. Features by the GCN model.
[112]	Zhang *et al.*	2022	√			Using the long-tail strategy to encode propagation features based on GNN.
[113]	Liu *et al.*	2023	√	√		Using word2vec to represent textual features, capturing propagation features by GCN.
[114]	Luo *et al.*	2023	√			Adopting Transformer encoders to capture temporal features and GCN to get the propagation features.
[33]	Lin *et al.*	2020	√			GCN updates the features of nodes and creates reinforced features for each post based on its propagation path, by combining the features of its neighbors.
[115]	Ke *et al.*	2020	√	√		Utilizing the multi-head attention to represent microblog, and encoding propagation structure via GCN.
[75]	Yuan *et al.*	2019	√	√		Learning word representation by the multi-head attention module, and representing each node in the graph by the attention mechanism.

Table 7. The details of rumor detection based on propagation features (*continue*)

Ref.	Author	Year	Features: Propagation	Text	User	Extraction method description
[20]	Wang *et al.*	2019	√	√		Using RNN to cater to the sub-structure, and adopting Bi-GRU to learn representations for dynamic structures.
[37]	Huang *et al.*	2023	√		√	Encoding the words by a fixed-length vector, and using Recursive variational neural network (RvNN) to obtain propagation features and semantic features.
[22]	Liu *et al.*	2019	√	√	√	Calculating the tf*idf value to obtain the fixed-dimensional word vector for each term, extracting eight discriminating user characteristics, and utilizing the user's forwarding to get the propagation features on time series.
[64]	Lu and Li	2020	√	√	√	Employing texts and profiles to define user statistical features, utilizing GRU to learn the word sequence representation, and using GRU and CNN to learn propagation representations.
[116]	Kotteti *et al.*	2019	√			Utilizing LSTM, GRU, bidirectional recurrent neural network (Bi-RNN), and CNN to capture propagation features.
[117]	Poddar *et al.*	2018	√	√		Using GloVe and a self-attention mechanism to represent textual information, using a bi-directional RNN to encode the conversation sequence.
[118]	Chen *et al.*	2021	√			Using GCN to model data in graph structure

ACKNOWLEDGEMENTS

Author thanks the Shanxi Province Education Department Project(2020L0665) to support.

REFERENCES

[1] Edelman, "Edelman Trust Barometer 2022. Reporte regional. La confianza en latinoamérica," *Edelman América Latina*, 2022.

[2] Edelman, "Edelman Trust Barometer 2021, Country Report: Trust in Indonesia," *Edelman Trust Barometer*, p. 37, 2021.

[3] N. Rani, P. Das, and A. K. Bhardwaj, "A hybrid deep learning model based on CNN-BiLSTM for rumor detection," in *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021*, IEEE, pp. 1423–1427, 2021, doi: 10.1109/ICCES51350.2021.9489214.

[4] F. M. Dito, H. A. Alqadhi, and A. Alasaadi, "Detecting Medical Rumors on Twitter Using Machine Learning," in *2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies, 3ICT 2020*, IEEE, pp. 1–7, 2020, doi: 10.1109/3ICT51146.2020.9311957.

[5] C. M. M. Kotteti, X. Dong, and L. Qian, "Ensemble deep learning on time-series representation of tweets for rumor detection in social media," *Applied Sciences (Switzerland)*, vol. 10, no. 21, pp. 1–21, 2020, doi: 10.3390/app10217541.

[6] S. Shelke and V. Attar, "Rumor detection in social network based on user, content and lexical features," *Multimedia Tools and Applications*, vol. 81, no. 12, pp. 17347–17368, 2022, doi: 10.1007/s11042-022-12761-y.

[7] P. K. Roy and S. Chahar, "Fake Profile Detection on Social Networking Websites: A Comprehensive Review," *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 3, pp. 271–285, 2020, doi: 10.1109/TAI.2021.3064901.

[8] Z. Wei, X. Xiao, G. Hu, B. Zhang, Q. Li, and S. Xia, "A Novel and High-Accuracy Rumor Detection Approach using Kernel Subtree and Deep Learning Networks," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2021, doi: 10.1109/IJCNN52387.2021.9534311.

[9] S. Shelke and V. Attar, "Role of Various Features in Identification of Rumors in the Social Network," in *2021 12th International Conference on Computing Communication and Networking Technologies, ICCCNT 2021*, IEEE, pp. 01–06, 2021, doi: 10.1109/ICCCNT51525.2021.9579856.

[10] H. M. Jabir, M. A. Naser, and S. O. Al-Mamory, "Rumor Detection on Twitter Using Features Extraction Method," in *Proceedings of 2020 1st Information Technology to Enhance E-Learning and other Application Conference, IT-ELA 2020*, IEEE, pp. 115–120, 2020, doi: 10.1109/IT-ELA50150.2020.9253027.

[11] A. Zubiaga, M. Liakata, R. Procter, G. Wong Sak Hoi, and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PLoS ONE*, vol. 11, no. 3, p. e0150989, 2016, doi: 10.1371/journal.pone.0150989.

[12] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on Twitter," in *Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011*, New York, NY, USA: ACM, pp. 675–684, 2011, doi: 10.1145/1963405.1963500.

[13] M. E. Jaeger, S. Anthony, and R. L. Rosnow, "Who Hears What from Whom and with What Effect," *Personality and Social Psychology Bulletin*, vol. 6, no. 3, pp. 473–478, 1980, doi: 10.1177/014616728063024.

[14] Z. Wu, D. Pi, J. Chen, M. Xie, and J. Cao, "Rumor detection based on propagation graph neural network with attention mechanism," *Expert Systems with Applications*, vol. 158, p. 113595, 2020, doi: 10.1016/j.eswa.2020.113595.

[15] D. Varshney and D. K. Vishwakarma, "A review on rumour prediction and veracity assessment in online social network," *Expert Systems with Applications*, vol. 168, p. 114208, 2021, doi: 10.1016/j.eswa.2020.114208.

[16] D. Lin, B. Ma, D. Cao, and S. Li, "Chinese microblog rumor detection based on deep sequence context," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 23, 2019, doi: 10.1002/cpe.4508.

[17] S. Rastogi and D. Bansal, "A review on fake news detection 3T's: typology, time of detection, taxonomies," *International Journal of Information Security*, vol. 22, no. 1, pp. 177–212, 2023, doi: 10.1007/s10207-022-00625-3.

[18] M. R. Islam, S. Liu, X. Wang, and G. Xu, "Deep learning for misinformation detection on online social networks: a survey and new perspectives," *Social Network Analysis and Mining*, vol. 10, no. 1, p. 82, 2020, doi: 10.1007/s13278-020-00696-x.

[19] F. Xing and C. Guo, "Mining Semantic Information in Rumor Detection via a Deep Visual Perception Based Recurrent Neural

Networks," in *Proceedings-2019 IEEE International Congress on Big Data, BigData Congress 2019-Part of the 2019 IEEE World Congress on Services*, IEEE, pp. 17–23, 2019, doi: 10.1109/BigDataCongress.2019.00016.
[20] S. Wang, Q. Kong, Y. Wang, and L. Wang, "Enhancing rumor detection in social media using dynamic propagation structures," in *2019 IEEE International Conference on Intelligence and Security Informatics, ISI 2019*, IEEE, pp. 41–46, 2019, doi: 10.1109/ISI.2019.8823266.
[21] N. Xu, G. Chen, and W. Mao, "MNRD: A Merged Neural Model for Rumor Detection in Social Media," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–7, 2018, doi: 10.1109/IJCNN.2018.8489582.
[22] Y. Liu, X. Jin, and H. Shen, "Towards early identification of online rumors based on long short-term memory networks," *Information Processing and Management*, vol. 56, no. 4, pp. 1457–1467, 2019, doi: 10.1016/j.ipm.2018.11.003.
[23] W. Zhao, Q. Zhan, and Y. Yan, "Health rumors detection based on deep learning," in *Proceedings-2021 International Conference on Culture-Oriented Science and Technology, ICCST 2021*, IEEE, pp. 413–416, 2021, doi: 10.1109/ICCST53801.2021.00092.
[24] J. Ma *et al.*, "Detecting rumors from microblogs with recurrent neural networks," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 3818–3824, 2016.
[25] K. Tu, C. Chen, C. Hou, J. Yuan, J. Li, and X. Yuan, "Rumor2vec: A rumor detection framework with joint text and propagation structure representation learning," *Information Sciences*, vol. 560, pp. 137–151, 2021, doi: 10.1016/j.ins.2020.12.080.
[26] F. Xu, V. S. Sheng, and M. Wang, "Near real-time topic-driven rumor detection in source microblogs," *Knowledge-Based Systems*, vol. 207, p. 106391, 2020, doi: 10.1016/j.knosys.2020.106391.
[27] M. Bharti and H. Jindal, "Automatic rumour detection model on social media," in *PDGC 2020-2020 6th International Conference on Parallel, Distributed and Grid Computing*, IEEE, pp. 367–371, 2020, doi: 10.1109/PDGC50313.2020.9315738.
[28] H. Yang, Z. Xu, L. Liu, M. Guo, and Y. Zhang, "Dynamic Slide Window-Based Feature Scoring and Extraction for On-Line Rumor Detection with CNN," in *IEEE International Conference on Communications*, IEEE, pp. 1–6, 2019, doi: 10.1109/ICC.2019.8761288.
[29] X. Chen, F. Zhou, F. Zhang, and M. Bonsangue, "Catch me if you can: A participant-level rumor detection framework via fine-grained user representation learning," *Information Processing and Management*, vol. 58, no. 5, p. 102678, 2021, doi: 10.1016/j.ipm.2021.102678.
[30] X. Chen, F. Zhou, F. Zhang, and M. Bonsangue, "Modeling microscopic and macroscopic information diffusion for rumor detection," *International Journal of Intelligent Systems*, vol. 36, no. 10, pp. 5449–5471, 2021, doi: 10.1002/int.22518.
[31] C. Li *et al.*, "Joint Stance and Rumor Detection in Hierarchical Heterogeneous Graph," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 6, pp. 2530–2542, 2022, doi: 10.1109/TNNLS.2021.3114027.
[32] N. Bai, F. Meng, X. Rui, and Z. Wang, "Rumour Detection Based on Graph Convolutional Neural Net," *IEEE Access*, vol. 9, pp. 21686–21693, 2021, doi: 10.1109/ACCESS.2021.3050563.
[33] H. Lin, X. Zhang, and X. Fu, "A graph convolutional encoder and decoder model for rumor detection," in *Proceedings-2020 IEEE 7th International Conference on Data Science and Advanced Analytics, DSAA 2020*, IEEE, pp. 300–306, 2020, doi: 10.1109/DSAA49011.2020.00043.
[34] B. Malhotra and D. K. Vishwakarma, "Classification of Propagation Path and Tweets for Rumor Detection using Graphical Convolutional Networks and Transformer based Encodings," in *Proceedings-2020 IEEE 6th International Conference on Multimedia Big Data, BigMM 2020*, IEEE, pp. 183–190, 2020, doi: 10.1109/BigMM50055.2020.00034.
[35] K. Yu, H. Jiang, T. Li, S. Han, and X. Wu, "Data Fusion Oriented Graph Convolution Network Model for Rumor Detection," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2171–2181, 2020, doi: 10.1109/TNSM.2020.3033996.
[36] L. Zhang, J. Li, B. Zhou, and Y. Jia, "Rumor Detection Based on SAGNN: Simplified Aggregation Graph Neural Networks," *Machine Learning and Knowledge Extraction*, vol. 3, no. 1, pp. 84–94, 2021, doi: 10.3390/make3010005.
[37] Q. Huang, C. Zhou, J. Wu, L. Liu, and B. Wang, "Deep spatial–temporal structure learning for rumor detection on Twitter," *Neural Computing and Applications*, vol. 35, no. 18, pp. 12995–13005, 2023, doi: 10.1007/s00521-020-05236-4.
[38] K. Mishima and H. Yamana, "A Survey on Explainable Fake News Detection," *IEICE Transactions on Information and Systems*, vol. E105D, no. 7, pp. 1249–1257, 2022, doi: 10.1587/transinf.2021EDR0003.
[39] M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar, and M. S. Rahman, "A Comprehensive Review on Fake News Detection with Deep Learning," *IEEE Access*, vol. 9, pp. 156151–156170, 2021, doi: 10.1109/ACCESS.2021.3129329.
[40] N. Kotonya and F. Toni, "Explainable Automated Fact-Checking: A Survey," in *COLING 2020-28th International Conference on Computational Linguistics, Proceedings of the Conference*, Stroudsburg, PA, USA: International Committee on Computational Linguistics, pp. 5430–5443, 2020, doi: 10.18653/v1/2020.coling-main.474.
[41] Z. Guo, M. Schlichtkrull, and A. Vlachos, "A Survey on Automated Fact-Checking," *Transactions of the Association for Computational Linguistics*, vol. 10, pp. 178–206, 2022, doi: 10.1162/tacl_a_00454.
[42] A. R. Pathak, A. Mahajan, K. Singh, A. Patil, and A. Nair, "Analysis of Techniques for Rumor Detection in Social Media," *Procedia Computer Science*, vol. 167, pp. 2286–2296, 2020, doi: 10.1016/j.procs.2020.03.281.
[43] M. Ahsan and M. Kumari, "Rumors and their controlling mechanisms in online social networks: A survey," *Online Social Networks and Media*, vol. 14, p. 100050, 2019, doi: 10.1016/j.osnem.2019.100050.
[44] A. Bondielli and F. Marcelloni, "A survey on fake news and rumour detection techniques," *Information Sciences*, vol. 497, pp. 38–55, 2019, doi: 10.1016/j.ins.2019.05.035.
[45] M. Al-Sarem, W. Boulila, M. Al-Harby, J. Qadir, and A. Alsaeedi, "Deep learning-based rumor detection on microblogging platforms: A systematic review," *IEEE Access*, vol. 7, pp. 152788–152812, 2019, doi: 10.1109/ACCESS.2019.2947855.
[46] J. C. S. Reis, A. Correia, F. Murai, A. Veloso, and F. Benevenuto, "Explainable machine learning for fake news detection," in *WebSci 2019-Proceedings of the 11th ACM Conference on Web Science*, New York, NY, USA: ACM, pp. 17–26, 2019, doi: 10.1145/3292522.3326027.
[47] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Computing Surveys*, vol. 51, no. 2, pp. 1–36, 2018, doi: 10.1145/3161603.
[48] J. Cao, J. Guo, X. Li, Z. Jin, H. Guo, and J. Li, "Automatic Rumor Detection on Microblogs: A Survey," *arXiv*, vol. 1, no. c, pp. 1–14, 2018.
[49] S. M. Alzanin and A. M. Azmi, "Detecting rumors in social media: A survey," *Procedia Computer Science*, vol. 142, pp. 294–300, 2018, doi: 10.1016/j.procs.2018.10.495.
[50] L. Lin and Z. Chen, "Social rumor detection based on multilayer transformer encoding blocks," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 6, 2021, doi: 10.1002/cpe.6083.
[51] H. Zeng, R. Wang, Y. Huang, X. Cui, and Q. Jiang, "Scientific Rumors Detection in Short Online Texts," in *Conference Proceedings-IEEE International Conference on Systems, Man and Cybernetics*, IEEE, pp. 1233–1240, 2021, doi: 10.1109/SMC52423.2021.9659056.

[52] J. Chen, Z. Wu, Z. Yang, H. Xie, F. L. Wang, and W. Liu, "Multimodal Fusion Network with Contrary Latent Topic Memory for Rumor Detection," *IEEE Multimedia*, vol. 29, no. 1, pp. 104–113, 2022, doi: 10.1109/MMUL.2022.3146568.

[53] L. Wu, Y. Rao, H. Yu, Y. Wang, and N. Ambreen, "A Multi-semantics Classification Method Based on Deep Learning for Incredible Messages on Social Media," *Chinese Journal of Electronics*, vol. 28, no. 4, pp. 754–763, 2019, doi: 10.1049/cje.2019.05.002.

[54] Y. Peng and J. Wang, "Rumor detection based on attention cnn and time series of context information," *Future Internet*, vol. 13, no. 11, p. 267, 2021, doi: 10.3390/fi13110267.

[55] C. Yuan, W. Qian, Q. Ma, W. Zhou, and S. Hu, "SRLF: A Stance-aware Reinforcement Learning Framework for Content-based Rumor Detection on Social Media," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2021, doi: 10.1109/IJCNN52387.2021.9533444.

[56] H. Zhang, S. Qian, Q. Fang, and C. Xu, "Multi-Modal Meta Multi-Task Learning for Social Media Rumor Detection," *IEEE Transactions on Multimedia*, vol. 24, pp. 1449–1459, 2022, doi: 10.1109/TMM.2021.3065498.

[57] L. Ke, X. Chen, Z. Lu, H. Su, and H. Wang, "A Novel Approach for Cantonese Rumor Detection based on Deep Neural Network," in *Conference Proceedings-IEEE International Conference on Systems, Man and Cybernetics*, IEEE, pp. 1610–1615, 2020, doi: 10.1109/SMC42975.2020.9283056.

[58] Z. Zhang, Z. Dan, F. Dong, Z. Gao, and Y. Zhang, "A Rumor Detection Method Based on Adaptive Fusion of Statistical Features and Textual Features," *Information (Switzerland)*, vol. 13, no. 8, p. 388, 2022, doi: 10.3390/info13080388.

[59] Y. Zuo, W. Zhu, and G. Cai, "Continually Detection, Rapidly React: Unseen Rumors Detection based on Continual Prompt-Tuning," *Proceedings-International Conference on Computational Linguistics, COLING*, vol. 29, no. 1, pp. 3029–3041, 2022.

[60] A. G. Kim and S. Yoon, "Detecting Rumor Veracity with Only Textual Information by Double-Channel Structure," in *SocialNLP 2022-10th International Workshop on Natural Language Processing for Social Media, Proceedings of the Workshop*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 24–33, 2022, doi: 10.18653/v1/2022.socialnlp-1.3.

[61] X. Cai, T. Tohti, and A. Hamdulla, "A Rumor Detection Method Incorporating Correlation Features," in *2022 3rd International Conference on Pattern Recognition and Machine Learning, PRML 2022*, IEEE, pp. 328–333, 2022, doi: 10.1109/PRML56267.2022.9882233.

[62] H. Han, Z. Ke, X. Nie, L. Dai, and W. Slamu, "Multimodal Fusion with Dual-Attention Based on Textual Double-Embedding Networks for Rumor Detection," *Applied Sciences (Switzerland)*, vol. 13, no. 8, p. 4886, 2023, doi: 10.3390/app13084886.

[63] L. Wu, Y. Rao, Y. Zhao, H. Liang, and A. Nazir, "DTCA: Decision tree-based co-attention networks for explainable claim verification," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 1024–1035, 2020, doi: 10.18653/v1/2020.acl-main.97.

[64] Y. J. Lu and C. Te Li, "GCAN: Graph-aware co-attention networks for explainable fake news detection on social media," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 505–514, 2020, doi: 10.18653/v1/2020.acl-main.48.

[65] A. M. Almars, M. Almaliki, T. H. Noor, M. M. Alwateer, and E. Atlam, "HANN: Hybrid Attention Neural Network for Detecting Covid-19 Related Rumors," *IEEE Access*, vol. 10, pp. 12334–12344, 2022, doi: 10.1109/ACCESS.2022.3146712.

[66] Z. Wang and Y. Guo, "Rumor events detection enhanced by encoding sentimental information into time series division and word representations," *Neurocomputing*, vol. 397, pp. 224–243, 2020, doi: 10.1016/j.neucom.2020.01.095.

[67] S. Han, J. Gao, and F. Ciravegna, "Neural language model based training data augmentation for weakly supervised early rumor detection," in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019*, New York, NY, USA: ACM, pp. 105–112, 2019, doi: 10.1145/3341161.3342892.

[68] T. Zou, Z. Qian, and P. Li, "Multi-level Interaction Network for Multi-Modal Rumor Detection," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2023, doi: 10.1109/IJCNN54540.2023.10191639.

[69] X. Chen, L. Ke, Z. Lu, H. Su, and H. Wang, "A novel hybrid model for cantonese rumor detection on twitter," *Applied Sciences (Switzerland)*, vol. 10, no. 20, pp. 1–12, 2020, doi: 10.3390/app10207093.

[70] A. Aker, A. Sliwa, F. Dalvi, and K. Bontcheva, "Rumour verification through recurring information and an inner-attention mechanism," *Online Social Networks and Media*, vol. 13, p. 100045, 2019, doi: 10.1016/j.osnem.2019.07.001.

[71] M. Al-Sarem, A. Alsaeedi, F. Saeed, W. Boulila, and O. Ameerbakhsh, "A novel hybrid deep learning model for detecting covid-19-related rumors on social media based on lstm and concatenated parallel cnns," *Applied Sciences (Switzerland)*, vol. 11, no. 17, p. 7940, 2021, doi: 10.3390/APP11177940.

[72] M. S. Akhtar, A. Ekbal, S. Narayan, and V. Singh, "No, That Never Happened!! Investigating Rumors on Twitter," *IEEE Intelligent Systems*, vol. 33, no. 5, pp. 8–15, 2018, doi: 10.1109/MIS.2018.2877279.

[73] A. P. Ben Veyseh, M. T. Thai, T. H. Nguyen, and D. Dou, "Rumor detection in social networks via deep contextual modeling," in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019*, New York, NY, USA: ACM, pp. 113–120, 2019, doi: 10.1145/3341161.3342896.

[74] T. Ma, H. Zhou, Y. Tian, and N. Al-Nabhan, "A novel rumor detection algorithm based on entity recognition, sentence reconfiguration, and ordinary differential equation network," *Neurocomputing*, vol. 447, pp. 224–234, 2021, doi: 10.1016/j.neucom.2021.03.055.

[75] C. Yuan, Q. Ma, W. Zhou, J. Han, and S. Hu, "Jointly embedding the local and global relations of heterogeneous graph for rumor detection," in *Proceedings-IEEE International Conference on Data Mining, ICDM*, IEEE, pp. 796–805, 2019, doi: 10.1109/ICDM.2019.00090.

[76] Z. Guo, K. Yu, A. Jolfaei, A. K. Bashir, A. O. Almagrabi, and N. Kumar, "Fuzzy Detection System for Rumors through Explainable Adaptive Learning," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 12, pp. 3650–3664, 2021, doi: 10.1109/TFUZZ.2021.3052109.

[77] T. Lan, C. Li, and J. Li, "Mining Semantic Variation in Time Series for Rumor Detection Via Recurrent Neural Networks," in *Proceedings-20th International Conference on High Performance Computing and Communications, 16th International Conference on Smart City and 4th International Conference on Data Science and Systems, HPCC/SmartCity/DSS 2018*, IEEE, pp. 282–289, 2019, doi: 10.1109/HPCC/SmartCity/DSS.2018.00068.

[78] X. Li and Z. Qian, "Employing Adversarial Training to Detect Rumors on Social Media," in *2021 International Conference on Asian Language Processing, IALP 2021*, IEEE, pp. 84–89, 2021, doi: 10.1109/IALP54817.2021.9675245.

[79] H. Zhang, S. Qian, Q. Fang, and C. Xu, "Multimodal Disentangled Domain Adaption for Social Media Event Rumor Detection," *IEEE Transactions on Multimedia*, vol. 23, pp. 4441–4454, 2021, doi: 10.1109/TMM.2020.3042055.

[80] W. Chen, Y. Zhang, C. K. Yeo, C. T. Lau, and B. S. Lee, "Unsupervised rumor detection based on users' behaviors using neural networks," *Pattern Recognition Letters*, vol. 105, pp. 226–233, 2018, doi: 10.1016/j.patrec.2017.10.014.

[81] Y. Zhang, W. Chen, C. K. Yeo, C. T. Lau, and B. S. Lee, "Detecting rumors on Online Social Networks using multi-layer

autoencoder," in *2017 IEEE Technology and Engineering Management Society Conference, TEMSCON 2017*, IEEE, pp. 437–441, 2017, doi: 10.1109/TEMSCON.2017.7998415.
[82] Y. Luo, J. Ma, and C. K. Yeo, "BCMM: A novel post-based augmentation representation for early rumour detection on social media," *Pattern Recognition*, vol. 113, p. 107818, 2021, doi: 10.1016/j.patcog.2021.107818.
[83] Y. Xu, C. Wang, Z. Dan, S. Sun, and F. Dong, "Deep recurrent neural network and data filtering for rumor detection on Sina Weibo," *Symmetry*, vol. 11, no. 11, p. 1408, 2019, doi: 10.3390/sym11111408.
[84] C. Song, C. Yang, H. Chen, C. Tu, Z. Liu, and M. Sun, "CED: Credible Early Detection of Social Media Rumors," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3035–3047, 2021, doi: 10.1109/TKDE.2019.2961675.
[85] S. Tarnpradab and K. A. Hua, "Attention based neural architecture for rumor detection with author context awareness," in *2018 13th International Conference on Digital Information Management, ICDIM 2018*, IEEE, pp. 82–87, 2018, doi: 10.1109/ICDIM.2018.8847052.
[86] J. Li, P. Bao, H. Shen, and X. Li, "MiSTR: A Multiview Structural-Temporal Learning Framework for Rumor Detection," *IEEE Transactions on Big Data*, vol. 8, no. 4, pp. 1007–1019, 2022, doi: 10.1109/TBDATA.2021.3107481.
[87] Z. Luo, Q. Li, and J. Zheng, "Deep Feature Fusion for Rumor Detection on Twitter," *IEEE Access*, vol. 9, pp. 126065–126074, 2021, doi: 10.1109/ACCESS.2021.3111790.
[88] N. Zhong, G. Zhou, W. Ding, and J. Zhang, "A Rumor Detection Method Based on Multimodal Feature Fusion by a Joining Aggregation Structure," *Electronics (Switzerland)*, vol. 11, no. 19, p. 3200, 2022, doi: 10.3390/electronics11193200.
[89] F. Zeng and W. Gao, "Early Rumor Detection Using Neural Hawkes Process with a New Benchmark Dataset," in *NAACL 2022-2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 4105–4117, 2022, doi: 10.18653/v1/2022.naacl-main.302.
[90] Z. Wang, Y. Guo, J. Wang, Z. Li, and M. Tang, "Rumor Events Detection from Chinese Microblogs via Sentiments Enhancement," *IEEE Access*, vol. 7, pp. 103000–103018, 2019, doi: 10.1109/ACCESS.2019.2928044.
[91] P. Zhang, H. Ran, C. Jia, X. Li, and X. Han, "A lightweight propagation path aggregating network with neural topic model for rumor detection," *Neurocomputing*, vol. 458, pp. 468–477, 2021, doi: 10.1016/j.neucom.2021.06.062.
[92] Z. Wang, Y. Guo, Z. Li, M. Tang, T. Qi, and J. Wang, "Research on Microblog Rumor Events Detection via Dynamic Time Series Based GRU Model," in *IEEE International Conference on Communications*, IEEE, pp. 1–6, 2019, doi: 10.1109/ICC.2019.8761457.
[93] J. Ma, J. Li, W. Gao, Y. Yang, and K. F. Wong, "Improving Rumor Detection by Promoting Information Campaigns with Transformer-Based Generative Adversarial Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 3, pp. 2657–2670, 2023, doi: 10.1109/TKDE.2021.3112497.
[94] Q. Huang, J. Yu, J. Wu, and B. Wang, "Heterogeneous Graph Attention Networks for Early Detection of Rumors on Twitter," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2020, doi: 10.1109/IJCNN48605.2020.9207582.
[95] M. R. Islam, S. Muthiah, and N. Ramakrishnan, "Rumorsleuth: Joint detection of rumor veracity and user stance," in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019*, New York, NY, USA: ACM, pp. 131–136, 2019, doi: 10.1145/3341161.3342916.
[96] C. Bing, Y. Wu, F. Dong, S. Xu, X. Liu, and S. Sun, "Dual Co-Attention-Based Multi-Feature Fusion Method for Rumor Detection," *Information (Switzerland)*, vol. 13, no. 1, p. 25, 2022, doi: 10.3390/info13010025.
[97] L. Tian, X. Zhang, and J. H. Lau, "DUCK: Rumour Detection on Social Media by Modelling User and Comment Propagation Networks," in *NAACL 2022-2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 4939–4949, 2022, doi: 10.18653/v1/2022.naacl-main.364.
[98] Z. Huang, Z. Lv, X. Han, B. Li, M. Lu, and D. Li, "Social Bot-Aware Graph Neural Network for Early Rumor Detection," *Proceedings-International Conference on Computational Linguistics, COLING*, vol. 29, no. 1, pp. 6680–6690, 2022.
[99] S. Song, Y. Huang, and H. Lu, "Rumor Detection on Social Media with Out-In-Degree Graph Convolutional Networks," in *Conference Proceedings-IEEE International Conference on Systems, Man and Cybernetics*, IEEE, pp. 2395–2400, 2021, doi: 10.1109/SMC52423.2021.9659106.
[100] C. Bai, H. Wang, L. Wang, and J. Yao, "Rumor Detection Based on Bi-directional Graph Attention Network," *Proceeding-2021 China Automation Congress, CAC 2021*, pp. 1728–1733, 2021, doi: 10.1109/CAC53003.2021.9728123.
[101] J. Li, S. Ni, and H. Y. Kao, "Birds of a Feather Rumor Together? Exploring Homogeneity and Conversation Structure in Social Media for Rumor Detection," *IEEE Access*, vol. 8, pp. 212865–212875, 2020, doi: 10.1109/ACCESS.2020.3040263.
[102] Z. Lingyu, S. Chenguang, W. Bin, and W. Bai, "SMAM: Detecting rumors from microblogs with stance mining assisting task," in *Proceedings-2019 IEEE 4th International Conference on Data Science in Cyberspace, DSC 2019*, IEEE, pp. 242–249, 2019, doi: 10.1109/DSC.2019.00044.
[103] S. Dong, Z. Qian, and P. Li, "Rumor Detection with Adversarial Training and Supervised Contrastive Learning," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2022, doi: 10.1109/IJCNN55064.2022.9892819.
[104] S. Ni, J. Li, and H. Y. Kao, "HAT4RD: Hierarchical Adversarial Training for Rumor Detection in Social Media," *Sensors*, vol. 22, no. 17, p. 6652, 2022, doi: 10.3390/s22176652.
[105] H. Lin, J. Ma, L. Chen, Z. Yang, M. Cheng, and G. Chen, "Detect Rumors in Microblog Posts for Low-Resource Domains via Adversarial Contrastive Learning," in *Findings of the Association for Computational Linguistics: NAACL 2022-Findings*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 2543–2556, 2022, doi: 10.18653/v1/2022.findings-naacl.194.
[106] L. Chen *et al.*, "A Progressive Framework for Role-Aware Rumor Resolution," *Proceedings-International Conference on Computational Linguistics, COLING*, vol. 29, no. 1, pp. 2748–2758, 2022.
[107] Z. Nanjiang, Z. Guomin, D. Weijie, and Z. Jiawen, "A Rumor Detection Method Based on Multimodal Information Fusion," in *2022 IEEE 5th International Conference on Electronics Technology, ICET 2022*, IEEE, pp. 1032–1037, 2022, doi: 10.1109/ICET55676.2022.9824021.
[108] W. Zhang, T. Zhong, C. Li, K. Zhang, and F. Zhou, "CausalRD: A Causal View of Rumor Detection via Eliminating Popularity and Conformity Biases," in *Proceedings-IEEE INFOCOM*, IEEE, pp. 1369–1378, 2022, doi: 10.1109/INFOCOM48880.2022.9796678.
[109] Z. Luo, X. Zhu, Z. Qian, and P. Li, "Employing Temporal Information and Propagation Structure to Detect Rumors," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2022, doi: 10.1109/IJCNN55064.2022.9892725.
[110] M. Yan, W. Yang, B. Sun, and Y. Zhu, "Heterogeneous Graph Attention Networks with Bi-directional Information Propagation

for Rumor Detection," in *2022 7th International Conference on Big Data Analytics, ICBDA 2022*, IEEE, pp. 236–242, 2022, doi: 10.1109/ICBDA55095.2022.9760328.

[111] L. Wei, D. Hu, W. Zhou, X. Wang, and S. Hu, "Modeling the Uncertainty of Information Propagation for Rumor Detection: A Neuro-Fuzzy Approach," *IEEE transactions on neural networks and learning systems*, vol. 35, no. 2, pp. 2522–2533, 2024, doi: 10.1109/TNNLS.2022.3190348.

[112] G. Zhang, R. Liang, Z. Yu, and S. Zhang, "Rumour Detection on Social Media with Long-Tail Strategy," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2022, doi: 10.1109/IJCNN55064.2022.9892019.

[113] X. Liu, Z. Zhao, Y. Zhang, C. Liu, and F. Yang, "Social Network Rumor Detection Method Combining Dual-Attention Mechanism with Graph Convolutional Network," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 5, pp. 2350–2361, 2023, doi: 10.1109/TCSS.2022.3184745.

[114] Z. Luo, P. Li, Z. Qian, and X. Zhu, "TEH-GCN: Topic-Event Based Hierarchical Graph Convolutional Networks for Rumor Detection," in *Proceedings of the International Joint Conference on Neural Networks*, IEEE, pp. 1–8, 2023, doi: 10.1109/IJCNN54540.2023.10191237.

[115] Z. Ke, Z. Li, C. Zhou, J. Sheng, W. Silamu, and Q. Guo, "Rumor detection on social media via fused semantic information and a propagation heterogeneous graph," *Symmetry*, vol. 12, no. 11, pp. 1–14, 2020, doi: 10.3390/sym12111806.

[116] C. M. M. Kotteti, X. Dong, and L. Qian, "Rumor Detection on Time-Series of Tweets via Deep Learning," in *Proceedings-IEEE Military Communications Conference MILCOM*, IEEE, pp. 1–7, 2019, doi: 10.1109/MILCOM47813.2019.9020895.

[117] L. Poddar, W. Hsu, M. L. Lee, and S. Subramaniyam, "Predicting stances in twitter conversations for detecting veracity of rumors: A neural approach," in *Proceedings-International Conference on Tools with Artificial Intelligence, ICTAI*, IEEE, pp. 65–72, 2018, doi: 10.1109/ICTAI.2018.00021.

[118] X. Chen, D. Zhu, D. Lin, and D. Cao, "Rumor knowledge embedding based data augmentation for imbalanced rumor detection," *Information Sciences*, vol. 580, pp. 352–370, 2021, doi: 10.1016/j.ins.2021.08.059.

[119] B. Liu *et al.*, "Nowhere to Hide: Online Rumor Detection Based on Retweeting Graph Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2022, doi: 10.1109/TNNLS.2022.3161697.

[120] Z. Zhou, Y. Qi, Z. Liu, C. Yu, and Z. Wei, "A C-GRU Neural Network for Rumors Detection," in *Proceedings of 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems, CCIS 2018*, IEEE, pp. 704–708, 2019, doi: 10.1109/CCIS.2018.8691263.

BIOGRAPHIES OF AUTHORS

Lifan Zhang received the Master's degree in Computer Software and Theory from the School of Computer and Information Technology, Shanxi University, Taiyuan, China, in 2013. She is currently pursuing the Ph.D. degree in Computer Science from the College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Shah Alam, Malaysia. Her research interest includes text classification, rumor detection in the field of natural language processing, and artificial intelligence. She can be contacted at email: zhanglifan@xmphdsss.cn.

Shafaf Ibrahim is an associate professor in College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Shah Alam, Malaysia. She holds a Diploma, Bachelor's degree, Masters, and Ph.D. in Computer Science. Her research interests are artificial intelligence, evolutionary algorithms, DL, machine learning, and image processing. She can be contacted at email: shafaf2429@uitm,edu.my.

Ahmad Firdaus Ahmad Fadzil is a senior lecturer in College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Malacca Campus, Malaysia. He holds a Diploma, Bachelor's Degree, Masters, and Ph.D. in Computer Science. His research interests are parallel processing, evolutionary computing, DL, and image processing. He can be contacted at email: firdausfadzil@uitm.edu.my.