



Evaluating the effectiveness of publishers' features in fake news detection on social media

Ali Jarrahi¹ · Leila Safari¹

Received: 20 October 2021 / Revised: 21 December 2021 / Accepted: 21 February 2022 /
Published online: 11 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

With the expansion of the Internet and attractive social media infrastructures, people prefer to follow the news through these media. Despite the many advantages of these media in the news field, the lack of control and verification mechanism has led to the spread of fake news as one of the most critical threats to democracy, economy, journalism, health, and freedom of expression. So, designing and using efficient automated methods to detect fake news on social media has become a significant challenge. One of the most relevant entities in determining the authenticity of a news statement on social media is its publishers. This paper examines the publishers' features in detecting fake news on social media, including *Credibility*, *Influence*, *Sociality*, *Validity*, and *Lifetime*. In this regard, we propose an algorithm, namely *CreditRank*, for evaluating publishers' credibility on social networks. We also suggest a high accurate multi-modal framework, namely FR-Detect, for fake news detection using user-related and content-related features. Furthermore, a sentence-level convolutional neural network is provided to properly combine publishers' features with latent textual content features. Experimental results show that the publishers' features can improve the performance of content-based models by up to 16% and 31% in accuracy and F1, respectively. Also, the behavior of publishers in different news domains has been statistically studied and analyzed.

Keywords Fake news detection · CreditRank algorithm · Social media · Deep neural network · Machine learning · Text classification

✉ Ali Jarrahi
jarrahi@znu.ac.ir

Leila Safari
lsafari@znu.ac.ir

¹ Electrical and Computer Engineering, University of Zanjan, Zanjan, Iran

1 Introduction

Nowadays, the Internet has made up a significant part of the human lifestyle. The role of traditional news channels, such as newspapers and television, has been diminished and weakened dramatically in the news reception. In particular, the expansion of social media infrastructures, such as Facebook and Twitter, has had a significant role in undermining traditional media. People use social media to connect with friends, relatives and gather information and news worldwide. The reason for this behavior can be traced back to the nature of these media. First, it is much faster and less expensive to get the news through these media than traditional media. Second, it is easy to share the news with friends and other people for further discussions. As of August 2018, around 68% of Americans received the news via social media, compared to 62% in 2016 and 49% in 2012.¹

However, these benefits of social media are not costless. The lack of control and verification of the news releases has made social media a fertile ground for disseminating false or unverified information [71]. An attractive news headline is often enough for an article to be shared thousands of times despite its inaccurate or unapproved content.

Fake news is not a new phenomenon. Before the advent of the Internet, journalists investigated and verified their news and sources [43]. However, the impact of fake news on public opinion was minimal and therefore insignificant. Today, with the expansion of social media, the spread of inaccurate or unverified information among many people, regardless of geographical boundaries, has been facilitated. As a result, public perceptions of events can be profoundly affected by fake news [71]. The 2016 US Presidential Election is one of the prominent examples of the impact of spreading fake news [1].

Fake news is now recognized as one of the most significant threats to democracy, journalism, health, and freedom of expression, which can even undermine public confidence in governments [68]. The economy is also not immune to the spread of fake news. Significant fluctuations occur with the propagation of fake news related to the stock market [40]. The importance of fake news has led to the term “fake news” being chosen as the word of the year by Macquarie and Oxford dictionaries in 2016.

Social and psychological factors play an essential role in gaining public trust and spreading fake news. For example, it has been shown that when humans are overly exposed to misleading information, they become vulnerable and irrational in recognizing truth and falsehood [6]. Studies in social and communication psychology have also shown that human ability to detect deception is slightly better than chance, with a mean accuracy of 54% obtained over 1000 participants in over 100 experiments [42]. This situation is more critical for fake news because of its unique features. Therefore, it is crucial to provide methods for the automatic detection of fake news on social media.

The most critical challenges in fake news detection are accuracy and early detection. In general, models for automatically detecting fake news on social media can take advantage of news content or social context data. Utilizing the right combination of these data types is essential to meet the challenges because each data has its strengths and weaknesses. Despite the usefulness of social context data in improving the accuracy of methods, many of them cause considerable delays in detection. So, the proper use of social context data and news content remains a significant challenge. One of the most relevant entities in determining the authenticity of a news statement in the real world is its narrator. Therefore, news publishers on

¹ <https://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/>

social media can be considered and studied as the most relevant entities in fake news detection. Another advantage of using publisher-related data is that it does not delay detection. So, the primary objective of this paper is to investigate the effectiveness of publishers' features in detecting fake news on social media. For this purpose, the most important features related to news publishers on social media and their relevant algorithms have been introduced. Furthermore, a sentence-level convolutional neural network is provided to combine these features with latent textual content features properly. Table 1 includes symbols used throughout the paper to assist researchers when encountering issues due to symbols. The novelties of the paper are as follows:

- A comprehensive study of publisher-related features from different aspects to evaluate their applicability and effectiveness in detecting fake news on social media
- Development of an algorithm (CreditRank) to assess the credibility of publishers (as a complex feature) on social media
- Development of a novel CNN with 3D input (SLCNN) for text classification, which allows simultaneous learning at the word and sentence level; it also enables developers to integrate additional features at the sentence level
- Provide an efficient multi-modal framework (FR-Detect) for detecting fake news on social media utilizing news content and publishers' features with early detection capability and state-of-the-art results

The rest of paper is structured as follows. The related concepts for studying fake news on social media are presented in the next section. The previous works have been summarized in Section 3. The details of the proposed methods are described in Section 4. We have evaluated our approach on a comprehensive fake news detection benchmark dataset. The experimental results are presented in Section 5. Finally, the paper concludes with future research directions in Section 6.

2 Fake news on social media

This section provides concepts and definitions related to fake news on social media to give readers and researchers a better understanding of its features. Although there is no comprehensive definition for fake news [68], a clear definition can help distinguish related concepts and better analyze and evaluate fake news. The definition of *news* in the Oxford Dictionary is as follows: new information about something that has happened recently. In social media, the most related concept to fake news is rumor. A *rumor* is an unverified claim or information created by users on social media and can potentially spread beyond their private network [7]. This unverified information could be accurate, partly accurate, completely false, or even unverified [71]. Similar to fake news, spreading false rumors can cause severe damages, even in a short time.

Researchers in [68] have distinguished related terms and concepts, like rumor and satire news, based on three characteristics: Authenticity (false or not), Intention (bad or not), and Type of information (news or not). For example, a rumor is a piece of information that all these characteristics are *unknown*. In contrast, fake news is false news presented with a bad intention to mislead the general public or a particular group. So, fake news can be defined as follows: fake news is *intentionally* and *verifiably* false news published by a news outlet [48, 68].

According to the definitions and the characteristics provided, the relationship between the concepts of news, fake news, and rumors can be considered as Fig. 1.

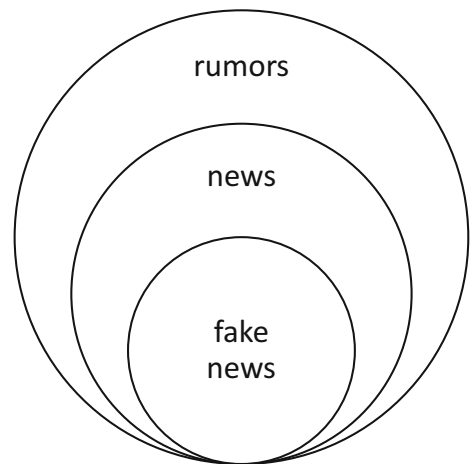
In addition to definitions, determining the life cycle of fake news and its related components in social media is essential for the proper study of fake news in this context. Zhou et al. [68] have considered the life cycle of fake news based on three stages: creation, publication, and propagation. However, given that fake news is verifiable, we believe there is a detection stage in the life cycle, and eventually, all fake news is detected. Therefore, we have modified the life cycle of fake news, as shown in Fig. 2. Each stage of the life cycle is described below.

Creation At this stage, fake news content is created by one or more authors for specific purposes. Creating fake news can be done in the context of social media or outside. The main parts of the news include the headline and the body. Other optional sections may include images, authors, and news sources.

Table 1 The table contains the symbols used in this paper

Symbol	Meaning
ML	Machine Learning
DL	Deep Learning
NLP	Natural Language Processing
CNN	Convolutional Neural Network
SLCNN	Sentence-Level Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRUs	Gated Recurrent Units
HAN	Hierarchical Attention Network
3HAN	Three-level Hierarchical Attention Network
BERT	Bidirectional Encoder Representations from Transformers
POS	Part of Speech
LIWC	Linguistic Inquiry and Word Count
RR	Rhetorical Relations
RST	Rhetorical Structure Theory
BoW	Bag-of-Words
TF	Term Frequency
TF-IDF	Term Frequency-Inverse Document Frequency
ReLU	Rectified Linear Unit
NLTK	Natural Language Toolkit
OOV	Out-Of-Vocabulary
HCN	Horizontal Convolutional Block
FR-Detect	Fake-Real Detector
dEFEND	Explainable Fake News Detection
TCNN-URG	Two Level Convolutional Neural Network with User Response Generator
SAFE	Similarity-Aware Fake news detection
OPCNN-FAKE	Optimized Convolutional Neural Network model to detect fake news
T_s	The threshold for the number of words in the sentences
T_d	The threshold for the number of sentences in the news body
PTN	The total number of news published by the publisher
PFN	The number of fake news published by the publisher
PCR	The credibility rank of the publisher on the social network
NTN	The average number of news published by publishers of the news
NFN	The average number of fake news published by publishers of the news
NCR	The average credibility rank of publishers of the news
UI	The influence of the user
NI	The average influence of publishers of the news
numP	The number of publishers of the news

Fig. 1 The relationship between the concepts of news, fake news, and rumors



Publication After creating fake news, one or more publishers must inject the created news into social media. Here, the publisher is a user of that social media. Each user on social media has a specific identity that can be defined through features such as friends, followers, history of activities, etc. The followers of each publisher primarily receive the published news on social media. This stage is called the publication phase.

Propagation After the publication stage, each news article enters a phase that depends entirely on the recipients' behavior. After receiving the news, each recipient may share, comment, or like the news or leave it without any action. In general, the news recipients can be divided into three categories:

- **Malicious User:** A user who intentionally endorses and shares fake news for specific purposes while being aware that the news is fake.
- **Conscious User:** A user who carefully tries to avoid sharing fake or suspicious news as much as possible.
- **Naïve User:** A user who unintentionally shares fake news due to the deception of malicious users and social effects. Naïve users participate in the fake news propagation process because of their prior knowledge (as expressed by *confirmation bias*² [33]) or the peer pressure (as explained by the *bandwagon effect*³ [27]).

After some news recipients share the fake news, their followers also receive fake news, and this process continues. This stage is called the propagation phase.

Detection As stated in the fake news definition, the authenticity of the news can be verified using existing evidence, and therefore, its falsity can be detected. Of course, it will take a while to determine if the news is fake. The longer this period lasts, the more people on social media will be affected by fake news. Therefore, the detection must be

² Individuals tend to trust information that confirms their preexisting beliefs or hypotheses.

³ Individuals do something primarily because others are doing it.

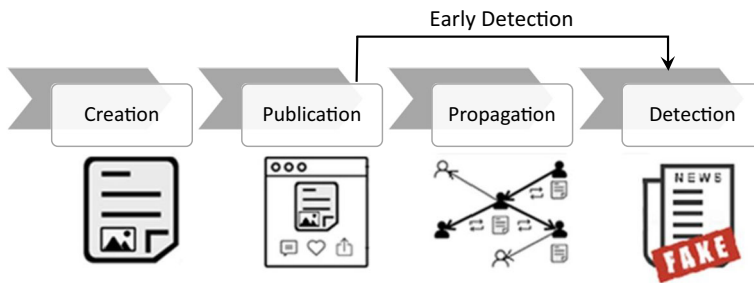


Fig. 2 The life cycle of fake news on social media

made as soon as possible (ideally before the propagation stage, as shown in Fig. 2), known as early detection in the fake news field. After the news is detected as fake, the propagation phase ends.

The process of spreading fake news on social media is summarized in Fig. 3, and an example of fake news on Facebook is shown in Fig. 4. Given the process and its components, there are helpful features to help fake news detection. As summarized in Fig. 5, these features can be divided into four general categories, described below.

Content-related features Some features are directly related to the news content. Structurally, a news story includes headline, body, image(s), source, and author(s). Each of these parts or the relationships between them may contain useful features that can be extracted and utilized.

Writing style features can be used to determine the author's intent (bad or not) [68]. These features can be extracted based on existing theories, such as the complexity of the text (e.g., the average number of words in sentences) and the features that measure the sentiment of the text (e.g., the amount of positive and negative words), Or features extracted from the structure of the text, e.g., *bigram* [38], *POS* (Part of Speech) [69], *LIWC* (Linguistic Inquiry and Word Count) [28, 38] and *RR* (Rhetorical Relations) [44].

Regarding the writing style features, it is important to note that fake news is generally about important events with financial or political benefits. Therefore, its authors are so motivated to write the news in such a way that it is not detectable by current fake news detection methods. Therefore, developing a real-time representation and learning of writing style features is essential. Deep learning methods can help extract the news content's latent features. Therefore, current writing style-based fake news detection methods mainly rely on deep learning techniques [53, 56].

Other news content features include image-related features such as image forgery and how the image relates to the news body. Another feature is the headline credibility and its relevance to the news body, which is similar to the *clickbait* recognition problem. Authors' credibility, as well as news sources, can also help detect fake news. Analyzing fake news content is not sufficient to create an effective and reliable identification system. So, other important aspects, such as the social context information of news, should also be considered [66].

User-related features Regardless of name or account, a social media user is an identity associated with a human or robot interacting with other users and components in social media. Users have significant features that can be used in fake news detection. Some of these features are listed below:

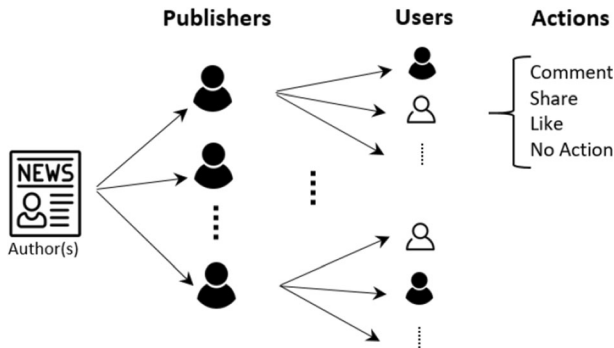


Fig. 3 The process of spreading fake news on social media. After the news was created by the authors, some publishers started publishing the news on social media, which led to actions by followers and users

- *Validity*: This feature indicates whether the user matches the original identity associated with him/her in the real world or not. In some social media, it is known as the *blue verified badge*.
- *Lifetime*: This feature indicates the time elapsed since the creation of the user on social media.
- *Influence*: This feature indicates the average impact of the news published by the user on social media. In other words, how many social media users receive the news published by this user on average? This feature can easily be considered equal to the number of followers, although the influence of each follower can also be significant in determining the user’s influence.
- *Sociality*: This feature shows how much the user interacts with other users. It can be considered equivalent to the number of friends.
- *Partisan bias*: This feature indicates the user’s political orientation.



Fig. 4 An example of fake news on Facebook

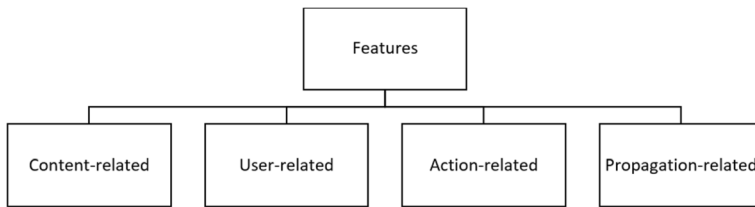


Fig. 5 Types of features available in the fake news life cycle on social media

- *Activity credibility*: In the news field, this feature indicates how much of the news published by that user was fake or real. This feature can be calculated from the user's activity history on social media.
- *Activity level*: This feature indicates the amount of user activity (such as comments, shares, and likes) on the received news.

Propagation-related features These features determine how the news propagates on social media. There are different patterns in spreading fake and real news on social media [57]. So, by extracting features related to propagation patterns, such as depth and level in the fake news cascade [68], we can estimate the possibility that the news is fake.

Action-related features Some other features are related to the actions performed on received news by the users. For example, the liking rate or the comments polarity of a news article can provide helpful information about the authenticity of the news. To use these features effectively, it is necessary to consider the credibility of the user who created the action, because, for example, positive polarity in a comment can create different meanings depending on the user's credibility.

Using these features, the issue of fake news detection can be considered a classification problem. According to the availability of the content and user-related features at the publication stage, utilizing these features does not delay the detection. In contrast, propagation and action-related features require time to be created, resulting in delayed detection.

3 Related works

This section provides a brief review of research on fake news detection. Fake news detection methods generally use news content and/or social context information. News content features can be extracted from text, images, and news sources such as authors and websites writing or publishing the news. News textual information can be used to extract features related to writing style at different language levels [41], i.e., lexicon-level [38, 60, 67, 69], syntax-level [69], semantic-level [38] and discourse-level [24]. These features can be explicitly obtained using methods like n-grams [38], Bag-of-Words (BoWs) [69], Part-Of-Speeches (POSS) [69], Linguistic Inquiry and Word Count (LIWC) [28, 38], Rhetorical Structure Theory (RST) [44], etc.; or implicitly using deep neural networks with word embedding (for example word2vec [29]) to extract appropriate latent features that have shown good performance [21, 24, 34, 53]. One of the most important networks in the text classification area is the Hierarchical Attention Network (HAN) [63]. In this network, which is based on Gated

Recurrent Units (GRUs), two levels of attention are used at the word and sentence levels. Signhania et al. [53] have provided a version of HAN, called 3HAN, specifically for detecting fake news, in which a layer of attention has been added at the Headline-Body level. Recently, convolutional neural networks (CNNs) have been successfully utilized in fake news detection [14, 21, 46]. Visual features extracted from visual elements such as images and videos have also been used alongside textual features to detect fake news [52, 60, 64]. Zhou et al. [70] used the relationship (similarity) between the textual and visual information in news articles to predict authenticity. Sitaula et al. [54] evaluated the credibility of the news using authors and content, and Baly et al. [3] detected fake news by their source websites. Also, a deep diffusive network model has been used to simultaneously learn the representations of news articles, creators and subjects [67]. Recently, hybrid deep learning models have been considered in various fields [62]. A hybrid CNN-RNN based deep learning is also proposed for fake news detection [32].

Moreover, the use of social context information to detect fake news has recently become very attractive [50]. For example, Vosoughi et al. [57] have shown that fake news spreads faster, farther, and more widely than true news. Utilizing user comments to detect fake news has recently been considered as well. For example, Cui et al. [11] applied user comments to identify important sentences in the news body. However, since the use of user comments causes delays in detecting fake news, recent research has focused on the issue of early detection by, for example, adversarial learning [60] and user response generating [39], and unsupervised detection [17, 65]. Other social context information, like user profiles [49] and social connections [45], have also been used. Using the information of neighbors is common in many algorithms in computer science. For instance, [4] presents an algorithm for link prediction based on mutual influence nodes and their neighbors. A similar idea is considered in current research to compute scores to show the credibility and influence of publishers in spreading fake news based on their followers' information on social media. Sentiment analysis has also been applied to detect fake news [5, 10] and rumors [59].

Authors in [20] proposed a Recurrent Neural Network with an attention mechanism (att-RNN) to combine multi-modal features for rumor detection. This network incorporates image features into the joint text and social context features, obtained with an LSTM network, to create a reliable fused classifier. The neural attention from the outputs of the LSTM is used when fusing with the visual features.

DeepFakE [22] uses the news content and the presence of echo chambers (community of social media-based users with similar views) on a social network to detect fake news. The correlation between user-profiles and news articles is formed as a tensor by combining news, user, and community information. The news content is merged with the tensor, and coupled matrix-tensor factorization is used to represent news content and social context. Factors obtained after decomposition were used as features to the news classification. A deep neural network model is utilized for classification.

Authors in [36] aim to present an insight into the characterization of news text, together with the differential content types of the news story and its effect on readers. Existing text-based fake news detection techniques and several fake news datasets, together with four critical open research challenges, are provided in this survey. The challenges in fake news detection mainly focus on incomplete multi-modal datasets (not having datasets with full features), need to multi-modal verification methods (in addition to the text, images, audio, embedded content, and hyperlinks have also been considered), considering the source of news in evaluating fake news stories, and also author's credibility.

Authors in [13] have provided a review of trends and challenges on fake news detection. The main focus of this survey was the definitions of fake news, the traditional methods for identification, the available datasets, and their features to characterize the fake news. In addition, the primary methods for converting natural language text into vectors to be used in fake news detection and the research opportunities and initiatives on fake news detection are considered in this paper. Also, the main challenges, including the circulation of fake news on multilingual platforms, large volumes of real-time unlabeled data, complex and dynamic network structure, and early detection of rumors, are explained in the paper.

A deep neural network architecture [31] is proposed for fake news detection on Twitter data, allowing various input modes, including the word embeddings of both news headers and bodies, linguistic features, and network account features (user profiles). It lets the fusion of input at various network layers. One significant contribution of this work is developing a new Twitter data set with real/fake news regarding the Hong Kong protests.

FakeBert [23] proposes a BERT-based (Bidirectional Encoder Representations from Transformers) approach. Bert is used for context representation or generating sentence embedding vectors. The generated vectors were then fed to three parallel blocks of the single-layer CNN, followed by concatenation, convolution, dense, and flatten layers. Due to the transformer-based nature of BERT, their proposed model outperformed other models like LSTM, CNN, and classical machine learning models that used Glove/Word2Vec for context representation. Only context features are used in the paper, and other features like user credibility and news proliferation methods are not considered in this work. Similarly, BerConvoNet [9] has used BERT for contextual representation of news text which was then fed to a multi-scale feature block that consists of multiple kernels of varying sizes and aims to extract various features from word embeddings followed by a fully connected layer for classification. Utilizing the BERT transformer model in BerConvoNet, word tokens of the input sentence and position and segment embeddings corresponding to the input tokens were used to represent the input sentences.

Authors in [19] have reported the performance of five ML (Machine Learning) models and three DL (Deep Learning) models on two datasets with different sizes. TF and TF-IDF were used as tokenization methods for ML-based models, and embedding techniques were used to obtain text representation for deep learning models. Using McNemar's test, they evaluated the significance of the difference between the performance results of all models. They proposed a stacking method based on training another Random Forest model using the prediction results of all individual models.

A linguistic model [8] is suggested to find out content features, mainly syntactic, grammatical, sentimental, and readability features of news text, then used in a neural-based sequential learning model for fake news detection. Similarly, Hakak et al. have proposed an ensemble classification model for fake news detection based on linguistic features [15]. They extracted 26 linguistic features from text which were then fed into the ensemble model of Decision Tree, Random Forest, and Extra Tree Classifier.

In addition, as mentioned earlier, the spread of fake news has a huge impact on various aspects of today's life. Significantly since the outbreak of (COVID-19) in the last two years, the proliferation of false news concerning coronavirus disease has increased on social media [2, 18]. As a result, in addition to the political and social aspects, fake news propagation has also affected public health. So, the research on effective fake news detection techniques and various theoretical aspects of fake news is growing very fast.

Varma et al. [55] survey the existing machine learning-based and deep learning-based fake news detection techniques, pre and post corona pandemic. Available databases, pre-processing steps, feature extraction approaches, and evaluation criteria for current fake news identification techniques have been studied in this work. The authors mentioned that the ML algorithms like Naive Bayes, support vector machine, and logistic regression are the most successful solutions for fake news detection; however, the solutions are shifting toward the use of ensemble approaches like random forest and DL based approaches. Especially following the COVID-19 pandemic, researchers primarily focus on building hybrid ensemble models and using both text and author features extracted manually for the ML-based techniques or automatically for DL algorithms. By the way, the study could not establish a universal methodology for successful fake news detection.

In this paper, we examine the effectiveness of publishers' features in detecting fake news on social media, including credibility, as a complex feature and suggest a high accurate multi-modal framework with early detection capability.

4 The proposed framework

This section introduces our proposed method, namely FR-Detect (**F**ake-**R**eal **D**etector), to detect fake news on social media before the propagation stage. As illustrated in Fig. 6, the method uses content-related and publisher-related features simultaneously to improve the overall performance. Among the publisher-related features that we introduced in the previous section, the following features are considered for evaluation: *Credibility*, *Influence*, *Sociality*, *Validity*, and *Lifetime*. As shown in the figure, the framework consists of three main parts, including *Feature Extractors*, *Integrator*, and *Classifier*, described in the following subsections.

4.1 Feature extractors

To evaluate the role of publishers' features in fake news detection, introduced features and their combination are considered along with a basic content-based model to measure their effectiveness. So, a proper latent linguistic features extractor has been designed to combine features efficiently. Each of the feature extraction modules is described below.

4.1.1 Latent linguistic features extractor

Due to the importance of real-time representation and learning of content-related features in the scope of fake news detection, this part is designed based on deep learning methods. CNN is commonly applied to analyze visual imagery [61]. These networks aim to extract local features from the input tensors of images for image classifications. However, CNNs are also gaining popularity in other areas like the NLP techniques. A convolutional neural network consists of an input layer, hidden layers, and an output layer. Middle layers are called hidden in any feed-forward neural network because the activation function and final convolution mask their inputs and outputs. The hidden layers in a CNN include layers that perform convolutions. Generally, this includes a layer that performs a dot product of the convolution filter (or kernel) with the layer's input matrix. This product is usually the Frobenius inner product, and its activation function is commonly Rectified Linear Unit (ReLU), $f(x) = \max(0, x)$. As the convolution filter slides along the input matrix for the layer, the convolution operation generates a feature

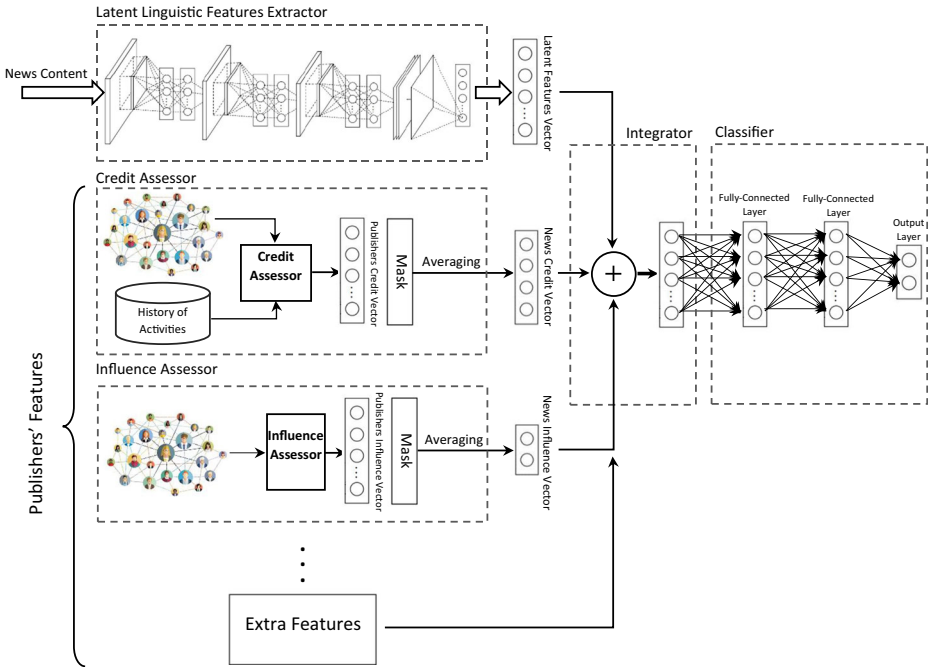


Fig. 6 The framework of FR-Detect. The features of news content and publishers are extracted and efficiently integrated, and used in the learning process

map, which contributes to the input of the next layer. This is followed by other layers such as pooling, fully-connected, and normalization layers. In this research, we have designed a novel sentence-level convolutional neural network (SLCNN). In this network, the news headline and body are transformed into a three-dimensional (3D) tensor, illustrated in Fig. 7. As shown in the figure, the headline and sentences of the body form the first dimension of the tensor. In the same way, the words of the sentences shape the second dimension, while the third dimension represents the word vectors of the words. The pre-trained word embedding, e.g., *word2vec* [29] or *GloVe* [37], could be used for representing the word vectors.

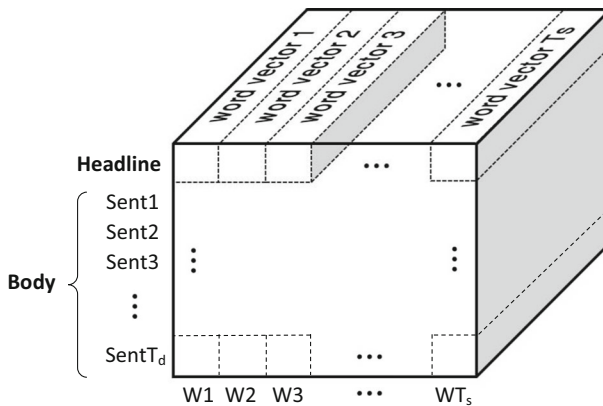


Fig. 7 Shape of the transformed news content. One dimension represents the sentences of news body, and the other is the words of the sentences, while the third dimension is related to the word vectors of the words

Since the input size of the network must be fixed, two thresholds are considered to adjust the different sizes of both texts and sentences (one for the number of sentences in the texts, T_d , and the other for the number of words in the sentences, T_s). The texts and the sentences longer than the thresholds would be cropped, and shorter ones would be padded by zeros.

After some statistical analysis on the datasets in our experiments and considering the structure of the SLCNN, we chose $T_s = 46$ (about 2% of sentences have more than 46 words). In the same way, the threshold for the number of sentences in the news body is calculated by the following equation:

$$T_d = \lceil \mu + \sigma \rceil \tag{1}$$

where μ is the average number of sentences in the news body, and σ is the standard deviation. As a result, the performance of the model is significantly improved by ignoring the outlier sizes and preventing the construction of very large and sparse tensors. To better understand, the distribution of the number of sentences and their number of words in a news dataset is plotted in Fig. 8. As a result of applying thresholds, the size of the 3D tensor is dropped from $1881 \times 4119 \times$ (the size of word vectors) to $85 \times 46 \times$ (the size of word vectors), i.e., more than 99% reduction. The reduction rate is almost the same for different datasets.

The architecture of the SLCNN is illustrated in Fig. 9. Overall, the news articles are provided in the shape of the introduced 3D tensor for the input layer. Then, using four horizontal convolutional blocks (HCB), one feature vector is extracted for each sentence individually. The main advantages of the SLCNN over traditional CNN for text classification [25] are: 1) the positional information of the sentences (sent1, sent2, ..., sent n) is used in the learning process. In other words, the role and importance of each sentence in the falsity of the news is also learned, and 2) the SLCNN enables us to combine other extra features at the sentence level.

Looking at the details of the HCB, as shown in Fig. 10, there are two sequential convolution layers, each one followed by a ReLU activation function. A convolution operation consists of a filter $w \in \mathbb{R}^s \times t \times d$, which is applied to each possible window of $s \times t$ features from its input feature map, X , to produce a new feature map by Eq. 3:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix} \tag{2}$$

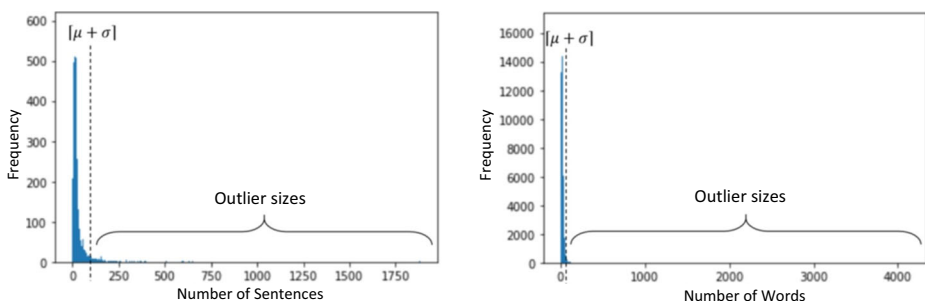


Fig. 8 The distribution of the number of sentences in a news dataset and the number of words in the sentences

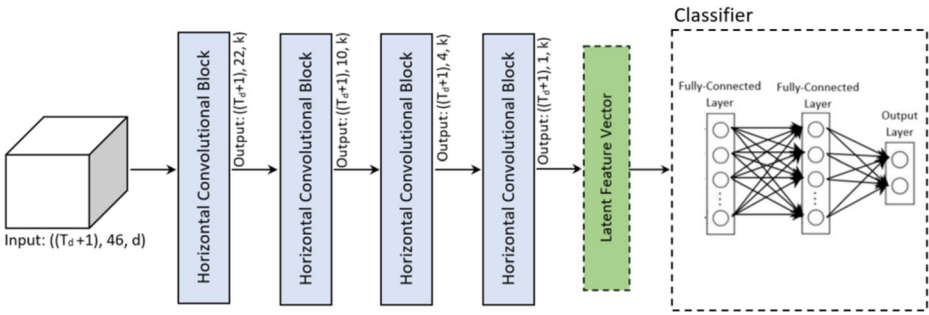


Fig. 9 The architecture of the SLCNN. d is the size of word vectors and k is the number of filters

$$\tilde{x}_{i,j} = f(w \cdot x_{i,j:i+s-1,j+t-1} + b) \tag{3}$$

where $x_{i,j:y,z}$ is the concatenation of features within the specified interval, $b \in \mathbb{R}$ is a bias term, and f is a non-linear function such as the ReLU. For our purpose, we consider $s = 1$ and $t = 2$. In the first convolution layer of the first HCB, d (the third dimension of the filters) is equal to the size of the word vectors, and in other cases, $d = 1$. At the end of the blocks, a max-pooling operation, with the pooling size = 2, is applied over the generated intermediate feature map to select the maximum value from any two adjacent features as a more important feature. The new feature map is calculated by the following equation:

$$\tilde{x}_{i,j} = \max\{x_{i,2j-1}, x_{i,2j}\} \tag{4}$$

The process of extracting one feature from one filter was described. The model uses multiple filters to obtain multiple features. The final extracted features are passed to the fully connected layers (the Classifier) that end to a *softmax* output layer which is the probability distribution over labels.

4.1.2 Publishers’ features extractor

Since this paper aims to evaluate the effectiveness of publishers’ features in fake news detection, several modules are required to extract them. Some of these features, such as

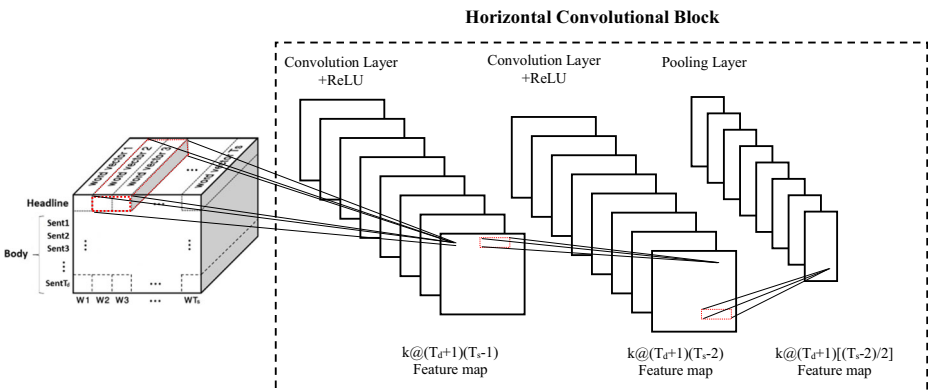


Fig. 10 The horizontal convolutional blocks. k is the number of filters. The size of the filters for the first convolution layer of the first HCB is equal to $1 \times 2 \times$ (the size of word vectors), and in other cases, it is 1×2

Validity, *Lifetime*, and *Sociality*, can be easily extracted through user profiles, but others, i.e., *Credibility* and *Influence*, require some calculations. So, we have developed algorithms for these purposes, which are described below.

Credit assessor Due to the importance of publishers' credibility in determining the authenticity of the news, this module is responsible for calculating the *news credit vector* based on its publishers' credibility. Since credible people generally follow credible people, the publishers' credibility can be studied from two aspects: 1) their history in publishing news, 2) their credit rank on the social network. Unlike the activity history, the credit rank on the social network cannot be manipulated by publishers. So, it is essential to consider credit rank in the algorithm. Therefore, the calculated credit will be more reliable for each publisher. As shown in Fig. 6, the *Credit Assessor* module determines the credibility of publishers by considering both of these aspects. Figure 11 shows the *CreditRank* algorithm that we have developed for this purpose. As shown, the algorithm generates a triple vector (*PTN*, *PFN*, *PCR*) for each publisher called the *publisher credit vector*, which *PTN* is the total number of news published by the publisher, *PFN* is the number of fake news published by the publisher, and *PCR* is credibility rank of the publisher on the social network. Then, the mask function selects the relevant publishers for the news article and creates the *news credit vector* (*NTN*, *NFN*, *NCR*, *numP*) by averaging, where *NTN* is the average number of news published by the news publishers, *NFN* is the average number of fake news published by the news publishers, *NCR* is the average credibility rank of the publishers and *numP* is the number of the news publishers. All the values are normalized by min-max normalization.

In the *CreditRank* algorithm, which is inspired by the PageRank algorithm [35], publishers' credibility is initialized by their activity history. Then, it is updated in several iterations based on the credibility of its followers. Since the credibility of publishers with more followers is more reliable and valuable, the effect of the credibility of each follower is considered in proportion to the number of its followers. As shown in the algorithm, two parameters must be specified according to the application: 1) *iteration*, which indicates how many levels of followers should be considered. This amount should not be more than the diameter of the social network, and 2) $0 \leq \alpha \leq 1$, which determines how much the publishers' credibility depends on their activity history and how much it depends on the credibility of their followers. The closer this value is to 1, the less the followers' credibility is considered.

Algorithm *CreditRank*

```

for each  $P_i$  in Publishers on social media do
  Create a triple vector ( $PTN_i, PFN_i, PCR_i$ )
   $PTN_i \leftarrow$  total number of news published by  $P_i$ 
   $PFN_i \leftarrow$  total number of fake news published by  $P_i$ 
   $PCR_i \leftarrow (PTN_i - PFN_i) / PTN_i$ 
end for
while  $iteration > 0$  do
  for each  $P_i$  in Publishers on social media do
     $PCR_i \leftarrow \alpha PCR_i + (1 - \alpha) \left( \sum_{k \in f(i)} \frac{|f(i)|}{|\bigcup_{j \in f(i)} f(j)|} PCR_k \right)$  //  $f(i)$  is set of followers  $P_i$ 
  end for
   $iteration \leftarrow iteration - 1$ 
end while
end Algorithm

```

Fig. 11 The *CreditRank* Algorithm. The algorithm creates a triple credit vector for each publisher on social media

Influence assessor As mentioned before, another important feature of the news publishers on social media is their reputation or influence. It means the news published by a more famous publisher can affect more users on social media. This feature also seems to help detect fake news. By providing a definition and calculation formula for the publishers' influence on social media, its usefulness in detecting fake news has been investigated in the FR-Detect framework.

Definition (User influence on social media): user influence is the average impact of the news published by the user on social media.

According to the definition, the user's influence on social media equals the average ratio of users receiving the news published by that user. Considering an example of a social network, shown in Fig. 12, we propose the following equation to calculate a user's influence on social media:

$$UI(u) = \frac{1}{N-1} \left(|f_1(u)| + \sum_{i=2}^d p^{i-1} \left| f_i(u) - \bigcup_{j=1}^{i-1} f_j(u) \right| \right) \quad (5)$$

Where N is the total number of users on social media, d is the diameter of the social network, p is the average probability of sharing news by users, and $f_i(u)$ is the set of the level i followers of user u on the network, which is calculated by the following equation:

$$f_i(u) = \begin{cases} \text{set of followers of } u, & i = 1 \\ \bigcup_{x \in f_{i-1}(u)} f_1(x) & , i \geq 2 \end{cases} \quad (6)$$

According to Eq. 5, the first-level followers receive the news published by the publisher directly. Whereas the second-level followers receive the news if the recipient of the previous level share/retweet it with a probability of p . The same goes for higher levels.

For simplicity, the influence of users can be estimated by the number of followers. As shown in Fig. 6, after calculating the users' influence (UI), the mask function selects the relevant publishers for the news article. Then, it creates the news influence vector (NI , $numP$) by averaging, where NI is the average influence of the news publishers and $numP$ is the number of the news publishers. All the values are normalized by min-max normalization.

4.2 Integrator

Once the desired features are ready, they must be integrated to enter the classifier. As shown in Fig. 13, the integrator concatenates features of the news publishers to the latent linguistic features at the sentence level. Then, using the correct number of HCBs, one new feature vector with size k (k is equal to the number of filters) is extracted for each row of the feature map. Finally, the final integrated feature vector is prepared by flattening the vectors and sent to the classifier.

4.3 Classifier

Once the integrator integrates the required features, the Classifier is ready for learning and classifying the news articles based on provided features. This module includes two hidden fully-connected layers that end to a *softmax* output layer for classification. For regularization, a dropout module [16] is employed after each fully connected layer.

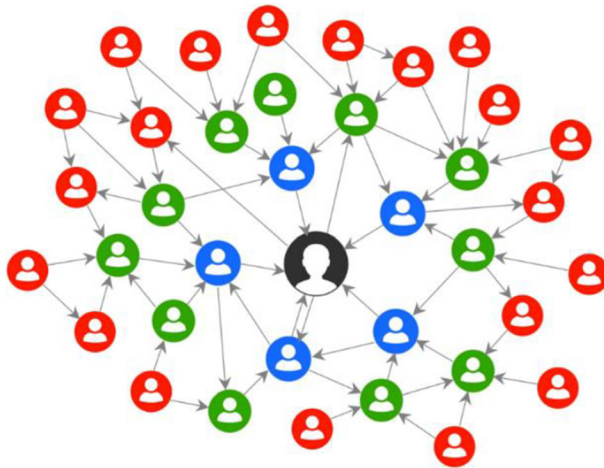


Fig. 12 An example of a social network. Each arrow from A to B means A follows B on social media. The blue users are first-level followers, the green users are second-level followers, and the red users are third-level followers of the black user

5 Experiments

5.1 Experimental settings

In this section, we introduce the settings used in our experiments. The proposed framework is implemented in python with Keras.⁴ For the SLCNN, the Natural Language Toolkit (NLTK)⁵ was used to tokenize words and sentences. As mentioned before, a pre-trained word-embedding is used in the input layer to convert the words into the corresponding word vectors. The 100-dimensional GloVe vectors have been used in our experiments. Out-Of-Vocabulary (OOV) words are initialized from a uniform distribution with range $[-0.01, 0.01]$. We set the number of filters to 8 for all the convolutional blocks. We have also set the size of the fully connected layers to 64, and both the dropout rates are set to 0.5. The model's parameters were trained by the Adam Optimizer [26], with an initial learning rate of 0.001. The batch size is set to 8. Note that these network parameters are adjusted to prevent overfitting due to the small number of samples in the datasets. However, to maintain the same conditions in different experiments, these values are not necessarily optimal.

Due to the limitations of the available datasets, we considered the number of followers as the influence of the publishers. All the values for the news credit vector and the news influence vector are normalized using the min-max normalization method.

5.2 Benchmark datasets

Several datasets are available for fake news detection with different characteristics [12]. For instance, LIAR [58], CRED BANK [30], and IFND [47]. Due to the need for social context data along with news content to conduct our experiments, we utilize a comprehensive fake news detection benchmark dataset called *FakeNewsNet* [51]. The dataset is collected from two

⁴ <https://keras.io/>

⁵ <https://www.nltk.org/>

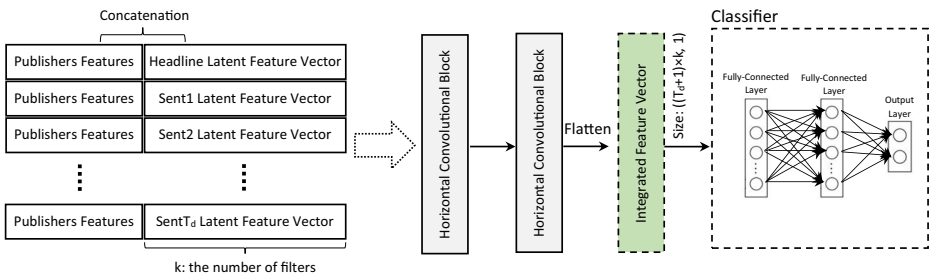


Fig. 13 The architecture of the Integrator. The publishers’ feature vector is concatenated to the feature vector of each sentence. Then, using the correct number of HCBs, the new feature vector is provided to the classifier

fact-checking platforms: *GossipCop* (news related to celebrities) and *PolitiFact* (political news), both containing labeled news content and related social context information in Twitter. The detailed statistics of the datasets are listed in Table 2. Since many experiments should be performed to evaluate the effectiveness of each feature, initially, 20% of samples in each dataset are uniformly separated for fair tests (unseen data).

5.3 The CreditRank algorithm parameters

As mentioned before, the CreditRank algorithm has two parameters (*iteration* and α) that must be specified. So, we have performed appropriate experiments to find the optimal values; the results are illustrated in Figs. 14. As shown, by setting $\alpha = 0.5$, the algorithm has achieved better results in one iteration. In other words, better results are obtained by considering the credibility of each publisher and its first-level followers to evaluate the final credibility. On the other hand, $\alpha = 0.5$ has shown the best result, which indicates that the history of activity and credit rank (followers credibility) have an equal share in determining the publisher’s credit.

5.4 Results

To evaluate the performance of fake news detection methods, we use the following metrics, which are commonly used to evaluate classifiers in related areas: Accuracy, Precision, Recall, and F1. The experiments have been conducted under the same conditions as follows. First, we compare the performance of the SLCNN (our base model) with traditional CNN for text classification [25]. As shown in Fig. 15, the SLCNN has achieved significantly better results than the traditional text-CNN in all metrics for both datasets due to having extra information from the text.

Table 2 Statistics of the datasets

Platform	PolitiFact		GossipCop	
	Real	Fake	Real	Fake
# Train samples	192	188	9342	3162
# Test samples	49	47	2336	790
T_d	280		85	
# Publishers	512,370			



Fig. 14 Parameters analysis for the CreditRank Algorithm. The best result is obtained with iteration = 1 and $\alpha = 0.5$ for both datasets

Then, to examine the effectiveness of publishers’ features, i.e., *Credibility (C)*, *Influence (I)*, *Sociality (S)*, *Validity (V)*, and *Lifetime (L)*, on the performance of fake news detection models, we have prepared comprehensive experiments that enable evaluating the impact of each feature and its combinations. As mentioned in section 4–2 (Integrator), one or more features have been added to SLCNN in each experiment to analyze its impact on overall performance. For simplicity, we used SLCNN (XYZ) as a notation to indicate which features are used in the FR-Detect framework. Thus, SLCNN (XYZ) means the framework involves the SLCNN and features X, Y, and Z of the publishers.

The performance analysis for the publishers’ features is summarized in Table 3 and compared in Fig. 16. We make the following observations from the results: The CreditRank feature has dramatically increased the accuracy, more than other features (around 0.16 in PolitiFact and 0.14 in GossipCop datasets). On the other hand, the Sociality feature had the weakest performance; it also reduced the accuracy of the base model. In summary, the effectiveness of publishers’ features in PolitiFact and GossipCop is *Credibility* >> *Lifetime* > *Validity* > *Influence* > *Sociality* and

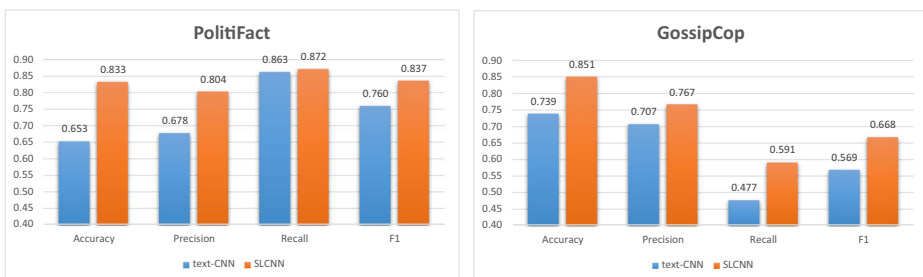


Fig. 15 Comparison of the performance of the SLCNN and the traditional text-CNN

Table 3 Classification results using different publishers’ features in the FR-Detect framework

Model	Politifact				GossipCop			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
SLCNN()	0.833	0.804	0.872	0.837	0.851	0.767	0.591	0.668
SLCNN(C)	0.990	0.979	1.000	0.989	0.988	0.985	0.966	0.975
SLCNN(I)	0.813	0.784	0.851	0.816	0.883	0.774	0.761	0.767
SLCNN(S)	0.729	0.769	0.638	0.697	0.782	0.566	0.590	0.578
SLCNN(V)	0.823	0.841	0.787	0.813	0.919	0.869	0.803	0.835
SLCNN(L)	0.896	0.863	0.936	0.898	0.896	0.838	0.728	0.779
SLCNN(CI)	0.990	0.979	1.000	0.989	0.985	0.975	0.967	0.971
SLCNN(CS)	0.990	0.979	1.000	0.989	0.983	0.985	0.955	0.970
SLCNN(CV)	0.979	0.959	1.000	0.979	0.985	0.978	0.963	0.970
SLCNN(CL)	0.979	0.959	1.000	0.979	0.987	0.990	0.958	0.974
SLCNN(CIS)	0.990	0.979	1.000	0.989	0.985	0.983	0.958	0.970
SLCNN(CIL)	0.990	0.979	1.000	0.989	0.987	0.986	0.962	0.974
SLCNN(CIV)	0.979	0.959	1.000	0.979	0.986	0.986	0.958	0.972
SLCNN(CSV)	0.990	0.979	1.000	0.989	0.987	0.990	0.957	0.973
SLCNN(CSL)	0.979	0.979	0.979	0.979	0.987	0.988	0.958	0.973
SLCNN(CVL)	0.990	0.979	1.000	0.989	0.986	0.991	0.953	0.972
SLCNN(CIVL)	0.990	0.979	1.000	0.989	0.988	0.986	0.966	0.976
SLCNN(CIVS)	0.990	0.979	1.000	0.989	0.986	0.975	0.971	0.973
SLCNN(CISL)	0.979	0.979	0.979	0.979	0.986	0.983	0.961	0.972
SLCNN(CSVL)	0.979	0.959	1.000	0.979	0.986	0.988	0.956	0.972
SLCNN(CISVL)	0.979	0.959	1.000	0.979	0.987	0.987	0.962	0.974

Credibility > > *Validity* > *Lifetime* > *Influence* > *Sociality* respectively. Also, it has been observed that combining other features with the *Credibility* feature has not been able to improve the model’s overall performance. This indicates that the credibility of publishers plays a crucial role in verifying the authenticity of the news.

Accuracy and cross-entropy loss of different features for PolitiFact and GossipCop are shown in Figs. 17 and 18, respectively. From the figures, the training loss decays faster and more with credibility than other features.

Finally, we also compared the performance of FR-Detect (SLCNN, C), as our winner model, with state-of-the-art methods for fake news detection. The algorithms used for comparison are listed as follows:

- **3HAN [53]:** 3HAN utilizes a hierarchical attention neural network framework on news textual contents for fake news detection. It encodes textual contents using a three-level hierarchical attention network for words, sentences, and headlines.

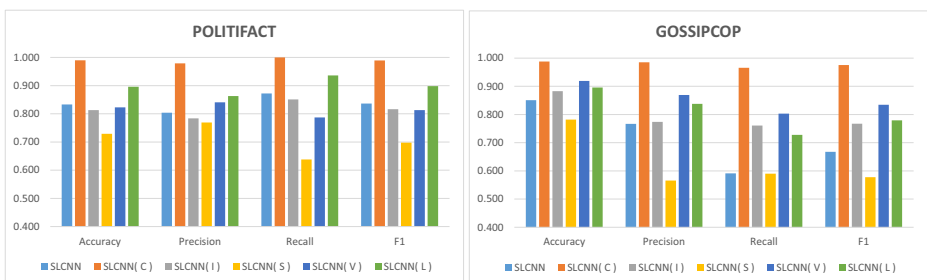


Fig. 16 The test performance comparison of publishers’ features. Credibility outperforms the other features

- **TCNN-URG [39]:** TCNN-URG utilizes a Two-Level Convolutional Neural Network with User Response Generator (TCNN-URG) where TCNN captures semantic information from textual content by representing it at the sentence and word level, and URG learns a generative model of user response to news contents from historical user responses to generate responses for new incoming articles and use them in fake news detection.
- **dDEFEND [11]:** dDEFEND utilizes a sentence-comment co-attention sub-network to exploit both news contents and user comments to jointly capture top-k check-worthy sentences and user comments for fake news detection.
- **SAFE [70]:** SAFE uses multi-modal (textual and visual) information of news articles. First, neural networks are adopted to extract textual and visual features for news representation separately. Then the relationship between the extracted features is investigated across modalities. Finally, news textual and visual representations and their relationship are jointly learned and used to predict fake news.
- **OPCNN-FAKE [46]:** it represents an optimized Convolutional Neural Network model to detect fake news. Grid search and hyperopt optimization techniques have been used to optimize the parameters of the network.

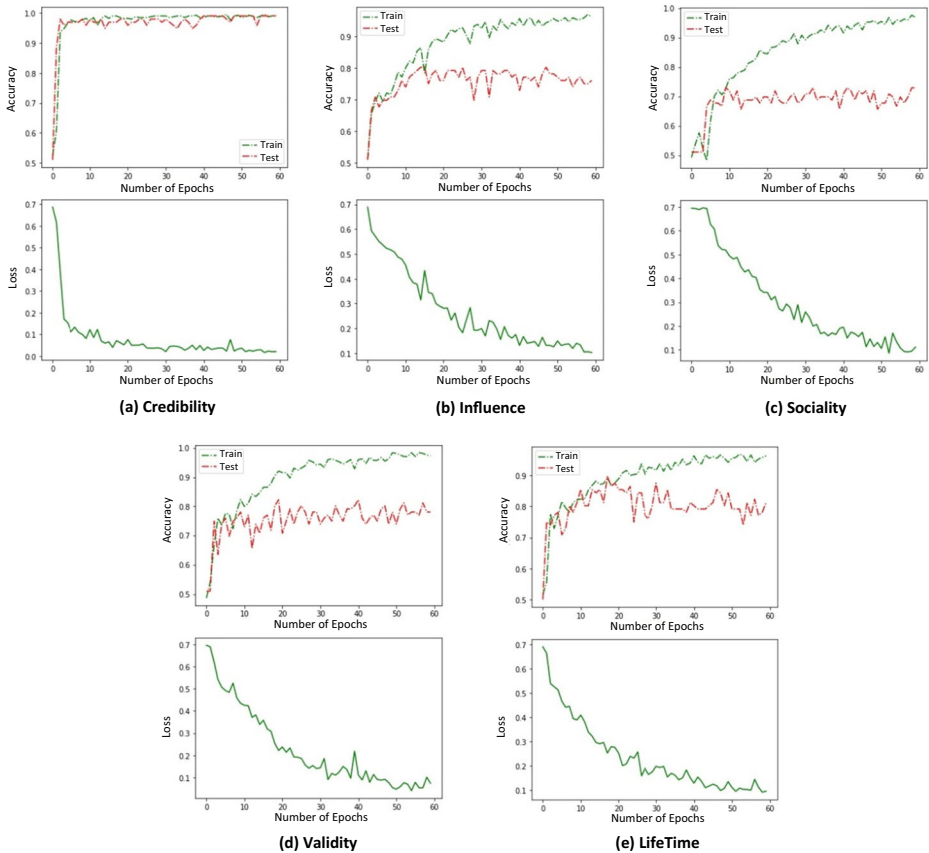


Fig. 17 Accuracy and cross-entropy loss of different features for PolitiFact

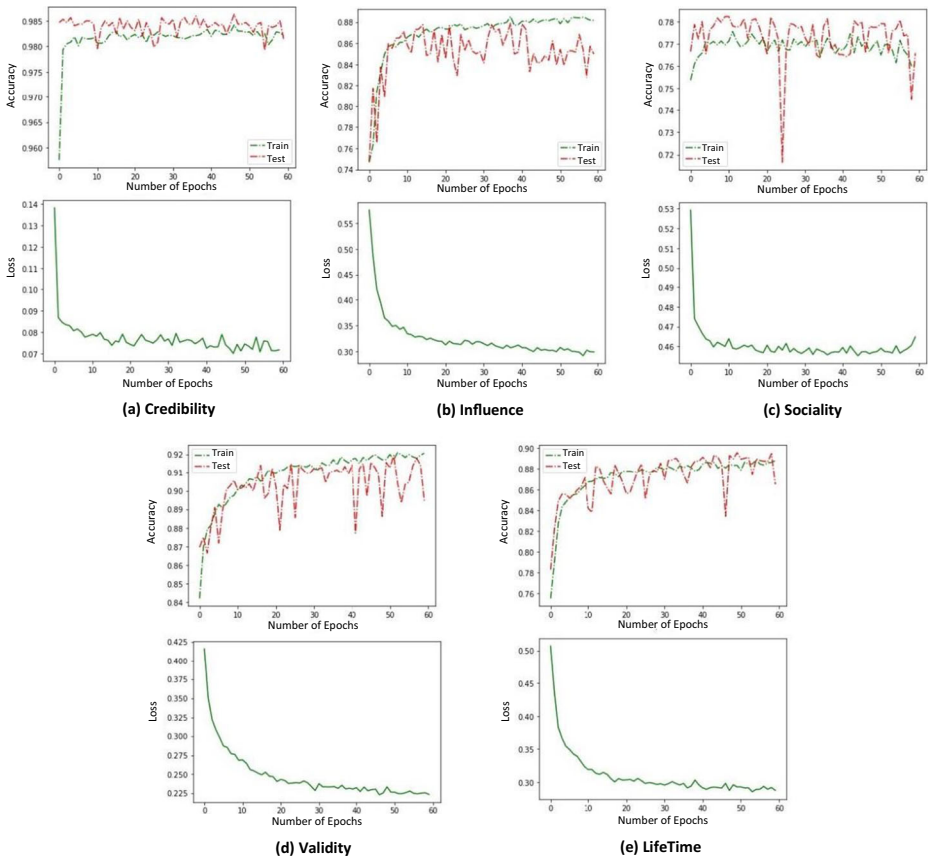


Fig. 18 Accuracy and cross-entropy loss of different features for GossipCop

Note that all the models used in this comparison, except DEFEND (because of using real comments), have the early detection property. The results are shown in Table 4. The results reveal that FR-Detect has managed to achieve by far the best result for both datasets in all metrics.

5.5 Discussion

In this section, we discuss three issues:

11. Characteristics of the user-related features
12. Statistical analysis of the publishers' features
13. The computational complexity for extracting the features

Cold start and unreliability are the most important issues of some user-related features that should be considered in real-world applications. Cold start means that little information may be available about that feature because the user is a newcomer. Among the features discussed in this paper, *Credibility*, *Influence*, and *Sociality* have the cold start issue. Due to the lack of a

significant number of followers of the newcomers, this issue is not critical in the fake news detection area because the published news of these publishers cannot be widely disseminated on social media and, therefore, will not have much impact. In contrast, unreliability is very important and effective in fake news detection. Unreliability means that the feature can be manipulated by the user. Publishers can use this manipulation to mislead the model. Among all features discussed in this paper, just *Sociality* is unreliable. So, *Sociality* is not a suitable feature for fake news detection. Characteristics of the user-related features are summarized in Table 5.

The following is a statistical analysis of the publishers’ features to gain a deeper understanding of each of them and their relationship with the authenticity of the news. The correlation between publishers’ features is shown in Fig. 19. From the figure, we can have the following findings:

- Publishers’ credibility has a strong positive correlation with *Validity* and *Lifetime* in political news and a strong negative correlation in the news related to celebrities. This means that validated publishers have published less fake political news, while such publishers have published more fake news in the realm of celebrities. In other words, fake news related to celebrities is mainly published by validated people, while fake political news is published by unvalidated people.
- Fake news about celebrities is spread more by influencers, while fake political news is spread more by people with fewer followers.
- There is not much significant correlation between publishers’ credibility and their sociality.
- In general, older or validated publishers have more followers.
- Validated publishers generally have a longer lifetime.

As shown in Table 6, the average number of publishers for each news item varied in different news areas. In general, political news is published by more publishers. Also, fewer publishers publish fake political news, while fake celebrity news is published by more publishers. Therefore, it can be concluded that the behavior of publishers on social media is entirely different according to the news domain.

Another critical issue is the computational complexity of the features extraction. First, it should be noted that all the features introduced for publishers (*PTN*, *PFN*, *PCR*, *Influence*, *Sociality*, *Validity*, and *Lifetime*) can be maintained and updated in their user profiles. Hence, these features can be accessed when publishing news with $O(1)$. The computational complexity for updating each feature is as follows:

Table 4 The test performance of methods in fake news detection. Results of OPCNN-FAKE are reprinted from the reference. They merged both datasets and reported one result

Datasets	Metrics	3HAN	TCNN-URG	dDEFEND	SAFE	OPCNN-FAKE*	FR-Detect (SLCNN, C)
PolitiFact	Accuracy	0.844	0.712	0.904	0.874	0.952	0.990
	Precision	0.825	0.711	0.902	0.889	0.952	0.979
	Recall	0.899	0.941	0.956	0.903	0.952	1.000
	F1	0.860	0.810	0.928	0.896	0.952	0.989
GossipCop	Accuracy	0.750	0.736	0.808	0.838	0.952	0.988
	Precision	0.659	0.715	0.729	0.857	0.952	0.985
	Recall	0.695	0.521	0.782	0.937	0.952	0.966
	F1	0.677	0.603	0.755	0.895	0.952	0.975

Table 5 Characteristics of the user-related features

	<i>Credibility</i>	<i>Influence</i>	<i>Sociality</i>	<i>Validity</i>	<i>Lifetime</i>
Cold start	✓	✓	✓		
Unreliability			✓		

- *Credibility*: according to the CreditRank algorithm, the publisher credit vector has three components PTN, PFN, and PCR. Components PTN and PFN for publishers can be updated with $O(1)$ when he/she publishes a new piece of news. By considering $iteration = 1$, component PCR can be updated on-demand or periodically, e.g., weekly or monthly, with $O(n)$, where n is the number of publishers on social media.
- *Influence*: we have proposed two options for calculating Influence: 1) Accurate calculation using Eq. 5, which can be updated on-demand or periodically, e.g., weekly or monthly, with $O(n^d)$, where n is the number of publishers on social media and d is the diameter of the social network. 2) Estimation using the number of followers, which can be updated with any change in the number of followers, with $O(1)$.
- *Validity*, *Lifetime*, and *Sociality* (the number of friends) are simple features in the user profile; their updating can be done with any change with $O(1)$.

Finally, the computational complexity of the Mask function is entirely related to its implementation. For example, if the publishers’ list is maintained for each news, the selection can be made with $O(1)$ and otherwise with $O(m)$, where m is the number of news.

6 Conclusion and future works

Fake news detection has received growing attention in recent years. One of the most relevant entities in assessing the authenticity of a news story in the real world is its narrator. So, this paper investigated the effectiveness of publishers’ features in detecting fake news on social media. In this regard, we introduced some main features for news publishers on social media, including *Credibility*, *Influence*, *Sociality*, *Validity*, and *Lifetime*. One of the most important advantages of publishers’ features is that they do not delay the detection process because they are available at the publication time. *Credibility* is a complex feature that requires a suitable algorithm for calculation. Therefore, we proposed the CreditRank algorithm, which considers the activity history and credit rank of publishers in the network. We have also presented a novel sentence-level convolutional neural network (SLCNN) that can be used generally in text classification. One of the advantages of SLCNN is that it enables us to combine other extra features at the sentence level. By statistical analysis, we found that the behavior of publishers on social media is completely different according to the news domain. Experiments on real-

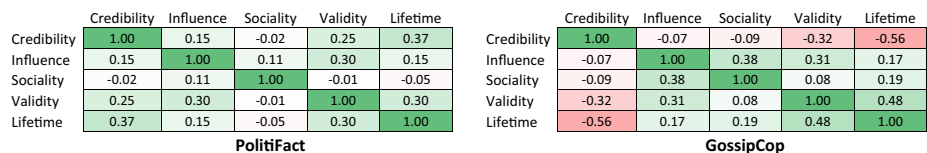


Fig. 19 The correlation heatmap of publishers’ features

Table 6 Average number of publishers for each news item in different news domains

	Fake	Real	Total
PolitiFact	376	739	560
GossipCop	100	52	64

world datasets demonstrate that the credibility of publishers plays a crucial role in verifying the authenticity of the news. The results have shown that the SLCNN with CreditRank of publishers outperforms the state-of-the-art methods. In other words, our proposed model has succeeded in detecting fake news with around 99% accuracy. As future work, we intend to extract and study more features from publishers and their interconnections.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

- Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. *J Econ Perspect* 31(2): 211–236
- Apuke OD, Omar B (2021) Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics Inform* 56:101475
- Baly R et al (2018) Predicting factuality of reporting and bias of news media sources. arXiv preprint arXiv: 1810.01765
- Berahmand K et al (2021) A preference random walk algorithm for link prediction through mutual influence nodes in complex networks. *J King Saud Univ Comput Inf Sci*
- Bhutani B et al (2019) Fake news detection using sentiment analysis. In: 2019 twelfth international conference on contemporary computing (IC3). IEEE
- Boehm LE (1994) The validity effect: a search for mediating variables. *Personal Soc Psychol Bull* 20(3): 285–293
- Bondielli A, Marcelloni F (2019) A survey on fake news and rumour detection techniques. *Inf Sci* 497:38–55
- Choudhary A, Arora A (2021) Linguistic feature based learning model for fake news detection and classification. *Expert Syst Appl* 169:114171
- Choudhary M, Chouhan SS, Pilli ES, Vipparthi SK (2021) BerConvoNet: a deep learning framework for fake news classification. *Appl Soft Comput* 110:107614
- Cui L, Lee SWD (2019) SAME: Sentiment-Aware Multi-Modal Embedding for Detecting Fake News
- Cui L et al (2019) dEFEND: a system for explainable fake news detection. In: Proceedings of the 28th ACM international conference on information and knowledge management. ACM
- D’Ulizia A, Caschera MC, Ferri F, Grifoni P (2021) Fake news detection: a survey of evaluation datasets. *Peer J Comput Sci* 7:e518
- de Oliveira NR, Pisa PS, Lopez MA, de Medeiros DSV, Mattos DMF (2021) Identifying fake news on social networks based on natural language processing: trends and challenges. *Information* 12(1):38
- Goldani MH, Safabakhsh R, Momtazi S (2021) Convolutional neural network with margin loss for fake news detection. *Inf Process Manag* 58(1):102418
- Hakak S, Alazab M, Khan S, Gadekallu TR, Maddikunta PKR, Khan WZ (2021) An ensemble machine learning approach through effective feature extraction to classify fake news. *Futur Gener Comput Syst* 117: 47–58
- Hinton GE et al (2012) Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580
- Hosseinimotlagh, S. and E.E. Papalexakis. Unsupervised content-based identification of fake news articles with tensor decomposition ensembles. In Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2). 2018.

18. Huynh TL (2020) The COVID-19 risk perception: a survey on socioeconomics and media attention. *Econ Bull* 40(1):758–764
19. Jiang T, Li JP, Haq AU, Saboor A, Ali A (2021) A novel stacking approach for accurate detection of fake news. *IEEE Access* 9:22626–22639
20. Jin Z, et al. (2017) Multimodal fusion with recurrent neural networks for rumor detection on microblogs. in *Proceedings of the 25th ACM international conference on Multimedia*
21. Kaliyar RK, Goswami A, Narang P, Sinha S (2020) FNDNet—A deep convolutional neural network for fake news detection. *Cogn Syst Res* 61:32–44
22. Kaliyar RK, Goswami A, Narang P (2021) DeepFakeE: improving fake news detection using tensor decomposition-based deep neural network. *J Supercomput* 77(2):1015–1037
23. Kaliyar RK, Goswami A, Narang P (2021) FakeBERT: fake news detection in social media with a BERT-based deep learning approach. *Multimed Tools Appl* 80(8):11765–11788
24. Karimi H, Tang J (2019) Learning hierarchical discourse-level structure for fake news detection. *arXiv preprint arXiv:1903.07389*
25. Kim Y (2014) Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*
26. Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*
27. Leibenstein H (1950) Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *Q J Econ* 64(2):183–207
28. Mihalcea R and Strapparava C (2009) The lie detector: explorations in the automatic recognition of deceptive language. In *Proceedings of the ACL-IJCNLP 2009 conference short papers*. Association for Computational Linguistics
29. Mikolov T et al (2013) Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*
30. Mitra T, Gilbert E (2015) Credbank: A large-scale social media corpus with associated credibility annotations. In: *Ninth International AAAI Conference on Web and Social Media*
31. Mouratidis D, Nikiforos MN, Kermanidis KL (2021) Deep learning for fake news detection in a pairwise textual input Schema. *Computation* 9(2):20
32. Nasir JA, Khan OS, Varlamis I (2021) Fake news detection: a hybrid CNN-RNN based deep learning approach. *Int J Inf Manage Data Insights* 1(1):100007
33. Nickerson RS (1998) Confirmation bias: a ubiquitous phenomenon in many guises. *Rev Gen Psychol* 2(2): 175–220
34. Ozbay FA, Alatas B (2020) Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications* 540:123174
35. Page L et al (1999) The PageRank citation ranking: bringing order to the web. *Stanford InfoLab*
36. Parikh SB, Atrey PK (2018) Media-rich fake news detection: a survey. In: *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*. IEEE
37. Pennington J, Socher R, Manning C (2014) Glove: Global vectors for word representation. In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*
38. Pérez-Rosas V et al (2017) Automatic detection of fake news. *arXiv preprint arXiv:1708.07104*
39. Qian F, et al (2018) Neural User Response Generator: Fake News Detection with Collective User Intelligence. in *IJCAI*
40. Rapoza K (2017) Can 'fake news' impact the stock market
41. Reis JC et al (2019) Supervised learning for fake news detection. *IEEE Intell Syst* 34(2):76–81
42. Rubin VL (2010) On deception and deception detection: content analysis of computer-mediated stated beliefs. In *Proceedings of the 73rd ASIS&T Annual Meeting on navigating streams in an information ecosystem-volume 47*. American Society for Information Science
43. Rubin VL, Chen Y, and Conroy NJ (2015) Deception detection for news: three types of fakes. In *Proceedings of the 78th ASIS&T Annual Meeting: information science with impact: research in and for the community*. American Society for Information Science.
44. Rubin VL, Conroy NJ, and Chen Y (2015) Towards news verification: Deception detection methods for news discourse. in *Hawaii International Conference on System Sciences*
45. Ruchansky N, Seo S, and Liu Y (2017) Csi: a hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on conference on information and knowledge management*. ACM
46. Saleh H, Alharbi A, Alsamhi SH (2021) OPCNN-FAKE: optimized convolutional neural network for fake news detection. *IEEE Access* 9:129471–129489
47. Sharma DK, Garg S (2021) IFND: a benchmark dataset for fake news detection. In: *Complex & Intelligent Systems*, pp 1–21
48. Shu K, Sliva A, Wang S, Tang J, Liu H (2017) Fake news detection on social media: a data mining perspective. *ACM SIGKDD Explorations Newsletter* 19(1):22–36

49. Shu K, et al. (2019) The role of user profiles for fake news detection. in Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining
50. Shu K, Dumais S, Awadallah AH, Liu H (2020) Detecting fake news with weak social supervision. *IEEE Intell Syst* 36:96–103
51. Shu K, Mahudeswaran D, Wang S, Lee D, Liu H (2020) Fakenewsnet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big data* 8(3):171–188
52. Singhal S et al (2019) SpotFake: a multi-modal framework for fake news detection. In: 2019 IEEE fifth international conference on multimedia big data (BigMM). IEEE
53. Singhanian S, Fernandez N, Rao S (2017) 3han: a deep neural network for fake news detection. In: International conference on neural information processing. Springer
54. Sitaula N et al (2020) Credibility-based fake news detection. In: Disinformation, Misinformation, and Fake News in Social Media. Springer, pp 163–182
55. Varma R et al (2021) A systematic survey on deep learning and machine learning approaches of fake news detection in the pre-and post-COVID-19 pandemic. *International Journal of Intelligent Computing and Cybernetics*
56. Verma A, Mittal V, Dawn S (2019) FIND: fake information and news detections using deep learning. In: 2019 twelfth international conference on contemporary computing (IC3). IEEE
57. Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359(6380):1146–1151
58. Wang WY (2017) "liar, liar pants on fire": A new benchmark dataset for fake news detection. arXiv preprint arXiv:1705.00648
59. Wang Z, Guo Y (2020) Rumor events detection enhanced by encoding sentimental information into time series division and word representations. *Neurocomputing* 397:224–243
60. Wang Y et al (2018) Eann: event adversarial neural networks for multi-modal fake news detection. In: Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining. ACM
61. Wang L, Qian X, Zhang Y, Shen J, Cao X (2019) Enhancing sketch-based image retrieval by cnn semantic re-ranking. *IEEE Transactions on Cybernetics* 50(7):3330–3342
62. Xu Y, Li Z, Wang S, Li W, Sarkodie-Gyan T, Feng S (2021) A hybrid deep-learning model for fault diagnosis of rolling bearings. *Measurement* 169:108502
63. Yang Z et al (2016) Hierarchical attention networks for document classification. In: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies
64. Yang Y et al (2018) TI-CNN: Convolutional neural networks for fake news detection. arXiv preprint arXiv:1806.00749
65. Yang S et al (2019) Unsupervised fake news detection on social media: A generative approach. In: Proceedings of the AAAI Conference on Artificial Intelligence
66. Zhang X, Ghorbani AA (2020) An overview of online fake news: characterization, detection, and discussion. *Inf Process Manag* 57(2):102025
67. Zhang J, Dong B, Philip SY (2020) Fakedetector: effective fake news detection with deep diffusive neural network. In: 2020 IEEE 36th international conference on data engineering (ICDE). IEEE
68. Zhou X, Zafarani R (2018) Fake news: A survey of research, detection methods, and opportunities. arXiv preprint arXiv:1812.00315
69. Zhou X, Jain A, Phoha VV, Zafarani R (2020) Fake news early detection: a theory-driven model. *Digital Threats: Research and Practice* 1(2):1–25
70. Zhou X, Wu J, Zafarani R (2020) SAFE: Similarity-aware multi-modal fake news detection. arXiv preprint arXiv:2003.04981
71. Zubiaga A, Aker A, Bontcheva K, Liakata M, Procter R (2018) Detection and resolution of rumours in social media: a survey. *ACM Computing Surveys (CSUR)* 51(2):32–36