Technological Forecasting & Social Change 195 (2023) 122746

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

COVID-19 fake news detection: A hybrid CNN-BiLSTM-AM model

Huosong Xia [a,b,c], Yuan Wang [a], Justin Zuopeng Zhang [d,*], Leven J. Zheng [e], Muhammad Mustafa Kamal [f], Varsha Arya [g,h,i]

[a] School of Management, Wuhan Textile University, Wuhan 430073, China
[b] Research Center of Enterprise Decision Support, Key Research Institute of Humanities and Social Sciences in Universities of Hubei Province, China
[c] Research Institute of Management and Economics, Wuhan Textile University, Wuhan 430073, China
[d] Department of Management, Coggin College of Business, University of North Florida, Jacksonville, FL 32224, USA
[e] Department of Global Business and Marketing, Lee Shau Kee School of Business and Administration, Hong Kong Metropolitan University, Ho Man Tin, Hong Kong
[f] School of Strategy and Leadership, Coventry University, Coventry CV1 5FB, United Kingdom
[g] Insights2Techinfo, India, & Lebanese American University, Beirut 1102, Lebanon
[h] University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India
[i] Chandigarh University, Chandigarh, India, & Immersive Virtual Reality Research Group, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

ARTICLE INFO

Keywords:
Deep learning
COVID-19
Fake news
Convolutional Neural Networks
Bidirectional Long Short-Term Memory Networks
Attention Mechanism

ABSTRACT

With the rapid development of technology, social media as a communication platform has caused a significant increase in the dissemination of false information and fake news. We propose an outlier knowledge management framework of "generation–spread–identification–refutation" for detecting fake news in emergencies based on the theory of complex adaptive systems and information transmission. We extract and acquire outlier knowledge of COVID-19 fake news, incorporate it into the outlier knowledge base of major emergencies for knowledge sharing and transformation, classify the acquired knowledge from three dimensions (people, organization, and technology), and develop wisdom according to the extracted knowledge. Our proposed hybrid model is based on Convolutional Neural Network, Bidirectional Long Short-term Memory Network, and Attention Mechanism (AM) for fake news detection, thereby improving the evaluation indicators of Loss, Accuracy, F1-score, and Recall by at least 1 %. The multi-head AM performs better in fake news detection when models with different AMs are adjusted, and significant differences in sentence length and topic distribution can be observed between real and fake news. The model provides a think tank and a platform for public opinion guidance to deal with major emergency news detection.

1. Introduction

The COVID-19 pandemic has had varying and equally unprecedented and unexpected challenges and effects that resulted in a significant global death toll, economic conditions, and social behaviors (Kamal, 2020; Shankar et al., 2021). However, the poor spread of fake news also contributes to the threat. The Internet and social media enable people to access information more directly than before while exposing them to less rigorous news coverage (Laato et al., 2020). Muhammad et al. (2022) reported in their study on digital footprints on social media reported that users spend approximately 144 min on social media platforms, and around 463 EB of data are created each day globally. Disinformation also rapidly spreads through social media channels due to this colossal span of data sharing and is widely described as "fake news" (Di Domenico et al., 2021). According to recent research, several rumors and false news reports concerning COVID-19 are circulating (Rodríguez et al., 2020; Carmine et al., 2021; Chiang et al., 2022; Narayan et al., 2022). Distinguishing fake news from stories whose veracity is beyond dispute has become increasingly challenging (Huynh, 2020). Accordingly, misinformation on social media has intensified public panic about the COVID-19 pandemic, and the government has urged citizens to lean the ability to distinguish between real and fake news (Huynh, 2020; Apuke and Omar, 2021). Although experts believe that the spread of fake news has made COVID-19 worse, the features of COVID-19 fake news have not been investigated to help with fake news detection (Carrion-Alvarez and Tijerina-Salina, 2020; Nam and Kabutey, 2021; Tembhurne et al., 2022; Zhang et al., 2023). An effective fake news detection algorithm is crucial for analyzing fake news during the COVID-

* Corresponding author.
E-mail addresses: justin.zhang@unf.edu (J.Z. Zhang), lzheng@hkmu.edu.hk (L.J. Zheng), ad2802@coventry.ac.uk (M.M. Kamal).

https://doi.org/10.1016/j.techfore.2023.122746
Received 10 June 2022; Received in revised form 30 March 2023; Accepted 7 June 2023
Available online 28 July 2023

19 pandemic and prevent negative hot topics and heightened public apprehension.

Fake news is nothing new, but the speed and scale of its spread reached a new level during the COVID-19 pandemic. 2020 was a pivotal year of the online information war, with COVID-19-related fake news emerging as the most pervasive issue on a global scale (Balakrishnan et al., 2021). For example, nearly half (46 %) of the population in the United Kingdom was found to have been exposed to COVID-19-related fake news (Ofcom, 2020). Meanwhile, over 25 % of the most viewed videos on YouTube about COVID-19 featured false or misleading information, with over 62 million views globally (Li et al., 2020). The spread of fake news has reportedly greatly influenced political elections around the world (e.g., the election of Donald Trump as president) and caused public health problems (Waszak et al., 2018; Schwitzer, 2017). The proliferation of fake news on social media increased during the COVID-19 pandemic, with many experts arguing that fake news contributed to the spread of the epidemic by drowning out the voices that official public health authorities put out online.

The disruptive and destructive nature of the COVID-19 outbreak has presented a major challenge to the healthcare system in our society (Depoux et al., 2020; Kamal, 2020). Considering this uncertainty, new information on viruses, clinical manifestations, transmission, and disease prevention is needed (Eysenbach, 2020; Zheng et al., 2021a, 2021b,2021c; Zheng et al., 2022c). As COVID-19 spreads across the globe, the spread of misinformation has created chaos on a global scale. Accordingly, understanding the spread of fake news and taking appropriate measures are extremely important for the well-being of the society (Kim et al., 2021; Xing et al., 2021; Zheng et al., 2021a, 2021b, 2021c). Videos and photos of the origins, spread, vaccines, and deaths of COVID-19 are circulating on social media, and 30 % to 35 % of the news, videos, and photos disseminated on social media platforms are fake (Khan et al., 2022). Fake news spreads faster than the virus, affecting residents' judgment of the epidemic and fueling public panic (Pennycook et al., 2020). Fake news is specifically designed to harm or promote a non-existing phenomenon and is misleading and propagandistic. Online social media, represented by Weibo and WeChat, have provided a platform for netizens to obtain information and express their opinions on the pandemic (Carvajal-Miranda et al., 2020; Meadows et al., 2021). The spread of fake news is rampant in today's digital world (Apuke and Omar, 2021; Wang et al., 2022). Fake news touches almost every aspect of people's lives (Kucharski, 2016; Varshney and Vishwakarma, 2021). Detection of fake news about COVID-19 can provide policymakers a reliable foundation on which to base their decisions. By analyzing the COVID-19 microblog news, capturing the typical characteristics of true and false COVID-19 news, and sharing this knowledge, the effectiveness of decision-making can be improved (Al-Qirim et al., 2022; Cui et al., 2022; Huang and Wang, 2022; Tang and Zhong, 2022; Xing et al., 2022). Several calls for digital literacy have been voiced by researchers, governments, and public interest groups who have been developing an array of resources for the public (DeJong, 2023).

Manually checking fake news has become ineffective due to the rapid spread of online misinformation (i.e., not obvious, difficult, and time-consuming) (Mondal and Chakrabarti, 2021; Sample et al., 2020). Some scholars studied fake news detection techniques during COVID-19. Traditional machine learning algorithms do not consider the representation of semantic context, so feature selection is difficult to manually carry out. The neural network achieves better results compared with other machine learning methods (Haran and Niederman, 2022; Jin et al., 2022; LeCun et al., 2015; Li et al., 2022b; Li et al., 2022a; Yu et al., 2022; Almomani et al., 2022). Deep learning algorithms are superior to machine learning algorithms in fake news detection (Cheng et al., 2021; Varma et al., 2021). Deep learning model fusion has a better performance in detecting fake news compared with a single model (Sahoo and Gupta, 2021; Aslam et al., 2021). The predictive effect of supervised learning can be improved with the help of domain knowledge (Zheng et al., 2022a, 2022b; K. Zheng et al., 2022; W. Zheng et al., 2022). However, domain knowledge is not embedded in the model in the research on detecting COVID-19 fake news. Therefore, this work proposes the Convolutional Neural Network–Bidirectional Long Short-term Memory Network–Attention Mechanism (CNN–BiLSTM–AM) hybrid model fake news detection method based on domain knowledge.

As researchers work to find a permanent cure for COVID-19, the proliferation of fake news on social media has exacerbated the pandemic threat (Lampos et al., 2021). Fake news recognition has become an important research area in natural language processing (NLP) (Goel and Donaldson, 2021; Trueman et al., 2021). At present, knowledge of the predictors of the spread of fake news is limited, and studies on COVID-19 fake news are scarce (Apuke and Omar, 2020). This work effectively works in solving major emergencies based on the outlier knowledge management paradigm, "data-information–outlier knowledge–wisdom" (Xia et al., 2022). The complex adaptive system and information transmission theories consider that information dissemination is "top-down" and view fake news in the generation–dissemination–identification–refutation process (Uhl-Bien and Arena, 2018; Silva and Guerrini, 2018). On the basis of this backdrop, we construct a fake news detection model under major emergencies by combining complex adaptive systems and information transfer theories with heterogeneous knowledge management embedded. This work makes contributions by (1) using the NLP and deep learning techniques to improve the accuracy of fake news detection, (2) utilizing intelligent analysis technology to identify the characteristics of fake news and enhance the recognition ability of public fake news, and (3) extending the theory of outlier knowledge management in the context of COVID-19 to better serve the practice. In particular, this work addresses the following three research questions:

RQ1: *What are the typical characteristics of fake news (from the fake-news generation, spread, identification, and refutation perspectives)? What are the unique characteristics of producing and spreading fake news in major emergencies?*
RQ2: *How does machine learning identify COVID-19 fake news?*
RQ3: *From the perspective of machine learning algorithm design and evaluation, how can the efficiency of COVID-19 fake news identification be improved?*

To answer these proposed questions, we adopt the NLP technology to analyze specific COVID-19 news, mainly focusing on true and false COVID-19 news, and establish a COVID-19 news domain vocabulary. The NLP technology helps in analyzing news during the outbreak to obtain true and false information about COVID-19 news. We discuss how to deal with fake news from three aspects: *"people"*, *"organizations"*, and *"technology"*, providing a new direction for people, organizations, and governments to make informed decisions during major emergencies.

The remainder of this paper is structured as follows. Section 2 reviews previous relevant literature and proposes research hypotheses. Section 3 describes the research process, including the data acquisition and the construction of the COVID-19 fake news detection model based on deep learning. Section 4 analyzes the experimental results. Section 5 concludes the paper and outlines the contribution and future research directions. The meanings of the notations and abbreviations are summarized in the Appendix.

2. Literature review

2.1. Generation of fake news

Fake news refers to content that is deliberately created to resemble genuine news, with the aim of misleading the public into believing that it is factual but contains false information, and involves various forms of false news and misinformation (Duffy et al., 2020). A fake news story contains physical content (such as headline, body, picture, and video) and non-physical content (such as purpose, mood, and news topic) (Zhang and Ghorbani, 2020). According to Molina et al. (2023), one of the major issues resulting from disinformation is the manifestation of fabricated anecdotes that mainly impedes the healthcare sector, which causes distress and anguish among the masses. Malicious online accounts express deceptive messages by deliberately confusing their writing styles or trying to imitate other users. Fake news was used as a synonym for "satire, parody, fabrication, manipulation, propaganda, and advertising" in an analysis of papers related to fake news from 2003 to 2017 (Tandoc Jr. et al., 2018). Some scholars expanded their list of fake news keywords to include "hoaxes, rumors, conspiracy theories, fabricated stories, and bait headlines" (Shao et al., 2018).

Recent approaches attempt to extract explanatory factors from user reviews and Web documents to explain how fake news is generated (Shu et al., 2020). The most prominent feature of fake news is its syntactic and semantic features (Choudhary and Arora, 2021). With the success of deep learning and deep generation models, the machine-generated text could become a new type of fake news, smooth, readable, and catchy, for example, generative adversarial network (Hanshal et al., 2022; Jing et al., 2023). Fake news can be effectively detected from cause to effect by combining CNN and RNN.

2.2. Fake news spread

During the COVID-19 outbreak, fake news flooded social media, creating an environment where it could quickly spread (Ram and Zhang, 2021; Yang and Tian, 2021). Social media played an essential role during the COVID-19 pandemic, enabling people to disseminate news in real time across the globe (Islam et al., 2020). The reasons why people share false information related to COVID-19 must be understood, and interventions to improve the quality of the information people share online must be developed. An empirical study on fake news sharing factors among social media users shows that altruism is a critical factor affecting the sharing of COVID-19 fake news (Apuke and Omar, 2021; Balakrishnan et al., 2021). Not all people spread fake news, and young consumers rely less on official and trusted sources of information, such as TV, newspaper, and radio news, indicating that they prefer social media-based sources of information (Peng et al., 2021; Pennycook and Rand, 2020; Singla et al., 2022; Rouibah et al., 2022). When people share real news to enhance social cohesion, the same practices can also lead them to spread fake news (Duffy et al., 2020). The key stages in the spread of fake news are network creation, analysis, content generation, and information dissemination (Almomani et al., 2021; Ng and Taeihagh, 2021; Wang et al., 2021).

Given that policy interventions rely on verifiable evidence, pandemics require governments to not only consider expert input but also ensure that science is translated into public understanding to disprove and disrupt the spread of fake news (Hartley and Vu, 2020). However, research focused on understanding fake news sharing during the pandemic is only just emerging (Islam et al., 2020; Laato et al., 2020; Apuke and Omar, 2021), suggesting that more studies are needed. In this context, a COVID-19 fake news detection technology is particularly important to lessen the panic caused by information asymmetry. Moreover, the social instability brought by fake news causes social unrest. Thus, fake news must be promptly identified to prevent it from being further disseminated.

2.3. Fake news detection

Authenticity is compromised by the occurrence of deliberate deception. Online news publication has changed, and traditional fact-checking and censorship, which may be deceptive, have become impossible to fight against the flood of online news (Conroy et al., 2015). News must be published in authentic form, which is frequently found in adulterated versions, leaving us with an urgent need to identify real news from any possible fake news (Sharma and Sharma, 2019). Many individuals rely on identifying user, content, and context features that indicate false information to identify fake news (Zhang and Ghorbani, 2020). For instance, COVID-19-related fake news can be detected using feature extraction to avoid computational overhead when processing all the features in the dataset (Khan et al., 2022). Fake news is characterized by large quantities, diversity, and rapid dissemination (Zhang and Ghorbani, 2020). New tools, including deep learning technologies and NLP, must be created to develop systems that can be easily integrated into Internet browsers and social networks, provide immediate proof that the claim, news, or content is true or possibly false information, and generate an identifier (Marco-Franco et al., 2021; Mutambik et al., 2021). In the context of major emergencies, the research on COVID-19 fake news can improve the accuracy of fake news detection in a single scenario.

Fake news detection techniques for COVID-19 have been studied. A dataset was constructed using manual annotation due to the lack of datasets for fake news detection in the early stages of COVID-19 research (Kim et al., 2021). The conversion of text to vector enables the conversion of text information into a machine-readable form (Gaurav et al., 2021; Srivastava and Eachempati, 2021). In word vector construction, Word2vec is a commonly used encoding algorithm in addition to the basic one-hot encoding. This algorithm considers the upper and lower semantics to generate embeds more suitable for tasks involving syntax (Ji et al., 2019). Meanwhile, doc2vec considers the order of upper and lower sentences, and it is better used in paragraphs (Jeon et al., 2022). Although deep learning word vector construction is more complex than machine learning word vector construction, it uses the encoder–decoder framework to specify the output dimensions for text vector encoding (Bai et al., 2021). This work adopts the Word2vec algorithm to construct the word vector of fake news, combining the complexity and effect of the word vector algorithm.

The choice of the fake news prediction model is analyzed using a deep learning algorithm. The semi-supervised end-to-end neural attention model deals with unlabeled data, primarily using external knowledge to analyze fake news with defects in accuracy and universality (Al-Azad et al., 2022; Liang et al., 2022a, 2022b; Paka et al., 2021). The use of CNN and Long Short-term Memory (LSTM) Networks to detect fake news concerning COVID-19 is well-established (Wani et al., 2021). However, the requirement for fake news detection by using a single model has not been met in reality. Model fusion will affect the detection accuracy of COVID-19 fake news, and excellent model fusion can improve the accuracy (Amine et al., 2019). A hybrid approach is based on a deep learning classifier to better use more datasets for decision-making (Sahoo and Gupta, 2021). Accordingly, some scholars have fused the CNN and BiLSTM models to detect fake news on the public social platform, and the detection accuracy has been improved. Moreover, feature extraction has been added to the detection process, and more attention has been paid to the characteristics of fake news (Kaliyar et al., 2021; Wani et al., 2021; Amine et al., 2019; Rahman et al., 2022). In addition, AMs play an important role in NLP (Trueman et al., 2021). AM integrated into deep learning models can discover the most relevant information (Karnyoto et al., 2022; Kesharwani et al., 2021; Shih et al., 2021; Zhang et al., 2021). The fusion of deep learning models can significantly improve the detection effect of fake news and the applicability of multiple datasets (Sahoo and Gupta, 2021; Aslam et al., 2021). Predictions based on domain knowledge can improve the ability to detect fake news, but research that embeds the domain knowledge of COVID-19 into the model is limited (Pham et al., 2022). The existing project has yet integrated CNN, BiLSTM, and AM models to verify that multi-model fusion can improve the detection effect of COVID-19 fake news. According to previous research, only technical methods are innovated during fake news detection (Zheng et al., 2021b; Zheng et al., 2021a; Zheng et al., 2022a). Moreover, domain knowledge is not embedded during COVID-19 fake news detection in combination with management theories. This work proposes the method of the CNN–BiLSTM–AM hybrid model to detect fake COVID-19 news with domain knowledge. Therefore, this work adopts an advanced deep-learning network for model fusion to improve detection accuracy.

2.4. Outlier knowledge management

From the perspective of knowledge management, knowledge can be divided into general and outlier knowledge. Outlier knowledge management is similar to knowledge management, and outlier data can be managed through "data management, information management, outlier knowledge acquisition, and knowledge sharing" (Al-Taie et al., 2021; van Rensburg, 2021; Xia et al., 2022). Data mining and deep learning techniques are used to identify, extract, mine, and analyze the existing data, which has been well applied in accurately identifying collective abnormal behavior (Belhadi et al., 2021). The researchers developed and applied machine learning methods to extract the changing characteristics of data and define accurate threshold boundaries to successfully detect outliers (Shao et al., 2022). The detection and correction of outliers are important steps in analyzing various types of data in the financial, economic, industrial, geographic, and medical fields (Erkuş and Purutçuoğlu, 2021; Wahid and Annavarapu, 2021). However, fake news knowledge has not been widely studied in theory and practice in news detection.

The major emergencies are outlier events requiring quick decision response time. Fake news on social platforms under the influence of the pandemic has a significant negative effect on people's decision-making, reflecting laws that are different from ordinary knowledge and may be more valuable (Xia et al., 2017). Extracting outlier information that is considered useful data, rather than purging useless or unwanted data, can improve data quality for further analysis (Blázquez-García et al., 2021). This work uses the NLP technology to analyze and mine data from the process of outlier knowledge management, namely, "data management, information management, outlier knowledge acquisition, and knowledge sharing", for obtaining outlier knowledge of news during COVID-19 and improve the accuracy of fake news detection.

2.5. Complex adaptive system theory and information transmission theory

The basic concept of complex adaptive system theory is the interaction and adaptability of the subject in the environment, which adopts the "bottom-up" research route (Uhl-Bien and Arena, 2018; Silva and Guerrini, 2018; Tchuente, 2022; ZareRavasan and Krčál, 2021). Management of information dissemination not only involves the dissemination of misinformation but also needs an understanding of the potential factors driving the dissemination of misinformation (Bok et al., 2021). We proposed a fake news identification process of "generation–spread–identification–refutation" during fake news identification. This work summarizes the above-mentioned theoretical basis based on existing theories and proposes a new outlier knowledge management framework by making up the existing theoretical gaps in the process of "generation–spread–identification–refutation". Furthermore, this work aims to analyze COVID-19 fake news from the perspective of outlier knowledge management, extract the knowledge of identifying fake news, promote strategies to combat fake news, and solve decision-making problems under different circumstances.

In the process of outlier knowledge discovery, outlier knowledge is first discovered and refined into outlier knowledge using domain knowledge, which is an interactive spiraling process to knowledge. This work combines the theories of complex adaptive systems and information dissemination, adopts the "bottom-up" research route, and proposes the process of "generation–spread–identification–refutation" to discover the dissimilarity of fake news. The research framework of this work is the combination of fake news identification and outlier knowledge management processes.

2.6. Research framework

Contrary to the typical scenario, people can enrich their knowledge by reading news on social media platforms. However, the emergence of fake news on these platforms has a significant negative effect on knowledge acquisition. In the context of COVID-19, we used advanced the NLP technology to study COVID-19-related news from keywords extraction, text analysis, target detection, etc., to acquire knowledge of COVID-19 outliers, build a knowledge base of major emergency outliers, and carry out knowledge sharing and transformation. The relevant knowledge of detecting fake news is extracted and classified from three dimensions: people, organization, and technology, and wisdom is formed according to the extracted knowledge. The COVID-19 fake news detection research framework based on an outlier knowledge management framework and algorithm principle together with the process of fake news "generation–spread–identification–refutation" is shown in Fig. 1. Using the data processing flow of data management, information management, knowledge management, and knowledge sharing, the former is the foundation of the latter, and the latter is the extension and development of the former. According to heterogeneous knowledge management in information theory, the data processing flow is used to construct the knowledge base of extreme public health events and fake news outliers based on the basic paradigm of data–information–outlier knowledge–wisdom, and the knowledge base is refined into wisdom for knowledge sharing and transformation.

3. Model construction and evaluation (research design)

3.1. Problem description

Assuming that each news item in the training data has a modal data form: text (represented by S), a group of target news that needs to identify the authenticity of the content can be expressed as Formula 1:

$$D = \{(S)\} = \{(S_1), (S_2), \dots (S_i), \dots\}, \tag{1}$$

where S_i is the i^{th} news document containing a group of words. The decision problem is to accurately judge whether news i is true according to the content (S_i) of news i.

Consider a training set D_M, where label information Y is in each news, as shown in Formula 2:

$$D_M = \{(S, Y)\} = \{(S_1, Y_1), (S_2, Y_2), \dots (S_i, Y_i), \dots\}, \tag{2}$$

where Y is the domain name tag of the i^{th} news and is the news tag (Y_i=1 indicates that the news is true, while Y_i=0 denotes that the news is false). Given that the D_M news category labels are clear and accurate, we can learn a classification model as FM: D_M- > Y. Given a dataset D_M with label information, the research problem of this work is how to learn a model F that can be used to identify the labels of news in unknown dataset D, as shown in Formula 3:

$$F : (D|D_M) \rightarrow Y. \tag{3}$$

3.2. Data

A COVID-19 fake news dataset is selected from a variety of sources in social media applications, such as Twitter, Facebook, and Fact-Checking, to assess the generality of the training model test setup (Patwa et al., 2021). The real news comes from verified accounts, such as CDC,[1] WHO, and ICMR.[2] The dataset contains 10,700 real and fake social media posts and articles, all in English. The tweets are fine-grained and tagged with COVID-19-related misinformation based on the interests of the different parties involved in the outbreak information. We mapped the true and false labels in our experiment according to the binary settings made by the data publisher instead of using multiple labels. In our experiment, we split the dataset into validation and test

[1] https://www.cdc.gov/.

[2] https://www.icmr.gov.in/.

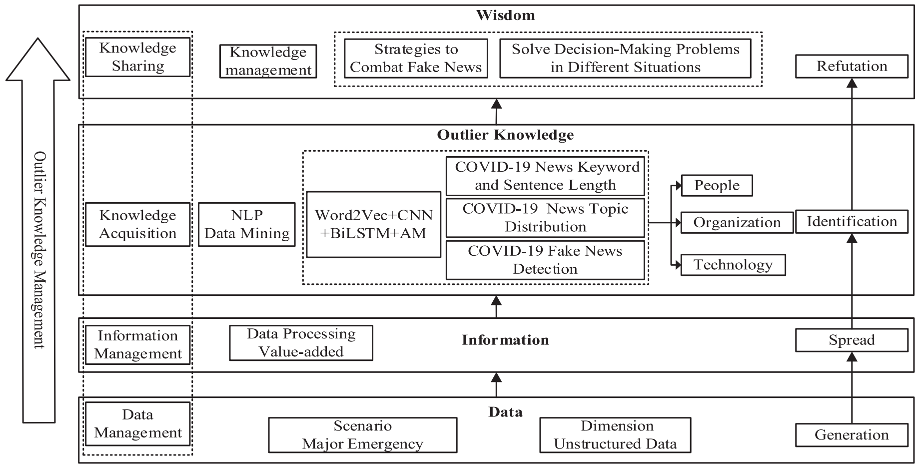


Fig. 1. Outlier knowledge management for COVID-19 fake news research framework.

Table 1
Data distribution across labels and splits.

Label	Training	Validation	Test	Total
Real	3360	1120	1120	5600
Fake	3060	1020	1020	5100
Total	6420	2140	2140	10,700

Table 2
Ten most common words in the COVID-19 Fake news dataset.

Dataset	Most frequent words
Real	Cases, covid19, new, covid, tests, people, states, deaths, total, and testing
Fake	Covid, coronavirus, people, virus, vaccine, coronavirus, trump, says, new, and covid19

sets with the same label distribution. The details are shown in Table 1. Table 2 lists the words that most frequently occur after the stop words are removed from each dataset. The *k*-fold cross-validation is a commonly used performance evaluation method for improving the generalization ability of the model, but it is frequently ignored in the research. A higher value of *k* results in a smaller bias, but a higher variability. The lower values of *k* lead to greater bias, so $k = 10$ is adopted and supported by many practical applications (Marcot and Hanea, 2021; Jain and Gupta, 2019; Chui et al., 2016).

3.3. Pre-processing

3.3.1. Data pre-processing

(1) First, the news is segmented by using jieba. cut ().
(2) Stop words are words, such as "and", "the", and "him", that are considered to provide no information in terms of the content of the text and can be removed to avoid affecting the prediction effect. We remove all links, non-alphanumeric characters, and English stop words. This work uses Baidu stop words table to remove news stop words. Table 2 lists the words that most frequently occur after the stop words are removed from each dataset.

(3) We use the COVID-19 domain knowledge base combined with Word2vec to construct a word vector and output a 614-word vector. Every word is a 1256 vector.

3.3.2. Text alignment

This work analyzes and visualizes the sentence-length distribution of news. Python is used to visualize the length distribution of comments, as shown in Figs. 2 and 3. According to Fig. 2, the length of news presents an aggregation distribution, with data mainly distributed in 150–300 characters. The minimum and maximum lengths of news are 18 and 8856 words, respectively.

Fig. 3 shows the cumulative frequency distribution of text length. The median sentence length at the coordinate (168, 0.5) is 168. Given the aggregation of the length of each news item, the general solutions, such as resampling or oversampling, are used to set the average sentence length for each batch. This work omits or sets a sentence truncation value for the over-long sentences. In each batch, the relatively short text is padded with a specific encoding, and the long text is truncated. In the CNN model, the inputs must be of the same length and vector dimension. Therefore, the consistency of each input dimension is guaranteed for text truncation, which is convenient for matching the input dimensions of the CNN model. The text matrix algorithm is shown in Table 3.

3.3.3. Text matrix

CNN can be used for not only images but also NLP. Given that our model input uses the CNN model, which is characterized by the input data being a matrix, the news data text must be matrixed. The operation algorithm is shown in Table 4. Assuming that the text is T | T | in length, each word is expressed as *d* dimension vector; typically taking 50, 100, and 200, T can be represented as a 2D matrix based on the number of lines for word count and the number of columns for *d*, making the text information similar to the image structure.

3.4. Self-attention mechanism and multi-attention mechanism

The inputs of $x_1, x_2, \dots, x_T$ represent the input sequence data, which can mean the vector corresponding to the word of a sentence. First, the data are preliminarily added via the embedding layer to obtain $\alpha_1\alpha_2, \dots, \alpha_T$ as Eq. (4) shows.

$$\alpha_i = Wx_i \tag{4}$$

Matrices W^Q, W^K, and W^V are multiplied to obtain k_i,q_i,v_i,i∈(1,2,3…, T), as shown in Eqs. (5)–(7).

$$q_i = W^Q \alpha_i \tag{5}$$

$$k_i = W^K \alpha_i \tag{6}$$

$$v_i = W^V \alpha_i \tag{7}$$

$q_1$and $k_1, k_2, \dots, k_T$ are used to take the dot product $a_{(1,1)}a_{(1,2)}\dots a_{(1,T)}$. Then, $a_{(1,1)}a_{(1,2)}\dots a_{(1,T)}$ are inputted into the softmax layer to determine the attention weight values $\hat{a}_{(1,1)}, \hat{a}_{(1,2)}\dots\hat{a}_{(1,T)}$, as shown in Eqs. (8)–(9).

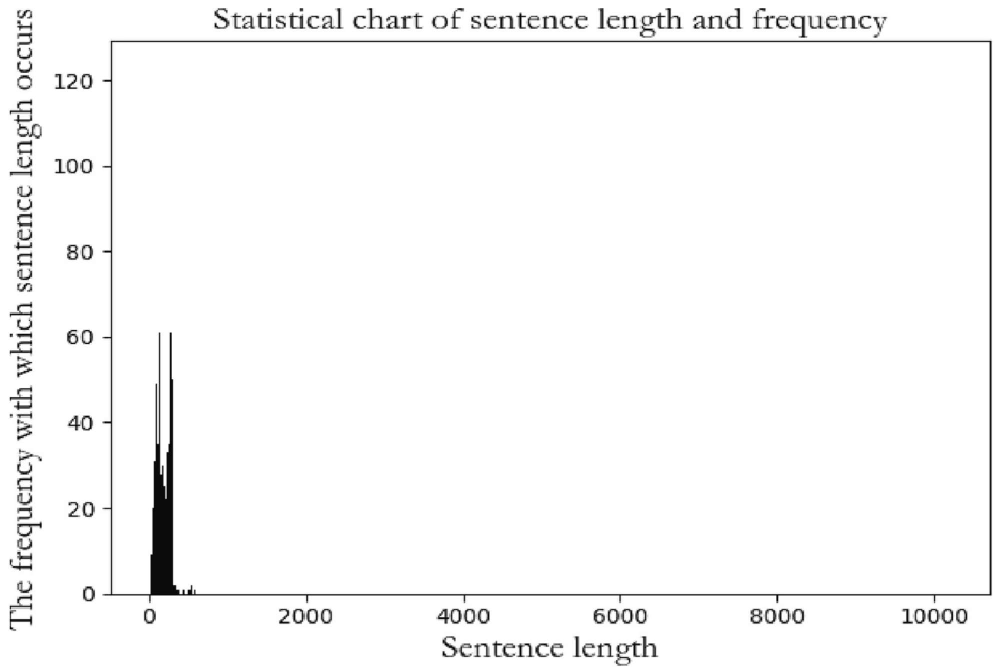


Fig. 2. Distribution of sentence length and frequency.

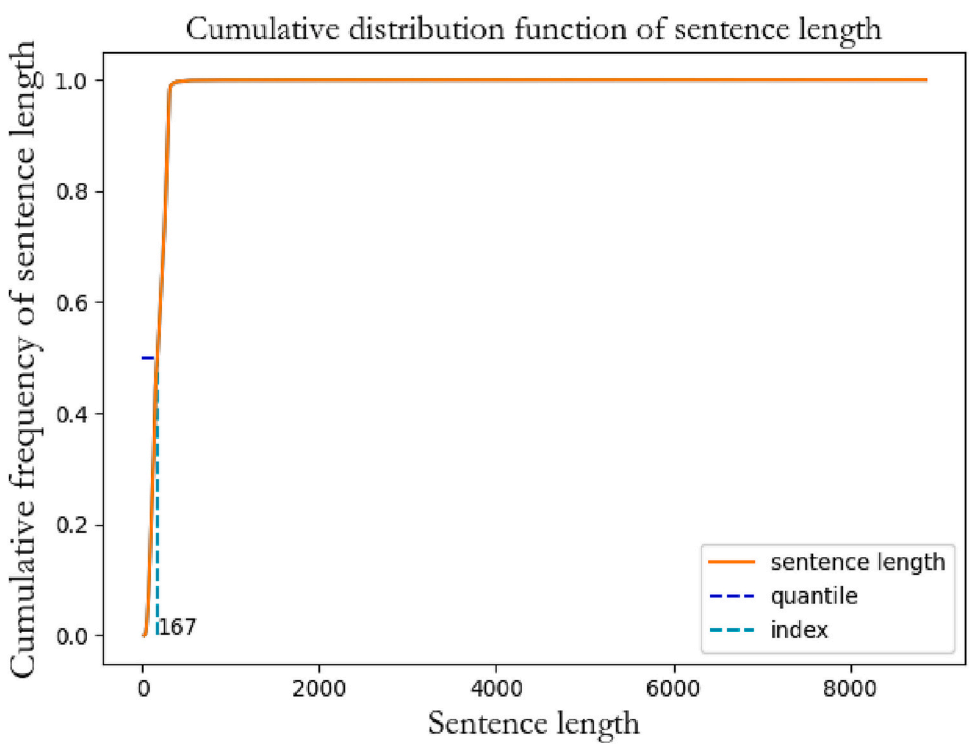


Fig. 3. Cumulative distribution of sentence length.

$$a_{(2,i)} = \frac{q_2^{Transpose} * k_i}{\sqrt{d_{q,k}}} \tag{8}$$

$$\hat{a}_{(2,1)} = Softmax\left(a_{(2,i)}\right) \tag{9}$$

$\hat{a}_{(1,1)}, \hat{a}_{(1,2)} \ldots \hat{a}_{(1,T)}$ and corresponding position $v_1 v_2, \ldots, v_T$ are multiplied and summed to obtain output b_1 corresponding to input x_1, as shown in Eq. (10).

$$b_1 = \sum_i^T \hat{a}_{2,i} {}^* v_i \tag{10}$$

Output b_2 corresponding to input x_2 is obtained by using a similar process; however, in this case, it is obtained by using the corresponding q_2 corresponding to b_2 and $k_1, k_2, \ldots, k_T$ to calculate the vector dot product. The process of the self-attention mechanism is shown in Fig. 4.

If k_i, q_i, and v_i obtained above are regarded as a "head" as a whole, then "multi-head" means that multiple sets of W^Q, W^K, and W^V must be

Table 3
Text completion algorithm.

Algorithm 1: text completion
1:Input: N news items are added to the batch
2:Output: News of the same length
3:for i in Batch:
4: if len(i)<=Mean:
5: i. extend (['PAD'] * (Mean−len(i)))
6: elif len(i])>Mean:
7: i=i[0:Mean]
8: end for

Table 4
Text matrix algorithm.

Algorithm 2: text matrix
1:Input: News text data_text;
2:Output: News matrix;
3: for m in range (len(data_text)):
4: for n in range (len(data_text[m])):
5: data_text[m][n]=vocab[data_text[m][n]]
6: end for
7: end for

multiplied by them for a particular word vector (x_i) to obtain multiple sets of k_i, q_i, and v_i, as shown in Fig. 5.

The defect of the self-attention mechanism is that when the model encodes the information of the current position, it will excessively focus on its position. Some scholars have concluded that when the convolution kernel of the CNN model is 3 × 3, the hyperparameter head = 3 of the multi-head AM model will make the prediction effect of the model reach the peak (Tang et al., 2022). Accordingly, head = 3 is used in this work to experiment with the mechanism of multi-head attention. Taking input α1 in the figure as an example, three outputs, namely, b1head, b2head, and b3head, are obtained through the multi-head mechanism. In the multi-head self-attention, we will spline b^1_{head}, b^2_{head}, and b^3_{head} obtained here (the vectors are connected end to end) and then obtain b_1 through linear transformation (that is, a single-layer fully connected neural network without nonlinear activation layer) to acquire output b_1

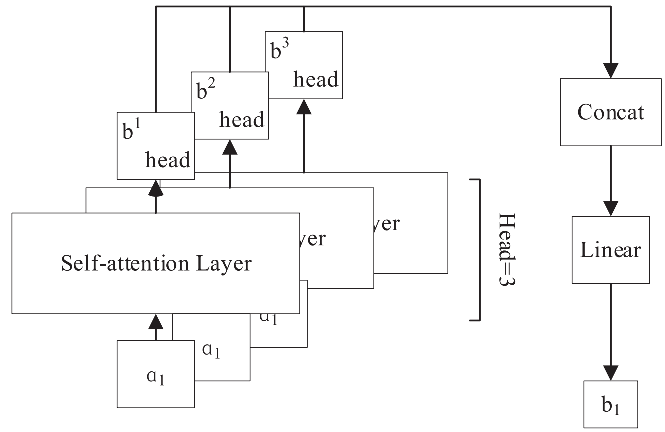


Fig. 5. Schematic of the multi-attention mechanism.

corresponding to α_1. The same process is performed for the other inputs in the sequence, and they share the parameters of these networks.

3.5. Model construction

This work proposes that the CNN–BiLSTM–AM hybrid model is obtained by serial fusion of single models, and the Word2vec word vector construction algorithm that retains the information before and after is used, as shown in Fig. 6. The hybrid model can better extract the semantic information of news and improve the accuracy of predicting COVID-19 fake news detection compared with a single deep learning model. Moreover, the AM integrated into the model CNN–BiLSTM makes the model focus on the information that is more critical to the current task in many input information, reduces the attention to other information, and filters out irrelevant information, which can solve the problem of information overload and improve the efficiency and accuracy of task processing. Accordingly, the fake news detection first constructs the word vector of the text and then inputs it into the CNN–BiLSTM–AM model for training. Finally, the detection results of fake news are attained through the output of the full connection layer,

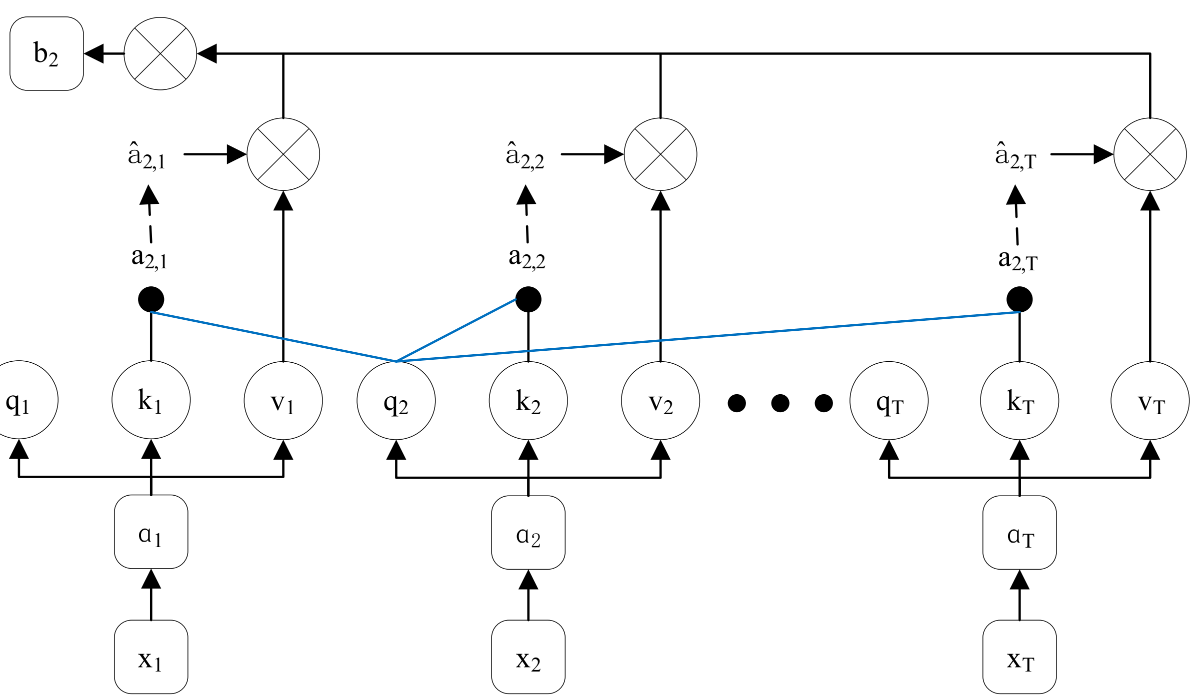


Fig. 4. Self-attention mechanism diagram.

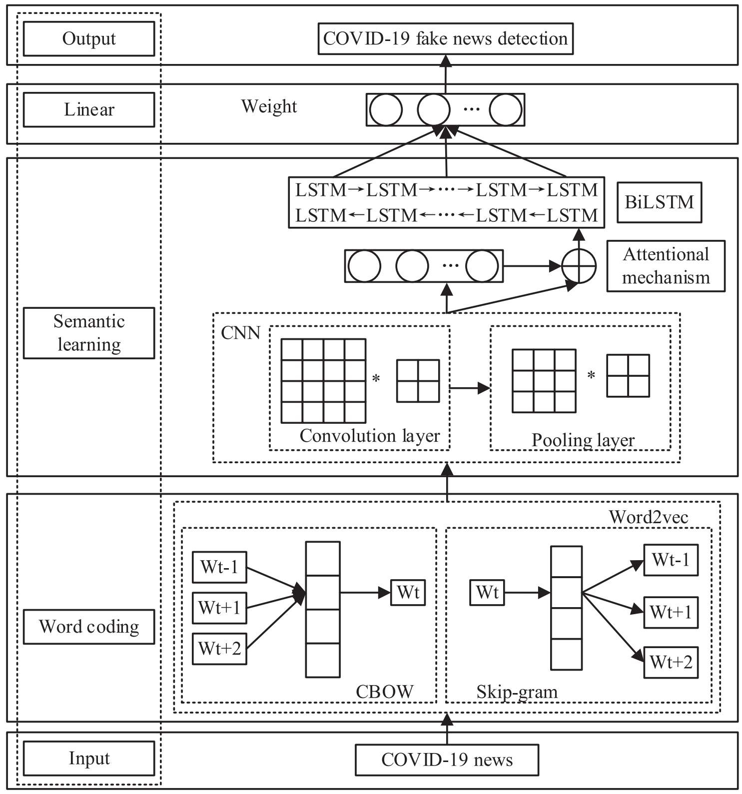


Fig. 6. COVID-19 fake news detection model of CNN–BiLSTM–AM.

and the characteristics of true and false news are analyzed.

3.6. Evaluation index

Machine learning techniques have emerged as a promising solution to the challenges of automatic malware detection and analysis. The basic metrics for evaluating the performance of machine learning algorithms are True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), as shown in Eqs. (11)–(14), respectively.

$$TP = \left\{ \sum \text{Correctly classified as fake news in the sample} \right\} \tag{11}$$

$$TN = \left\{ \sum \text{Correctly classified as true news in the sample} \right\} \tag{12}$$

$$FP = \left\{ \sum \text{Misclassified as fake news in the sample} \right\} \tag{13}$$

$$FN = \left\{ \sum \text{Misclassified as true news in the sample} \right\} \tag{14}$$

We can use accuracy, recall, accuracy, and F1-score using the above-mentioned indicators to quantify the performance of the machine learning algorithm in fake news detection.

1. Precision: Precision quantifies the accuracy of the model and is defined as the ratio of correctly predicted news (TP) to the total number of predicted news (TP + FP) in any class C, which may be positive or negative, as shown in Formula 15.

$$P = \frac{TP}{TP + FP} \tag{15}$$

2. Recall: The completeness of the recall calculation model, which is defined as the ratio of the correctly predicted number of news (TP) to the actual total number of news (TP + FN), is shown in Formula 16.

$$R = \frac{TP}{TP + FN} \tag{16}$$

3. Accuracy: Accuracy evaluates the correctness of a model and calculates the ratio of correct predictions to total news theories, as shown in Formula 17.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

4. F1-measure: An excellent model considers accuracy (P) and recall (R), and the evaluation index combining them is the harmonic mean of accuracy value and recall, as shown in Formula 18.

$$F - measure = \frac{(a^2 + 1)P^*R}{a^2(P + R)} \tag{18}$$

When $a = 1$, precision = recall, as shown in Eq. (19):

$$F1 = \frac{2P^*R}{P + R} = \frac{2TP}{2TP + FP + FN}. \tag{19}$$

4. Experiment and model validation analysis

4.1. Dataset and parameter design

This section will verify the effectiveness of the CNN–BiLSTM–AM method through public data and the effect of the AM on COVID-19 fake news detection by comparing the self-attention mechanism with the multi-attention mechanism. The attributes of the dataset include “id”, “tweet”, and “label” tags, and each story has a “fake” or “real” tag.

All experiments were performed using Python 3.8, and PyTorch 1.11.0 implements the construction of the CNN–BiLSTM–AM model. All experiments were conducted on a Windows 10 operating system, 8 GB RAM, and 4 GB NVIDIA GeForce MX 330 GPU. The parameters of this

Table 5
Parameters of the CNN–BiLSTM–AM model.

Parameters	Value
Embedding	256
Conv1d	100
MaxPool1d	(2,2)
Linear	500
BiLSTM	256
SelfAttentionLayer	100
MultiHeadAttentionAttLayer	(100,3)
Batch size	32
Maxlen	200
Learning_rate	2e-5
Optimizer	Adam
Loss	BCEWithLogitsLoss

model are shown in Table 5.

4.2. Model complexity

(1) Time complexity of the CNN model

The training and reasoning speed of the model is determined by the "operational quantity" (i.e., FLOPs), which represents the time complexity of the model. The bigger FLOPs, the slower model training and reasoning will be. The amount of computation for each convolutional layer in CNN is shown in Eq. (20) as follows:

$$\mathrm{FLOPs} = [(\mathrm{Ci} \times \mathrm{k2}) + (\mathrm{Ci} \times \mathrm{k2} - 1) + 1] \times \mathrm{Co} \times \mathrm{W} \times \mathrm{H} \tag{20}$$

where Ci represents the number of input channels, Co represents the number of output channels, k represents the size of convolution kernel, W and H represent the size of feature maps, the (Ci × k2) part in the above equation is the multiplication operation, the (Ci × k2 − 1) part is the addition operation, and +1 is the bias operation. Papers in the CV field frequently treat the "Multi-Add" combination as a single floating-point operation, as shown in Eq. (21):

$$\mathrm{FLOPs} = \mathrm{Ci} \times \mathrm{k2} \times \mathrm{C0} \times \mathrm{W} \times \mathrm{H} \tag{21}$$

In the FC layer (Eq. (22)):

$$\mathrm{FLOPs} = [\mathrm{I} + (\mathrm{I} - 1) + 1] \times \mathrm{O} = 2 \times \mathrm{I} \times \mathrm{O} \tag{22}$$

where I and O stand for the input and output neurons, respectively. In the activation layer parameter calculation:

1) Sigmoid: FLOPs = 3n, and
2) ReLU: FLOPs = n,

where n represents the n neurons, but the amount of computation in the activation layer is frequently ignored.

(2) Time complexity of the LSTM model

LSTM will maintain a total of four sets of parameters corresponding to the input gate, output gate, forget gate, and candidate state. The model complexity is shown in Eq. (23),

$$\mathrm{FLOPs} = 4^{*}\mathrm{Hidden\_size}^{*}(\mathrm{Input\_size} + \mathrm{Bias} + \mathrm{Output\_size}) \tag{23}$$

4.3. Experimental results and analysis

4.3.1. COVID-19 news sentence length analysis

When we analyze the sentence length, we need to clean news data by (i) deleting empty, invalid, and duplicate data and (ii) removing the link to the news. We first obtained the distribution of real and fake news keywords, as shown in Fig. 7. The glossary of real and fake COVID-19 news is shown in Table 6. According to the keyword graph and table comparison, some words have different meanings from the general situation in the context of COVID-19. For example, "death" is often used in real news, and "died" and "death(s)" in fake news. The representation of real news for words is relatively uniform, while that of fake news for words is particularly rich, and a word is expressed by verbs, nouns, etc. The frequency of use of "coronavirus" in fake news is much higher than that in true news, which more frequently uses COVID-19. Fake news will report unknown news, such as "vaccines". This situation indicates that fake news frequently takes advantage of people's ignorance and fear of new things to increase users' panic and trust. The term Wuhan is frequently used in fake news, and the word COVID-19 virus is referred to as the Wuhan virus. Given that countries have just started to name Wuhan as the source of the virus, the "Wuhan virus" itself is false information, not to mention the news content. Thus, some signs can help us distinguish real news from fake news.

The clean () function is constructed to clear the original "Tweet" column with a function that removes punctuation, converts everything to lowercase, removes the stop word, and returns a clean list of tweets. The same technical choice is Python's Collections and WordCloud libraries after applying the clean function to the tweet columns of the real and fake datasets. In Fig. 8, "covid-19", "coronavirus", and "covid-19" indicate that the spread of news about COVID-19 is still the mainstream whether true or false, which has different degrees of effects on the politics, economy, and culture of the society and people's health and life. Accordingly, real and fake news about COVID-19 remains at the center of major emergencies and poses a major threat to humanity. Words, such as "hospital", indicate a shortage of resources during COVID-19. During the pandemic, many people fell ill and were hospitalized because of the infection, straining hospital resources. The fake news caused panic among the masses, resulting in a large inventory of masks, disinfectants, and other products, which will break the supply of social resources. "Patients", "recovered", "infected", "flu", and "deaths" mean that a large number of people were infected or died. The public's lack of understanding about COVID-19 has led to a state of fear, largely driven by the increasing number of infections or fatalities attributed to the disease. The WordCloud map of fake news shows that the proportion of reports on infection and other related news is higher than that of real news. This notion suggests that fake news uses false news to attract attention and achieve its goal of spreading misinformation.

We calculated the sentence length by summing up the words in each sentence, as shown in Fig. 9. The minimum and maximum lengths of news are 34 and 2149 words, respectively. In particular, the minimum and maximum lengths of fake news are 18 and 8856 words, respectively. The median sentence length of authentic news was higher than that of fake news, which was mostly published as a simulation of fake news and shorter because of "more talk, more mistakes". Nevertheless, fake news can be extreme, with the longest length of fake news being about four times longer than the longest of real news. Strict requirements for length and format must be followed when writing real news, while fake news does not care about these details. To attract attention, the grammar and sentence pattern are not rigorous enough to cause sentence redundancy, making fake news length considerably long.

4.3.2. COVID-19 news topic distribution analysis

Topic variance is the degree of deviation between the probability distributions of subject words. Formula 24 is used to calculate topic variance:

$$Var(T) = \frac{\sum_{j=1}^{n} KL(T_i \| \Omega)}{N}, \tag{24}$$

where T_i is the theme of the extraction, N is the theme of the extraction, Ω represents the "theme-word" probability distribution after the normalization of the mean, and KL ($T_i || \Omega$) sets the similarity between the main body. The larger the topic variance, the higher the degree of differentiation between topics. The number of topics corresponding to

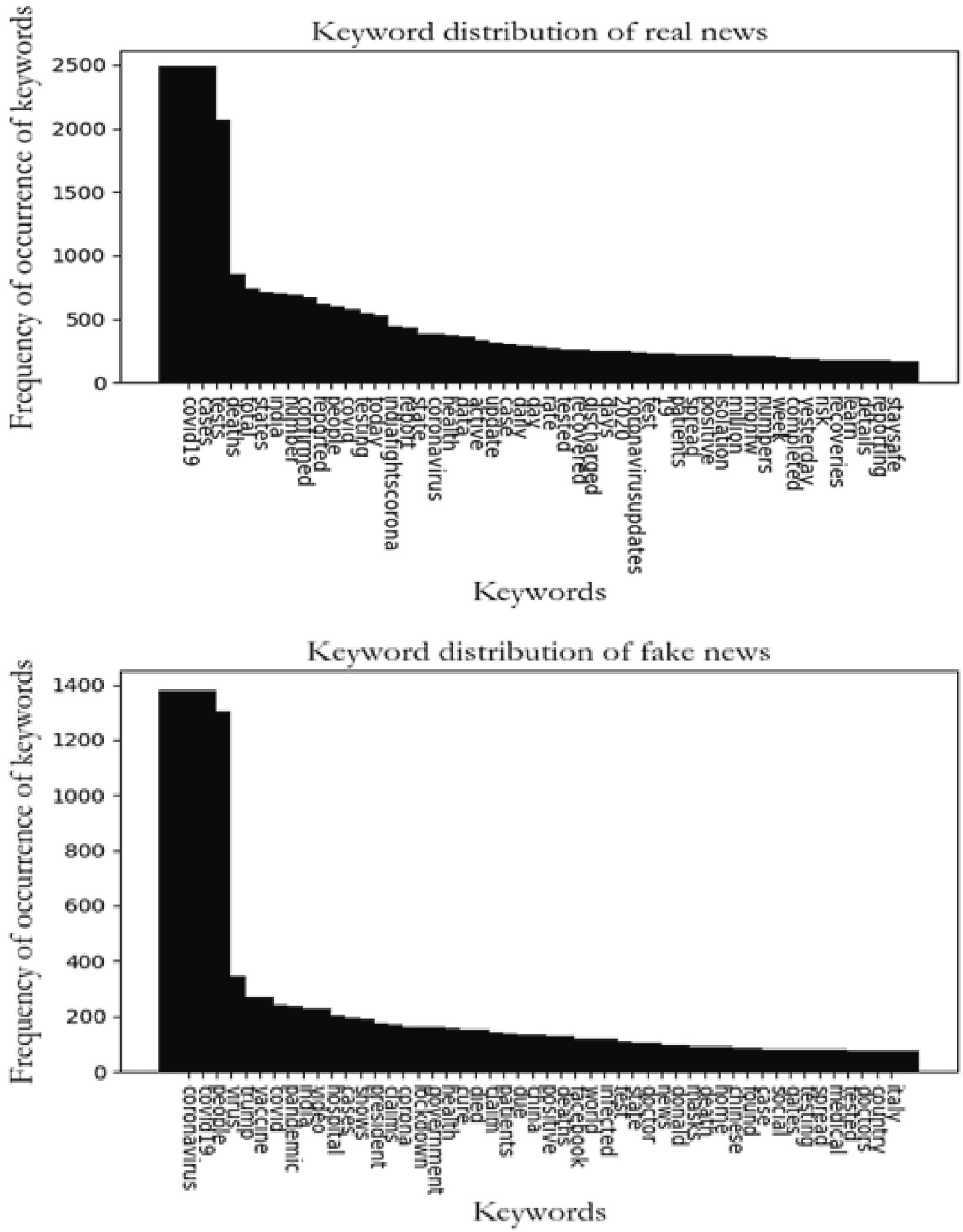


Fig. 7. Distribution of real and fake COVID-19 news keywords.

the largest topic variance is the best.

The LDA library is used to determine the number of topics of true and false news. In Fig. 10(a), in true news, the number of topics 3 and 5 has the largest variance and the highest discrimination between topics. In Fig. 10(b), the number of topic 5 of fake news has the largest variance. In this experiment, the number of true and false news topics is five.

Table 7 shows the five main topics of real and fake news. A difference in the distribution of subjects can be observed between the real news and the fake news, which includes not only novel coronavirus but also laboratory tests and how different situations can be solved, whereas, all the subjects for the fake news are novel coronavirus, the origin of novel Coronavirus, and novel Coronavirus is COVID-19. The subject variance of real news is higher than that of fake news. The research scope of real news is wide, and the subject is distinguished. We can use the LDA library to classify news subject distribution, as shown in Fig. 11. Starting with 1 as the first topic, and so on, we can identify which topic the news belongs to, and which is convenient for the masses to distinguish between true and false news.

4.3.3. Fake news detection model

The results are shown in Fig. 12 to verify the effectiveness of the self-attention mechanism. After 100 iterations, the results of our study

Table 6
Important features of true and false news.

Real news				Fake news			
Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency
Covid19	2489	Cases	1952	Coronavirus	1381	Covid19	1303
Tests	854	Deaths	660	People	345	Virus	268
Total	707	States	598	Trump	267	Vaccine	242
India	691	Number	572	Covid	235	Pandemic	225
Confirmed	619	Reported	492	India	225	Video	204
People	582	Covid	475	Hospital	196	Cases	190
Testing	525	Today	377	Shows	177	President	170
India-fights-corona	431	Report	365	Claims	161	Corona	160
State	379	Coronavirus	327	Lockdown	159	Government	155
Health	363	Data	320	Health	150	Cure	149
Active	307	Update	282	Died	144	Claim	135
Case	287	Daily	264	Patients	134	Due	132
Day	271	Rate	254	China	129	Positive	127
Tested	258	Recovered	235	Deaths	120	Facebook	118
Discharged	249	Days	227	World	117	Infected	108
2020	235	Coronavirus updates	216	Test	102	State	102
Test	223	19	210	Doctor	96	News	95
Patients	218	Spread	206	Donald	92	Masks	92
Positive	215	Isolation	202	Death	91	Home	87
Million	202	mohfw	198	Chinese	87	Found	82
Numbers	193	Week	192	Case	82	Social	80
Completed	185	Yesterday	185	Gates	79	Testing	78
Risk	180	Recoveries	177	Spread	78	Medical	77
Learn	172	Details	169	Tested	76	Doctors	75
Reporting	170	Staysafe	167	Country	75	Italy	74

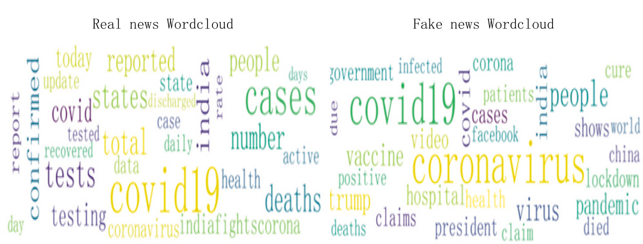


Fig. 8. COVID-19 WordCloud of real and fake news.

showed that Loss, Accuracy, F1-score, and Recall had good indicators in the training and test sets. The model has a good ability to distinguish the truth and falsehood of the news with a good fit, and the loss continues to decline and achieves relatively stable results. The feature extraction ability of recall using the self-attention mechanism is poor compared with the multi-attention mechanism, but a large difference can be observed between the evaluation indexes of the test set and the training set model. The high degree of model fitting results in the inferior performance of recall in the training results on the test set. A possible reason for this result may be the simple characteristics of true and false news, and the use of self-attention mechanisms leads to overfitting problems.

Fig. 13 verifies the effectiveness of the multi-head attention machine. After 100 iterations, our research results show that Loss, Accuracy, F1-score, and Recall have good indicators in the training and test sets. The multi-attention mechanism has a good feature extraction ability, and the model has a good ability to distinguish true and false news with

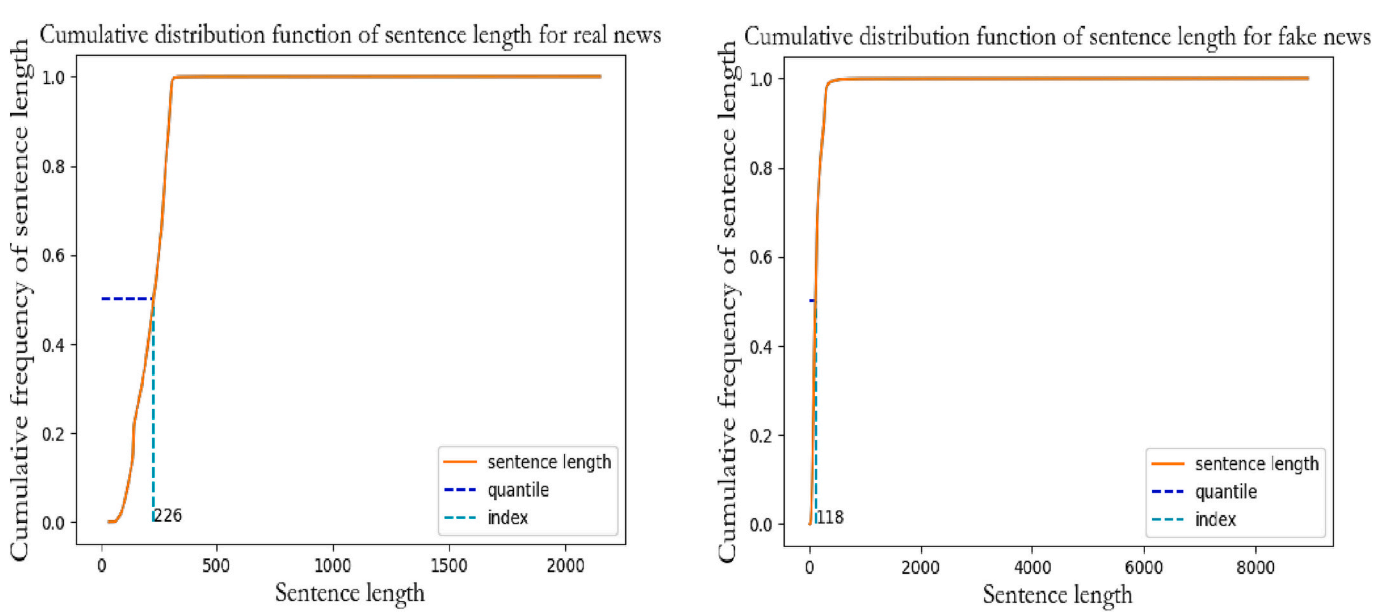
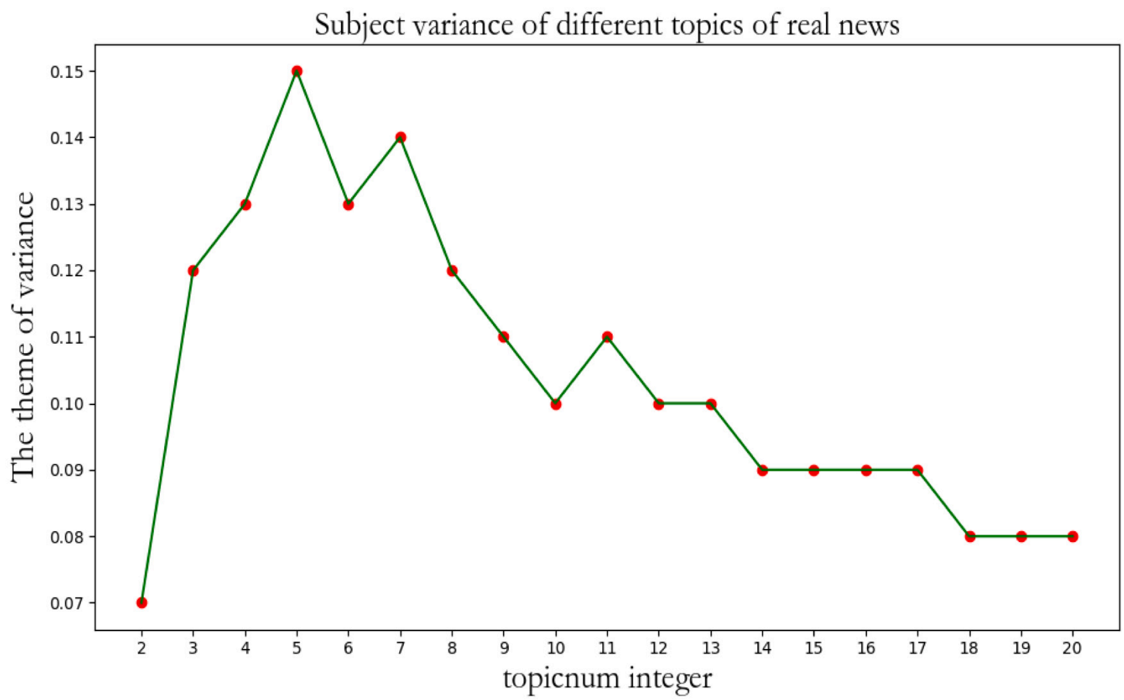
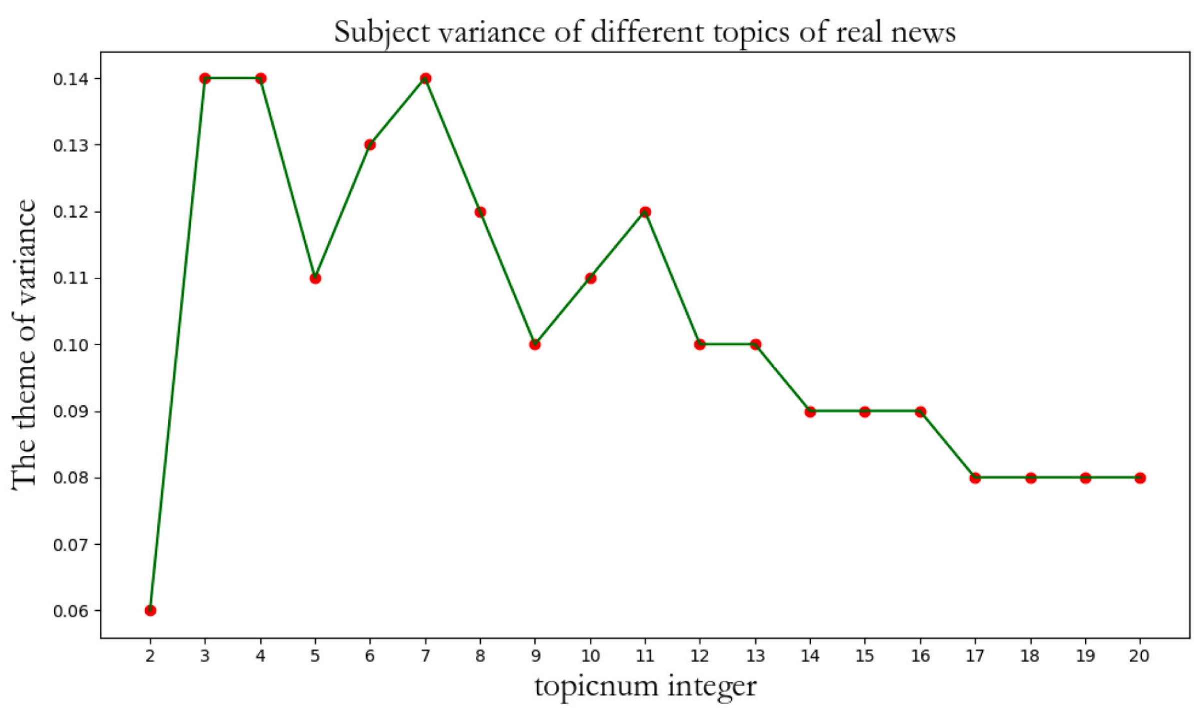


Fig. 9. Sentence length of true and fake COVID-19 news.



(a) Number of optimal topics for real news



(b) Number of optimal topics for fake news

Fig. 10. (a) Number of optimal topics for real news.
(b) Number of optimal topics for fake news.

a good fit. The loss continues to decline and achieve relatively stable results. Fake news detection using the self-attention mechanism model fusion is more stable compared with the recall value in Fig. 12. Therefore, the experimental results show that CNN–BiLSTM combined with the multi-attention mechanism can better explain the decision-making ability of fake news detection.

We adjusted the model by changing the self-attention mechanism and multi-attention mechanism and found that the multi-attention

Table 7
Distribution of true and false news topics.

Number of topics	Real news	Fake news
Topic 1	Tests states cases reported deaths	Covid19 coronavirus India hospital patients
Topic 2	Cases covid19 confirmed number total	Coronavirus covid19 to vaccine gates
Topic 3	Cases covid19 indiafightscorona India total	Covid19 coronavirus trump president pandemic
Topic 4	Covid19 amp spread health people	Covid19 coronavirus virus cure vaccine
Topic 5	Covid19 coronavirus testing test covid	Coronavirus covid19 video claims claim

mechanism was superior to the self-attention mechanism to verify the performance of the AM in COVID-19 fake news detection. This work uses the same dataset to compare with the advanced model and verifies the proposed model's effectiveness by comparing the evaluation indicators, such as Loss, Accuracy, F1-score, and Recall. The experimental results are the predicted mean after a 10-fold cross-validation, as shown in Table 8. Loss, Accuracy, F1-score, and Recall values of the CNN–BiLSTM–Self-attention model are higher than those of the CNN–BiLSTM–MultiHeadAttention model. Therefore, CNN–BiLSTM–AM performs better when using the multi-attention mechanism for fake news detection. Furthermore, the evaluation indicators of fake news detection by k-NN, CNN, BERT, and LSTM models have been improved compared with those of other advanced models in detecting COVID-19 fake news, and the maximum can be increased by 20 %. Our model improves the Accuracy, F1-score, and Recall by at least 1 % compared with CNN–LSTM and BiLSTM–attention fusion models.

4.4. Robustness test of the CNN–BiLSTM–AM fake news detection model

We used CHECKED datasets to verify the robustness of our model. The CHECKED dataset is a Chinese COVID-19 dataset containing fact-checked tweets from Weibo (Yang et al., 2021). The experimental results of fake news detection of the CHECKED dataset in the CNN–BiLSTM–AM model are shown in Figs. 14 and 15. The model still has good performance in terms of the Loss, Accuracy, F1-score, and Recall of the CHECKED dataset, indicating that the model is robust.

The evaluation indexes of Accuracy, F1-score, and Recall are shown in Table 9 to verify the performance of the AM in detecting COVID-19 fake news. The self-attention mechanism performed better than the multi-attention mechanism in detecting Chinese COVID-19 news, but the difference was insignificant, indicating that our proposed CNN–BiLSTM–AM fake news detection has good robustness.

However, the model has the problem of overfitting in the process of increasing the complexity of the model to improve the fake news detection ability. Accordingly, we need to optimize the model in future work, which can be taken as follows: (1) reduce the model, (2) complexity L1 and L2 regularization, (3) drop out, (4) early stopping, (5) data cleaning, (6) use ensemble learning methods, and (7) batch normalization. Preventing the model from having different expected outputs on the test samples makes the model highly sensitive to the noise and individual differences of the sample data and does not reflect the non-intuitive data generation process.

4.5. Discussions

We propose a framework based on NLP and heterogeneous knowledge management, combined with domain knowledge, to detect and analyze fake news and reach the following conclusions.

First, we conduct keyword extraction and WordCloud map construction for real and fake news and analyze the difference between real and fake news. Second, we calculate the sentence length by adding up the words of each sentence. The median sentence length of authentic news is higher than that of fake news. Fake news is mainly published to simulate fake news.

Then, the difference in topic distribution between real and fake news is discussed. Differences in topic distribution can be observed between real and fake news. The topics in real news include not only the novel coronavirus but also the experimental test of the novel coronavirus and how to solve its different situations. Meanwhile, all the topics in fake news are COVID-19, such as the origin of the novel coronavirus, which is COVID-19. The topic variance of real news is higher than that of fake news. The research of real news has a wide range and distinct topics. We can judge real and fake news according to the different topics.

Finally, the model is adjusted by changing the self-attention mechanism and multi-attention mechanism. The result showed that the model of the self-attention mechanism is better than the multi-attention mechanism to verify the performance of the attention mechanism in COVID-19 fake news detection. Loss, Accuracy, F1-score, and Recall values of the CNN–BiLSTM–Self-attention model are higher than those of the CNN–BiLSTM–MultiHeadAttention model. Accordingly, CNN–BiLSTM–AM performs better when using the multi-attention mechanism for fake news detection. In the context of big data, fake news about COVID-19 is increasing; hence, the sustainable development of big data becomes the basis of fake news governance (Wu et al., 2016; Zheng et al., 2022b). The advanced intelligent analysis technology provides technical support for the detection of fake news. Information and communication technology will help in producing and disseminating new knowledge, improving people's ability to distinguish fake news of COVID-19, reducing the negative impression brought by social communication, and accelerating the all-round development of mankind (Wu et al., 2018).

5. Conclusions

This work detects fake COVID-19 news with domain knowledge and analyzes the sentence length and topic distribution of real and fake news of major emergencies by combining the process of "generation–spread–identification–refutation" of fake news from the perspective of outlier knowledge management. New outlier knowledge is added to the knowledge base of major emergencies, and wisdom is formed by analyzing outlier knowledge from three aspects: human, organization, and technology. This work can provide a reference for refuting false news in major emergencies in the future. Moreover, this work provides a new research perspective for fake news detection.

First, users spread fake news out of the following types of motivations:

- First, users think that fake news is fun, and they use them to attract others' attention and enhance communication. These users can become "tool people" and "accomplices" for those who want to spread fake news. The second type is the purposeful promotion of fake news to generate negative hot topics, causing public emotional tension and social unrest (Duffy et al., 2020). Our analysis is conducted based on the second type of motivation.
- Second, public social platforms lack relevant verification measures for fake news, such as validating news submissions and monitoring news sources. Laws, regulations, supervision policies to control fake news are lacking, and a severe crackdown is necessary to prevent fake news dissemination and create a social atmosphere of "Do not fabricate or spread rumors".
- Third, from the technical perspective, publishers of fake news frequently forge personal information to distribute fake news from the government's perspective. Flaws in public social platforms have led to the rapid spread of fake news, resulting in a sharp increase in negative hot topics and public tension. Fake news has a specific language and theme, and we can detect them from different technical perspectives.

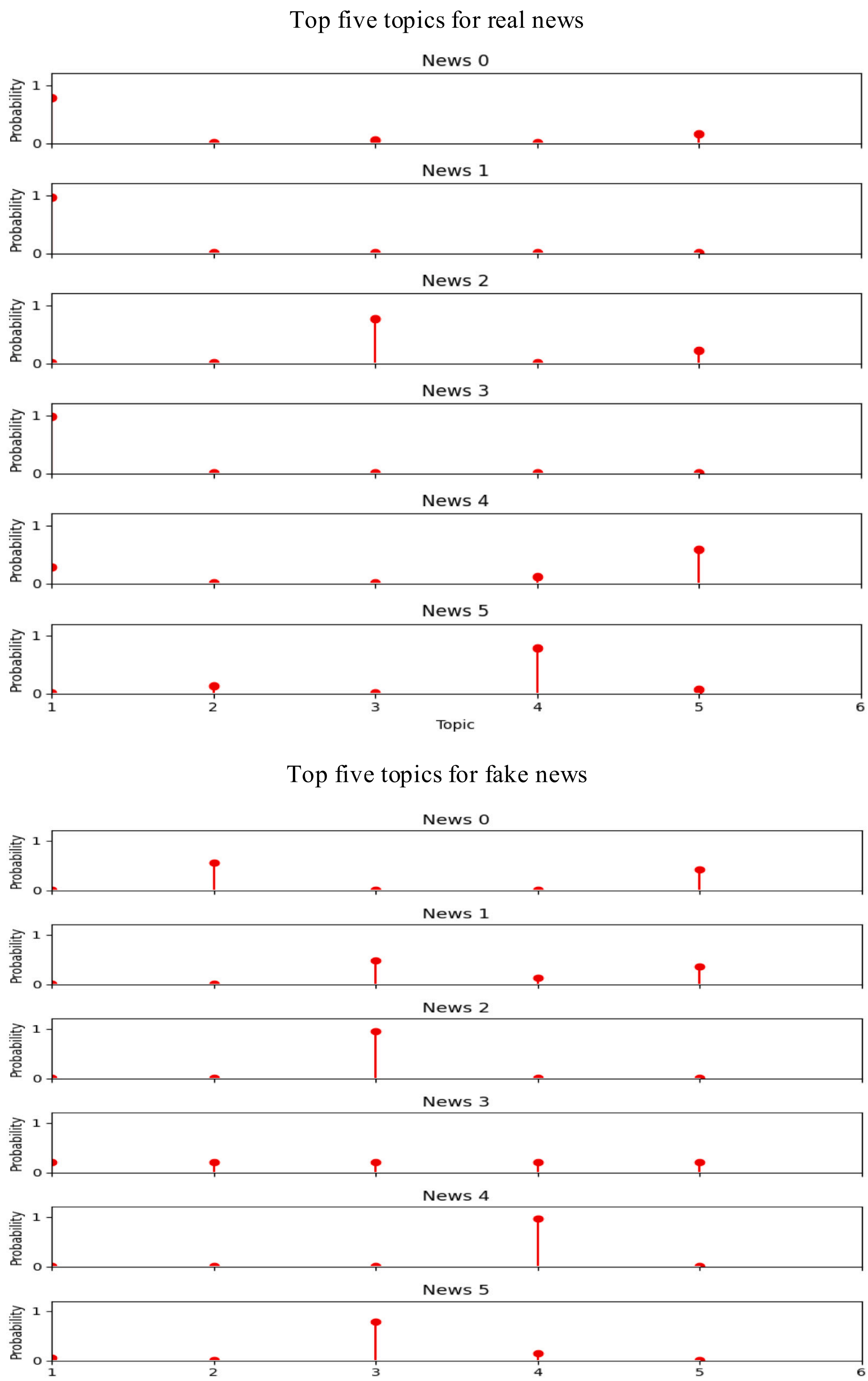


Fig. 11. Visualization of the first five news topics of true and false news.

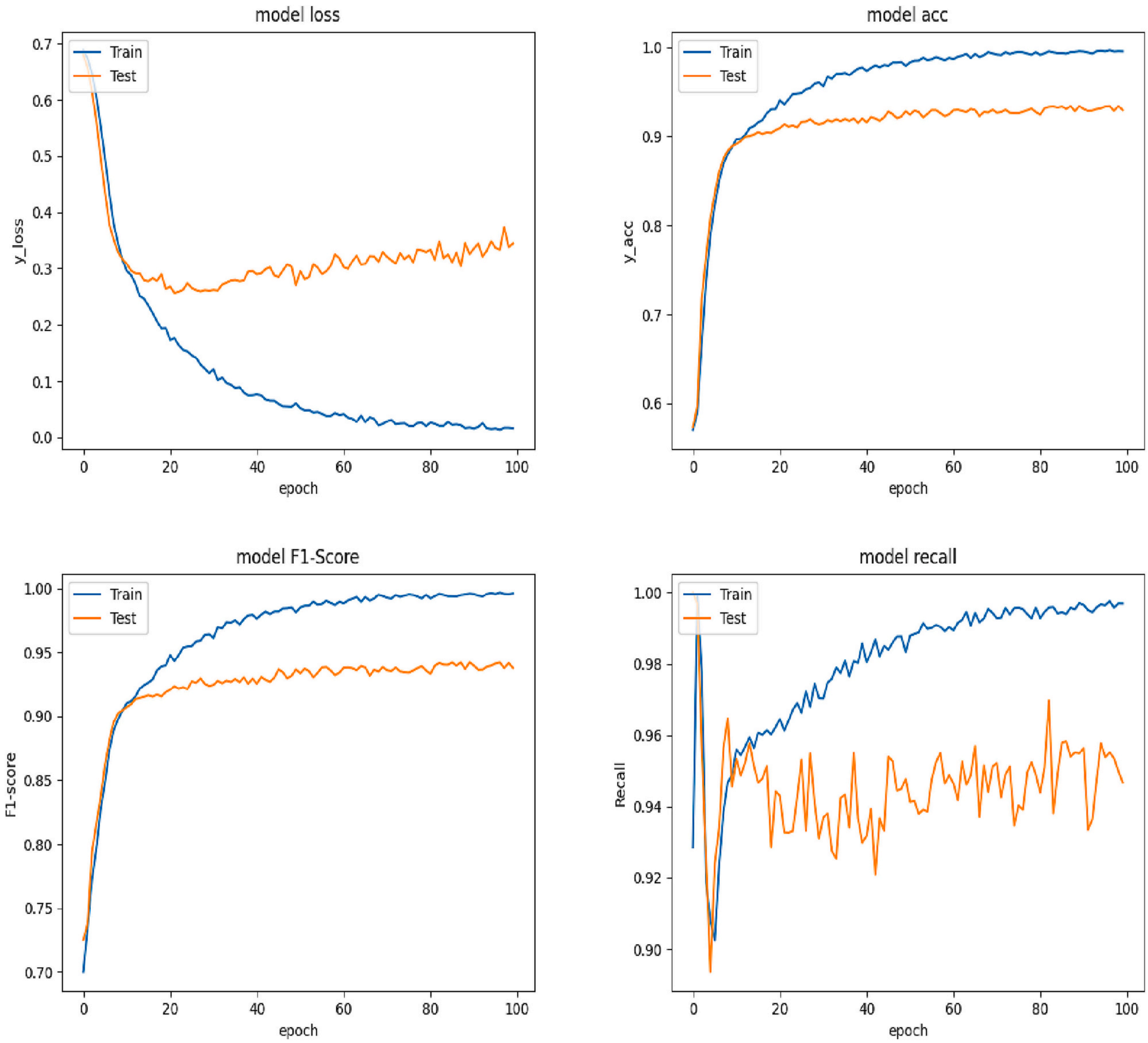


Fig. 12. Fake news detection and evaluation indicators of the CNN–BiLSTM–Self Attention model.

5.1. Theoretical contribution

The research framework of heterogeneous knowledge management was improved, and the detection of fake news in major emergencies was dealt with in combination with the process of fake news "generation–-spread–identification–refutation" (Xia et al., 2022). The outlier knowledge is obtained from four stages to avoid the loss of knowledge, especially outlier knowledge. In the context of extreme public health events, we obtain outlier knowledge by analyzing true and fake news and propose a think tank for extreme public health events prevention, control, and management based on COVID-19 outlier knowledge.

The complex adaptive system theory and information dissemination theory are used to help and improve the detection and decision analysis of fake news under major emergencies. The theoretical support of fake news detection is improved, and the theory is integrated into heterogeneous knowledge management, which enriches the application scope of heterogeneous knowledge management. The theory of complex adaptive systems emphasizes the initiative and adaptability of agents and analyzes it from the bottom up (Uhl-Bien and Arena, 2018; Silva and Guerrini, 2018). On the basis of this theory, this work obtains the outlier knowledge of fake news under the background of the COVID-19 epidemic and analyzes it from two aspects: news length and word usage, thereby providing a think tank for the prevention and control of fake news in major emergencies. Information diffusion theory emphasizes the process of information diffusion. We consider the cycle of fake news "generation–spread–identification–refutation" in detecting COVID-19 fake news. In major emergencies, fake news detection is an important data source to obtain outlier knowledge. This work uses the advanced NLP technology to perform data mining on COVID-19 content on fake news, obtain the outlier knowledge of COVID-19 fake news, and enrich and strengthen the prevention and control of extreme public health events based on the above-mentioned two theories.

The CNN–BiLSTM–AM model is constructed, and the domain knowledge is used to effectively identify COVID-19 fake news and improve the accuracy of fake news detection. When dealing with domain-specific (COVID-19) customized data, multi-attention mechanisms outperform self-attention mechanisms in single-text news detection (Trueman et al., 2021). Experiments show that the multi-attention mechanism has a robust feature retrieval ability. In the text, fake news detection, the Loss, Accuracy, F1, and recall indicators of the multi-attention mechanism are higher than those of the self-attention mechanism.

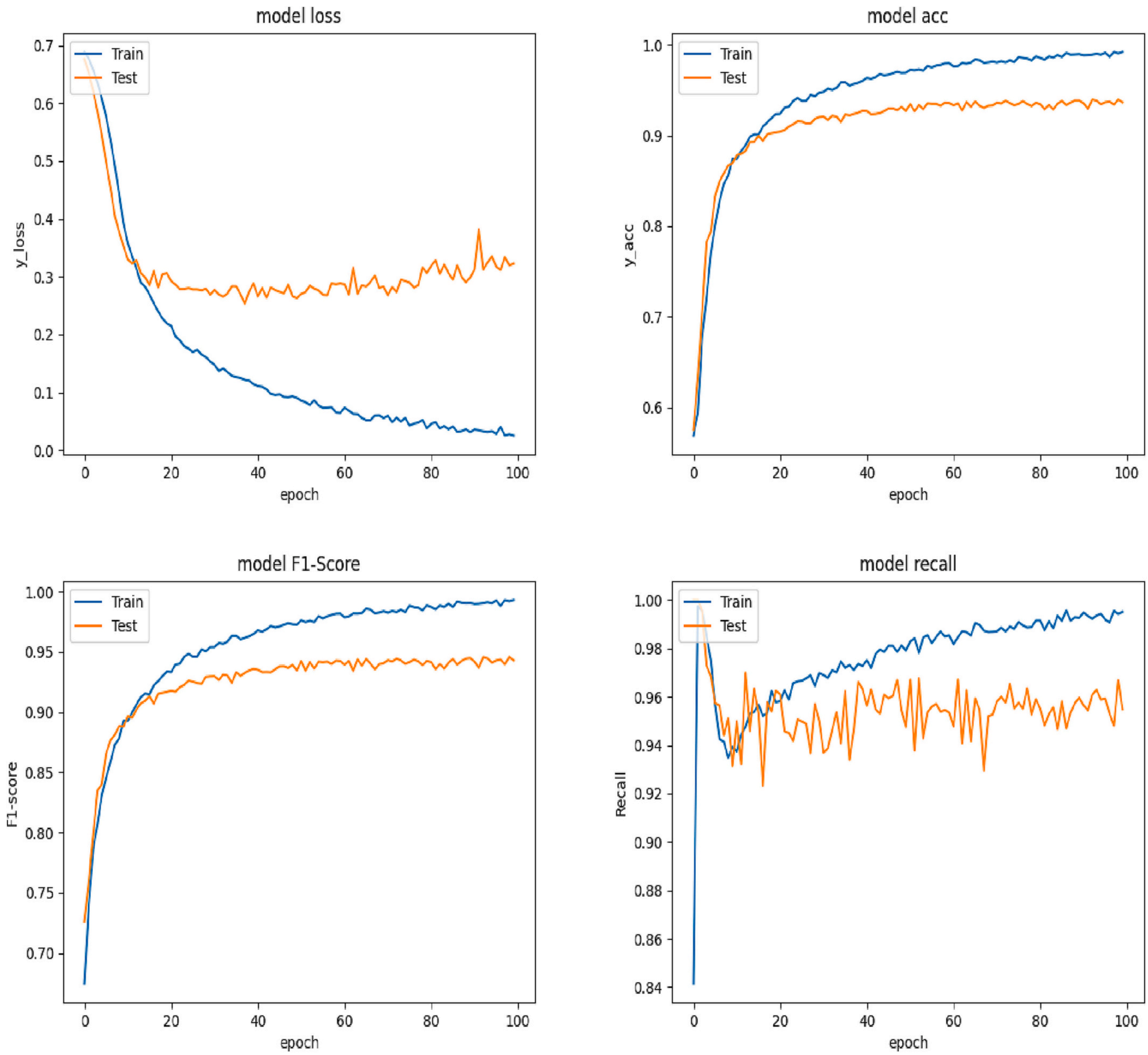


Fig. 13. Fake news detection and evaluation indicators of the CNN–BiLSTM–MultiHeadAttention model.

Table 8
Comparison between CNN–BiLSTM–AM and other advanced models in COVID-19 fake news detection.

Evaluation index	Accuracy_mean	F1-Score_mean	Recall_mean
K-NN (Al-Ahmad et al., 2021)	70.65 %	57.85 %	59.36 %
BERT (Heidari et al., 2021)	86.76 %	82.58 %	82.58 %
CNN (Srivastava et al., 2022)	91.98	91.95	91.96
LSTM (Li et al., 2021)	92.4 %	92.4 %	92.6 %
CNN–LSTM (Kaliyar et al., 2021)	90.2 %	90.3 %	90.4 %
BiLSTM–Attention (Karnyoto et al., 2022)	91.96 %	91.13 %	91.93 %
CNN–BiLSTM–SelfAttention	**92.926 %**	**93.772 %**	**94.679 %**
CNN–BiLSTM–MultiHeadAttention	**93.689 %**	**94.301 %**	**95.483 %**

5.2. Practical contribution

The research findings highlight that, in the case of major emergencies, scientific refutation of false news can promote good interaction between the government and the public, combined with the process of "generation–spread–identification–refutation" of fake news. We highlight the specific practical value of this study as follows.

Building a reliable and timely news release system is crucial to enhancing the interaction between the government and the public. The government must establish an official and reliable information release

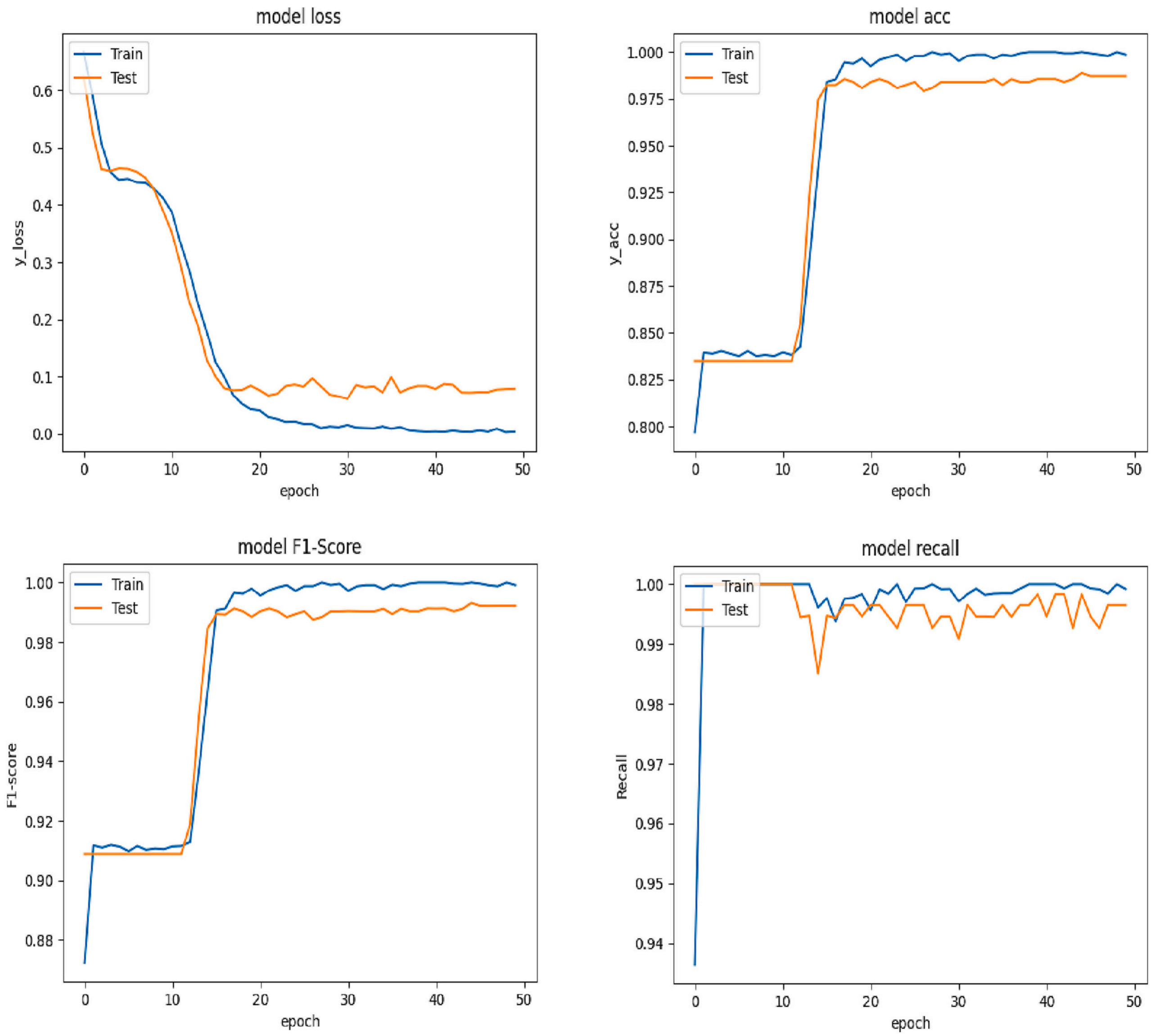


Fig. 14. CNN–BiLSTM–Self-attention robustness test evaluation index.

platform to reduce the harm of fake news to the public. When the outlier knowledge base is established, the value of outlier knowledge can be mined to provide the necessary basis for detecting fake news or other similar emergencies. The establishment of intelligent retrieval technology gives technical support to the public to identify fake news and acquire scientific defense and epidemic prevention and control knowledge.

Social media, as a platform for spreading fake news, must constantly improve its technology and use data mining to form outlier knowledge graphs for fake news analysis. When users upload or forward news, news detection services are provided to prevent users from becoming disseminators of fake news without their knowledge. Moreover, social platforms provide detection during news transmission. If fake news is found, then rumors will be immediately refuted to reduce the harm brought by fake news to society and the masses. Meanwhile, an automated intelligent epidemic fake news system can be established in organizational decision-making to provide the scientific basis for government decision-making.

A good ecosystem of public opinion can guide and educate people's behavior, improve people's understanding of extreme public events, and help people keep learning new knowledge. Individuals find it difficult to determine fake news because major emergencies happen without a scientific mechanism to explain them. Such outlier knowledge enables people to face challenges with scientific optimism and guide and tolerate the spread of fake news caused by ignorance and panic. People can improve their ability to scientifically identify true and false news and improve how science uses new media to express its claims. Phased governance measures will help in promptly eliminating fake news without the need for mass screening and removal of posts, saving social costs. Progress information should be promptly released to prevent rumors that need time to be verified from spreading and causing panic. We need to track down the sources of rumors with uncertain consequences, listen to opinions from all sides, promptly block and delete posts, and release persuasive positive information.

5.3. Limitations and future research directions

This works some limitations. For example, the analysis granularity is not detailed enough, and the analysis relies on domain knowledge. Meanwhile, our CNN–BiLSTM–AM model only identifies text when detecting fake news. We can consider expanding the data dimension to analyze fake news from pictures, sounds, videos, etc. Moreover, a certain degree of overfitting still occurs even though the effect of fake news detection is improved when the AM is added. Therefore, the model should be simplified in the future research process, and the

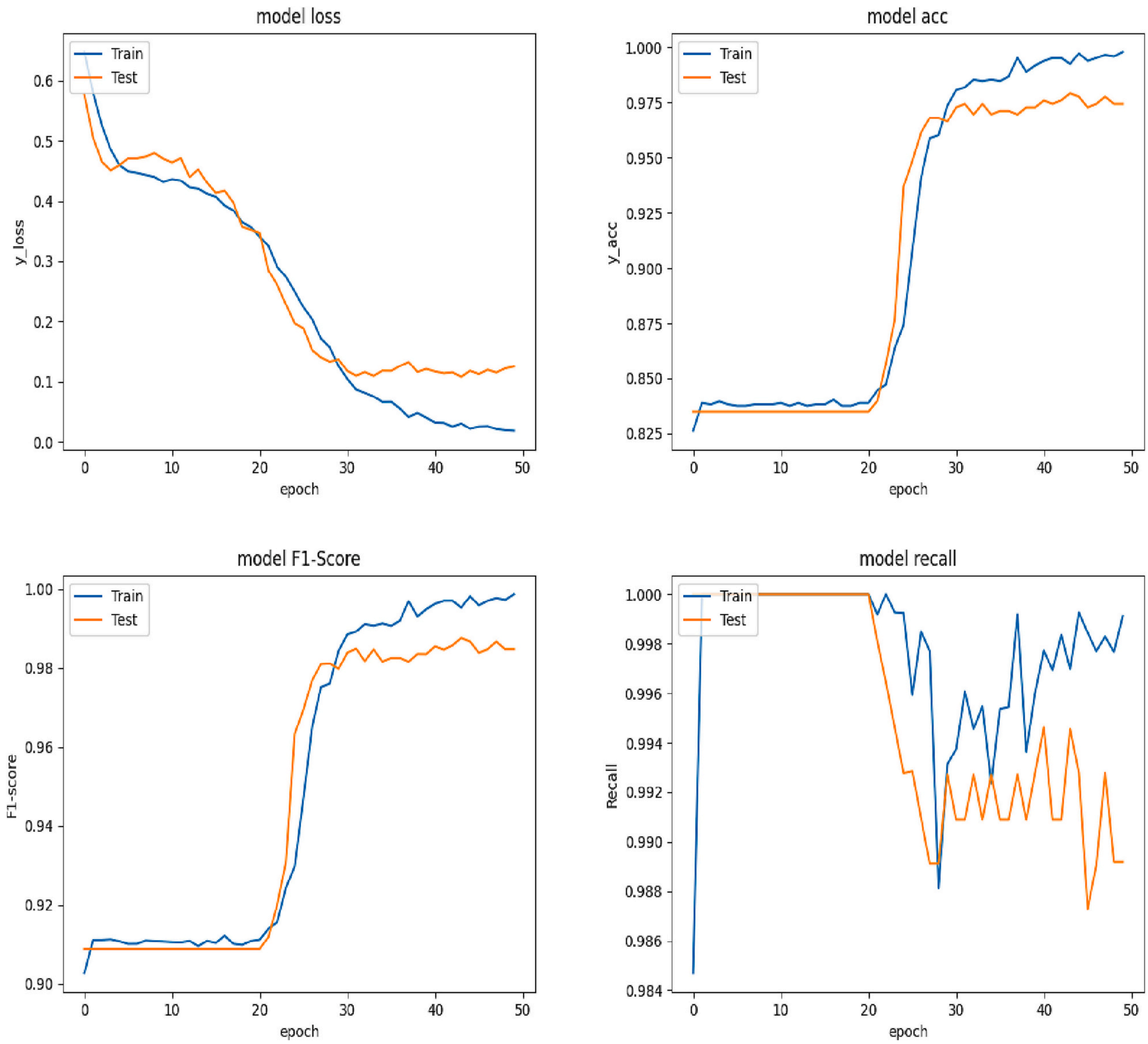


Fig. 15. CNN–BiLSTM–MultiHeadAttention robustness test evaluation index.

Table 9
CNN–BiLSTM–AM robustness test.

Evaluation index	Accuracy	F1-Score	Recall
CNN–BiLSTM–SelfAttention	98.718 %	99.214 %	99.651 %
CNN–BiLSTM–MultiHeadAttention	97.436 %	98.478 %	98.918 %

generalization ability of the model should be enhanced.

CRediT authorship contribution statement

Huosong Xia: Conceptualization, Supervision, Funding acquisition, Investigation, Methodology, Resources, Validation, Writing – original draft, Writing – review & editing, Project administration. **Yuan Wang:** Formal analysis, Software, Visualization, Investigation, Writing – original draft, Writing – review & editing. **Justin Zuopeng Zhang:** Supervision, Investigation, Validation, Writing – original draft, Writing – review & editing, Project administration. **Leven J. Zheng:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Muhammad Mustafa Kamal:** Conceptualization, Investigation, Validation, Writing – original draft, Writing – review & editing. **Varsha Arya:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgments

This research has been supported by the National Natural Science Foundation of China (NSFC, 72171184, title: Grey Private Knowledge model of security and trusted BI on the federal Learning Perspective); (NSFC, 71871172, title: Model of Risk knowledge acquisition and Platform governance in FinTech based on deep learning). We deeply appreciate the suggestions from fellow members of Xia's project team and Research center of Enterprise Decision Support, Key Research Institute of Humanities and Social Sciences in Universities of Hubei Province (DSS2023).

Appendix A. Notations

Notation	Mean	Notation	Mean
S	Each news item in the training data has a modal data form: Text	Y	Y is the domain name tag of the i news and is the news tag
D_M	A training set	D	Unknown data set
k	Number of cross-validation parameters	Head	Number of a head of multi-head AM
x_i	Word vector	N	Theme of the extraction
Ω	"Theme-word" probability distribution after the normalization of the mean	d	Dot product of the vectors

Appendix B. Abbreviations

Full name	Abbreviation	Full name	Abbreviation
Sars-Cov-2 Novel Coronavirus	COVID-19	Convolutional Neural Networks	CNN
Bidirectional Long Short-term Memory Networks	BiLSTM	Attention Mechanisms	AM
Convolutional Neural Network–Bidirectional Long Short-term Memory Network–Attention Mechanisms	CNN–BiLSTM–AM	Long Short-term Memory Networks	LSTM
Natural Language Processing	NLP	Artificial Intelligence	AI

References

Al-Ahmad, B., Al-Zoubi, A.M., Abu Khurma, R., Aljarah, I., 2021. An evolutionary fake news detection method for COVID-19 pandemic information. Symmetry 13 (6), 2–16.

Al-Azad, M.S., Mohiuddin, M., Su, Z., 2022. The client and service provider relationship in IT outsourcing project success: the moderating effects of organizational attitudes on knowledge sharing and partnership quality. J. Glob. Inf. Manag. 30 (1), 1–27.

Almomani, A., Alauthman, M., Shatnawi, M.T., Alweshah, et al., 2022. Phishing website detection with semantic features based on machine learning classifiers: a comparative study. Intern. J. Sem. Web Inform. Sys. 18 (1), 1–24.

Almomani, A., Al-Nawasrah, A., Alomoush, W., Al-Abweh, M., Alrosan, A., et. al., 2021. Information management and IoT technology for safety and security of smart home and farm systems. J. Glob. Inf. Manag. 29 (6), 1–23.

Al-Qirim, N., Rouibah, K., Abbas, H., Hwang, Y., 2022. Factors affecting the success of social commerce in Kuwaiti microbusinesses: a qualitative study. J. Glob. Inf. Manag. 30 (1), 1–31.

Al-Taie, M.Z., Nour, M.A., Cater-Steel, A., 2021. An empirical assessment of the relationship between information intensity and IT leaders' role and structural power. J. Glob. Inf. Manag. 29 (6), 1–25.

Amine, B.M., Drif, A., Giordano, S., 2019. Merging deep learning model for fake news detection. In: 2019 International Conference on Advanced Electrical Engineering (ICAEE). IEEE, pp. 1–4. November.

Apuke, O.D., Omar, B., 2020. Fake news proliferation in Nigeria: consequences, motivations, and prevention through awareness strategies. Humanit. Soc. Sci. Rev. 8 (2), 318–327.

Apuke, O.D., Omar, B., 2021. Fake news and COVID-19: modeling the predictors of fake news sharing among social media users. Telematics Inform. 56, 101475 https://doi.org/10.1016/j.tele.2020.101475.

Aslam, N., Ullah Khan, I., Alotaibi, F.S., Aldaej, L.A., Aldubaikil, A.K., 2021. Fake detect: a deep learning ensemble model for fake news detection. Complexity 2021. https://doi.org/10.1155/2021/5557784.

Bai, Y., Yi, J., Tao, J., Tian, Z., Wen, Z., Zhang, S., 2021. Fast end-to-end speech recognition via non-autoregressive models and cross-modal knowledge transferring from BERT. IEEE/ACM Trans. Audio Speech Lang. Process. 29, 1897–1911.

Balakrishnan, V., Ng, K.S., Rahim, H.A., 2021. To share or not to share–the underlying motives of sharing fake news amidst the COVID-19 pandemic in Malaysia. Technol. Soc. 66, 101676 https://doi.org/10.1016/j.techsoc.2021.101676.

Belhadi, A., Djenouri, Y., Srivastava, G., Djenouri, D., Lin, J.C.W., Fortino, G., 2021. Deep learning for pedestrian collective behavior analysis in smart cities: a model of group trajectory outlier detection. Inf. Fusion 65 (2021), 13–20.

Blázquez-García, A., Conde, A., Mori, U., Lozano, J.A., 2021. A review on outlier/anomaly detection in time series data. ACM Comput. Surv. 54 (3), 1–33.

Bok, S., Martin, D.E., Acosta, E., Lee, M., Shum, J., 2021. Validation of the COVID-19 transmission misinformation scale and conditional indirect negative effects on wearing a mask in public. Int. J. Environ. Res. Public Health 18 (2021), 11319. https://doi.org/10.3390/ijerph182111319.

Carmine, V., Pierluigi, T., Manta, F., Marco, B., 2021. The more you know, the better you act? Institutional communication in COVID-19 crisis management. Technol. Forecast. Soc. Chang. 170, 120929 https://doi.org/10.1016/j.techfore.2021.120929.

Carrion-Alvarez, D., Tijerina-Salina, P.X., 2020. Fake news in COVID-19: a perspective. Health Promot. Perspect. 10 (4), 290–291.

Carvajal-Miranda, C., Mañas-Viniegra, L., Liang, L., 2020. Online discourse in the context of COVID-19, the first health crisis in China after the advent of mobile social media: a content analysis of China's Weibo and Baidu. Soc. Sci. 9 (10), 167. https://doi.org/10.3390/socsci9100167.

Cheng, L., Hu, H., Wu, C., 2021. Spammer group detection using machine learning technology for observation of new spammer behavioral features. J. Glob. Inf. Manag. 29 (2), 61–76.

Chiang, T.A., Che, Z.H., Huang, Y.L., Tsai, C.Y., 2022. Using an ontology-based neural network and DEA to discover deficiencies of hotel services. Intern. J. Sem. Web Inform. Sys. 18 (1), 1–19.

Choudhary, A., Arora, A., 2021. Linguistic feature based learning model for fake news detection and classification. Expert Syst. Appl. 169, 114171 https://doi.org/10.1016/j.eswa.2020.114171.

Chui, K.T., Tsang, K.F., Chi, H.R., Ling, B.W.K., Wu, C.K., 2016. An accurate ECG-based transportation safety drowsiness detection scheme. IEEE Trans. Ind. Inf. 12 (4), 1438–1452.

Conroy, N.K., Rubin, V.L., Chen, Y., 2015. Automatic deception detection: methods for finding fake news. Proc. Assoc. Inf. Sci. Technol. 52 (1), 1–4.

Cui, Y., Zhu, J., Liu, Y., 2022. Exploring the social and systemic influencing factors of mobile short video applications on the consumer urge to buy impulsively. J. Glob. Inf. Manag. 30 (1), 1–23.

DeJong, S., 2023. Playing with fake news: state of fake news video games. Int. J. Games Soc. Impact 1 (1), 94–111.

Depoux, A., Martin, S., Karafillakis, E., Preet, R., Wilder-Smith, A., Larson, H., 2020. The pandemic of social media panic travels faster than the COVID-19 outbreak. J. Travel Med. 27 (3), 1–2.

Di Domenico, G., Sit, J., Ishizaka, A., Nunan, D., 2021. Fake news, social media and marketing: a systematic review. J. Bus. Res. 124 (2021), 329–341.

Duffy, A., Tandoc, E., Ling, R., 2020. Too good to be true, too good not to share: the social utility of fake news. Inf. Commun. Soc. 23 (13), 1965–1979.

Erkuş, E.C., Purutçuoğlu, V., 2021. Outlier detection and quasi-periodicity optimization algorithm: frequency domain based outlier detection (FOD). Eur. J. Oper. Res. 291 (2), 560–574.

Eysenbach, G., 2020. How to fight an infodemic: the four pillars of infodemic management. J. Med. Internet Res. 22 (6), e21820 https://doi.org/10.2196/21820.

Gaurav, A., Gupta, B.B., Hsu, C.H., Castiglione, A., Chui, K.T., 2021. Machine learning technique for fake news detection using text-based word vector representation. In: International Conference on Computational Data and Social Networks. Springer, Cham, pp. 340–348. November.

Goel, L., Donaldson, J., 2021. Social media exploration and exploitation by small and medium enterprises for business continuity. J. Glob. Inf. Manag. 29 (6), 1–15.

Hanshal, O.A., Ucan, O.N., Sanjalawe, Y.K., 2022. Hybrid deep learning model for automatic fake news detection. Appl. Nanosci. 1-11 https://doi.org/10.1007/s13204-021-02330-4.

Haran, V.V., Niederman, F., 2022. Social context of turnover-mixed methods study of Indian IT professionals. J. Glob. Inf. Manag. 30 (1), 1–24.

Hartley, K., Vu, M.K., 2020. Fighting fake news in the COVID-19 era: policy insights from an equilibrium model. Policy. Sci. 53 (4), 735–758.

Heidari, M., Zad, S., Hajibabaee, P., Malekzadeh, M., HekmatiAthar, S., Uzuner, O., Jones, J.H., 2021. Bert model for fake news detection based on social bot activities in the COVID-19 pandemic. In: 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, pp. 0103–0109. December.

Huang, S., Wang, L., 2022. An exploration study on the consequence of the COVID-19 pandemic on online Q&A communities. J. Glob. Inf. Manag. 30 (1), 1–22.

Huynh, T.L., 2020. The COVID-19 risk perception: a survey on socioeconomics and media attention. Econ. Bull. 40 (1), 758–764.

Islam, A.N., Laato, S., Talukder, S., Sutinen, E., 2020. Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective. Technol. Forecast. Soc. Chang. 159, 120201 https://doi.org/10.1016/j.techfore.2020.120201.

Jain, A.K., Gupta, B.B., 2019. A machine learning based approach for phishing detection using hyperlinks information. J. Ambient. Intell. Humaniz. Comput. 10 (5), 2015–2028.
Jeon, D., Ahn, J.M., Kim, J., Lee, C., 2022. A doc2vec and local outlier factor approach to measuring the novelty of patents. Technol. Forecast. Soc. Chang. 174, 121294.
Ji, S., Satish, N., Li, S., Dubey, P.K., 2019. Parallelizing word2vec in shared and distributed memory. IEEE Trans. Parallel Distrib. Syst. 30 (9), 2090–2100.
Jin, X., Wang, J., Tan, J., Li, Q., 2022. Media platforms and stock performance: evidence from internet news. J. Glob. Inf. Manag. 30 (1), 1–25.
Jing, J., Wu, H., Sun, J., Fang, X., Zhang, H., 2023. Multimodal fake news detection via progressive fusion networks. Inf. Process. Manag. 60 (1), 103120.
Kaliyar, R.K., Goswami, A., Narang, P., 2021. A hybrid model for effective fake news detection with a novel COVID-19 dataset. In: Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART), 2, pp. 1066–1072. January.
Kamal, M.M., 2020. The triple-edged sword of COVID-19: understanding the use of digital technologies and the impact of productive, disruptive and destructive nature of the pandemic. Inf. Syst. Manag. 37 (4), 310–317.
Karnyoto, A.S., Sun, C., Liu, B., Wang, X., 2022. Transfer learning and GRU-CRF augmentation for COVID-19 fake news detection. Comput. Sci. Inf. Syst. 19 (00), 53. https://doi.org/10.2298/CSIS210501053K.
Kesharwani, A., Mani, V., Gaur, J., Wamba, S.F., Kamble, S.S., 2021. Service quality measurement in information systems: An expectation and desire disconfirmation approach. J. Glob. Inf. Manag. 29 (6), 1–19.
Khan, S., Hakak, S., Deepa, N., Prabadevi, B., Dev, K., Trelova, S., 2022. Detecting COVID-19-related fake news using feature extraction. Front. Public Health 9, 788074. https://doi.org/10.3389/fpubh.2021.788074.
Kim, J., Aum, J., Lee, S., Jang, Y., Park, E., Choi, D., 2021. FibVID: comprehensive fake news diffusion dataset during the COVID-19 period. Telematics Inform. 64, 101688 https://doi.org/10.1016/j.tele.2021.101688.
Kucharski, A., 2016. Study epidemiology of fake news. Nature 540 (7634), 525.
Laato, S., Islam, A.N., Islam, M.N., Whelan, E., 2020. What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? Eur. J. Inf. Syst. 29 (3), 288–305.
Lampos, V., Majumder, M.S., Yom-Tov, E., Edelstein, M., Moura, S., Hamada, Y., Cox, I. J., 2021. Tracking COVID-19 using online search. NPJ Digit. Med. 4 (1), 1–11.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Li, C., Chu, J., Zheng, L.J., 2022a. Better not let me know: consumer response to reported misuse of personal data in privacy regulation. J. Glob. Inf. Manag. 30 (1), 1–22.
Li, H.O.Y., Bailey, A., Huynh, D., Chan, J., 2020. YouTube as a source of information on COVID-19: a pandemic of misinformation? BMJ Glob. Health 5 (5), e002604. https://doi.org/10.1136/bmjgh-2020-002604.
Li, Q., Shaw, N., Shi, V., 2022b. Exploring communication and networks using social network analysis: a case study of a utility company's asset management programme. J. Glob. Inf. Manag. 30 (1), 1–25.
Li, X., Xia, Y., Long, X., Li, Z., Li, S., 2021. Exploring text-transformers in AAAI 2021 shared task: Covid-19 fake news detection in English. In: International Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation. Springer, Cham, pp. 106–115. February.
Liang, X., Ruan, W., Xu, Z., Liu, J., 2022a. Analysis of safe storage of network information data and financial risks under blockchain combined with edge computing. J. Glob. Inf. Manag. 30 (11), 1–20.
Liang, X., Zhang, X., Paulet, R., Zheng, L.J., 2022b. A literature review of the COVID-19 pandemic's effect on sustainable HRM. Sustainability 14 (5), 2579.
Marco-Franco, J.E., Pita-Barros, P., Vivas-Orts, D., González-de-Julián, S., Vivas-Consuelo, D., 2021. COVID-19, fake news, and vaccines: should regulation be implemented? Int. J. Environ. Res. Public Health 18 (2), 744. https://doi.org/10.3390/ijerph18020744.
Marcot, B.G., Hanea, A.M., 2021. What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis? Comput. Stat. 36 (3), 2009–2031.
Meadows, C.Z., Tang, L., Zou, W., 2021. Managing government legitimacy during the COVID-19 pandemic in China: a semantic network analysis of state-run media Sina Weibo posts. Chin. J. Commun. https://doi.org/10.1080/17544750.2021.2016876.
Molina, L., Arroba, G., Viteri-Guevara, X., Tigua-Moreira, S., Clery, A., Orrala, L., Figueroa-Martínez, E., 2023. Spread of fake news about COVID: the Ecuadorian case. In: Proceedings of SAI Intelligent Systems Conference. Springer, Cham, pp. 672–678.
Mondal, A., Chakrabarti, A.B., 2021. Information and communication technology adoption strategies of emerging multinationals from India. J. Glob. Inf. Manag. 29 (5), 161–175.
Muhammad, S.S., Dey, B.L., Syed Alwi, S.F., Kamal, M.M., Asaad, Y., 2022. Consumers' willingness to share digital footprints on social media: the role of affective trust. Inf. Technol. People. https://doi.org/10.1108/ITP-10-2020-0694.
Mutambik, I., Almuqrin, A., Liu, Y., Alhossayin, M., Qintash, F.H., 2021. Gender differentials on information sharing and privacy concerns on social networking sites: perspectives from users. J. Glob. Inf. Manag. 29 (3), 236–255.
Nam, T., Kabutey, R., 2021. How does social media use influence the relationship between emotional labor and burnout?: the case of public employees in Ghana. J. Glob. Inf. Manag. 29 (4), 32–52.
Narayan, N., Jha, R.K., Singh, A., 2022. A differential epidemic Model for Information, Misinformation, and disinformation in online Social Networks: CoVId-19 Vaccination. Intern. J. Sem. Web Inform. Sys. 18 (1), 1–20.
Ng, L.H., Taeihagh, A., 2021. How does fake news spread? Understanding pathways of disinformation spread through APIs. Policy Internet 13 (4), 560–585.
Ofcom, 2020. Half of UK adults exposed to false claims about coronavirus. DOI. https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/half-of-uk-adults-exposed-to-false-claims-about-coronavirus.
Paka, W.S., Bansal, R., Kaushik, A., Sengupta, S., Chakraborty, T., 2021. Cross-SEAN: a cross-stitch semi-supervised neural attention model for COVID-19 fake news detection. Appl. Soft Comput. 107, 107393 https://doi.org/10.1016/j.asoc.2021.107393.
Patwa, P., Sharma, S., Pykl, S., Guptha, V., Kumari, G., Akhtar, M.S., Chakraborty, T., 2021. Fighting an infodemic: COVID-19 fake news dataset. In: International Workshop on Combating Online Hostile Posts in Regional Languages During Emergency Situation. Springer, Cham, pp. 21–29. February.
Peng, G., Chen, S., Chen, X., Liu, C., 2021. An investigation to the Industry 4.0 readiness of manufacturing enterprises: the ongoing problems of information systems strategic misalignment. J. Glob. Inf. Manag. 29 (6), 1–20.
Pennycook, G., Rand, D.G., 2020. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. J. Pers. 88 (2), 185–200.
Pennycook, G., McPhetres, J., Zhang, Y., Lu, J.G., Rand, D.G., 2020. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol. Sci. 31 (7), 770–780.
Pham, T., Tao, X., Zhang, J., Yong, J., Li, Y., Xie, H., 2022. Graph-based multi-label disease prediction model learning from medical data and domain knowledge. Knowl.-Based Syst. 235, 107662 https://doi.org/10.1016/j.knosys.2021.107662.
Rahman, S., Islam, M.Z., Hossain, M., Abdullah, D.A., Jasimuddin, S.M., 2022. Organizational factors, ICT support, and affective commitment: the case of Bangladesh-based service organizations. J. Glob. Inf. Manag. 30 (1), 1–18.
Ram, J., Zhang, C., 2021. Examining the role of social media analytics in providing competitive intelligence: the impacts and limitations. J. Glob. Inf. Manag. 29 (6), 1–18.
Rodríguez, C.P., Carballido, B.V., Redondo-Sama, G., Guo, M., Ramis, M., Flecha, R., 2020. False news around COVID-19 circulated less on Sina Weibo than on Twitter. How to overcome false information? Int. Multidiscip. J. Soc. Sci. 9 (2), 107–128.
Rouibah, K., Qurban, H., Al-Qirim, N., 2022. Impact of risk perceptions and user trust on intention to re-use E-government: a mixed method research. J. Glob. Inf. Manag. 30 (1), 1–29.
Sahoo, S.R., Gupta, B.B., 2021. Multiple features based approach for automatic fake news detection on social networks using deep learning. Appl. Soft Comput. 100, 106983 https://doi.org/10.1016/j.asoc.2020.106983.
Sample, C., Jensen, M.J., Scott, K., McAlaney, J., Fitchpatrick, S., Brockinton, A., Ormrod, A., 2020. Interdisciplinary lessons learned while researching fake news. Front. Psychol. 11, 537612 https://doi.org/10.3389/fpsyg.2020.537612.
Schwitzer, G., 2017. Pollution of health news. BMJ 356.
Shankar, K., Perumal, E., Elhoseny, M., Taher, F., et al., 2021. Synergic deep learning for smart health diagnosis of COVID-19 for connected living and smart cities. ACM Trans. Internet Technol. 22 (3), 1–14.
Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., 2018. The spread of low-credibility content by social bots. Nat. Commun. 9 (1), 1–9.
Shao, C., Zheng, S., Gu, C., Hu, Y., Qin, X., 2022. A novel outlier detection method for monitoring data in dam engineering. Expert Syst. Appl. 193, 116476 https://doi.org/10.1016/j.eswa.2021.116476.
Sharma, S., Sharma, D.K., 2019. Fake news detection: a long way to go. In: 2019 4th International Conference on Information Systems and Computer Networks (ISCON). IEEE, pp. 816–821. November.
Shih, H., Lai, K., Guo, X., Cheng, T.C., 2021. Believe it or not: employees intend to comply with information security policy because of the desire for trade-offs of ISP compliance. J. Glob. Inf. Manag. 29 (6), 1–18.
Shu, K., Wang, S., Lee, D., Liu, H., 2020. Mining disinformation and fake news: concepts, methods, and recent advancements. In: Disinformation, Misinformation, and Fake News in Social Media. Springer, Cham, pp. 1–19.
Silva, A.L., Guerrini, F.M., 2018. Self-organized innovation networks from the perspective of complex systems: a comprehensive conceptual review. J. Organ. Chang. Manag. 31 (5), 962–983.
Singla, A., Gupta, N., Aeron, P., Jain, A., Sharma, D., Bharadwaj, S.S., 2022. Decentralized identity management using blockchain: cube framework for secure usage of IS resources. J. Glob. Inf. Manag. 31 (2), 1–24.
Srivastava, P.R., Eachempati, P., 2021. Intelligent employee retention system for attrition rate analysis and churn prediction: an ensemble machine learning and multi-criteria decision-making approach. J. Glob. Inf. Manag. 29 (6), 1–29.
Srivastava, S., Raj, R., Saumya, S., 2022. COVID-19 fake news identification using multi-layer convolutional neural network. In: Advanced Computational Paradigms and Hybrid Intelligent Computing. Springer, Singapore, pp. 149–157.
Tandoc Jr., E.C., Lim, Z.W., Ling, R., 2018. Defining "fake news": a typology of scholarly definitions. Digit. Journal. 6 (2), 137–153.
Tang, J., Liu, D., Jin, X., Peng, Y., Zhao, Q., Ding, Y., Kong, W., 2022. BAFN: bi-direction attention based fusion network for multimodal sentiment analysis. IEEE Trans. Circuits Syst. Video Technol. 1–13. https://doi.org/10.1109/TCSVT.2022.3218018.
Tang, X., Zhong, T., 2022. Research on COVID-19 prevention and control model based on evolutionary games. J. Glob. Inf. Manag. 30 (10), 1–17.
Tchuente, D., 2022. User modeling and profiling in information systems: a bibliometric study and future research directions. J. Glob. Inf. Manag. 30 (1), 1–25.
Tembhurne, J.V., Almin, M.M., Diwan, T., 2022. Mc-DNN: Fake news detection using multi-channel deep neural networks. Intern. J. Sem. Web Inform. Sys. 18 (1), 1–20.
Trueman, T.E., Kumar, A., Narayanasamy, P., Vidya, J., 2021. Attention-based C-BiLSTM for fake news detection. Appl. Soft Comput. 110, 107600 https://doi.org/10.1016/j.asoc.2021.107600.
Uhl-Bien, M., Arena, M., 2018. Leadership for organizational adaptability: a theoretical synthesis and integrative framework. Leadersh. Q. 29 (1), 89–104.

van Rensburg, S.K.J., 2021. End-user perceptions on information security: pragmatic lessons on social engineering attacks in the workplace in Gauteng, South Africa. J. Glob. Inf. Manag. 29 (6), 1–16.
Varma, R., Verma, Y., Vijayvargiya, P., Churi, P.P., 2021. A systematic survey on deep learning and machine learning approaches of fake news detection in the pre-and post-COVID-19 pandemic. Int. J. Intell. Comput. Cybern. 14 (4), 617–646. https://doi.org/10.1108/IJICC-04-2021-0069.
Varshney, D., Vishwakarma, D.K., 2021. A review on rumour prediction and veracity assessment in online social network. Expert Syst. Appl. 168, 114208 https://doi.org/10.1016/j.eswa.2020.114208.
Wahid, A., Annavarapu, C.S.R., 2021. NaNOD: a natural neighbour-based outlier detection algorithm. Neural Comput. & Applic. 33 (6), 2107–2123.
Wang, H., Zheng, L.J., Xu, X., Hung, T.H.B., 2022. Impact of financial digitalization on organizational performance: a look at the dark side. J. Glob. Inf. Manag. 30 (1), 1–35.
Wang, X.W., Riaz, M., Haider, S., Alam, K.M., Sherani, Yang, M., 2021. Information sharing on social media by multicultural individuals: experiential, motivational, and network factors. J. Glob. Inf. Manag. 29 (6), 1–25.
Wani, A., Joshi, I., Khandve, S., Wagh, V., Joshi, R., 2021. Evaluating deep learning approaches for COVID19 fake news detection. In: International Workshop on Combating Online Hostile Posts in Regional Languages During Emergency Situation. Springer, Cham, pp. 153–163. February.
Waszak, P.M., Kasprzycka-Waszak, W., Kubanek, A., 2018. The spread of medical fake news in social media–the pilot quantitative study. Health Policy Technol. 7 (2), 115–118.
Wu, J., Guo, S., Li, J., Zeng, D., 2016. Big data meet green challenges: big data toward green applications. IEEE Syst. J. 10 (3), 888–900.
Wu, J., Guo, S., Huang, H., Liu, W., Xiang, Y., 2018. Information and communications technologies for sustainable development goals: state-of-the-art, needs and perspectives. IEEE Commun. Surv. Tutorials 20 (3), 2389–2406.
Xia, H., An, W., Li, J., Zhang, Z.J., 2022. Outlier knowledge management for extreme public health events: understanding public opinions about COVID-19 based on microblog data. Socio Econ. Plan. Sci. 80, 100941 https://doi.org/10.1016/j.seps.2020.100941.
Xia, H.S., Long, J., Li, F., He, T.T., 2017. Outlier detection and outlier knowledge based on high frequency keywords: a study from the literature analysis perspective. J. Intell. 36 (5), 181–186.
Xing, F., Peng, G., Wang, J., Li, D., 2022. Critical obstacles affecting adoption of industrial big data solutions in smart factories: an empirical study in China. J. Glob. Inf. Manag. 30 (1), 1–21.
Xing, Y., Wang, X., Wang, F., Shi, Y., He, W., Chang, H., 2021. Ranking influential nodes of fake news spreading on mobile social networks. J. Glob. Inf. Manag. 29 (4), 93–130.
Yang, C., Zhou, X., Zafarani, R., 2021. CHECKED: Chinese COVID-19 fake news dataset. Soc. Netw. Anal. Min. 11 (1), 1–8.
Yang, J., Tian, Y., 2021. "Others are more vulnerable to fake news than I Am": third-person effect of COVID-19 fake news on social media users. Comput. Hum. Behav. 125, 106950 https://doi.org/10.1016/j.chb.2021.106950.
Yu, Z., He, L., Luo, W., Tse, R., Pau, G., 2022. Deep learning hybrid models for COVID-19 prediction. J. Glob. Inf. Manag. 30 (10), 1–20.
ZareRavasan, A., Krčál, M., 2021. A systematic literature review on 30 years of empirical research on information systems business value. J. Glob. Inf. Manag. 29 (6), 1–37.
Zhang, X., Ghorbani, A.A., 2020. An overview of online fake news: characterization, detection, and discussion. Inf. Process. Manag. 57 (2), 102025 https://doi.org/10.1016/j.ipm.2019.03.004.
Zhang, Y., Ge, L., Xiao, L., Zhang, M., Liu, S., 2021. A bibliometric review of information systems research from 1975–2018: setting an agenda for IS research. J. Glob. Inf. Manag. 29 (6), 1–24.
Zheng, K., Zheng, L.J., Gauthier, J., Zhou, L., Xu, Y., Behl, A., Zhang, J.Z., 2022c. Blockchain technology for enterprise credit information sharing in supply chain finance. J. Innov. Knowl. 7 (4), 100256.
Zheng, L.J., Fan, Y., Wang, H., Liu, W., 2021a. Born innovator? How founder birth order influences product innovation generation and adoption in entrepreneurial firms. J. Bus. Res. 136, 414–430.
Zheng, L.J., Xiong, C., Chen, X., Li, C.S., 2021b. Product innovation in entrepreneurial firms: how business model design influences disruptive and adoptive innovation. Technol. Forecast. Soc. Chang. 170, 120894.
Zhang, Q., Guo, Z., Zhu, Y., Vijayakumar, P., Castiglione, A., et al., 2023. A deep learning-based fast fake news detection model for cyber-physical social services. Pattern Recognit. Lett. 168, 31–38.
Zheng, L.J., Bai, T., Cross, A.R., 2021c. Signaling information management in entrepreneurial firms' financing acquisition: an integrated signaling and screening perspective. J. Glob. Inf. Manag. 29 (6), 1–31.
Zheng, L.J., Zhang, Y., Zhan, W., Sharma, P., 2022a. How B2B relationships influence new product development in entrepreneurial firms? The role of psychological tension. J. Bus. Res. 139, 1451–1462.
Zheng, L.J., Zhang, J.Z., Wang, H., Hong, J.F., 2022b. Exploring the impact of Big Data Analytics Capabilities on the dual nature of innovative activities in MSMEs: a Data-Agility-Innovation Perspective. Ann. Oper. Res. 1–29.
Zheng, W., Cheng, J., Wu, X., Sun, R., Wang, X., Sun, X., 2022d. Domain knowledge-based security bug reports prediction. Knowl.-Based Syst. 241, 108293 https://doi.org/10.1016/j.knosys.2022.108293.

Huosong Xia is a professor in the School of Management of the Wuhan Textile University. He graduated from Huazhong University of Science and Technology of China and was a visiting scholar at Eller College of Management of the University of Arizona from 2006 to 2007. His main research interests are knowledge management, data mining, e-commerce, and logistics information systems. His publications have appeared in over 100 papers in refereed journals, book chapters, and conferences. He has received 5 projects of research funding, such as the National Social Science Foundation of China and the National Science Foundation of China.

Yuan Wang is a graduate student in the School of Management of Wuhan Textile University. His main research interests are fintech, knowledge management, data mining, and e-commerce.

Justin Zuopeng Zhang is a faculty member in the Coggin College of Business at the University of North Florida. He received his Ph.D. in Business Administration with a concentration on Management Science and Information Systems from Pennsylvania State University, University Park. His research interests include economics of information systems, knowledge management, electronic business, business process management, information security, and social networking. He is the Editor-in-Chief of the *Journal of Global Information Management*, an ABET programme evaluator, and an IEEE senior member. He has published in *Technological Forecasting & Social Change*, *International Journal of Information Management*, *Decision Support Systems*, *Industrial Marketing Management*, *Information & Management*, *Knowledge Management Research & Practice*, *Journal of Knowledge Management*, *IEEE Transactions on Engineering Management*, *Expert Systems and Applications*, *Electronic Markets*, *Production Planning and Control*, *Annals of Operations Research*, *Journal of the Operational Research Society*, and many others leading journals.

Leven J. Zheng is an Assistant Professor in Management at the Department of Global Business and Marketing, Lee Shau Kee School of Business and Administration, Hong Kong Metropolitan University. Leven was awarded a Ph.D. in Innovation and Entrepreneurship (Management) from the Management School at the University of Liverpool. His research examines entrepreneurship and innovation at both privately entrepreneurial and newly public stages of entrepreneurial firms, with a particular focus on new product innovation, firm growth rate and growth actions, and strategic management of resources and innovation. His research has appeared in some good-quality international journals such as *Journal of Business Research*, *Technovation*, *Technological Forecasting & Social Change*, and *Asian Business & Management*. He uses both qualitative and quantitative research methods.

Muhammad Mustafa Kamal is an Associate Professor in Supply Chain Management, Curriculum Lead (Subject Head) in Business Analytics and Digital Transformation, and Director of the Structured PhD (Online) Programme at the School of Strategy and Leadership, Coventry University. Prior to joining Coventry University, he was a Senior Lecturer in Operations and Supply Chain Management, Director of UG Programmes and Director of Alumni at the Brunel Business School, Brunel University London. He earned his Doctorate in Information Systems from Brunel University London. His areas of specialism are Supply Chain Digitalisation. Other research interests include Circular Economy, Operations Management, Industry 4.0/5.0, Disruptive Technologies, Information Systems and Technology Management, Social Media, Big Data and Business Analytics. Currently, he is the Deputy Editor for the *Journal of Enterprise Information Management*, Senior Editor for *Information Technology and People* and *Information Systems Management*, Editorial Board Member for *International Journal of Information Management* and member of the International Advisory Board for *Journal of Global Information Management*. He has published over 80 papers in refereed academic journals, conference proceedings, book chapters, and magazine articles. His research work has appeared in several leading ABS ranked journals, such as *International Journal of Production Research*, *Journal of Business Research*, *Information and Management, Production Planning and Control*, *Expert Systems with Applications*, *Government Information Quarterly*, *Supply Chain Management: An International Journal*, *Journal of the Operational Research Society*, *International Journal of Information Management*, and *Information Systems Management*. He has also presented his research papers at globally esteemed conferences such as BAM, ECIS, AMCIS, and HICSS. He has worked on several internal and external research grant (including EU Framework 7, Horizon 2020, Qatar National Research Foundation, Erasmus+, and British Council).

Varsha Arya obtained a Master's degree from Rajasthan University, India in 2015 and has been working as a researcher for the last 7 years. She published several papers in journals and conferences. Her research interests include business administration, technology management, Cyber-physical systems, cloud computing, healthcare, and networking.