Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **2**, May **2024**

False Information Detection Based on Deep Learning Model in News Communication

Jian-Qiang Wang*

Henan Institute of Economics and Trade
Zhengzhou 450018, P. R. China
hnjmwjq@163.com

Jing-Jing Guo

Henan Institute of Economics and Trade
Zhengzhou 450018, P. R. China
guojingjing202306@163.com

Tai-Ping Liu

College of Science and Technology
Adamson University
Ermita 1000, Manila, Philippines
uw5520@163.com

*Corresponding author: Jian-Qiang Wang

Received February 21, 2023, revised March 24, 2023, accepted May 20, 2023.

Abstract. *In response to the problems of existing news detection methods not considering their own effective features and low efficiency, this article offers a counterfeit message detection algorithm on the ground of deep learning models in news communication. Firstly, for the purpose of preventing the deep learning model from overly relying on the training set, this article offers a method of re dividing the training set on the ground of news information features and combining them. Then, a multi feature false information detection framework is constructed in news communication, utilizing the probable information generated during the dissemination process to provide enhanced detection methods for traditional multi feature false information detection. Secondly, self attention mechanism and graph neural network are used to further enhance the learning and attention to long tail nodes. The feature vectors obtained through convolution operations are then extracted to the maximum value through maximum pooling. On the ground of memory enhancement, the final event is entered into the counterfeit message detector for prediction. Finally, the performance of the model was verified on the Twitter 15 dataset. The experimental outcome manifested that the accuracy, precision , and recall of the suggested model in detecting false information were 0.938, 0.952, and 0.943, respectively. Contrasted with other algorithms, it greatly enhanced the accuracy, precision, and recall of false information detection in news dissemination.*

Keywords: News communication; deep learning; information detection; neural network; attention mechanism

1. **Introduction.** In modern years, Internet technology and news media platforms are developing rapidly, and hundreds of millions of social applications have spread throughout the entire social production and life. Social apps, represented by Weibo, Facebook and Twitter, have become important news media platforms for users to achieve, publish and share information. While the majority of users have adapted to browsing and accessing

the latest news and events through online social media, it also creates fertile soil for the emergence and spread of false information [1, 2]. However, owing to the diversification of communication subjects, contents and information transmission channels, the quality of information content on social media is uneven, and it is hard to distinguish between true and false, which provides extremely convenient conditions for the breeding and dissemination of false information [3, 4, 5]. False information on online social platforms is convenient for access and operate, and people are more inclined to achieve and share information and express and exchange opinions through social media platforms. Secondly, as news metier turning on, the amount of human and the complication of message roots, all counterfeit message continues to expand. If the government and relevant departments do not control and verify it, false information will continue to expand. The spread of false information may cause panic among the masses, have a more negative social impact, and even hinder the growth of social economy.

1.1. **Related Work.** Owing to the increasing development of deep learning, researchers have suggested a variety of methods to extract corresponding features in the area of deep learning, which not only keeps off the complexity of manual feature extraction, but also excellently enhances the detection effect of the model.

Zhou and Zafarani [6] used politifact website to manually verify the authenticity and reliability of information. Zaghouani [7] uses crowd sourcing to identify false information. Jin et al. [8] used the information content of the image itself, the amount and attributes of the image and other features to detect false information, but this method also needed human participation in feature design. Subsequently, researchers began to apply deep learning technology to false information detection and use models to extract features from information. Liu et al. [9] used convolutional neural networks to construct deeper text features for false information detection, which enhanced the achievement of false information detection. Ye et al. [10] incorporated the heterogeneity of text content into the communication structure and declared a network-based multivariate cyclic event model to infer the hidden spatial communication network from the data. Thota et al. [11] offered a convolutional neural network model on the ground of text and image information, which uses convolutional neural network to learn its potential features and uses the learned features to detect false information. Zeng et al. [12] pointed out that emotion is an significant indicator in the study of false information detection, and extracting emotional representations from published articles and comments will help identify false information. The development of picture eliminating software for example Photoshop and great success in generative adversarial networks [13, 14] in the area of image synthesis are lowering the technical threshold of image forgery, and the detection of forged images has also caused extensive research recently. On the ground of the above problems, Song et al. [15] suggested a multi-modal false information detection framework (CARMN) on the basis of cross-modal attention residual network and multi-channel convolutional neural network. Meel and Vishwakarma [16] used the writing style analysis technique of multi-level attention network learning false information to canvass the authenticity of images, but image feature extraction was insufficient. Tong and Koller [17] used linear support vector machine to arrange counterfeit message on Twitter, and obtain space structure to model changes in mixer characteristics. Zhou et al. [18] offered a theory-driven model for counterfeit message detection, which resorts to the mature theories of social psychology and forensic psychology to realize the detection of false information within the framework of supervised machine learning. Wang et al. [19] proposed a new SemSeq4FD model on the basis of graph neural networks to deal with the problem that existing methods are not easy to be generalized to cross-domain data. Hakaket al. [20] proposed an integrated classification model to classify and identify false information, extract features in false news, and then use decision trees, random forests and extra trees classifiers to build classification models, but the classification accuracy is not high. Zhao et al. [21] offered a counterfeit message detection model in accordance with refined likelihood model, which uses manual coding method to preprocess data, but the processing efficiency is low. Ahmad et al. [22] used machine learning to identify false news and selected different data sets to verify the effectiveness of the model, but the efficiency was low.

1.2. **Motivation and contribution.** In view of the low efficiency of the current fake information detection methods, this paper suggests a false information detection method on the ground of deep learning model in news communication.

In accordance with in-depth study on the feature of news information, this paper dissects the content characteristics of false news, builds a false information detection framework on the ground of the characteristics of news information, and extends the matching between news headlines and news categories to the matching between news text feature groups and news category feature groups, so as to excavate the potential information contained in news information.

Then LSTM is used to encode the news information for the purpose of preserving its hidden status value. By means of data enhancement, the generalization capability of the model is enhanced, the generalization error is reduced, and the false information detection performance is enhanced. The experimental outcome indicates that the false information detection method designed in this paper has good performance.

2. Relevant theoretical analysis.

2.1. **Definition of fake news.** Since the beginning of news communication activities, society has been facing a struggle against false news. However, until now, the academic community has not reached a consensus on the term "fake news". Current studies generally summarize fake news into narrow sense fake news and broad sense fake news [23]. In a narrow sense, false news is defined as news that can be proven false and is intended to mislead readers [24]. This definition has two characteristics: authenticity and purpose. First, in terms of authenticity, fake news is generally information that can be verified to be false. Secondly, in terms of purpose, fake news generally aims at deceiving readers. The definition of false news in a broad sense pays more attention to the authenticity of its reports. For example, some newspapers regard satirical news as fake news. Although satirical news is usually intended to enhance the entertainment effect and expose its deceptive behavior to news readers, the content is fake [25].

To distinguish the existing study of fake news from some other terms and concepts, such as: false news, fake news, satire news, disinformation, misinformation, rumor, etc., Three attributes of authenticity, intention and whether the information is news are used to distinguish the differences between them. Table 1 indicates the specific differences. On the basis of the above analysis, this paper adopts the concept of false news in a narrow sense, and defines false news as news published by social media websites that can be verified as false and have bad intentions.

2.2. **BERT model.** BERT [26] is a self-coding language model (Autoencoder LM) that primarily uses attention mechanisms to capture correlations between different parts of an input sequence. In BERT, the input sequence is represented as a vector sequence and encoded by the encoder. The encoder consists of multiple hidden layers, each of which is used to capture different features in the input sequence. Specifically, BERT's Self-attention process consists of representing each element in the input model as a sequence

Table 1. The concept of fake news.

Type	Authenticity	Intention	Whether it is news or not
Fake news	False	Bad	Yes
False news	False	Unknown	Yes
Satirical news	Unknown	Unknown	Yes
False information	False	Bad	No
Error information	False	Unknown	No
Rumor	Unknown	Unknown	No

of vectors, converting it into a one-dimensional vector and feeding it into the BERT model. This input is then passed into a self-attention layer containing multiple attention heads. For each attentional head, the model calculates the Euclidean distance between the sequence of input vectors and that attentional head and calculates the weight on the ground of the distance. The resulting weight information is used to calculate the importance of each input vector in relation to the output vector. With the Self-attention mechanism, the model can capture correlations between different parts of the input sequence and generate a high-dimensional vector representation.

The BERT model chiefly computes the association among words by the multi-head concern apparatus, each input sequence element is associated with all other input sequence elements, and calculates the weight to determine the importance of each element in the output vector. The grade of correlation among every word with other words are able to obtain using the dot product operation, which can be expressed by the Equation (1). BERT mainly uses transformer architecture, which can not only reduce the number of parameters, but also ensure the accuracy of the model, thereby enhancing the efficiency and feasibility of the model.

$$Atten(P, L, G) = soft\max\left(\frac{PL^T}{\sqrt{c_l}}\right)G \tag{1}$$

3. **Classification of news information text features.** Aiming at preventing the overdependence of the deep learning model on the training set, this paper proposes a method to repartition the training set and combine it. This method improves the generalization ability of the model and reduces the generalization error by means of data enhancement.

Assuming that the size of the dataset is W, the dataset is divided into m subsets W_j ($j \in [1, m]$) of the same size by random sampling, and each subset is taken as a sample set. The specific steps are as follows. Firstly, all data set W is pre-trained to obtain the initial model v_0. Then $W - W_j$ is used as a new data set and the training is continued on the basis of v_0. m subsets of it are trained for m v_j. At last, in the prediction part, the test set is mixed and the mean value of the final m results is taken as its detection precision.

Each model v_j is equivalent to a weak classifier, and for a model with m weak classifiers, $L(\cdot)$ can be used to represent the training model. $L(\cdot)$ is shown in Equation (2).

$$L_j(\cdot) = \sum_{j=1}^{m} v_j \times h_j(\cdot) \tag{2}$$

Since this method adopts balanced sampling, $h_j(\cdot) = \frac{1}{m}$, for the balanced random sampling, the results trained by each weak classifier have an approximate distribution.

Integrators are added to the current integration model one by one, and the specific implementation process is shown in Equation (3).

$$L_m(\cdot) = L_m(\cdot) + v_j \times h_j(\cdot) = L_m(\cdot) + \frac{v_j}{m} \tag{3}$$

For solving this model, the objective function is to maximize the test result for this $L(\cdot)$ on the test set, that is, to minimize the error on each classifier.

$$v_m = argminE(v + v_j/m) = argmin\sum_{j=1}^{m} \varepsilon(x_m, v_{m-1}(y_m) + v/m) \tag{4}$$

where $E(\cdot)$ is the fitting error of a given model and $\varepsilon(\cdot)$ is the loss or error function. Local optimization is achieved by adding classifiers one by one throughout the model.

Given the initial data set, different data subsets can be generated from it, and different base learners can be trained using different data subsets. The diversity of base classifiers can be increased mainly through sample perturbation.

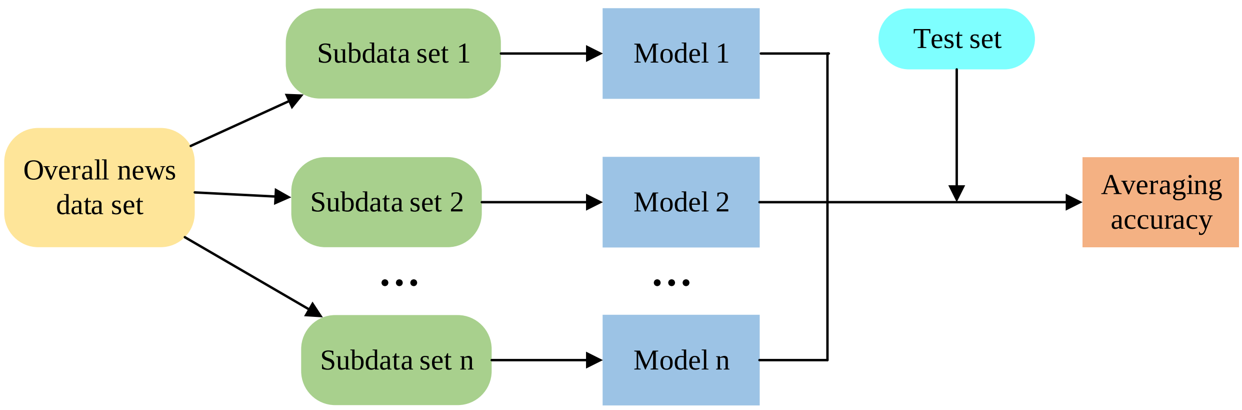

Figure 1. Flow diagram

4. False information detection based on deep learning model in news communication.

4.1. News information presentation and interaction. Firstly, the multi-feature false information detection framework in news communication is constructed, that is, the feature of news title, the feature of the first sentence of news content, the feature of the last sentence of news content and the theme of news content are extracted to form the feature group of news text, and the feature group of news category is selected to form the feature group of news multiple features and supplementary news interpretation features.

The matching between news headlines and news categories is extended to the matching between news text feature groups and news category feature groups, so as to mine the potential information contained in news information, and provide enhanced detection methods for traditional multi-feature false information detection with the help of potential information generated in the communication process, thus improving the performance of false information detection. The false information detection method on the ground of deep learning model suggested in this paper contains five main modules: information representation, information interaction, information feature extraction, communication structure feature extraction, and false information detection, as shown in Figure 2. Finally, the integrated learning mechanism redivides the message set and generates a base classifier, which further improves the generalization ability of the model, keeps off the overfitting problem and enhances the accuracy of news false information detection.

Let $N = \{n_1, n_2, \dots, n_{|n|}\}$ be a set of news false information data sets, where each information n_j consists of q reposts $\{s_1, \dots, s_q\}$. $M = \{m_1, m_2, \dots, m_p\}$ is a set of false information data sets different from M, in which each news is composed of text and pictures, that is, $m_j = \prec s_j | g_j \succ$, s_j represents the text of the j-th post, and g_j represents the picture corresponding to the corresponding text. The goal of this article can study a routine $f(N, M|\theta) \to x$ to predict whether the content of the source text is false information. x represents the category label and $x \in \{0, 1\}$, in which 0 means true message and 1 means counterfeit message. θ represents complete the quantity of the facsimile.

The data set is pre-trained by BERT model, and BERT word vector and LDA topic vector are fused to increase the topic semantic information. Firstly, word segmentation is applied to news text information and news category information, and stop words are removed. The news title vector F^{MS}, the first sentence of news content vector F^{CT}, the last sentence of news content word vector F^{MT}, and the theme word vector F^{ME} are trained respectively. The news title vector and the subject word vector are connected respectively to get the word embeddings representing F^M and F^D.

$$F^{MS,CT,MT,ME} = \left[F_1^{MS,CT,MT,ME}, ..., F_j^{MS,CT,MT,ME}, ..., E_n^{MS,CT,MT,ME}\right]^T, \forall j \in [1, 2, ..., n] \tag{5}$$

$$F^M = concat[F^{MS}, F^{CT}, F^{MT}, F^{ME}], F^M \in R^{n\times c} \tag{6}$$

$$F^D = concat[F^{MT}, F^{DF}], F^D \in R^{m\times c} \tag{7}$$

In order to enhance the contribution of local information, the self-attention method is adopted, and the input representation layer F^M and F^D are TB_j^M, TB_i^D through the self-attention mechanism representation layer. Then LSTM is used to encode the news information in order to preserve its hidden status value.

$$TB_j^M = \sum_{i=1}^{n} \frac{\exp(\beta_{ji})}{\sum\limits_{l=1}^{n} \exp(\beta_{jl})} F_i^M, \forall j, i \in [1, 2, ..., n] \tag{8}$$

$$TB_i^D = \sum_{i=1}^{n} \frac{\exp(\mu_{ji})}{\sum\limits_{l=1}^{m} exp(\mu_{jl})} F_j^D, \forall j, i \in [1, 2, ..., n] \tag{9}$$

4.2. **News information feature extraction.** The feature extraction of news information includes the feature extraction of text and pictures. In this paper, CNN model is adopted, whose input is made up with word series of the information, along with the word transmitter of every word is corresponded with the word transmitter space through advance treatment on the obtained data set. The j-th vocabulary transmitter in text is implied through the vocabulary vector $S_j \in m^l$. The dimension from the word vector is represented by I. Then a message is represented as: $S_{1m} = S_1 \oplus S_2 \oplus \dots \oplus S_n$.

The eigenvector S_j obtained through the swirl process is extracted with the greatest weight of S_j through the method of greatest pond. Behind the pond process, the row transmitter $R_{S_D} \in R^{(d \cdot m_w)}$ with property of whirl kernel size is obtained, which is connected with the greatest weight of every whirl kernel. Pooling can make up for the defect of not padding the sentence. By padding, sentences of different lengths have the same length, so that the dimensions of each column vector obtained through convolution are equal. Therefore, Pooling can be used to eliminate the difference caused by sentence

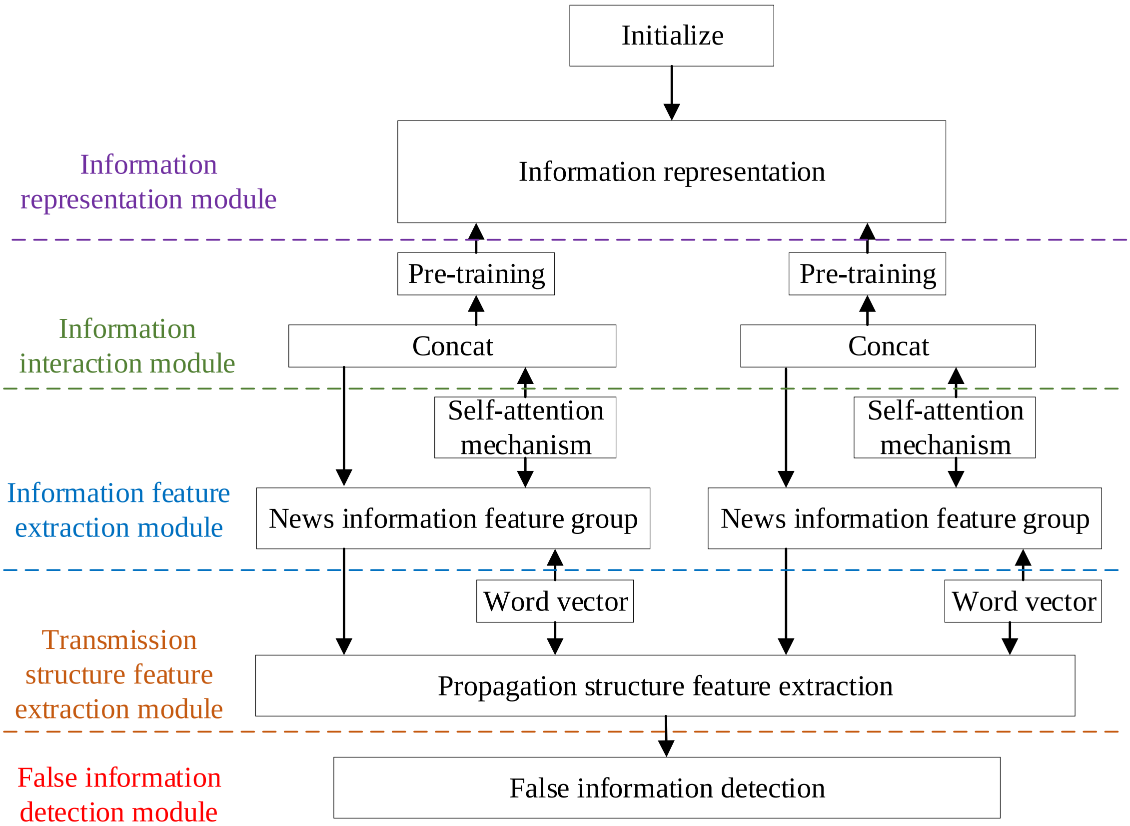

Figure 2. The framework of FIDB

length between the two. The final text features will be obtained through a fully connected layer, expressed as: $R_s = \psi(H_{su} \cdot R_{S_D})$, where H_{su} is the heavy substance of overall related layer, D is the magnitude of the available gliding showcase, and s means dimension.

4.3. **Propagation structure feature extraction.** For the purpose of making full use of the relationship among each node in the propagation process and take native predicted information with overall framework message into account, the figure courtesy network is devoted to study the allocated delegacy of every knob in the heterogeneous figure. Thus, the heterogeneous figure must be established first. The global heterogeneous graph consists of two nodes, information m and user u. The information knobs include source information and forwarding information nodes: $m_i' = m_i^s + \tilde{m}_i$, $u_i' = u_i^s + \tilde{u}_i$, where $m_i^s \in R^c$ and $u_i^s \in R^{c_u}$. Since m_i' and u_i' are not in the same dimension, the two types of nodes ought to be adapted to the identical semantic space: $m_i' = h_n m_i'$, $u_j' = h_f u_j'$, where $h_n \in R^{c \times c}$ and $h_f \in R^{c_f \times c}$ are parameters that can be learned.

Then, the courtesy mechanism is devoted to compute the importance of information m and user's neighbor node to it, and to normalize it. As shown in Equation (10) and Equation (11).

$$\beta_{ij}' = sofmax(LkyReLU(b^T[m_i'; u_j'])) \tag{10}$$

$$\mu_{js}' = sofmax(LkyReLU(d^T[u_j'; m_s])) \tag{11}$$

Finally, normalization is performed and the attention mechanism is used to obtain the final characteristic vector as shown in Equation (12) and Equation (13). Where, m_i^g is the end global text characteristic vector; $\|$ stands for vector concatenation operation; ψ

is the ELU activation function.

$$m_i^{gl} = \overset{L}{\underset{l=1}{\Big\|}} \psi\Big(\sum_{i\in M(n_i)} \beta_{ij}^l h_f^l u_j'\Big) \tag{12}$$

$$u_j^{gl} = \overset{L}{\underset{l=1}{\Big\|}} \psi\Big(\sum_{i\in M(u_j)} \mu_{is}^l h_n^l m_i'\Big) \tag{13}$$

4.4. **False information detection.** The input is a fusion representation of news information features and a new event representation stored in the event memory network. The ultimate incident table Y on the ground of memory improvement is entered to the counterfeit message detector for prediction, which arranges a plenty related layer with matched initiation affair to forecast the truth or falsity of information in a presumption data set, represented by $\hat{x} = E(y)$.

For a certain batch of incidents and true label $y = \{Y_1, Y_2, \ldots, Y_M\}$, the traverse asymptotic entropy loss function is devoted to calculate the final total loss, which can be expressed by Equation (14).

$$F(\delta) = \sum_{m=1}^{M} -\left[x_m \log(x_m) + (1 - x_m)\log(1 - x_m)\right] \tag{14}$$

At last, the detection loss is minimized through detecting the greatest parameter, that is to be explicit through Equation (15).

$$\delta^* = \arg\min_{\delta} \tau(\delta) \tag{15}$$

5. Performance testing and analysis.

5.1. **Comparative analysis of experimental results.** This article applies the public data set Twitter15, whose node features are the text features of news information. In the experiment in this chapter, the types of Information in the data set are divided into four groups, namely True Information (TI), Misinformation (MI), and misinformation (MI). Distorted information (DI) and Unverified Information (UI). The distorted information refers to the fact that the source of the content is true, but there are distorted and exaggerated phenomena in the process of transmission. In this chapter, the TF-IDF method is used to extract text features and the vocabulary size is set to 10000. The self-supervised pre-training process in this chapter iterates over 25 epochs and then over 100 epochs for the classifier.

In this paper, four indexes, including accuracy, precision, recall and F1 value, used in the classification were adopted to effectively evaluate the results generated by the experiment. For the purpose of verifying the performance of the model designed in this paper, it was compared with other existing models in this paper. All experiments were conducted under the Linux operating system and Python 3.7 programming environment, and GPU was used for accelerated training. For the convenience of analysis, reference [27] was recorded as Bert, [28] as ARDM, [29] as DNA, [30] as Botom, [31] as GAT, and the algorithm in this paper was recorded as FIDB.

A large number of experimental analyses were carried out between the model FIDB in this paper and a number of comparison models, and the experimental outcome and data were indicated in Table 2 and Figure 3. As can be seen from the experimental outcome in Table 2 and Figure 3 that the accuracy of FIDB model is higher than that of the comparison model, indicating that the feature extraction using data on the ground of news information features in this paper is superior to the existing model, which proves the effectiveness of the model designed in this paper. The experimental results were

Table 2. Comparative experimental results

Model	Accuracy	Real news			Fake news		
		Precision	Recall	F1	Precision	Recall	F1
Bert	0.718	0.736	0.695	0.715	0.722	0.695	0.708
ARDM	0.657	0.647	0.582	0.618	0.647	0.702	0.739
DNA	0.851	0.829	0.792	0.811	0.785	0.809	0.811
Botom	0.794	0.816	0.786	0.801	0.817	0.817	0.817
GAT	0.872	0.852	0.861	0.856	0.884	0.864	0.874
FIDB	0.938	0.945	0.947	0.946	0.952	0.943	0.947

analyzed from the feature perspective of news information. Since BERT model and ARDM model did not extract features from multiple angles, the experimental results on Twitter15 dataset showed that FIDB's accuracy was 22% higher than BERT's and 28.1% higher than ARDM's. It is proved that FIDB model enhances the deep and shallow feature information of text and image from the model level, and extracts more extensive and semantic text features and image features, thus improving the effectiveness of false news detection. From the perspective of information interaction, both the DNA model and the Botom model extract features directly from multiple fields of news, and neither contains functional modules to eliminate features in specific fields. The experimental outcome on the Twitter dataset also show that the accuracy of FIDB model is 8.7% higher than that of DNA model. It is 14.4% higher than the Botom model and 6.6% higher than the GAT model, which proves that the FIDB false information detection module can filter the features of news data from different fields, make the extracted features more universal, increase the generalization and robustness of the model, and thus improve the performance of the model. Therefore, the FIDB designed in this paper extracts more effective features from the text and image itself, deals with the problem of universality of domain features, and makes the extracted features more universal, thus improving the overall detection performance of false news detection.

5.2. **Ablation experiment.** Aiming at better verifying the influence of the two-branch module and the domain adversarial module in the model in this chapter on the experimental results, ablation experiments were conducted on two data sets about multimodal false news detection on Weibo and Twitter respectively, and two comparison models were designed for analysis.

(1) Remove the information interaction module. The text feature and image feature output are extracted from the model using the peer-to-peer news communication network, and then the features are spliced and input into the false detection module and domain antagonism module. This model is defined as FIDB-1.

(2) Remove the feature extraction module. Text and image content features are extracted using a two-way interactive network, which is connected and fed into the false-detection module. This model is defined as FIDB-2. The experimental results of model ablation are shown in Figure 4. In terms of the analysis of the experimental results, the accuracy of FIDB-2 model after removing the feature extraction module is lower than that of the original model on the Twitter15 dataset, which proves that there is a certain degree of loss of text feature information and image feature information when there is no feature extraction module, and once again verifies the effectiveness of FIDB model.

5.3. **Experiment on the influence of different imbalance ratios.** To build the detection task, this paper sets up a series of early detection time points and uses only news published before that time point for detection. This paper sets different time detection

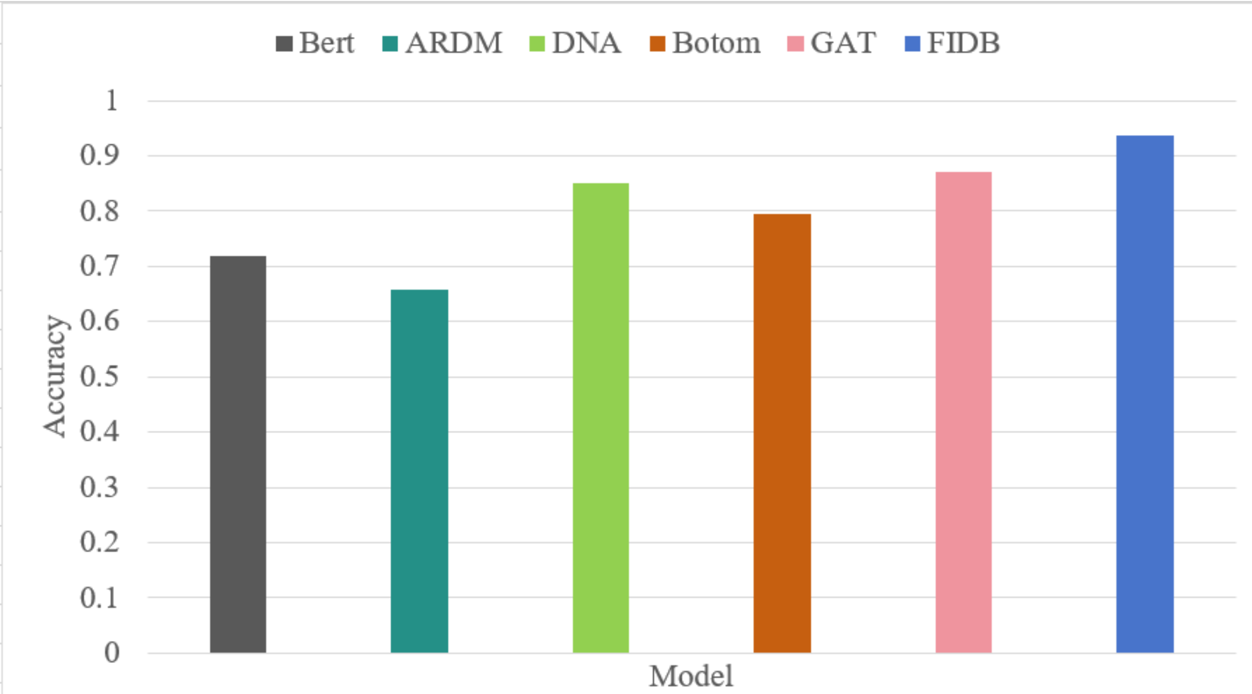

Figure 3. Comparison results of FIDB and comparison model accuracy

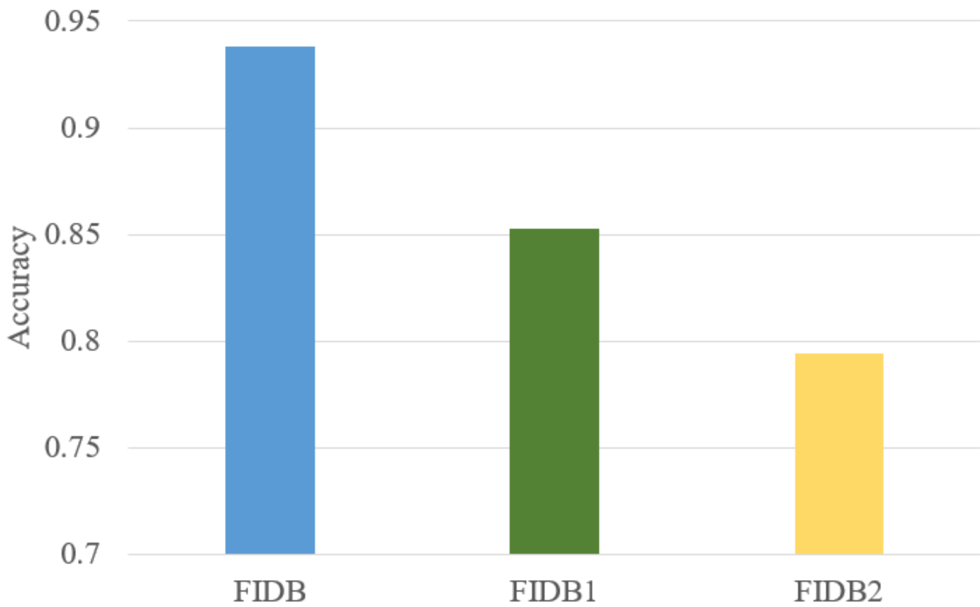

Figure 4. Results of model ablation experiment

points after the release of the original news, such as 0 minutes after the release, 10 minutes after the release, 20 minutes after the release and so on. Figure 5 shows the performance results of the early detection approach on the dataset. It can be seen that FIDB achieved high accuracy at every early detection point after the original post was published, and outperformed other comparison models throughout the process.

6. **Conclusion.** Aiming at the low accuracy of false information detection in current news dissemination, this article suggests a counterfeit message detection method on the

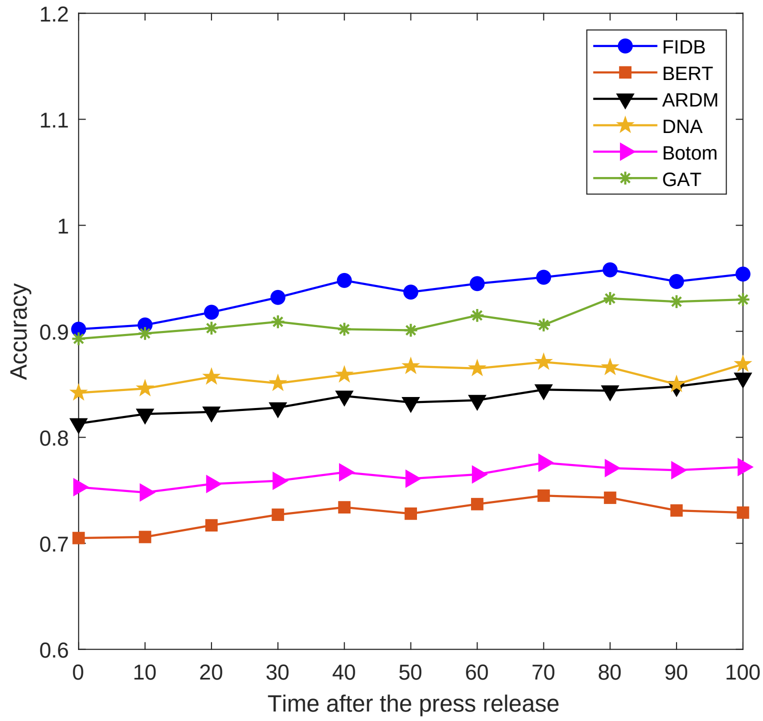

Figure 5. Performance outcome of early detection methods

ground of deep learning model. The method firstly generates multi-feature samples based on different features of news information, and carries out perturbations to increase the diversity among base classifiers. Then, multiple features of news and supplementary news interpretation features are selected to form a feature group of news categories to mine the potential information contained in news information. Secondly, the integrated learning mechanism redivides the message set and generates a base classifier, which further improves the generalization ability of the model, avoids the overfitting problem and enhances the accuracy of news false information detection. Finally, the experimental outcome indicates that the proposed method can effectively improve the accuracy rate, precision rate and recall rate of news false information detection.

REFERENCES

[1] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146-1151, 2018.

[2] J. Xue, Y. Wang, Y. Tian, Y. Li, L. Shi, and L. Wei, "Detecting fake news by exploring the consistency of multimodal data," *Information Processing & Management*, vol. 58, no. 5, 102610, 2021.

[3] H. Allcott, and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, vol. 31, no. 2, pp. 211-236, 2017.

[4] X. Li, S. Liu, S. Kumari, and C.-M. Chen, "PSAP-WSN: A Provably Secure Authentication Protocol for 5G-Based Wireless Sensor Networks," *Computer Modeling in Engineering & Sciences*, vol. 135, no. 1, pp. 711-732, 2023.

[5] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, 1-36, 2018.

[6] X. Zhou, and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys (CSUR)*, vol. 53, no. 5, 1-40, 2020.

[7] W. Zaghouani, "A large-scale social media corpus for the detection of youth depression (project note)," *Procedia Computer Science*, vol. 142, pp. 347-351, 2018.

[8] Z. Jin, J. Cao, Y. Zhang, J. Zhou, and Q. Tian, "Novel visual and statistical image features for microblogs news verification," *IEEE Transactions on Multimedia*, vol. 19, no. 3, pp. 598-608, 2016.

[9] Z. Liu, Z. Wei, and R. Zhang, "Rumor detection based on convolutional neural network," *Journal of Computer Applications*, vol. 37, no. 11, 3053, 2017.

[10] X. Ye, W. Wang, X. Zhang, Z. Li, D. Yu, J. Du, and Z. Chen, "Reconstructing spatial information diffusion networks with heterogeneous agents and text contents," *Transactions in GIS*, vol. 25, no. 4, pp. 1654-1673, 2021.

[11] A. Thota, P. Tilak, S. Ahluwalia, and N. Lohia, "Fake news detection: a deep learning approach," *Smu Data Science Review*, vol. 1, no. 3, pp. 10, 2018.

[12] J. Zeng, Y. Zhang, and X. Ma, "Fake news detection for epidemic emergencies via deep correlations between text and images," *Sustainable Cities and Society*, vol. 66, 102652, 2021.

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[14] J. Zeng, X. Ma, and K. Zhou, "Photo-realistic face age progression/regression using a single generative adversarial network," *Neurocomputing*, vol. 366, pp. 295-304, 2019.

[15] C. Song, N. Ning, Y. Zhang, and B. Wu, "A multimodal fake news detection model based on cross-modal attention residual and multichannel convolutional neural networks," *Information Processing & Management*, vol. 58, no. 1, 102437, 2021.

[16] P. Meel, and D. K. Vishwakarma, "HAN, image captioning, and forensics ensemble multimodal fake news detection," *Information Sciences*, vol. 567, pp. 23-41, 2021.

[17] S. Tong, and D. Koller, "Support vector machine active learning with applications to text classification," *Journal of Machine Learning Research*, vol. 2, no. 11, pp. 45-66, 2001.

[18] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani, "Fake news early detection: A theory-driven model," *Digital Threats: Research and Practice*, vol. 1, no. 2, 1-25, 2020.

[19] Y. Wang, L. Wang, Y. Yang, and T. Lian, "SemSeq4FD: Integrating global semantic relationship and local sequential order to enhance text representation for fake news detection," *Expert Systems with Applications*, vol. 166, 114090, 2021.

[20] S. Hakak, M. Alazab, S. Khan, T. R. Gadekallu, P. K. R. Maddikunta, and W. Z. Khan, "An ensemble machine learning approach through effective feature extraction to classify fake news," *Future Generation Computer Systems*, vol. 117, pp. 47-58, 2021.

[21] Y. Zhao, J. Da, and J. Yan, "Detecting health misinformation in online health communities: Incorporating behavioral features into machine learning based approaches," *Information Processing & Management*, vol. 58, no. 1, 102390, 2021.

[22] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, "Fake news detection using machine learning ensemble methods," *Complexity*, 1-11, 2020.

[23] Y. Ma, Y. Peng, and T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121-2133, 2022.

[24] T.-Y. Wu, L. Wang, and C.-M. Chen, "Enhancing the Security: A Lightweight Authentication and Key Agreement Protocol for Smart Medical Services in the IoHT," *Mathematics*, vol. 11, no. 17, 3701, 2023.

[25] T.-Y. Wu, Q. Meng, L. Yang, S. Kumari, and M. Pirouz, "Amassing the Security: An Enhanced Authentication and Key Agreement Protocol for Remote Surgery in Healthcare Environment," *Computer Modeling in Engineering & Sciences*, vol. 134, no. 1, pp. 317-341, 2023.

[26] C.-M. Chen, Z. Li, S. Kumari, G. Srivastava, K. Lakshmanna, and T. R. Gadekallu, "A provably secure key transfer protocol for the fog-enabled Social Internet of Vehicles based on a confidential computing environment," *Vehicular Communications*, vol. 39, 100567, 2023.

[27] Y. Qiao, X. Zhu, and H. Gong, "BERT-Kcr: prediction of lysine crotonylation sites by a transfer learning method with pre-trained BERT models," *Bioinformatics*, vol. 38, no. 3, pp. 648-654, 2022.

[28] J. Ma, W. Gao, S. Joty, and K.-F. Wong, "An attention-based rumor detection model with tree-structured recursive neural networks," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 4, 1-28, 2020.

[29] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "DNA-inspired online behavioral modeling and its application to spambot detection," *IEEE Intelligent Systems*, vol. 31, no. 5, pp. 58-64, 2016.

[30] K.-C. Yang, E. Ferrara, and F. Menczer, "Botometer 101: Social bot practicum for computational social scientists," *Journal of Computational Social Science*, vol. 5, no. 2, pp. 1511-1528, 2022.

[31] R. Zhong, R. Wang, Y. Zou, Z. Hong, and M. Hu, "Graph attention networks adjusted Bi-LSTM for video summarization," *IEEE Signal Processing Letters*, vol. 28, pp. 663-667, 2021.