



ARTICLE

A Model for Detecting Fake News by Integrating Domain-Specific Emotional and Semantic Features

Wen Jiang^{1,2}, Mingshu Zhang^{1,2,*}, Xu'an Wang^{1,3}, Wei Bin^{1,2}, Xiong Zhang^{1,2}, Kelan Ren^{1,2} and Facheng Yan^{1,2}

¹College of Cryptography Engineering, Engineering University of People's Armed Police, Xi'an, 710086, China

²Key Laboratory of Network and Information Security, Engineering University of People's Armed Police, Xi'an, 710086, China

³Key Laboratory of CTC & Information Engineering, Engineering University of People's Armed Police, Ministry of Education, Xi'an, 710086, China

*Corresponding Author: Mingshu Zhang. Email: zms2099@163.com

Received: 09 May 2024 Accepted: 17 June 2024 Published: 15 August 2024

ABSTRACT

With the rapid spread of Internet information and the spread of fake news, the detection of fake news becomes more and more important. Traditional detection methods often rely on a single emotional or semantic feature to identify fake news, but these methods have limitations when dealing with news in specific domains. In order to solve the problem of weak feature correlation between data from different domains, a model for detecting fake news by integrating domain-specific emotional and semantic features is proposed. This method makes full use of the attention mechanism, grasps the correlation between different features, and effectively improves the effect of feature fusion. The algorithm first extracts the semantic features of news text through the Bi-LSTM (Bidirectional Long Short-Term Memory) layer to capture the contextual relevance of news text. Senta-BiLSTM is then used to extract emotional features and predict the probability of positive and negative emotions in the text. It then uses domain features as an enhancement feature and attention mechanism to fully capture more fine-grained emotional features associated with that domain. Finally, the fusion features are taken as the input of the fake news detection classifier, combined with the multi-task representation of information, and the MLP and Softmax functions are used for classification. The experimental results show that on the Chinese dataset Weibo21, the F1 value of this model is 0.958, 4.9% higher than that of the sub-optimal model; on the English dataset FakeNewsNet, the F1 value of the detection result of this model is 0.845, 1.8% higher than that of the sub-optimal model, which is advanced and feasible.

KEYWORDS

Fake news detection; domain-related emotional features; semantic features; feature fusion

1 Introduction

The 53rd China Internet Development Statistics Report (CNNIC), released in March 2024, showed that China had 1092 million Internet users at the end of December 2023, an increase of 24.8 million from December 2022, and an Internet penetration rate of 77.5%. With the further acceleration



of digital infrastructure construction, the continuous enrichment of resource applications, and the explosive growth of information, social media platforms such as Weibo, WeChat public accounts, and TikTok, which are particularly important components of social media, provide information and social platforms for Internet users, while also creating a breeding ground for the spread of fake news that confuses the public. Fake news such as “essential oil suppresses viral infections”, “five eggs a day prevents colds”, and “residents go out to shave their heads during the pandemic” have proliferated as the Internet has become more inundated with information. Fake news that disseminates widely and quickly often contains arbitrary and inflammatory language, is carefully crafted as “clickbait”, is often characterized by “lack of detail” and “incitement others’ emotions”, can mislead people’s judgment and can have serious negative consequences for society. Fake news is spreading and deepening because of the “echo chamber effect” on social media. Therefore, the detection of fake news has become a hot topic in the field of information processing [1].

Fake news detection is often defined as a classification problem, and feature extraction is a key step in capping the performance [2]. Therefore, how to explore the relationship between different characteristics has also become a subject of continuous research by scholars. Jin et al. [3] first proposed att-RNN, which integrates a variety of pattern features including text, visual features, and user portrait features and adjusts features using the attention mechanism. Khattar et al. [4] proposed a Multimodal Variational Autoencoder (MVAE), which trains users to find correlations between different patterns in a particular tweet. The model proposed by Zeng et al. [5] not only explores the semantic relationship between text content and accompanying diagrams, but also utilizes attention mechanism to highlight important parts of specific news articles.

Association rule technology is widely used in social media, and its greatest advantage is the ability to provide and summarize information. The output of this algorithm is easy to explain and can be inferred in all application fields. Delgado et al. [6] proposed one of the earliest papers to solve the problem of text association rules, which defined text transactions and applied fuzzy association rules to text transactions. Another interesting study by Feng et al. [7] explored the relationship between mood, colour and social labels through associative rules. In this study, association rule was used to encode images, linking pixel colors to specific emotions. Sheydaei et al. [8] proposed an association classification algorithm, including feature selection stages for selecting vital features and a clustering phase based on class labels. Rohidin et al. [9] proposed a text classification method combining the association method with a fuzzy soft set and obtained good classification accuracy. Cai et al. [10] proposed a fully shared multi-domain neuro emotional classification model, revealing the relationship between domains and emotions. These studies indicate that by linking object characteristics in social media and utilizing various of data to improve classification performance, classification problems can be better understood.

Therefore, this work clearly recognizes the significance of characteristics and suggests the DESF model (Domain-Emotion-Semantics Fusion), which combines semantic and domain-related sentiment characteristics. The accuracy and robustness of false information detection are enhanced by integrating the sentiment, domain, and semantic feature extraction modules. Ultimately, classification detection is accomplished by utilizing the Softmax function. The following are this paper’s primary contributions:

- (1) It makes complete use of the feature relevance criteria, concentrates on key terms in the text, determines the sentiment polarity of those words in particular domains, and extracts more detailed sentiment features in other domains.

(2) It improves the model's capacity to concentrate on and comprehend important information by skillfully fusing numerous features, including domain features, sentiment features, and semantic features, using an attention method.

(3) To confirm the validity and efficiency of the suggested approach, this work carries out in-depth tests using the suggested DESF model (Domain-Emotion-Semantics Fusion) on Weibo21 and FakeNewsNet, then contrasts it with nine baseline models, including EANN.

(4) This study investigates the identification of false news within the framework of the Chinese language, emphasizing the unique connection between text and emotional characteristics.

[Section 2](#) provides a literature review of relevant works in recent years, introducing the current status of fake news detection and the application of feature fusion methods. [Section 3](#) introduces the proposed model. [Section 4](#) details the datasets, experimental settings, and thorough analysis of experimental results, along with a sufficient discussion of the ablation study results. [Section 5](#) draws conclusions based on the research presented in this paper.

2 Related Work

2.1 Fake News Detection

Fake news detection has been a highly concerned field in recent years, aiming to detect and address the widespread dissemination of false information on the Internet. In recent years, traditional machine learning methods, such as Naive Bayes, Support Vector Machine (SVM), and Random Forest, have been widely applied in detection. Researchers have utilized various features such as text features, social media features, and semantic features for detection. As research progresses, deep learning methods like CNN and RNN have also achieved certain results. Ji et al. [11] utilized knowledge graphs to detect fake news, establishing a rich knowledge base in terms of propagation relationships, entity knowledge, and event timelines, and combining them with text features for analysis. Liu et al. [12] paid more attention to the fusion of multimodal features, including text, images, and videos. The fusion of multimodal features can provide more comprehensive information, contributing to the improvement of accuracy. Given the continuous evolution of fake news creators, Li et al. [13] established competition between models and those generating fake news through methods such as Generative Adversarial Networks (GAN), enhancing the model's ability to counter various false information. Sahoo et al. [14] used multiple features associated with Facebook account with some news content features to analyze the behavior of the account through deep learning. Zhang et al. [15] presented a deep learning-based fast fake news detection model for cyber-physical social services. Taking Chinese text as the objective, each character in Chinese text was directly adopted as the basic processing unit. Considering the fact that the news were generally short texts and could be remarkably featured by some keywords, convolution-based neural computing framework was adopted to extract feature representation for news texts.

Despite some challenges in fake news detection, researchers are constantly striving to improve the accuracy and robustness of the models.

2.2 Feature Fusion

Feature fusion involves combining multiple features or feature sets to create a more informative and discriminative data representation. It aims to enhance the overall performance of the model by utilizing the complementary information provided by different features [16]. Common feature fusion methods include concatenation, which appends features together to form a longer feature vector; weighted summation, where different features are combined with weights reflecting their importance;

and more complex techniques such as kernel-based fusion or deep learning-based fusion. Feature fusion enables the combination of multiple features to create a more discriminative data representation, playing a crucial role in improving the performance of predictive models and extracting meaningful information from raw data. Wang et al. [17] proposed a novel fine-grained multimodal fusion network (FMFN) to fully fuse textual features and visual features for fake news detection. Li et al. [18] proposed a novel framework named semantic-enhanced multimodal fusion network, which can better capture mutual features among events and thus benefit the detection of fake news. Guo et al. [19] proposed a mutual attention neural network (MANN) that can learn the relationship between each different modality.

Based on the different sources of network information, it can be divided into several domains (politics, history, culture, etc.), and online fake news has different characteristics in each domain [20]. Typically, genuine information tends to be more objective and neutral in sentiment, while fake news is often spread in an exaggerated and propagandistic manner, with a higher degree of emotional arousal. Through sentiment features, we can analyze the emotional intensity, sentiment polarity, and other characteristics of the text to determine whether there is fake news. Semantic features can reveal the semantic relationships between texts. However, single sentiment features or semantic features only serve as an auxiliary means, and their accuracy is limited by the accuracy of annotation and extraction. Therefore, fusing features can improve the accuracy of detection. The innovation pairs of different methods are shown in Table 1.

Table 1: Comparison of innovation points of different methods

| Method | Innovation point |
|-------------------|---|
| Ji et al. [11] | Knowledge graphs |
| Liu et al. [12] | The fusion of multimodal features, including text, images, and videos |
| Li et al. [13] | Generative Adversarial Networks (GAN) |
| Sahoo et al. [14] | Used multiple features associated with Facebook account with some news content features |
| Zhang et al. [15] | Convolution-based neural computing framework |
| Wang et al. [17] | A novel fine-grained multimodal fusion network (FMFN) |
| Li et al. [18] | Semantic-enhanced multimodal fusion network |
| Guo et al. [19] | A mutual attention neural network (MANN) |
| Our DESF | Multi-domain feature, semantic feature, emotion feature, feature fusion, Attention |

3 Method

3.1 Model Construction

Fake news detection is often modeled as a binary classification task, where the model input is news text and the output is a label indicating whether the news is true or fake. This paper proposes the DESF model by effectively fusing emotional features, domain features, and textual features. The model aims to identify the emotional polarity of important words in different domains, capture domain-specific information in the text, and then fuse the emotional features of domain words with semantic features to extract more fine-grained textual information.

DESF model is primarily composed of five parts: the first part is the domain feature extraction module, which extracts feature representations of different domains through domain encoding of nine domains such as politics, military, culture, and society. The second part is the emotional feature extraction module, which analyzes the emotional tendency of the text through the Senta-BiLSTM model and utilizes an attention mechanism to fuse the domain features and emotional features. The third part is the semantic feature extraction module, which employs Bi-LSTM to extract semantic features from the text. The fourth part is the feature fusion module, which fully integrates semantic features with domain-emotion relevance features to capture the global dependencies of the text. The fifth part is the prediction layer, which classifies the fused features through MLP and the Softmax function, and calculates the loss. The overall architecture is shown in Fig. 1, consisting mainly of an input layer, a feature representation layer, an attention layer, a feature fusion layer, and an output layer.

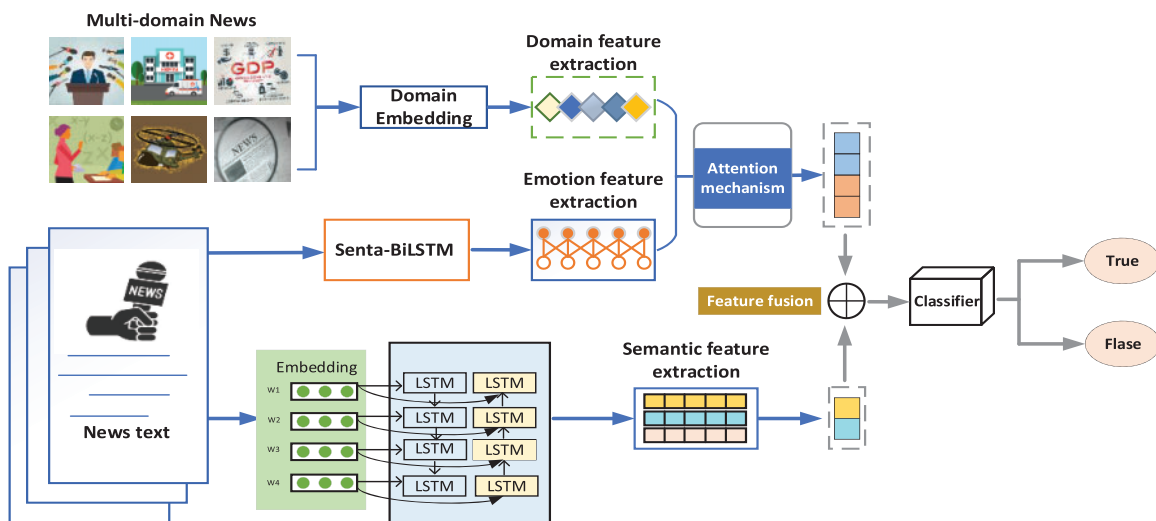


Figure 1: Overall architecture diagram of the DESF model

3.2 Input Layer

Assuming that N represents a news text composed of T words, where w represents a word, then a news text can be expressed as $N = \{w_1, w_2, w_3 \dots w_T\}$. Before input, the source data undergoes text cleaning, including removing duplicate data, stop words, and other unnecessary elements, to improve data quality and eliminate unnecessary noise. Feature representation layer.

3.3 Extracting Semantic Features with Bi-LSTM Layer

For a news text, it is often necessary to pay attention to contextual relevance, especially when dealing with texts from different domains. Combining the context can better determine the semantics. For instance, in the electronics domain, when commenting on battery performance, the adjective “fast” preceded by “charging” indicates that the battery has good performance with a short charging time, while if preceded by “power consumption”, it suggests poor battery performance with a shorter standby time.

In traditional recurrent neural network models and LSTM models, information can only propagate forward, resulting in the state at time t depending only on information before time t . To ensure that each moment contains contextual information, this paper utilizes the combination of the bidirectional

recurrent neural network (BiRNN) model and LSTM units, known as Bi-LSTM, to capture contextual information. Bi-LSTM can process data from two directions simultaneously using two independent hidden layers [21], combining the output data as input to the output layer. While preserving the relevance of the original text's words, it fully incorporates the context, effectively reducing semantic ambiguity. The working principle of the LSTM model is formulated as follows:

$$i_t = \text{sigmoid}(W_i \times [h_{t-1}, x_t] + b_i) \quad (1)$$

$$f_t = \text{sigmoid}(W_f \times [h_{t-1}, x_t] + b_f) \quad (2)$$

$$o_t = \text{sigmoid}(W_o \times [h_{t-1}, x_t] + b_o) \quad (3)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (4)$$

$$h_t = o_t \odot \tanh(c_t) \quad (5)$$

where x_t represents the input vector at time t ; The input door, the forgetting door and the output door are represented by i_t, f_t and o_t , respectively. W_i, W_f and W_o represent the weight matrix of input gate, forgetting gate and output gate, respectively. b_i, b_f and b_o represent bias vectors of input gate, forgetting gate and output gate, respectively. c_t represents the memory unit of the current moment; $t-1$ represents the moment after the current moment t ; W_c and b_c represent the weight matrix and bias vector of the current information; $\text{sigmoid}(\cdot)$ and $\tanh(\cdot)$ represent activation functions; h_t is the output vector.

Therefore, the output is:

$$\vec{h}_t = LSTM(\vec{h}_{t-1}, x_t) \quad (6)$$

$$\overleftarrow{h}_t = LSTM(\overleftarrow{h}_{t+1}, x_t) \quad (7)$$

$$y_t = \begin{bmatrix} \vec{h}_t, \overleftarrow{h}_t \end{bmatrix} \quad (8)$$

\vec{h} and \overleftarrow{h} represent the feature vector obtained by the forward and reverse LSTM models respectively, and y_t represents the semantic feature vector obtained by the Bi-LSTM model at the current time.

3.4 Sentiment Feature Extraction

In real-world scenarios, news publishers often release tweets with strong emotions to attract traffic, resonate with the audience, and promote widespread dissemination. Fake news often contains characteristics such as “exaggerated techniques” and “inflammatory vocabulary”, which affect readers’ subjective judgments. Therefore, sentiment features play a significant role.

After selecting a dataset with appropriate size and diverse categories, this paper uses the Senta-BiLSTM model to analyze the sentiment tendency of the text [22]. By capturing long-distance dependencies in the text and sentiment information in the context, it achieves accurate judgment of the sentiment polarity of the text. In specific experiments, the model is first fine-tuned by adjusting model parameters to adapt to the task and dataset. Then, the model is used to detect news texts, predicting

the probability of positive sentiment and the probability of negative sentiment of the text. Finally, a representation of the sentiment polarity features of the news text is obtained. In this paper, N is used to represent the news text, and y_e is used to represent the sentiment feature vector, then:

$$y_e = \text{Senta_BiLSTM}(N) \quad (9)$$

3.5 Domain Feature Embedding

The issue of domain shift and cross-domain relevance among different news domains poses significant challenges to the predictive capabilities of the model. The goal of domain feature extraction is to identify sources that represent the key information of the data, making the model discriminative, stable, and interpretable.

After extracting the semantic and sentiment features of the text, this paper employs domain encoding as an enhanced feature to make the feature extraction more fine-grained. During experimentation, manually classified domain labels are adopted, and for each domain, the encoded domain vector representation is obtained:

$$D = \{D_1, \dots, D_n\} \quad (10)$$

where D represents the domain vector.

3.6 Attention Layer

The same emotional word may have different emotional tendencies in different domain corpora. This paper adopts an attention mechanism to capture the “domain-sentiment” relationship. By assigning different weights to emotional words in different domains, the model can more accurately understand the emotional information in the text.

The experiment employs an additive attention mechanism to dynamically adjust the relationship between domain features and sentiment features by learning an attention weight vector. First, the attention score is calculated, and its formula is as follows:

$$\text{text score}(D, y_e) = \text{text} \tanh(W_D D + W_{y_e} y_e) \quad (11)$$

where W_D and W_{y_e} are the weight matrices for learning.

Based on the attention score, the attention weight α is calculated:

$$\alpha = \text{text soft max}(\text{text score}(D, y_e)) \quad (12)$$

Weighted summation of domain features using attention weights gives the final representation:

$$\text{text final representation} = f_r = \text{sum}(\alpha_n D_n) \quad (13)$$

where n represents the number of domains, and f_r represents the final representation.

3.7 Feature Fusion Layer

The enhanced domain features D , sentiment features y_e , and semantic features y_i are fused through an adaptive gate to form the final fused feature C . This process captures the global dependencies of the text and adaptively focuses on the key features, reducing redundant information introduced during model concatenation and improving the ability to understand news texts. As shown in the formula:

$$C = \alpha D + \beta y_e + y_i \quad (14)$$

In the formula, learnable parameters α and β are added, and the parameter before the semantic features y_i is represented as 1.

3.8 Prediction Layer

The fused features are input into MLP, and finally connected to a softmax layer for predicting false news, as shown in the formula:

$$y = \text{soft max}(W * MLP(C) + b) \quad (15)$$

where W and b represent the connection layer parameters and bias terms, respectively. When calculating the loss, this paper adopts the cross-entropy loss function for training, as shown in the formula:

$$L = -[y \log \hat{y} + (1 - y) \log (1 - \hat{y})] \quad (16)$$

where y represents the true label value, and \hat{y} represents the predicted probability value.

4 Experiment Process and Result Analysis

To prevent overfitting of the model, this paper introduces Dropout to ensure randomness during training, setting the Dropout parameter to 0.5. The learning rate is set to 0.001 to train the model, and 200 training iterations are performed. L2 regularization is applied to the fully connected layer with a value of 0.01, and the size is set to 50. To ensure the effectiveness, the Adam optimizer is selected to automatically adjust the learning rate and weight decay. Meanwhile, to reduce the instability of model training, different random numbers are used to initialize the model, and the average of the results from 3 runs is taken.

4.1 Datasets

This paper selects FakeNewsNet [23] and Weibo21 [24] to verify the effectiveness of DESF. Among them, FakeNewsNet is a data repository that contains news content, social context, and spatiotemporal information, mainly covering two domains of politics and entertainment. The Weibo21 dataset is a rearrangement of the publicly available Weibo dataset by a laboratory of the Institute of Computing Technology, Chinese Academy of Sciences, including false information officially disclosed by Weibo from December 2014 to March 2021. Each piece of data contains multi-dimensional information such as text content, comments, and timestamps, covering a total of 9 topic domains.

This paper collects and organizes the datasets, and the final statistical information of the English dataset is shown in Table 2, Chinese dataset is shown in Table 3.

Table 2: The statistics of FakeNewsNet datasets

| Domain | PolitiFact | GossipCop | All |
|--------|------------|-----------|--------|
| Real | 447 | 16,804 | 17,251 |
| Fake | 379 | 5067 | 5446 |
| Total | 826 | 21,871 | 22,697 |

Table 3: The statistics of Weibo21 datasets

| Domain | Science | Military | Education | Disasters | Politics |
|--------|---------|----------|---------------|-----------|----------|
| Real | 143 | 121 | 243 | 185 | 306 |
| Fake | 93 | 222 | 248 | 591 | 546 |
| Total | 236 | 343 | 491 | 776 | 852 |
| Domain | Health | Finance | Entertainment | Society | All |
| Real | 485 | 959 | 1000 | 1198 | 4460 |
| Fake | 515 | 362 | 440 | 1471 | 4488 |
| Total | 1000 | 1321 | 1440 | 2669 | 9128 |

This paper splits the dataset into training, validation, and test sets according to a 6:2:2 ratio, and fully leverages the enhancement effect of domain features through the 9 domains in the Chinese dataset and the 2 domains in the English dataset. Each news item is labeled with a specific domain.

4.2 Evaluation Metrics

The evaluation metrics includes accuracy, precision, recall, and F1 score. Their meanings are as follows:

(1) Accuracy represents the proportion of correctly classified samples to the total number of samples.

$$Accuracy = \frac{TP + FN}{TP + FP + TN + FN} \quad (17)$$

(2) Precision represents the proportion of correctly classified positive samples to all positive samples classified by the classifier.

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

(3) Recall represents the proportion of correctly classified positive samples to the total number of positive samples.

$$Recall = \frac{TP}{TP + FN} \quad (19)$$

(4) F1 score represents the weighted harmonic mean of precision and recall, where P stands for precision and R stands for recall.

$$F_1 = \frac{2 * P * R}{P + R} \quad (20)$$

4.3 Baseline Models

To fully validate the effectiveness of DESF, this paper adopts two types of baseline models for comparative verification and analyzes the experimental results.

The first type: We select the following baseline models from the perspective of how to consider domain feature factors:

FuDFEND [25]: A fuzzy domain approach for multi-domain fake news detection, which utilizes neural networks to fit the fuzzy reasoning process and constructs a fuzzy domain label for each news item.

MDFEND [24]: A multi-domain detection method based on a multi-expert framework, which aggregates the representations of multiple “experts” through “domain gates”.

SLFEND [26]: This model utilizes soft labels to extract multi-domain features of news and obtain the final overall feature representation.

M3FEND [27]: It uses a multi-view modeling approach to simultaneously model the semantic, style, and emotional information of news, and employs a cross-view interaction method to more fully model the interaction information between views.

CFPFND [28]: The method based on cross-feature perception fusion, captures multiple feature differences between news from different domains, explores the correlation between news, and controls the feature fusion strategy of the model in different domains from multiple dimensions.

The second type: The following baseline models are selected from the perspective of feature fusion:

BERT [29]: Uses the BERT pre-trained model to extract news text features and classify them.

BiGRU [30]: A multi-layer structure based on GRU, which adds a second GRU layer to capture high-level feature interactions between different time steps.

SpotFake [31]: A multimodal framework that utilizes both textual and visual features of an article to detect fake news.

EANN [32]: This model uses an event discriminator to measure the differences between different events and further learns event-invariant features.

4.4 Analysis of Experimental Results

The experiment utilized nine baseline models and the DESF model to train on both Chinese and English datasets, and the trained models were tested on the test set, using F1 score as the evaluation metric. Based on different baseline models, the results are presented as follows:

(1) Starting from the perspective of how the models combine domain features for detection, the results are shown in [Tables 4](#) and [5](#).

Table 4: Performance comparison between DESF and other baseline methods on Weibo21

| Method | Edu. | Disasters | Health | Finance | Politics | Military | Science | Society | Ent. |
|---------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| MDFEND | 0.883 | 0.878 | 0.943 | 0.875 | 0.883 | 0.936 | 0.833 | 0.894 | 0.909 |
| M3FEND | 0.884 | 0.882 | 0.952 | 0.899 | 0.888 | 0.928 | 0.838 | 0.904 | 0.929 |
| FuDFEND | 0.892 | 0.906 | 0.942 | 0.890 | 0.901 | 0.947 | 0.813 | 0.917 | 0.916 |
| SLFEND | 0.928 | 0.891 | 0.924 | 0.923 | 0.937 | 0.923 | 0.883 | 0.927 | 0.925 |
| CFPFND | 0.918 | 0.896 | 0.957 | 0.905 | 0.897 | 0.971 | 0.862 | 0.915 | 0.928 |
| DESF | 0.923 | 0.913 | 0.959 | 0.927 | 0.920 | 0.958 | 0.887 | 0.929 | 0.931 |

Table 5: A comprehensive view of performance across different models

| Method | Precision | Recall | F1 score |
|---------|--------------|--------------|--------------|
| MDFEND | 0.906 | 0.880 | 0.893 |
| M3FEND | 0.898 | 0.902 | 0.900 |
| FuDFEND | 0.901 | 0.905 | 0.903 |
| SLFEND | 0.910 | 0.916 | 0.918 |
| CFPFND | 0.904 | 0.930 | 0.917 |
| DESF | 0.929 | 0.925 | 0.927 |

To facilitate a clearer comparison of the results of DESF with other baseline models on the Weibo, this paper utilizes a combined graph to visualize the changes and differences in the data curves as shown in Fig. 2 below:

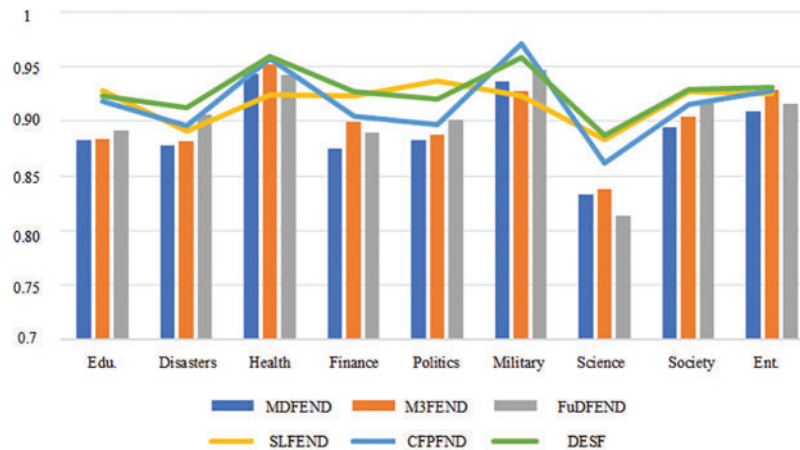
**Figure 2:** Results comparison of different models on Weibo21 datasets (F1 score)

Fig. 2 shows the proposed DESF model in this paper outperforms other models in multiple domains, especially in the fields of health, economy, science, society, accident disasters, and entertainment, demonstrating strong generalization ability and overall robustness. Analyzing the data from a single domain, DESF achieved an F1 score of 0.959 in the health domain, which is 0.002 higher than the second-best model CFPEND in this domain. In the accident disaster domain, the DESF model achieved an F1 score of 0.913, which is 0.007 higher than the F1 score of the second-best model FuDFEND, while the F1 scores of the other four models in this domain are around 0.89. In the economic domain, the DESF model achieved good performance with an F1 score of 0.927, which is 5.9% higher than MDFEND and 0.4% higher than the SLFEND model. In the science, society, and entertainment domains, the F1 scores of the DESF model are 0.887, 0.929, and 0.931, respectively, which are 0.5%, 0.2%, and 0.2% higher than the corresponding second-best models. In the military domain, DESF achieved an F1 score of 0.958, which is 0.013 lower than CFPFND, while FuDFEND and MDFEND also achieved good detection results with F1 scores of 0.947 and 0.936. In the political domain, the SLFEND model performed the best, while the DESF model, second only to SLFEND, achieved an F1 score of 0.920. There are a total of 491 news items in the education domain, with 243

real news and 248 fake news. Although the sample size is not large, it is able to detect fake news well. The DESF model achieved an F1 score of 0.923 in this domain. From the perspective of the nine domains included in this dataset, the green line representing the DESF model is generally evenly distributed, indicating this model is relatively stable and able to maintain a relatively accurate detection rate.

The FakeNewsNet dataset primarily consists of two domains: PolitiFact and GossipCop. As shown in Table 6, the proposed DESF outperforms the other five models. On the PolitiFact dataset, DESF achieved an F1 score of 0.862, representing a 0.001 improvement compared to the second-best model MDFEND's score of 0.861. On the GossipCop dataset, DESF achieved an F1 score of 0.828, surpassing the second-best model CFPFND by a margin of 0.2%, demonstrating good detection performance. Therefore, overall, the DESF model proposed in this paper shows some progress. Compared to the results on the Chinese dataset, the DESF model exhibits better fake news detection performance on the Chinese dataset than on the English dataset. This may be due to significant differences in grammar, vocabulary, and sentence structures between Chinese and English. These differences can lead to variations in the model's performance when processing the two languages. For example, Chinese words often have richer meanings and contextual dependencies, while English may focus more on word order and grammatical structures. The results on PolitiFact is shown in Fig. 3 and the results on GossipCop is shown in Fig. 4.

Table 6: Performance comparison between different methods on FakeNewsNet

| Method | PolitiFact | GossipCop | All |
|---------|--------------|--------------|-------|
| MDFEND | 0.861 | 0.806 | 0.822 |
| M3FEND | 0.856 | 0.822 | 0.820 |
| FuDFEND | 0.859 | 0.822 | 0.819 |
| SLFEND | 0.857 | 0.823 | 0.822 |
| CFPFND | 0.860 | 0.826 | 0.825 |
| DESF | 0.862 | 0.828 | 0.825 |

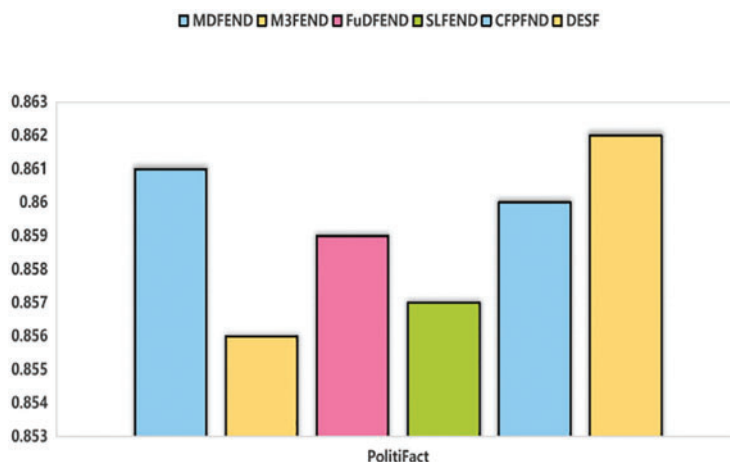


Figure 3: Results on PolitiFact

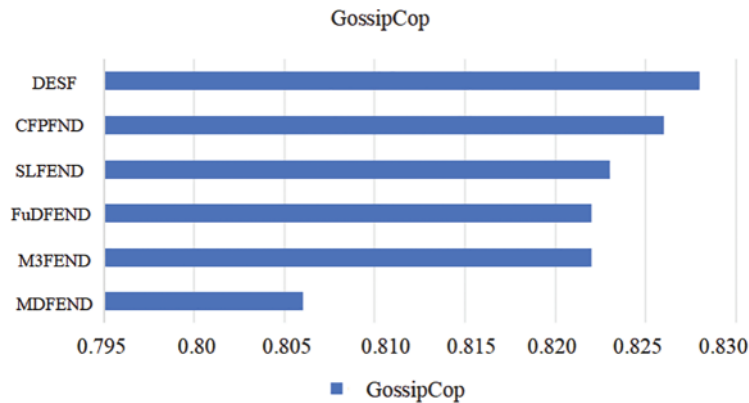


Figure 4: Results on GossipCop

From the perspective of whether the model combines feature fusion for false news detection, the experimental results are as follows:

Table 7 shows DESF exhibits outstanding performance compared to other baseline models. Among the nine domain datasets, the DESF achieved an F1 score of 0.958, which is 4.9% higher than the second-best model EANN and 6.3% higher than SpotFake. In the education domain, the DESF achieved an F1 score of 0.923, which is 6.3% higher than the second-best model EANN and 13.4% higher than BiGRU. Furthermore, in the areas of accident disasters, economy, health, entertainment, politics, and society, the DESF achieved F1 scores exceeding 0.9, demonstrating good performance. This suggests that feature fusion methods play an important role in the process of fake news detection, as combining different types of features through different models and algorithms can yield different results. DESF combines the relationship between domain-related sentiment features and semantic features, achieving good detection results on both datasets.

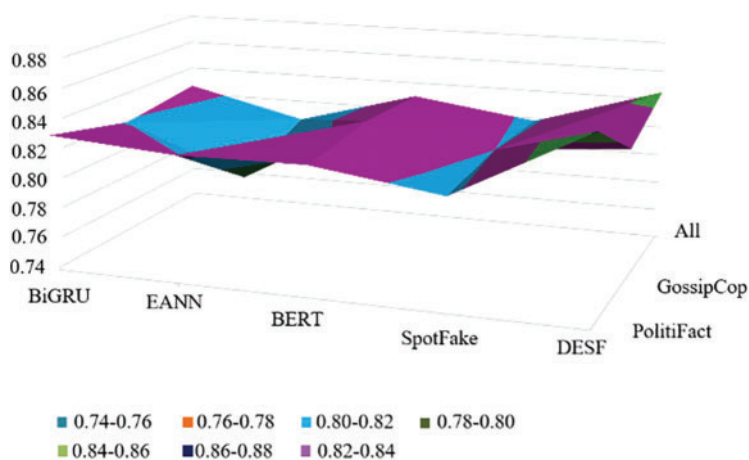
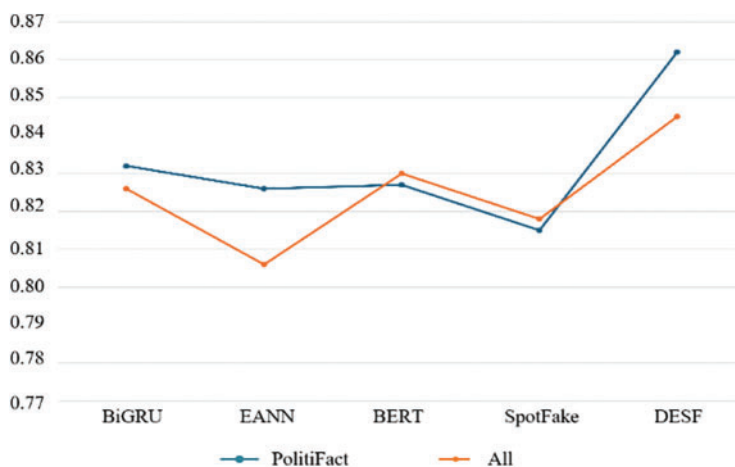
Table 7: Performance comparison between different methods on Weibo21

| Method | Edu. | Disasters | Health | Finance | Ent. | Politics | Military | Science | Society |
|----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| BiGRU | 0.814 | 0.794 | 0.887 | 0.829 | 0.863 | 0.836 | 0.872 | 0.727 | 0.849 |
| EANN | 0.868 | 0.894 | 0.935 | 0.879 | 0.919 | 0.869 | 0.913 | 0.826 | 0.904 |
| BERT | 0.833 | 0.851 | 0.909 | 0.875 | 0.877 | 0.921 | 0.907 | 0.777 | 0.858 |
| SpotFake | 0.857 | 0.893 | 0.964 | 0.847 | 0.903 | 0.819 | 0.901 | 0.932 | 0.898 |
| DESF | 0.923 | 0.913 | 0.959 | 0.927 | 0.931 | 0.920 | 0.958 | 0.887 | 0.929 |

Table 8 shows the performance of the five models: BiGRU, EANN, BERT, SpotFake, and DESF on FakeNewsNet. Through comparison, it is found that DESF achieved the best performance in the PolitiFact domain, obtaining an F1 score of 0.862, which is 3.6% higher than the second-best model BiGRU and 5.8% higher than the SpotFake model. In the GossipCop domain, DESF achieved the second-best performance. Overall, the DESF model, with an F1 score of 0.845, is 1.8% higher than the second-best model and 4.8% higher than EANN, demonstrating strong competitiveness and detection capabilities. To visually reflect the differences in model performance, this paper uses Figs. 5 and 6 to showcase the curve changes.

Table 8: Performance comparison between different methods on FakeNewsNet

| Method | PolitiFact | GossipCop | All |
|----------|--------------|--------------|--------------|
| BiGRU | 0.832 | 0.819 | 0.826 |
| EANN | 0.826 | 0.786 | 0.806 |
| BERT | 0.827 | 0.833 | 0.830 |
| SpotFake | 0.815 | 0.820 | 0.818 |
| DESF | 0.862 | 0.828 | 0.845 |

**Figure 5:** Performance between different models**Figure 6:** Performance on PolitiFact

Figs. 5 and 6 show the comparison of the results of different models. It can be clearly seen that DESF proposed in this paper has obvious advantages on the FakeNewsNet data set and can better realize the task of false news detection.

4.5 Ablation Experiment

To explore the contributions of each component in the proposed model, this paper conducts an ablation study to evaluate the impact of different features on the fake news detection algorithm. Specifically, the experiments are set up as follows: domain features, sentiment features, and semantic features are removed individually to observe the performance changes after the removal of these features.

Experiment ①: Remove the domain feature extraction step, and only fuse sentiment features and semantic features in the experiment.

Experiment ②: Remove the sentiment feature extraction step, and only fuse domain features and semantic features in the experiment.

Experiment ③: Remove the semantic feature extraction step, and only fuse sentiment features and domain features in the experiment.

Experiment ④: Remove the domain-related sentiment feature extraction step, only use semantic features for fake news detection.

These three experiments are validated on the Weibo21 and FakeNewsNet datasets separately to measure the performance and rationality of different modules, and are compared with the DESF model proposed in this paper. The experimental results are shown in [Table 9](#).

Table 9: Results of ablation experiments performed by DESF on two datasets

| | Method | Accuracy | Fake news | | | Real news | | |
|-------------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | | | Precision | Recall | F1 score | Precision | Recall | F1 score |
| Weibo21 | ① | 0.834 | 0.839 | 0.851 | 0.845 | 0.829 | 0.647 | 0.727 |
| | ② | 0.821 | 0.825 | 0.829 | 0.827 | 0.682 | 0.790 | 0.732 |
| | ③ | 0.817 | 0.819 | 0.824 | 0.821 | 0.547 | 0.698 | 0.613 |
| | ④ | 0.789 | 0.792 | 0.796 | 0.794 | 0.685 | 0.740 | 0.711 |
| | DESF | 0.896 | 0.899 | 0.913 | 0.906 | 0.891 | 0.784 | 0.834 |
| FakeNewsNet | ① | 0.823 | 0.803 | 0.834 | 0.818 | 0.832 | 0.849 | 0.840 |
| | ② | 0.827 | 0.799 | 0.863 | 0.828 | 0.829 | 0.837 | 0.833 |
| | ③ | 0.835 | 0.732 | 0.798 | 0.764 | 0.846 | 0.852 | 0.849 |
| | ④ | 0.799 | 0.657 | 0.721 | 0.689 | 0.801 | 0.805 | 0.803 |
| | DESF | 0.838 | 0.840 | 0.852 | 0.845 | 0.842 | 0.850 | 0.846 |

Through analysis, it can be found that DESF proposed in this paper, which combines domain-related sentiment features and semantic features for fake news detection, fully leverages the advantages of feature fusion and has absolute advantages. On Weibo21, it achieved an accuracy of 0.896, an F1 score of 0.906 for detecting fake news, and an F1 score of 0.727 for detecting real news. On the FakeNewsNet dataset, the accuracy is 0.838, the F1 score for detecting fake news is 0.845, and the F1 score for detecting real news is 0.834.

When Experiment ① removes domain features and only uses sentiment and semantic features on the Weibo21 dataset, the accuracy decreases by 7.4% compared to the DESF model, and the F1 score decreases by 7.2%. On the FakeNewsNet dataset, the accuracy decreases by 1.8% and the F1 score decreases by 3.3% compared to the DESF model, indicating that domain features are important.

When Experiment ② removes sentiment features and only uses domain and semantic features, the accuracy on the Weibo21 dataset is 0.821, and the F1 score is 0.827, representing a decrease of 9.1% and 9.6% respectively compared to DESF. On the FakeNewsNet dataset, the accuracy is 0.827 and the F1 score is 0.828, representing a decrease of 1.3% and 2.1% respectively compared to DESF, indicating that sentiment features contribute to fake news detection.

Comparing Experiments ① and ②, it is found that on the Weibo21 dataset, the detection performance after removing sentiment features declines faster than after removing domain features, indicating that sentiment features play a more prominent role.

In Experiment ③, when semantic features are removed and only sentiment and domain features are used for fake news detection, the accuracy decreases by 9.7% and the F1 score decreases by 10.4% compared to the DESF model, representing a larger performance decline than in Experiments ① and ②. This suggests that semantic features aid in understanding and judging a news item, and it is important to fully grasp the semantic information when detecting fake news to facilitate understanding of the news text content.

Observing the results of Experiment ④, which only uses semantic features and removes the attention mechanism to capture domain-related sentiment features, it achieves the lowest accuracy and F1 score among the five models. Compared to the DESF model, it is 13.6% lower in accuracy and 14.1% lower in F1 score on the Weibo21 dataset, and 4.9% lower in accuracy on the FakeNewsNet dataset. This demonstrates that the core module proposed in this paper (domain-related sentiment feature extraction) has a positive effect on fake news detection tasks. Visual comparisons of the ablation experiments are shown in Figs. 7 and 8.

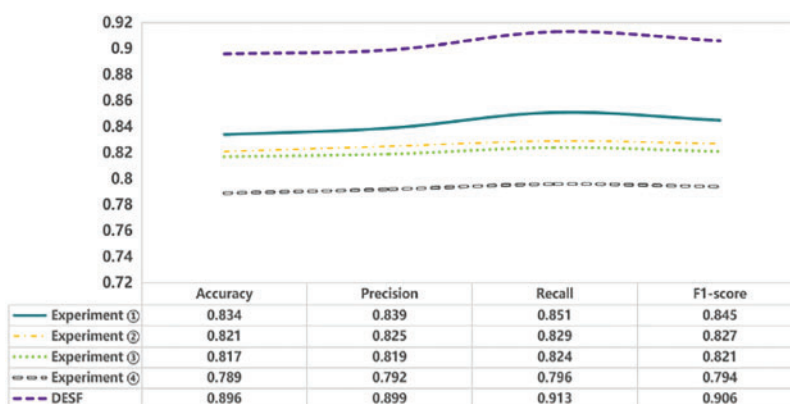


Figure 7: Ablation experiments on Weibo21

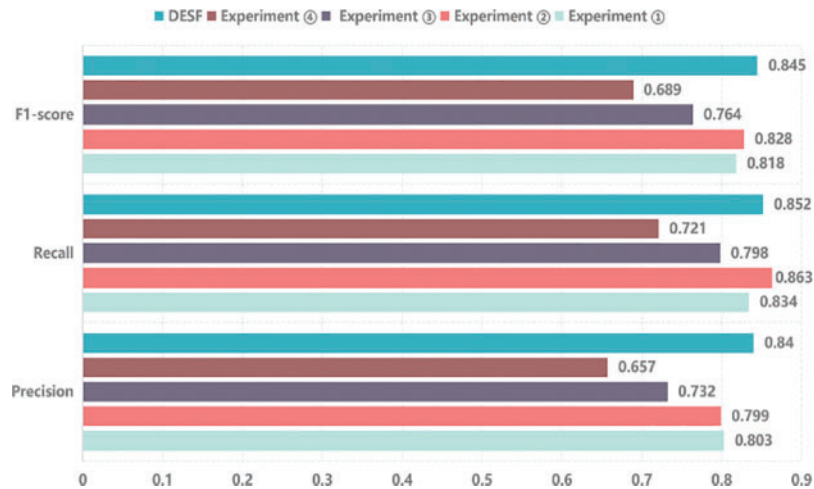


Figure 8: Ablation experiments on FakeNewsNet

5 Conclusion and Future Work

In the turbulent tide of the information age, fake news detection has become an urgent and critical task. Given its enormous social impact and potential harm, we delve into the field of fake news detection and come up with an innovative model. This model not only takes into account the semantic features of news texts, but also incorporates the emotional features and domain-specific expertise. By introducing the attention mechanism, we successfully achieve a deep integration of semantic features, emotional features and domain features. This fusion approach not only enhances the model's ability to understand news content, but also improves its sensitivity and accuracy in identifying fake news. The experimental results show that our model has achieved remarkable results in the task of fake news detection, which is not only more accurate than the existing methods, but also shows good robustness in dealing with complex situations. This result fully demonstrates the practical value and potential of our model in the field of fake news detection. In the future, we will continue to deepen the research on feature fusion methods and explore more efficient and accurate fusion strategies. At the same time, with the continuous development of information technology and the increasing richness of multi-modal data, we will also pay attention to the application of multi-modal and multi-domain data in fake news detection, and try to put forward new research directions and solutions.

Acknowledgement: We would like to give our heartfelt thanks to all the people who have ever helped us and we thank the reviewer for the positive and constructive comments regarding our paper.

Funding Statement: The authors are highly thankful to the National Social Science Foundation of China (20BXW101, 18XXW015). Innovation Research Project for the Cultivation of High-Level Scientific and Technological Talents (Top-Notch Talents of the Discipline) (ZZKY2022303). National Natural Science Foundation of China (Nos. 62102451, 62202496). Basic Frontier Innovation Project of Engineering University of People's Armed Police (WJX202316). This work is also supported by National Natural Science Foundation of China (No. 62172436) and Engineering University of PAP's Funding for Scientific Research Innovation Team, Engineering University of PAP's Funding for Basic Scientific Research, and Engineering University of PAP's Funding for Education and Teaching. Natural Science Foundation of Shaanxi Province (No. 2023-JCYB-584).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Wen Jiang, Mingshu Zhang, Xu'an Wang; data collection: Wei Bin, Kelan Ren; analysis and interpretation of results: Facheng Yan, Xiong Zhang; draft manuscript preparation: Wen Jiang. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data is not available..

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] H. T. Phan, N. T. Nguyen, and D. Hwang, "Fake news detection: A survey of graph neural network methods," *Appl. Soft Comput.*, vol. 139, no. 1, pp. 110235, May 2023. doi: [10.1016/j.asoc.2023.110235](https://doi.org/10.1016/j.asoc.2023.110235).
- [2] N. Capuano, G. Fenza, V. Loia, and F. D. Nota, "Content-based fake news detection with machine and deep learning: A systematic review," *Neurocomputing*, vol. 530, pp. 91–103, Apr. 2023. doi: [10.1016/j.neucom.2023.02.005](https://doi.org/10.1016/j.neucom.2023.02.005).
- [3] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, "Multimodal fusion with recurrent neural networks for rumor detection on microblogs," in *Proc. 25th ACM Int. Conf. Multimed.*, Mountain View, CA, USA, Oct. 2017, pp. 795–816.
- [4] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, "MVAE: Multimodal variational autoencoder for fake news detection," in *World Wide Web Conf.*, San Francisco, CA, USA, May 2019, pp. 2915–2921.
- [5] J. Zeng, Y. Zhang, and X. Ma, "Fake news detection for epidemic emergencies via deep correlations between text and images," *Sustain. Cities Soc.*, vol. 66, no. 3, pp. 102652, Mar. 2021. doi: [10.1016/j.scs.2020.102652](https://doi.org/10.1016/j.scs.2020.102652).
- [6] M. Delgado, N. Manín, M. J. Martín-Bautista, D. Sánchez, and M. A. Vila, "Mining fuzzy association rules: An overview," *Soft Comput. Inf. Process. Anal.*, vol. 164, pp. 351–373, 2005. doi: [10.1007/3-540-32365-1](https://doi.org/10.1007/3-540-32365-1).
- [7] H. Feng, M. J. Lesot, and M. Detyniecki, "Using association rules to discover color-emotion relationships based on social tagging," *Knowl.-Based Intell. Inf. Eng. Syst.*, vol. 6276, pp. 544–553, 2010. doi: [10.1007/978-3-642-15387-7](https://doi.org/10.1007/978-3-642-15387-7).
- [8] N. Sheydaei, M. Saraei, and A. Shahgholian, "A novel feature selection method for text classification using association rules and clustering," *J. Inf. Sci.*, vol. 41, no. 1, pp. 3–15, 2015. doi: [10.1177/0165551514550143](https://doi.org/10.1177/0165551514550143).
- [9] D. Rohidin, N. A. Samsudin, and M. M. Deris, "Association rules of fuzzy soft set based classification for text classification problem," *J. King Saud Univ.—Comput. Inf. Sci.*, vol. 34, no. 3, pp. 801–812, 2022. doi: [10.1016/j.jksuci.2020.03.014](https://doi.org/10.1016/j.jksuci.2020.03.014).
- [10] Y. Cai and X. Wan, "Multi-domain sentiment classification based on domain-aware embedding and attention," in *Proc. Twenty-Eighth Int. Joint Conf. Artif. Intell.*, Macao, China, 2019, pp. 4904–4910.
- [11] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 494–514, 2021. doi: [10.1109/TNNLS.2021.3070843](https://doi.org/10.1109/TNNLS.2021.3070843).
- [12] J. S. Liu, K. Feng, J. Pan, J. Z. Deng, and L. N. Wang, "MSRD: Multimodal network rumor detection method," *Comput. Res. Dev.*, vol. 57, no. 11, pp. 2328–2336, 2020.
- [13] A. Li, Z. P. Dan, F. M. Dong, L. W. Liu, and Y. Feng, "Rumor detection method based on improved generative adversarial network," *J. Chin. Inf. Sci.*, vol. 34, no. 9, pp. 78–88, 2020.
- [14] S. R. Sahoo and B. B. Gupta, "Multiple features based approach for automatic fake news detection on social networks using deep learning," *Appl. Soft Comput.*, vol. 100, no. 3, pp. 106983, 2021. doi: [10.1016/j.asoc.2020.106983](https://doi.org/10.1016/j.asoc.2020.106983).
- [15] Q. Zhang, Z. Guo, Y. Zhu, P. Vijayakumar, A. Castiglione and B. B. Gupta, "A deep learning-based fast fake news detection model for cyber-physical social services," *Pattern Recognit. Lett.*, vol. 168, no. 4, pp. 31–38, 2023. doi: [10.1016/j.patrec.2023.02.026](https://doi.org/10.1016/j.patrec.2023.02.026).

- [16] Z. Luo, Q. Li, and J. Zheng, "Deep feature fusion for rumor detection on Twitter," *IEEE Access*, vol. 9, pp. 126065–126074, 2021. doi: [10.1109/ACCESS.2021.3111790](https://doi.org/10.1109/ACCESS.2021.3111790).
- [17] J. Wang, H. Mao, and H. Li, "FMFN: Fine-grained multimodal fusion networks for fake news detection," *Appl. Sci.*, vol. 12, no. 3, pp. 1093, 2022. doi: [10.3390/app12031093](https://doi.org/10.3390/app12031093).
- [18] S. Li, T. Yao, S. Li, and L. Yan, "Semantic-enhanced multimodal fusion network for fake news detection," *Int. J. Intell. Syst.*, vol. 37, no. 12, pp. 12235–12251, 2022. doi: [10.1002/int.23084](https://doi.org/10.1002/int.23084).
- [19] Y. Guo, "A mutual attention based multimodal fusion for fake news detection on social network," *Appl. Intell.*, vol. 53, no. 12, pp. 15311–15320, 2023. doi: [10.1007/s10489-022-04266-w](https://doi.org/10.1007/s10489-022-04266-w).
- [20] A. Silva, L. Luo, S. Karunasekera, and C. Leckie, "Embracing domain differences in fake news: Cross-domain fake news detection using multi-modal data," *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 1, pp. 557–565, 2021. doi: [10.1609/aaai.v35i1.16134](https://doi.org/10.1609/aaai.v35i1.16134).
- [21] M. F. Mridha, A. J. Keya, M. A. Hamid, M. M. Monowar, and M. S. A. Rahman, "A comprehensive review on fake news detection with deep learning," *IEEE Access*, vol. 9, pp. 156151–156170, 2021. doi: [10.1109/ACCESS.2021.3129329](https://doi.org/10.1109/ACCESS.2021.3129329).
- [22] F. Liang, Y. F. Shu, M. Y. Lei, M. Zhou, J. Y. Kang and Q. L. Guo, "False news detection based on adaptive fusion of multimodal features," in *2024 IEEE 3rd Int. Conf. Electr. Eng., Big Data Alg. (EEBDA)*, Changchun, China, 2024, pp. 228–234. doi: [10.1109/EEBDA60612.2024.10485971](https://doi.org/10.1109/EEBDA60612.2024.10485971).
- [23] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020. doi: [10.1089/big.2020.0062](https://doi.org/10.1089/big.2020.0062).
- [24] Q. Nan, J. Cao, Y. Zhu, Y. Wang, and J. Li, "MDFEND: Multi-domain fake news detection," in *Proc. 30th ACM Int. Conf. Inf. Knowl. Manag.*, Queensland, Australia, Oct. 2021, pp. 3343–3347.
- [25] C. Liang, Y. Zhang, X. Li, J. Zhang, and Y. Yu, "FuDFEND: Fuzzy-domain for multi-domain fake news detection," *Nat. Lang. Process. Chin. Comput.*, vol. 13552, pp. 45–57, 2022. doi: [10.1007/978-3-031-17189-5](https://doi.org/10.1007/978-3-031-17189-5).
- [26] D. Wang, W. Zhang, W. Wu, and X. Guo, "Soft-label for multi-domain fake news detection," *IEEE Access*, vol. 11, no. 8, pp. 98596–98606, 2023. doi: [10.1109/ACCESS.2023.3313602](https://doi.org/10.1109/ACCESS.2023.3313602).
- [27] Y. Zhu *et al.*, "Memory-guided multi-view multi-domain fake news detection," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 7, pp. 7178–7191, Jul. 1, 2023. doi: [10.1109/TKDE.2022.3185151](https://doi.org/10.1109/TKDE.2022.3185151).
- [28] Z. Q. Wang, T. Chen, B. Y. Zhang, M. L. Zhang, C. Y. Sun and W. S. Zhang, "Multi-domain fake news detection based on cross-feature perception fusion," *Comput. Syst. Appl.*, vol. 33, no. 3, pp. 264–272, 2024. doi: [10.15888/j.cnki.csa.009439](https://doi.org/10.15888/j.cnki.csa.009439).
- [29] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NaacL-HLT*, 2009, vol. 1, pp. 4171–4186.
- [30] J. Ma *et al.*, "Detecting rumors from microblogs with recurrent neural networks," in *Proc. 25th Int. Joint Conf. Artif. Intell. (IJCAI 2016)*, New York, USA, 2016, pp. 3818–3824.
- [31] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru, and S. Satoh, "SpotFake: A multi-modal framework for fake news detection," in *2019 IEEE Fifth Int. Conf. Multimed. Big Data (BigMM)*, Singapore, 2019, pp. 39–47. doi: [10.1109/BigMM.2019.00-44](https://doi.org/10.1109/BigMM.2019.00-44).
- [32] Y. Wang *et al.*, "EANN: Event adversarial neural networks for multi-modal fake news detection," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, London, UK, Jul. 2018, pp. 849–857.