

RESEARCH

Open Access



# Research on sentiment analysis method of opinion mining based on multi-model fusion transfer learning

Zhongnan Zhao<sup>1,2\*</sup>, Wenjing Liu<sup>1</sup> and Kun Wang<sup>1</sup>

\*Correspondence:  
piconet@126.com

<sup>1</sup> School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China

<sup>2</sup> School of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China

## Abstract

With the popularity of social media, opinion mining has gradually become a popular research field. Among these fields, sentiment analysis mining is an important research direction in the field of opinion mining. It aims to reveal the public's sentiment tendency, and attitude towards specific topics or events by analyzing text data generated by users on online platforms and digital media. However, the large amount of opinion data usually lacks effective annotation, which limits the learning and construction of opinion models. Therefore, focusing on the problem of the scarcity of labeled data in opinion analysis, this paper proposes a mining method for public opinion sentiment analysis based on multi-model fusion transfer learning, that can make full use of the limited labeled data to improve the learning efficiency of sentiment features by integrating the advantages of different models. Additionally, it introduces a transfer learning strategy to enable the models of the target domains to perform better in the absence of labeled data. Furthermore, the attention mechanism is combined to strengthen the acquisition of key features of the emotional colors and improve the accuracy of sentiment analysis. Specifically, the paper uses the ERNIE model to generate dynamic representations of the text word vectors in the dataset. It also uses TextCNN and BiGRU to construct a joint model for extracting local and overall features of the text word vectors. The parameters of the feature layer of the trained model are migrated to the target domain through transfer learning. The attention mechanism is combined with the model to identify the extreme elements of the sentiment. Finally, the local and overall features are fused to achieve comprehensive mining of public opinion and emotional information. This method can effectively improve the accuracy and generalization of public opinion analysis in cases of data scarcity. In the experimental part, the paper conducts comparisons and analyses in eight aspects: word embedding model, model combination, attention mechanism, transfer learning, source domain dataset, target domain dataset, model training, and baseline model. The four indicators, namely accuracy, precision, recall, and F1-measure are used to evaluate the performance of the method. The experiments are thorough and detailed, demonstrating the effective improvement of opinion mining performance.

**Keywords:** Sentiment analysis, Transfer learning, Opinion mining, Attention mechanism

## Introduction

With the rapid development of social media and online forums, etc., there has been an explosive growth in the amount of public opinion information on online platforms. In order to gain a better understanding and make effective use of this vast amount of public opinion data, opinion mining has emerged as a popular research field. By analyzing and extracting large amounts of user-generated content such as social media posts and comments, opinion mining for sentiment analysis can provide valuable insights into public views and emotions. This is crucial for businesses, governments, and organizations in formulating strategies, identifying market trends, and effectively responding to challenges posed by public opinion.

Public opinion sentiment analysis mining aims to develop effective methods and techniques that help us understand and utilize the vast amount of online public opinion data. It seeks to accurately quantify the sentiment tendency, distribution of viewpoints, and identify hot topics. Traditional public opinion analysis relies on manual surveys and media monitoring, which are time-consuming and costly. However, by employing natural language processing, machine learning, and data mining technologies, it is possible to automate the collection, processing, and analysis of massive amounts of network data, enabling the quick extraction of information related to specific topics. Current research in this field shows a trend towards technological diversification. For example, methods based on sentiment dictionaries treat words in the text as sentiment labels and leverage semantic relations between words and syntactic rules to conduct analyses. There are studies using existing sentiment dictionaries to construct sentiment thesauri [1]. Another study involves studying video pop-ups to create a relevant sentiment dictionary for sentiment analysis [2]. Machine learning is also a critical research method. This type of approach utilizes a large amount of labeled or unlabeled data, extracts features using statistical machine learning algorithms, and generates sentiment analysis results. For example, the K-adaboost-based algorithm for network opinion influence is designed to calculate the degree of influence of different network opinions [3]. The combination of dictionaries and machine learning is used to express multiple emotions in a sentence [4]. The participle tool and the topic classification technique of the LDA model are employed to analyze the development trend of hot topics [5]. Furthermore, deep learning has emerged as a popular research approach. Deep learning can enhance the recognition accuracy of sentiment features by training complex neural network models to better capture meaningful and abstract semantic features, as well as optimize the model output automatically. For instance, an LSTM approach based on dependency embedding and attention mechanism is utilized for predicting sentiment polarity in a sentiment analysis task [6]. Another example is the design of a multi-task domain sentiment analysis model based on RoBERTa [7]. Additionally, a combination of multiple deep learning models is employed to construct a sentiment analysis method suitable for complex scenarios [8] and predict opinion trends [9].

However, there are still some problems and shortcomings with existing methods. Firstly, building a sentiment lexicon and formulating rules for judgment require extensive human effort. Secondly, most methods rely on traditional machine learning algorithms or rule-based approaches, which perform well on specific datasets but lack generalization capabilities and are dependent on the quality of labeled data. Furthermore, the

problem of a limited amount of labeled data in deep learning constrains the learning and construction of models, especially in the field of online opinion analysis and mining, where obtaining accurately labeled data is a complex task and there are very few publicly available large-scale labeled datasets. Therefore, these issues also make it challenging to solve the data scarcity problem.

Addressing the aforementioned concerns, this paper focuses on tackling the challenge of inadequately annotated data in public opinion. It proposes a novel approach, namely, the multi-model fusion transfer learning-based public opinion mining method. This method effectively integrates diverse model features to extract sentiment information, harnessing the advantages of different models. Additionally, it incorporates transfer learning to leverage existing sentiment models and mitigate the impact of data scarcity. Furthermore, an attention mechanism is employed to enhance the extraction of key emotional features. This approach utilizes the limited labeled data to improve the accuracy and generalization capability of sentiment analysis, effectively addressing the challenges posed by data scarcity.

The main contributions of this paper are:

- (1) Utilizing the ERNIE model as the word embedding layer to transform text into word vector representations. This allows for the utilization of rich information from multiple sources during pre-training and continuous learning to enhance the accuracy and reliability of the model.
- (2) Employing a joint model of TextCNN and BiGRU as a feature extractor, enabling the extraction of deep features across multiple dimensions. This facilitates a better understanding of emotional tendencies towards hot issues and sensitive topics.
- (3) Implementing a transfer learning approach using parameter Transfer, where model knowledge extracted from the source domain is applied in public opinion analysis. The attention mechanism is integrated to amplify the weight of sentiment keywords, thereby enhancing the model's accuracy and generalization ability for acquiring sentiment features.

The paper consists of five subsections. In the first section, the significance of studying sentiment analysis of Network public opinion, the current situation of sentiment analysis, and the research work of this paper are introduced; in the second section, the research work related to sentiment analysis of Network public opinion is discussed; the third section is the model framework, and the models used are introduced in detail; the fourth section is the experimental section, and eight sets of experiments are conducted to verify the feasibility of the model proposed in this paper; the fifth section Sect. “[Experimental analysis](#)” is the experimental section, which verifies the feasibility of the model proposed in this paper through eight sets of experiments; Sect. “Discussions” summarizes the sentiment analysis method of Network public opinion based on multi-model fusion transfer learning proposed in this paper and presents the outlook for the next work.

## Related work

Existing research on sentiment analysis and mining mainly focuses on three aspects: text word embedding, text feature extraction and model performance.

In-text word embedding. Processing natural language using machine learning algorithms requires the mathematization of language, and word embedding is a way to mathematize words in the language. The word embedding model used in the early studies of Network opinion is the Word2Vec model, and Wei et al. [10] used the Word2Vec model to achieve semi-automatic construction of product reviews, and on this basis, combined rule parsing and domain ontology to construct a feature-level sentiment analysis framework for product reviews. Li et al. [11] a sentiment analysis model for Network restaurant reviews is constructed by combining Word2vec, Bi-GRU, and Attention methods to address the current problem of consumers' difficulty in efficiently extracting information from the services provided by merchants. He et al. [12] build a deep learning-based analysis framework for university Network public opinion for the real-time problem of university Network public opinion, and use Word2Vec to embed words into text and use the LSTM-CFR method to classify Chinese words. The word vectors generated by the Word2Vec model have static characteristics and cannot well solve the situation of multiple words. Therefore, to meet these challenges, dynamic word vector models have been developed. In recent years, the BERT model has been used most frequently in the research of Network opinion sentiment analysis, and Li et al. [13] for the case that the BERT model cannot give contextual information, a method called GBCN is proposed, which sends the text to the BERT and context-aware embedding layer and then uses a gating mechanism to control the sentiment features of the BERT output. Liu et al. [14] proposed a Bert-BiGRU-Softmax deep learning model with an affective Bert model as the input layer to better understand consumers' emotions about goods. Li et al. [15] used BERT to construct a sentiment analysis model for Chinese stock market reviews to solve the problem of low accuracy of stock market sentiment analysis, and used the pre-trained sentiment analysis model to achieve effective recognition of stock market sentiment. The BERT model has more parameters, the model depth is deeper, and the problem of overfitting easily occurs when a small amount of data is trained. According to the above-mentioned situation faced by word embedding simulation, this paper adopts the ERNIE model as the word embedding template of this paper, which uses a continuous learning method to continuously draw information of words, structures, words, etc. in a large amount of information to achieve continuous improvement of simulation content.

In terms of text feature extraction. The existing feature extraction models used in Network opinion sentiment analysis are usually single or multiple models in series, which is difficult to extract more comprehensive feature information. Ren et al. [16] proposed a lexicon-enhanced attention network (LEAN) based on bidirectional LSTM for the relationship between words and sentences, which can not only capture sentiment words in sentences but also focus on important information in sentences. Wei et al. [17] constructed a BiLSTM model based on multi-pole orthogonal attention for the problem of an insufficient number of explicit emotion words, and applied it to implicit emotion analysis to achieve the distinction between words and emotional tendencies. Li et al. [18] investigated a bi-directional LSTM model with a self-attentive mechanism

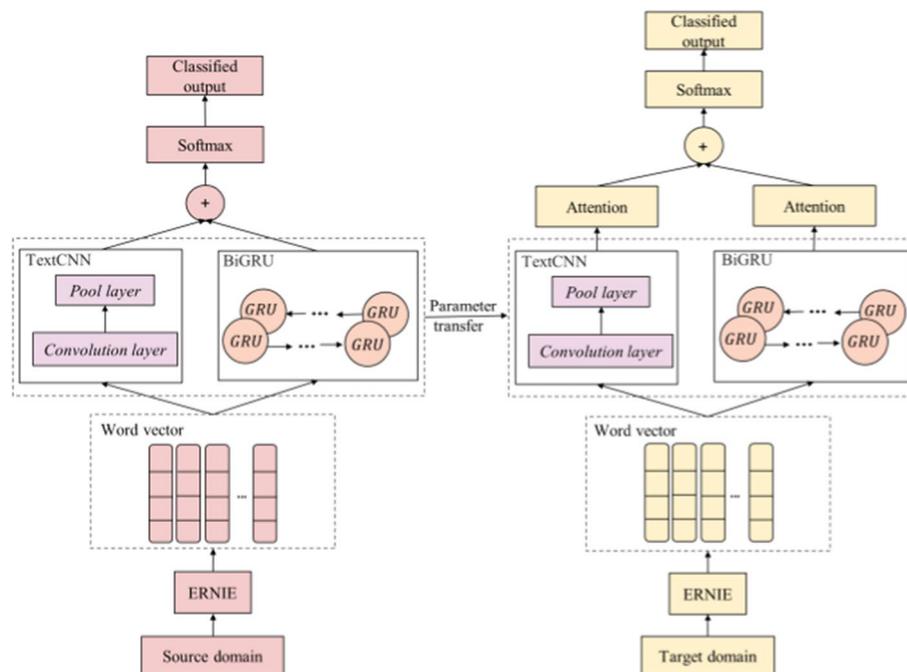
and multi-channel features that can fully explore the intrinsic connections between the polar components of the target words without relying on an artificially constructed sentiment dictionary. Huang et al. [19] proposed a lexicon-based contextual convolutional neural network to effectively extract the features and intensity of sentiment from words and texts. Gu et al. [20] to address the challenges faced in current sentiment analysis research, a new MBGCV method based on multi-granularity sentiment features is proposed by combining the advantages of the BiGRU model and the CNN model. Zhao et al. [21] to understand different sentiment polarities in comments, an aspect-based sentiment analysis model was constructed by combining convolutional neural networks with gated recurrent units in tandem, using the local feature extraction capability of CNN and combining the long-term dependent learning capability of GRU. Jain et al. [22] the proposed sentiment analysis method combines convolutional neural networks with long and short-term memory models, combining dropout, maximum pooling, and batch normalization to obtain the corresponding results. Lin et al. [23] proposed a sentiment analysis algorithm FAST-BiLSTM to optimize sentiment inference, which uses the FastText model to obtain word vectors, then uses a bi-directional long and short-term memory network (Bi-LSTM) to train word vectors, and then fuses them with FastText to perform comprehensive sentiment analysis. To address the shortcomings of the above studies, this paper uses TextCNN and BiGRU parallel structure models as text feature extractors to extract deep text features from multiple dimensions, to better understand people's sentiment tendency on hot issues and sensitive issues.

In terms of model performance. Most studies have been conducted to enhance the performance of models using a combination of transfer learning or attention mechanisms. On the one hand, relevant studies using attention mechanisms, Lv et al. [24] proposed a contextual and aspectual memory network (CAMN) approach based on deep memory networks, bidirectional long- and short-term memory networks, and multiple attention mechanisms to address the challenges of aspect-level sentiment analysis to better capture the sentiment characteristics of short texts. Sweidan et al. [25] proposed a hybrid ontology-XLNet approach for sentiment analysis classification at the sentence level, using the XLNet network to extract and associate the neighboring contexts in the text to generate more complete contextual information and improve the accuracy of feature extraction. Zhang et al. [26] established a sentiment analysis model based on BiTCN-Attention to better focus on sentiment words, which introduces different attention mechanisms in BiTCN to form BiTCN-SA and BiTCN-MHSA to improve the weight of sentiment words and the accuracy of feature extraction, and enhance the effect of sentiment analysis. Yang [27] combined BERT, CNN, and BiLSTM to construct the sentiment classification model BCBL; secondly, to address the problem that BCBL does not consider the distribution of word weights, the attention mechanism was introduced into the BCBL model and the BCBL-Att model was improved. The above studies improve the performance of the opinion analysis model by using the attention mechanism but lack the use of transfer learning; On the another hand, relevant studies using transfer learning, Tao et al. [28] proposed a sentiment analysis method based on the ABSA model by combining the ABSA model and transfer learning to address the shortcomings of the ABSA sentiment analysis method. Sanagar et al. [29] investigated an unsupervised sentiment dictionary that can transfer existing seed vocabulary from

category-level knowledge to the target domain and can be extended to new domains in the same category. Cao et al. [30] proposed a deep transfer learning mechanism (DTLM) for fine-grained cross-domain sentiment classification that better transfers sentiment across domains by combining BERT and KL divergence. Chandrasekaran et al. [31] used different migratory learning models including VGG-19, ResNet50V2, and DenseNet-121 models for image-based sentiment analysis and they were fine-tuned by freezing and unfreezing some layers. The above studies lacked applied research on the integration of transfer learning and attention mechanisms, and insufficient consideration was given to data as well as model synthesis factors. In addition, we conducted comprehensive research on the SOTA (State-of-the-Art) and compared them with our approach. Literature [32] utilizes the BERT model and incorporates CNN and BiLSTM to create an end-to-end model, achieving SOTA classification performance. Literature [33] also employs the BERT model and applies a BiLSTM model with an attention mechanism. Based on this model, two different applications are designed: pre-training and multi-task learning. In contrast, our paper utilizes the ERNIE model, an improved extension of BERT, which enhances pre-training effectiveness, semantic representation, and understanding. Meanwhile, this paper adopts the joint model of TextCNN and BiGRU. TextCNN is more specialized than CNN in the field of sentiment analysis. In addition to the advantages of higher computational efficiency and fewer parameters of BiGRU compared to BiLSTM, TextCNN and BiGRU employ parallel computing to capture local features and global semantic information of emotional data. This enables an improved comprehension of complex semantics and enhances computational efficiency, as opposed to the sequential computation approach employed by the BiLSTM model. Furthermore, our model incorporates attention mechanisms to capture crucial emotional features and filter out irrelevant noise, ultimately improving the accuracy of emotional information recognition. Our approach takes into consideration method selection, model design, and domain features in a comprehensive manner.

### **Description of the model framework**

To address the research problem of scarcity of labeled data in Network opinion sentiment analysis, this paper proposes a multi-model fusion transfer learning-based approach to Network opinion sentiment analysis (Transfer learning-ERNIE-TextCNN-BiGRU-Attention, TETBA), as shown in Fig. 1. The method incorporates the technical features of deep learning and transfer learning. The ERNIE model is used to generate the word embedding representation of the text of the dataset for model input, and the joint TextCNN and BiGRU model is used to extract the local features and overall features of the text word vector; based on the established opinion sentiment analysis model, news analysis data (source domain data) with rich annotation information is used to train the model; the parameters of the feature extraction layer of the trained model are transferred to the opinion sentiment analysis task in the target domain and combined with the attention mechanism to promote textual sentiment extremes, and finally the local features are fused with the overall features to complete the opinion sentiment analysis in the environment of scarce available labeled data.



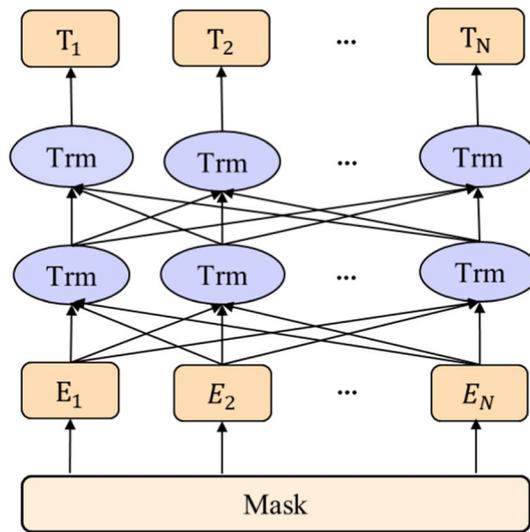
**Fig. 1** TETBA model

The model consists of the following components:

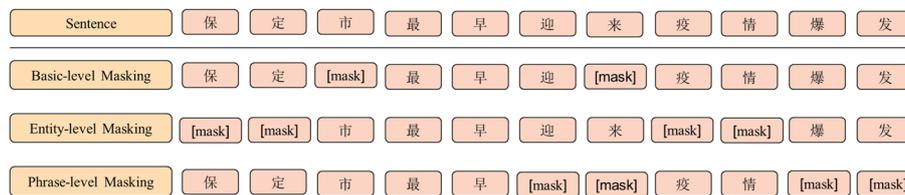
1. Word vector representation of public opinion data: the ERNIE model is used to transform the text data and extract the corresponding word vectors for input to the public opinion sentiment analysis model.
2. Joint model feature extraction: The joint model of TextCNN and BiGRU is used for feature extraction, where the TextCNN model is responsible for extracting local features and the BiGRU model is responsible for extracting overall features, and then the local features and overall features are merged.
3. Parameter transfer: the parameters of the feature extraction layer are transferred from the pre-trained model to the target model using the parameter transfer method to serve as the initial parameters of the feature extraction layer of the target model, and then the target model is trained using the network opinion data.
4. Attention and classification output: assign weights to the output matrix of the feature extraction layer to highlight the weights of emotional keywords, and then fuse the output features of different dimensions to achieve the classification of Network opinion data using the Softmax function to complete the emotional analysis of Network opinion.

**Word vector representation of public opinion data**

The task of Network opinion sentiment analysis requires language to be mathematical, and word vectors are an effective way to realize the conversion of input text to vector form. The ERNIE model is used in the paper to convert text data to word vector



**Fig. 2** ER NIE model framework diagram



**Fig. 3** Mask strategy

form. The greatest advantage of ER NIE is that it performs the initial fusion of a large amount of information from multiple sources and continuously draws new knowledge, such as vocabulary, structure, and semantics, from the vast amount of text data. The framework of the ER NIE model is shown in Fig. 2, including knowledge integration and Transformer coding. Knowledge integration is fusing semantic knowledge of entities and phrases into linguistic expressions, and Transformer can capture the contextual information of each sentence in a sentence through a self-attentive mechanism, producing a series of contextual embedding details as follows.

**Knowledge integration**

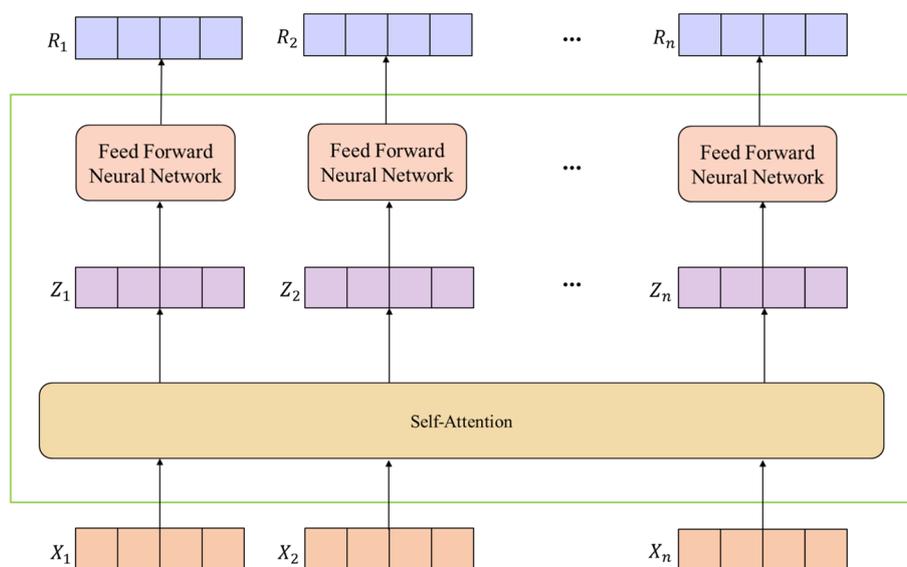
The ER NIE model is a modified version of the BERT model, and the improvement of BERT is mainly manifested in the mask (masking) strategy. The mask strategy is mainly to make the machine know the masked part of the input language by masking it and to achieve the training purpose in this way. Unlike the traditional BERT model which only uses a simple numeric word mask strategy, the ER NIE model replaces the Chinese word mask with consecutive entity word and phrase mask on this basis, thus greatly improving the understanding of the semantic content contained in the Chinese text, and the masking strategy of ER NIE model is shown in Fig. 3.

The masking strategy of the ERNIE model is divided into three stages, and the specific masking approach is as follows, taking “保定市最早迎来疫情爆发” as an example:

1. Phase 1: Basic level masking was used, where the sentence was viewed as a sequential basic language unit, with the basic unit being the word in English and the Chinese character in Chinese. In the pre-training process, the basic level mask is used to randomly mask a word in the sentence. In the example, the base level mask is used to cover “市、来”.
2. Second stage: entity-level masking is used, which can name places, people, items, etc. Firstly, we analyze the named entities in the sentence, and then mask part of them with the named entity-level mask, in the example sentence, we mask “保定、疫情”.
3. Phase 3: Phrase-level masking is used, where a phrase is used as a conceptual unit consisting of a series of characters. In this step, the named entity in the sentence is parsed first, and then the named entity-level mask is used to hide the part, which is “迎来、爆发” in the example sentence.

**Word vector dynamic representation**

After the masking strategy is the Transformer encoding part of the ERNIE model, the ERNIE model consists of twelve layers of Transformer encoder, although the ERNIE model is indeed based on the technical improvement of BERT, its internal network design is a two-way self-attentive mechanism layer of the encoder, the encoding device consists of two layers The encoder structure of the ERNIE model is shown in Fig. 4, where  $X_i(i = 1, 2, \dots, n)$  is the vector form of the text,  $Z_i(i = 1, 2, \dots, n)$  is obtained by the self-attentive mechanism, and  $R_i(1, 2, \dots, n)$  is the output of the ERNIE model and the input of the feature extraction layer model.



**Fig. 4** Encoder structure

The specific calculation process of the ERNIE model for handling public opinion data is as follows:

For a Network opinion text data, a corresponding segment of text sequence  $X$  is obtained after masking training in the knowledge integration stage:

$$X = [X_1, X_2, \dots, X_n] \tag{1}$$

to obtain the text vector for each input using the three weight matrices  $W_Q, W_K, W_V$  multiplied by the query matrix  $Q$ , the key matrix  $K$ , and the value matrix  $V$  on this basis.

$$Q = X \cdot W_Q \tag{2}$$

$$K = X \cdot W_K \tag{3}$$

$$V = X \cdot W_V \tag{4}$$

Then the self-attentive mechanism needs to be used to count all the attention of the target word to the words and phrases in the text sequence, where Softmax is performed to normalize the scores.

$$Z = Attention(Q, K, V) = \text{soft max} \left( \frac{Q \cdot K^T}{\sqrt{d_K}} \right) \cdot V \tag{5}$$

According to the self-attentive vector derived above, after mapping with the feedforward neural network, the next round of self-attentive operations will be performed, and the final obtained  $R_i(1, 2, \dots, n)$  is the output of ERNIE and the input of the feature layer model.

$$R = W \cdot Y + b \tag{6}$$

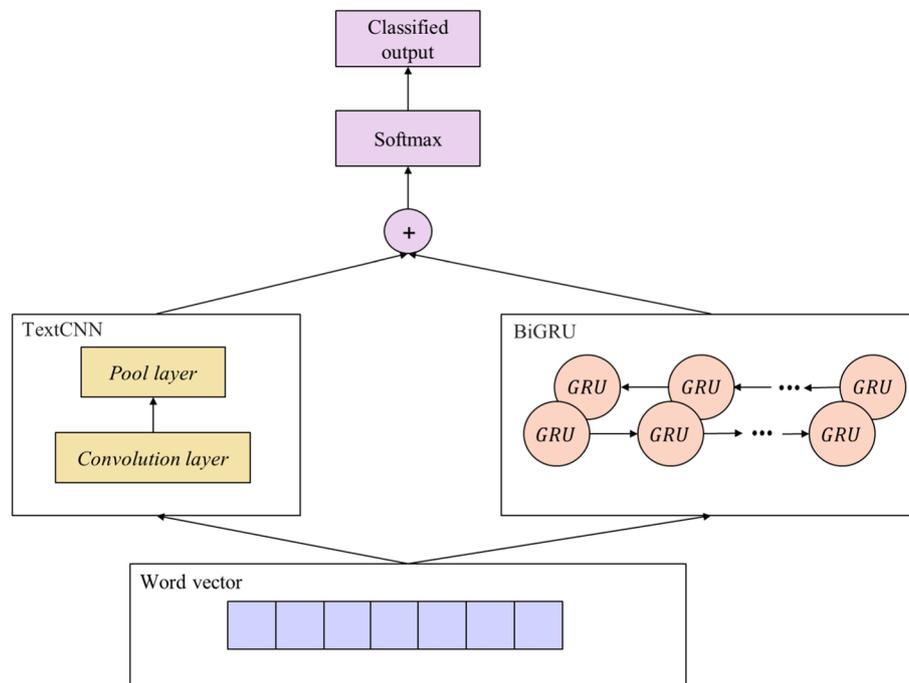
where,  $W$  is the weight matrix, and the bias value.

**Joint model feature extraction**

In this paper, we use the joint model of TextCNN and BiGRU to extract features from the word vectors generated by the ERNIE model, where the TextCNN model extracts the local features of the text and the BiGRU model extracts the overall feature information, and the joint model is shown in Fig. 5.

**Text CNN**

Convolutional neural networks can extract local features from a word vector matrix composed of multiple words, and can automatically synthesize and filter multiple words to obtain multi-level semantic information. By using multiple convolutional kernels, the TextCNN model can effectively extract important information in sentences and can capture local relevance more accurately. Compared with CNN models, TextCNN models present a more concise network structure, less computation, and faster training. textCNN models use multiple convolutional kernels of different sizes to extract important information in sentences simultaneously. This design allows the model to better



**Fig. 5** Joint model

capture local relevance and thus understand the meaning of the sentence more accurately. The model structure of the TextCNN used in this paper is shown in Fig. 6:

In Fig. 6, the feature vectors are obtained by convolutional pooling operation, where the size of convolutional kernels are 2, 3, and 4, respectively, and these feature vectors are composed of word vectors. The TextCNN model requires a series of complex computational processes when performing local feature extraction, and to ensure the accuracy and reliability of the method, a series of computations must be performed, which are shown below.

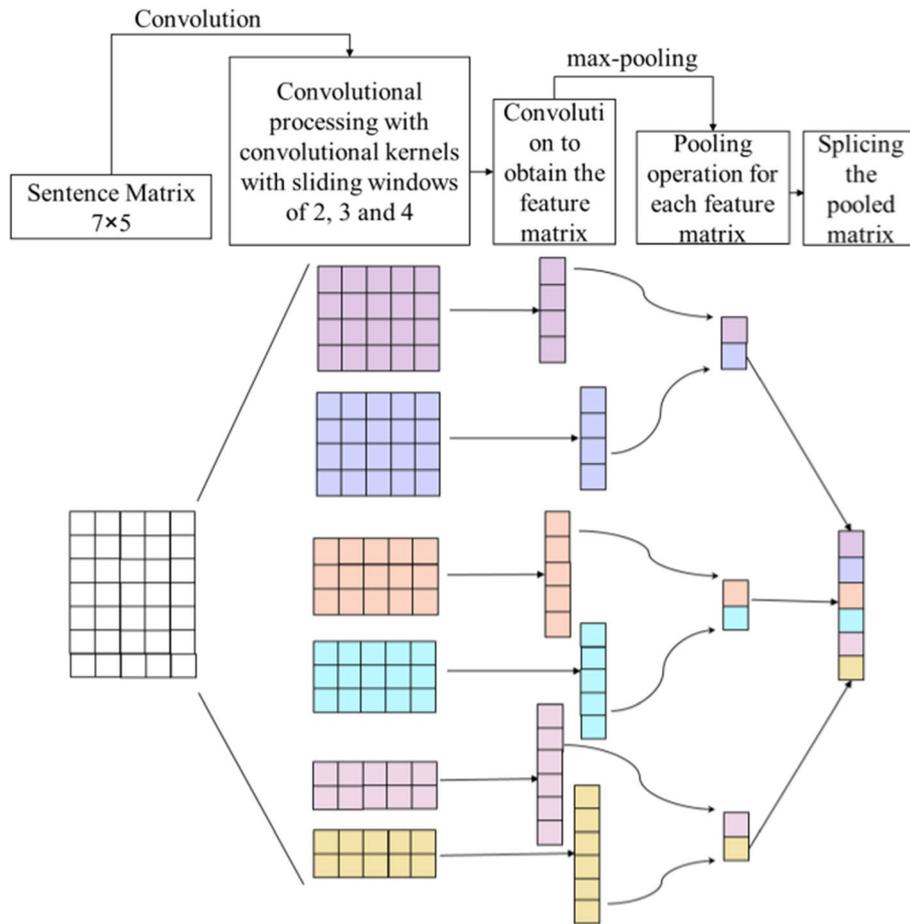
(1) Word vector input

For sentiment analysis of Network public opinion, the word vector matrix is used as model input data. In this chapter, the output of the ERNIE model is used as the input of Text CNN with word vector dimension 768. the pre-trained word vectors are embedded into the TextCNN model, and other corpora can also be used to obtain more prior knowledge. Through training, information closely related to the textual features of Network opinion is available, where the sequence length of the input text is  $n$ .

$$R = [R_1, R_2, \dots, R_n] \tag{7}$$

(2) Local feature extraction

TextCNN converts text into one-dimensional data by a one-dimensional convolution technique to extract features from the text, a process that can be implemented with word



**Fig. 6** TextCNN model structure

vectors. The convolution is consistent with the width and dimensionality of the word vector and can be self-determined in height. When the input sentence length reaches a certain value, a downward sliding trend occurs, which requires further analytical processing of the current output to obtain more accurate and complete information. For the downward sliding, different windows are defined to extract different feature vectors. In the paper, the convolutional kernels in the convolutional layer are of three sizes, and the width of the convolutional kernels is the same as the dimension of the word vector is  $d$ , and the height is 2, 3, and 4 respectively. There are multiple convolutional kernels of each size. For a web opinion text of length  $n$ , the convolutional layer can learn the attributes of the text by using  $h$  sliding windows of different sizes to convolve the text input vector and also to obtain the text convolutional feature values using the positions of the convolutional kernels. The computation of the feature  $C_i$  is represented by the following equation:

$$C_i = f(WR_{i:i+h-1} + b) \tag{8}$$

where  $W$  is the convolution kernel, which is a  $n \times d$  dimensional weight matrix,  $b$  is the bias,  $R_{i:i+h-1}$  is a sliding window consisting of rows  $i$  of the input matrix to extract

features from different locations in the text,  $h$  is the height of the convolution kernel,  $f$  is the activation function, and is the first  $i$  feature extracted by a convolution kernel. The entire convolutional kernel extracts the text features individually along the sliding window to obtain a feature vector extracted by the convolutional kernel at  $C$ .

$$C = [C_1, C_2, \dots, C_{n-h+1}] \tag{9}$$

(3) Maximum pooling

The pooling operation usually consists of two parts: average pooling and maximum pooling. The TextCNN model adopts a maximum pooling strategy, which can extract the most representative features from multiple sliding windows and combine them to form a complete vector representation. In the process of Network opinion sentiment analysis, each text contains some worthless information, and through maximum pooling, the most important keywords can be highlighted, thus assisting the model to locate the corresponding categories more easily. In this paper, the maximum pooling method is used, and the full feature vector is replaced by the maximum value in the feature vector  $C$  described above, and then the other same convolution kernels are used to convolve the obtained feature vector will also be implemented by maximum pooling to form a new feature vector  $S$ .

$$S_i = MAX\{C\} = MAX(C_1, C_2, \dots, C_{n-h+1}) \tag{10}$$

$$S = [S_1, S_2, \dots, S_m] \tag{11}$$

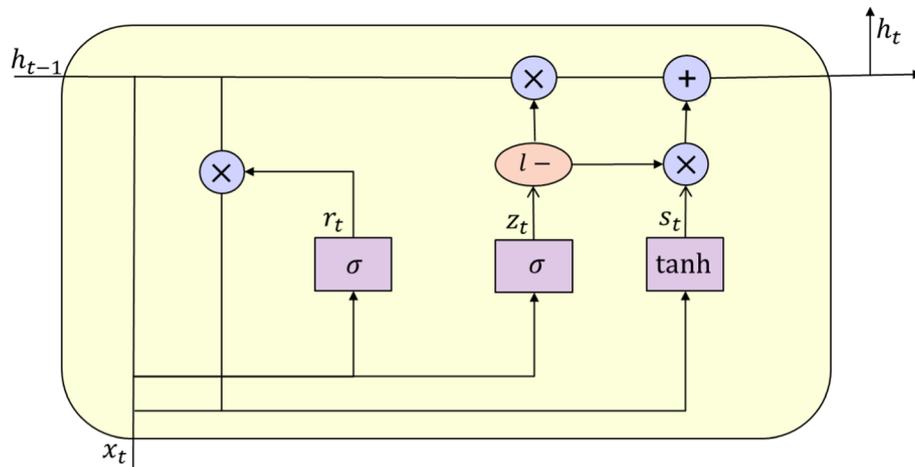
**BiGRU**

The traditional neural network has a unidirectional information transfer, and although this structure is useful for learning tasks, it constrains the capacity of the neural network when modeling. In many practical tasks, the output of a neural network is related to the input of the current moment in addition to the output of the previous moment. In addition, the length of ordered data such as video, speech, and text is often non-contrastive and fixed, whereas antecedent neural networks need to have fixed dimensions to process the data, and it is difficult for general neural networks to handle temporal data. Therefore, more efficient models are needed to solve the timing problem. GRU is a gated recurrent unit widely used in the field of timing models. The GRU model can fully utilize the power of LSTM while implementing complex architectures at a very low cost and providing excellent performance. The GRU model consists of an update gate and a reset gate, with the former controlling the process of combining new input information with an existing memory, while the latter defines the amount of existing memory in the current time step for more efficient information transfer.

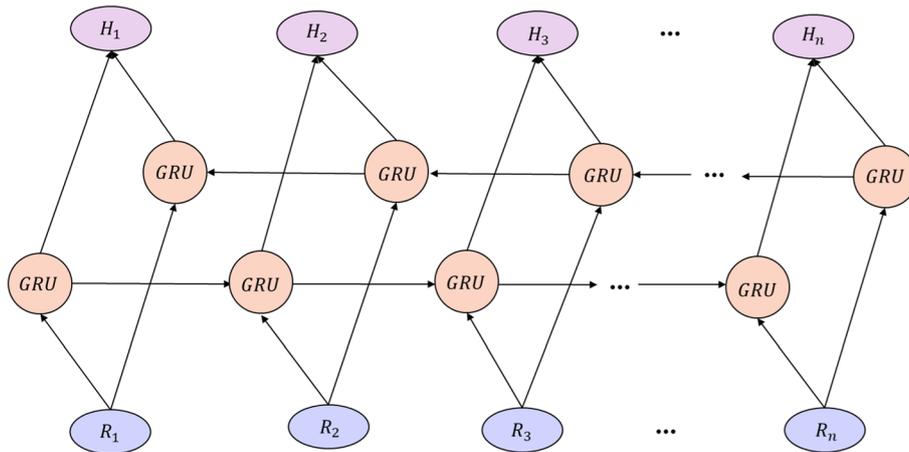
The structural diagram of the GRU model is shown in Fig. 7:

The parameter update of the GRU network is calculated as follows:

- 1) Compute the update gate  $z_t$ , where  $z_t$  is the update gate state,  $\sigma$  is the sigmoid activation function, and  $W_z$  is the weight parameter of the update gate.



**Fig.7** GRU model structure



**Fig.8** BiGRU structure diagram

$$z_t = \sigma(W_z \times [h_{t-1}, x_t]) \tag{12}$$

- 2) Compute the reset gate  $r_t$ , where  $r_t$  is the reset gate state,  $\sigma$  is the sigmoid activation function, and  $W_r$  is the weight parameter of the reset gate.

$$r_t = \sigma(W_r \times [h_{t-1}, x_t]) \tag{13}$$

- 3) Compute the candidate state  $s_t$ ,  $\tanh$  is the activation function and  $W_s$  is the weight parameter of the candidate state.

$$s_t = \tanh(W_s \times [r_t \times h_{t-1}, x_t]) \tag{14}$$

- 4) Calculate the current moment hidden status  $h_t$ .

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times s_t \tag{15}$$

However, the one-way GRU does not consider the global semantic information of the text and cannot analyze the emotion better, so this paper uses Bidirectional Gated Recurrent Unit (BiGRU) to extract the overall feature extraction, the structure diagram is shown in Fig. 8, it is composed of 4 parts, the input and output information of the input layer, at each moment The input and output information of the input layer is transmitted to the first hidden layer and the backward hidden layer synchronously at each moment, i.e., the information flows to the two GRU networks in opposite directions synchronously, and the information transmission order of the output layer is determined by these two GRU networks in opposite directions.

From Fig. 8, it can be concluded that since the BiGRU model is composed of two GRU networks with different directions, the hidden layer states of BiGRU are obtained by weighted summation of the forward hidden layer state  $\vec{h}_{i-1}$  and the reverse hidden layer state  $\overleftarrow{h}_{i-1}$  at time t to achieve feature extraction, where  $R_i$  is the dynamic word vector output from the ERNIE model, and the calculation process of the BiGRU model to obtain the overall features is shown below.

$$\vec{h}_i = GRU(R_i, \vec{h}_{i-1}) \tag{16}$$

$$\overleftarrow{h}_i = GRU(R_i, \overleftarrow{h}_{i-1}) \tag{17}$$

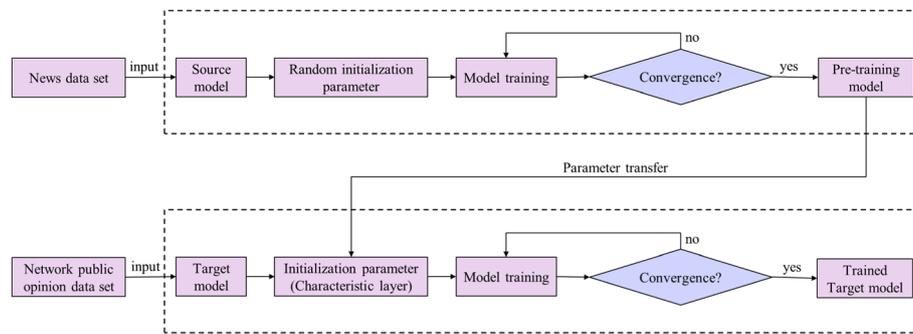
$$H_i = w_i \vec{h}_i + v_i \overleftarrow{h}_i + b_i \tag{18}$$

Through the joint model of TextCNN and BGRU, this paper obtains the local feature matrix  $S(S = [S_1, S_2, \dots, S_m])$  and the overall feature matrix  $H(H = [H_1, H_2, \dots, H_n])$  of the public opinion data.

**Parameter transfer**

Transfer learning is a machine learning technique that transfers knowledge from one domain (source domain) to another domain (target domain) to make the target domain better, and finding commonalities between the two domains and applying them skillfully is the key. According to the difference of what to perform the transfer, it is categorized into four types: instance-based, feature-based, model-based (parameter-based), and relationship-based. In this paper, we will introduce parameter-based transfer, which can find parameters that can be shared between source and target domains by constructing a parameter-sharing model, and migrate and transform these parameters effectively, because neural networks have a structure that can directly transfer parameters This method has been widely used in neural networks because of the structure of neural networks that can directly transfer parameters.

In this paper, we mainly divide into two parts: the source model and target model and use the TextCNN and BiGRU dual-channel model for feature extraction, where the TextCNN model extracts local features of text and the BiGRU model extracts overall



**Fig. 9** Transfer learning flow chart

feature information. Firstly, the label-rich news data are used to obtain feature parameters using TextCNN and BiGRU dual-channel models, and the pre-training model can obtain the common language of the text, and then the parameters of the feature extraction layer of the pre-training model are migrated to the target extraction layer of the target model, and these parameters will be used as the initial values of the feature extraction layer of the target task, and then the model is trained with the public opinion data. The flowchart of transfer learning is shown in Fig. 9.

The process of using transfer learning for sentiment analysis of Network public opinion is as follows:

- (1) Creation of neural network models using the ERNIE model and the joint model of TextCNN and BiGRU, i.e., source models pre-trained on a large news dataset, i.e., pre-trained models;
- (2) Build a target model which is designed like the pre-trained model except for the classification output layer, then add the attention mechanism and then add the attention mechanism and classification output layer whose output size is the number of categories in the web opinion dataset;
- (3) The parameters of the pre-trained model feature layer are migrated to the target model feature layer and used as the initial parameters of the target model feature layer, and the parameters of the pre-trained model ERNIE are kept unchanged, and next, the parameters of the classification output layer of the target model will be initialized randomly to achieve the best results;
- (4) Training by using labeled data in the target domain to optimize the model.

### Attention and classification output

In opinion sentiment analysis, the text often contains a large number of emotional words, which play a key role in the sentiment tendency of the text, and the application of the attention mechanism helps the emotional words to be captured, thus improving the sentiment polarity of the text. In this paper, the local feature matrix extracted by the TextCNN model  $S(S = [S_1, S_2, \dots, S_m])$  and the overall feature matrix extracted by the BiGRU model  $H(H = [H_1, H_2, \dots, H_n])$  is subsequently applied using the self-attention mechanism, and then the feature vectors obtained by the attention mechanism are merged and finally processed by applying the Softmax activation function to connect the whole system completely

and map them to the output classification results. For more details, please see the specific steps below.

- 1) First map each input to 3 different spaces to obtain 3 vectors: the query vector  $q_i$ , the key vector,  $k_i$  and the value vector  $v_i$ :

$$Q_S = S * W_i^Q \tag{19}$$

$$Q_H = H * W_i^Q \tag{20}$$

$$K_S = S * W_i^K \tag{21}$$

$$K_H = H * W_i^K \tag{22}$$

$$V_S = S * W_i^V \tag{23}$$

$$V_H = H * W_i^V \tag{24}$$

- 2) The attention evaluation function is scaled dot product to obtain a sequence of output vectors:

$$M_S = \text{soft max} \left( \frac{Q_S K_S^T}{\sqrt{d}} \right) V_S \tag{25}$$

$$M_H = \text{soft max} \left( \frac{Q_H K_H^T}{\sqrt{d}} \right) V_H \tag{26}$$

- 3) Fusion of the feature vectors obtained by the attention mechanism:

$$M = [M_S, M_H] \tag{27}$$

- 4) Classification results of Network public opinion using Softmax activation function mapping:

$$y = \text{soft max}(w^T * S + b) \tag{28}$$

where  $w$  is the fully connected layer weight and  $b$  is the bias term.

### Experimental analysis

To better evaluate the effectiveness of the proposed multi-model fusion transfer learning-based method for sentiment analysis of Network public opinion, this part will analyze the environment in which the experiments are conducted, the data sets in the source and target domains, and the evaluation indexes, etc. Eight kinds of currently used comparative experiments for opinion analysis, such as word embedding models, model combinations, with and without attention mechanism, and with and without transfer learning, are selected in the experimental part.

### Experimental environment and dataset

#### 1. Experimental environment

The experiment selects the PyCharm development platform and uses Pytorch1.1.0 deep learning framework for the implementation of a multi-model fusion transfer learning-based sentiment analysis method for Network public opinion, and the model mainly uses Python3.7 for code writing, and the experimental environment is shown in Table 1.

#### 2. Experimental dataset

The source dataset is derived from the THUCNews dataset published by Tsinghua University Laboratory, which contains a large amount of information and results. It contains 200,000 news headlines in ten categories, with 20,000 news items in each category. To perform machine learning and predictive analysis on this sample, a Chinese word separation method based on deep confidence networks is proposed. The data uses text as the input unit of the model, including 180,000 training sets, 10,000 validation sets, and 10,000 test sets.

The target dataset is a short-text microblogging dataset during the outbreak, provided by the official CCF BDCI platform, based on 230 keywords related to “novel coronary pneumonia”, and extracted from one million tweets, with 100,000 manually annotated, classified into positive, neutral, and negative categories to better understand the users’ situation.

Since the appeal dataset was not processed for data, the data may contain a large number of @ symbols, blank lines, and spaces. To improve the data quality, regular expressions are used in the text for denoising.

**Table 1** Experimental environment

Experimental environment	Environment configuration
Programming Languages	Python 3.7
Deep Learning Framework	Pytorch 1.1.0
CPU	AMD R7-5800H
GPU	NVIDIA RTX3060LP
Operating System	Windows 10

**Table 2** Hyperparameters of the ERNIE model

Parameter	Value
Attention_probs_dropout_prob	0.1
Activation Function	Relu
Hidden_dropout_prob	0.1
Hidden_size	768
Initializer_range	0.02
Max_position_embeddings	513
Num_attention_heads	12
Num_hidden_layers	12
Type_vocab_size	2
Vocab_size	18,000

**Table 3** Hyperparameters of the TextCNN model

Parameter	Value
Hidden_size	768
Filter_sizes	(2, 3, 4)
Num_filters	256
Activation Function	Relu
Dropout rate	0.1
Pooling	1-max pooling
Batch size	128
Learning rate	5e-5
Pad size	32
Optimizer	Adam
Criterion	CrossEntropyLoss
num_classes	class_list

**Table 4** Hyperparameters of the BiGRU model

Parameter	Value
Hidden_size	768
Bigru_hidden	768
Num_layers	2
Bidirectional	True
Batch_first	True
Dropout rate	0.1
Batch size	128
Learning rate	5e-5
Optimizer	Adam
Criterion	CrossEntropyLoss

### 3. Hyperparameters of the model

The specifics of the hyperparameters of the model are shown in Tables 2, 3, 4, respectively.

#### Evaluation indicators

To evaluate the effectiveness of the model proposed in this paper for Network opinion sentiment analysis, a series of tests were conducted, including the evaluation of accuracy, precision, recall, and F1-measure. The percentage of correctly predicted results in the total sample is called the accuracy rate. The proportion of positive samples that are correctly predicted in the actual predicted sample can be confirmed by the Precision. From the Recall, it can be seen that for positive samples, a method that can accurately predict the percentage of positive samples can be determined.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (29)$$

$$P = \frac{TP}{TP + FP} \quad (30)$$

$$R = \frac{TP}{TP + FN} \quad (31)$$

where TP describes a positive sample being correctly predicted, FN describes a positive sample being incorrectly predicted as a negative sample, FP describes a negative sample being incorrectly predicted as a positive sample, and TN describes a negative sample being incorrectly predicted as negative. If either accuracy or recall is considered separately as a measure of model performance, the F1-measure is a coordinated average between accuracy and recall, so the F1-measure is used to coordinate the two and is compatible with accuracy as well as recall.

$$F1 = \frac{2PR}{P + R} = \frac{2TP}{2TP + FP + FN} \quad (32)$$

#### Analysis of experimental results

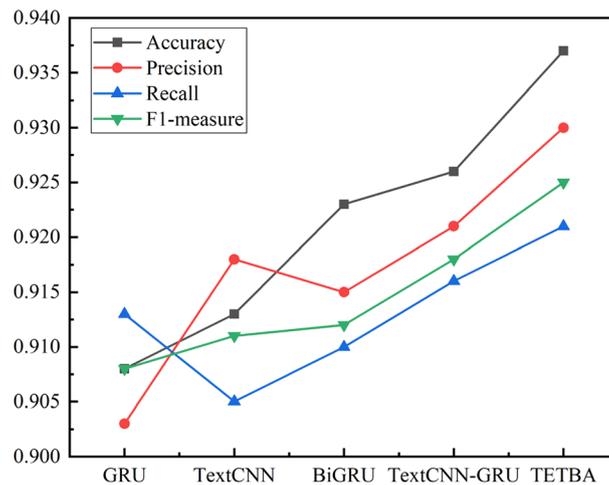
To better evaluate the effectiveness of the proposed multi-model fusion transfer learning-based sentiment analysis method for Network public opinion, eight currently used comparative experiments for public opinion analysis, such as word embedding model,

**Table 5** Experimental comparison results of ERNIE with other word embedding models

Method	Accuracy	Precision	Recall	F1-measure
Word2Vec	0.898	0.891	0.885	0.888
BERT	0.916	0.908	0.902	0.905
TETBA	0.937	0.930	0.921	0.925

**Table 6** Experimental comparison results of the TETBA model and its decomposition model

Method	Accuracy	Precision	Recall	F1-measure
GRU	0.908	0.903	0.913	0.908
TextCNN	0.913	0.918	0.905	0.911
BiGRU	0.923	0.915	0.910	0.912
TextCNN-BiGRU	0.926	0.921	0.916	0.918
TETBA	0.937	0.930	0.921	0.925



**Fig. 10** Experimental comparison results of the TETBA model and its decomposition model

model combination, with or without attention mechanism, and with or without transfer learning, were selected in this paper, and the specific experimental results and analysis are as follows.

**Comparison of word embedding models**

To study the effectiveness of the ERNIE model selected in this paper in terms of word embedding, two-word embedding models, Word2Vec and BERT, are selected for experimental comparison with the ERNIE model, and the word embedding ERNIE model in the TETBA model proposed in this paper is replaced by the Word2Vec model and the BERT model respectively for experiments, and the experimental results are shown in Table 5.

From the experimental results in Table 5, it can be intuitively seen that the four evaluation indexes using the ERNIE model as the word embedding model are the highest, and the four evaluation indexes using the WordVec model as the word embedding model is the lowest. the accuracy of the ERNIE model is improved by 3.9% and 2.1% compared with that of the Word2Vec model and the BERT model, respectively, and the F1-measure of the ERNIE model The main reasons are as follows: firstly, the Word2Vec model can only transform the text to a static word vector, which cannot solve the problem of multiple meaning words and cannot be dynamically optimized for specific tasks; secondly, the BERT model transforms the text into a dynamic word vector and uses the powerful

Secondly, the BERT model transforms text into a dynamic word vector and uses the powerful Transformer feature extractor for semantic representation, both from left to right and from right to left for bi-directional contextual representation; Finally, ERNIE is improved on the basis of BERT, which uses a large amount of data to model words, entities and entity relationships, compared with BERT, ERNIE learns more about the original language signal and ERNIE can directly construct a cross-semantic knowledge unit, so it can improve the semantic expression of the model. In summary, the ERNIE model is feasible as a word embedding model.

**Model combination comparison**

To better demonstrate the effectiveness of the combined approach of the model in this paper, the model is split into the form of GRU, TextCNN, BiGRU single model and the form of TextCNN, BiGRU tandem model, and the text data is vectorized using the ERNIE model as word embedding, and the transfer learning method is also used in conducting the experiments and the model proposed in this paper is comparison, the experimental results are shown in Table 6 and Fig. 10.

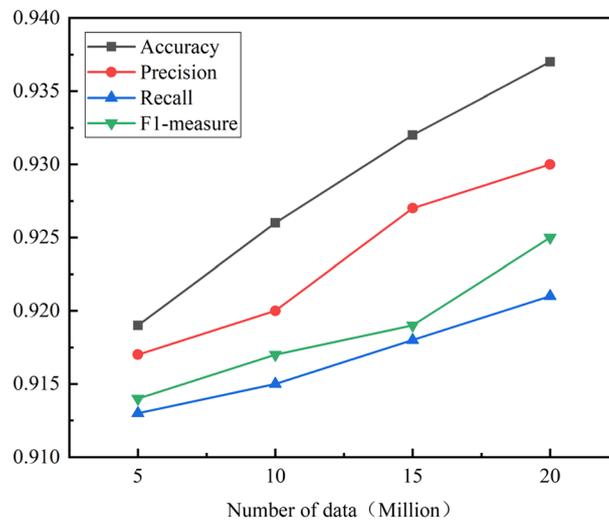
Table 6 and Fig. 10 show the results of the comparison experiments on the split part of the model, and it can be seen from the results in Table 6 that the BiGRU model works best among the three single models, GRU model, TextCNN model, and BiGRU model, with an accuracy rate of 92.3%, and when the BiGRU model is concatenated with the model of TextCNN, the accuracy rate of the concatenated model increases by 0.3 percentage points, and when the BiGRU model is concatenated with the model of TextCNN, the accuracy of the concatenated model increases by 1.4 percentage points, and the concatenated model works best. From the analysis of Fig. 10, it can be seen that the accuracy and F1-measure of all five models are on an increasing trend, the Precision is rising and then falling, and the Recall is falling and then rising. The reasons for this analysis are mainly as follows: first, the GRU model ignores the contextual information, and may make some adjustments to the hyperparameters to obtain the best performance; second, the TextCNN model uses different convolutional kernel sizes to extract N-gram information from the text and then uses the maximum pooling operation to highlight the most critical information extracted from each convolutional operation, which can better Third, the BiGRU model consists of opposing GRU networks, which can better extract the overall features of the text. Finally, the parallel model structure is more comprehensive than the series model structure for extracting text features. In summary, the parallel model of BiGRU and TextCNN is chosen as the model with stronger analysis ability and results.

**Table 7** Comparison results of experiments with and without attentional mechanisms

Method	Accuracy	Precision	Recall	F1-measure
TE-TextCNN-BiGRU	0.923	0.919	0.914	0.916
TETBA	0.937	0.930	0.921	0.925

**Table 8** Comparison results of experiments with and without transfer learning

Method	Accuracy	Precision	Recall	F1-measure
E-TextCNN-BiGRU	0.921	0.915	0.909	0.912
TETBA	0.937	0.930	0.921	0.925



**Fig. 11** Effect of source domain dataset size on Transfer Learning

**Comparison with and without attention mechanism**

To better demonstrate the effectiveness of the attention mechanism in the model of this paper, the ERNIE model was used as word embedding to vectorize the text data, and then combined with transfer learning, the model with and without the attention mechanism was compared with experiments, and the experimental results are shown in Table 7.

According to the experimental data in Table 7, all four evaluation metrics of the model reached the highest level after using the attention mechanism, which is a significant advantage compared with the model without the attention mechanism. Relative to the model without the attention mechanism, the accuracy rate increased by 1.4% and the Precision increased by 1.1% with the attention mechanism. This is because the attention mechanism can capture the emotional words of the text and improve their emotional polarity. In conclusion, the use of the attention mechanism in this paper can improve the accuracy of model sentiment analysis.

**Comparison with and without transfer learning**

To better demonstrate the effectiveness of transfer learning in this paper’s model, the ERNIE model was used to embed words into vectorized text data, and then the model with transfer learning and the model without transfer learning were compared and

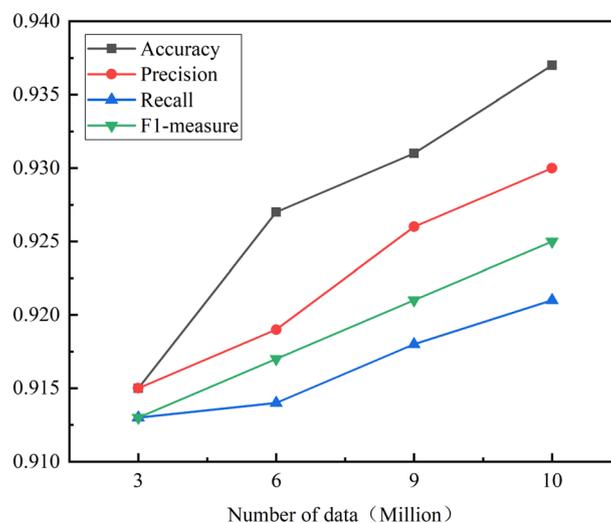
analyzed in conjunction with the attention mechanism, and the experimental results are shown in Table 8.

The experimental results in Table 8 show intuitively that the model using transfer learning has higher accuracy, precision, recall, and F1-measure. Compared with the model without transfer learning, the accuracy of the model with transfer learning increased by 1.6% and the precision rate increased by 1.5%. The reasons for this are analyzed as follows. To a certain extent, transfer learning solves the problem of the lack of large amounts of labeled data in the field of sentiment analysis. By learning the existing knowledge of similar domains with a large amount of labeled data, the learning rate is significantly improved and the computational effort is reduced. In summary, this paper uses transfer learning well to improve the performance of the model in sentiment analysis.

**Effect of source domain dataset size on transfer learning**

To investigate the effect of the size of the source domain dataset on transfer learning, experiments were conducted with 50,000, 100,000, 150,000, and 200,000 pieces of data from the source domain to observe the effect of different numbers of source datasets on transfer learning, and the comparison results are shown in Fig. 11.

Figure 11 shows the effect of the different number of source domain datasets on the transfer effect, from which it can be analyzed that all four evaluation metrics of the model tend to improve as the source domain data increases, the accuracy rate improves at approximately the same rate, the F1-measure improves slowly and then at a faster rate, and the four evaluation metrics are highest for 200,000 data and lowest for 50,000 data. It can be seen that better transfer learning was obtained using a large amount of annotated source data in the experiments.



**Fig. 12** Effect of the size of the target dataset on the transfer effect

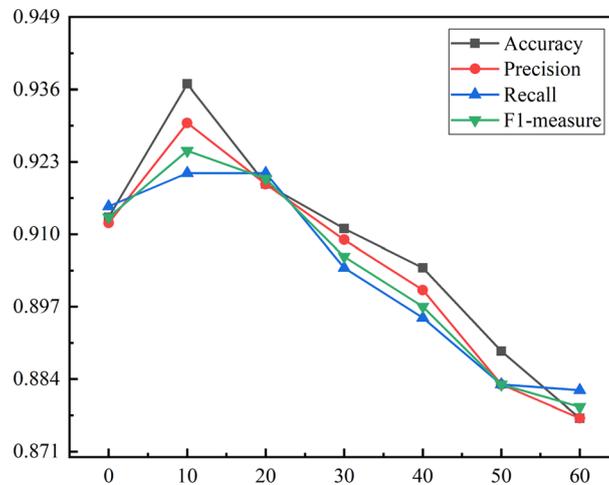


Fig. 13 Effect of different epochs on transfer effect

Table 9 Experimental comparison results with other baseline models

Method	Accuracy	Precision	Recall	F1-measure
BERT-BiGRU	0.901	0.882	0.874	0.878
CNN-GRU	0.909	0.901	0.903	0.902
BERT-DPCNN	0.913	0.909	0.905	0.907
RBR	0.916	0.910	0.910	0.910
TETBA	0.937	0.930	0.921	0.925

**Effect of target domain dataset size on transfer learning**

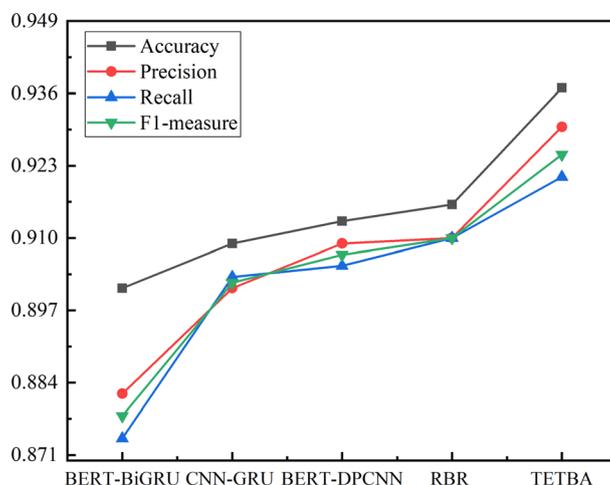
To investigate the effect of target domain dataset size on transfer learning, experiments were conducted with 30,000, 60,000, 90,000, and 100,000 data extracted from the target domain to observe the effect of different numbers of target datasets on transfer learning, and the comparison results are shown in Fig. 12.

Figure 12 shows the effect of different numbers of target data sets on transfer learning, from which we can analyze that as the target domain data increases, all four evaluation metrics of the model show an increasing trend, in which accuracy first rises rapidly and then slows down, and recall first rises slowly and then accelerates. The rate of increase of F1-measure is the same, with the highest four evaluation metrics for 100,000 items and the lowest four evaluation metrics for 30,000 items. The four evaluation indexes are the lowest. It can be seen that the larger the number of target datasets for conducting experiments, the better the effect of transfer learning.

**Effect of different epochs on transfer effect**

To study the effect of the source domain model on transfer effect under different epochs, the values of the epoch are selected as 0, 10, 20, ..., and 60 are trained on the source domain model, and the comparison results are shown in Fig. 13.

Figure 13 shows the effect of different epoch values on the transfer effect, from the Fig., it can be seen that with the increase of epoch value, all four evaluation indicators



**Fig. 14** Experimental comparison results with other baseline models

are not always increasing, the four evaluation indicators of epoch in the interval of [0,10] all show an increasing trend; the four evaluation indicators of epoch in the interval of [10,60] all show a decreasing trend, where the four evaluation indicators in the interval of [10,50] The decline rate of the four evaluation indicators in the [10,50] interval is the largest, the decline rate of the four evaluation indicators in the [50,60] interval is the smallest, and the four evaluation indicators in the epoch value of 10 is the highest. The reasons for this are as follows: the source domain model is trained too many times, and the training time is too long and too powerful so that it works too well on the current data that it cannot be generalized to the target domain data anymore, thus generating the overfitting phenomenon. It can be seen that the transfer effect does not get better as the epoch increases, and the transfer effect is best when the source domain model is trained for 10 epochs.

**Comparison with other baseline models**

To effectively prove the validity of the TETBA model in this paper, the BERT-BiGRU model is selected [14], the CNN-GRU model [21], BERT-DPCNN model and RBR model [34] The four baseline models are the experimentally compared with the model in this paper, and the experimental results are shown in Table 9 and Fig. 14.

Table 9 and Fig. 14 show the experimental results compared with the other baseline models, from the results in the table, it can be seen that the BERT-BiGRU model, the CNN-GRU model, the BERT-DPCNN model, and the RBR model have the best results among the four baseline models, with an accuracy of 91.6%, when using the BiGRU model in parallel with the TextCNN model and using transfer learning, the accuracy was improved by 2.1% and the F1-measure was improved by 1.5%. The analysis in Fig. 14 shows that all the evaluation metrics show an increasing trend, and it can be seen from the images that the rate of increase from the BERT-BiGRU model to the RBR model is smaller, the rate of increase from RBR to TETBA is larger, the rate of increase of F1-measure is a trend of accelerating, then slowing down and then accelerating, and the four evaluation metrics of TETBA are the highest, and

the analysis is mainly due to the following reasons. Firstly, the BERT-BiGRU model is using Bert model to extract multidimensional features and the bidirectional GRU model to obtain semantic code, CNN-GRU model is using local features generated by CNN and long-term dependency on GRU learning, therefore, the CNN-GRU model has higher recognition accuracy than BERT-BiGRU model. Secondly, DPCNN is a low-complexity word-level deep CNN for text classification, which can effectively model long-term dependencies in the text; the RBR model uses BERT as word embedding and then uses RCNN combined with attention mechanism for further extraction of contextual features of comments and combined with multi-task learning, so the CNN-GRU model, BERT- DPCNN model and RBR model the generalization ability of RBR model is a bit stronger. Finally, this paper uses the joint model of TextCNN and BiGRU for feature extraction and uses transfer learning to learn the existing knowledge, and the performance of the model is higher than that of the RBR model. In summary, the feature extraction model and the transfer learning analysis ability chosen for the model in this paper are stronger, which can better improve the accuracy and analysis performance of the model.

## Discussions

In order to expand the scope of this work and enhance the reader's comprehension of the research, this section will provide further explanations on the model analysis, data analysis, potential limitations analysis and other related aspects.

**Model analysis:** The computational patterns of deep learning models have expanded the realm of opinion analysis mining, offering a wide range of models applicable to research in this field. One such example is RoBERTa (A Robustly Optimized BERT Pre-training Approach). RoBERTa is an improved and optimized model based on BERT, enhancing performance and robustness through modifications in training methods and hyper-parameter settings. Another model worth mentioning is GPT (Generative Pre-trained Transformer). GPT is a Transformer-based language model primarily used for text generation, but it can also be utilized for sentiment analysis tasks. The latest version of GPT boasts powerful generative and linguistic comprehension capabilities. The selection of different models, customized to the problem scenarios, can aptly address various application requirements.

**Data analysis:** The selection of datasets and evaluation metrics plays a crucial role in opinion mining. In this paper, the selection of datasets from the source domain to the target domain takes into account various factors, including the relevance of the domain topic, the scale of the data, and the timeliness and reliability. As for performance evaluation metrics, Accuracy, Precision, Recall, and F1-measure are chosen. These metrics assess the prediction level of a large percentage of events, high-risk events, and the ability to accurately capture and predict high-risk events, respectively.

**Potential limitations analysis:** Among various factors influencing model performance, the data scale has a significant impact. Generally, larger data can enhance the model's ability to learn features, patterns, and generalization. However, if the data scale is too small, there is a risk of overfitting. In the case of traditional sentiment analysis models, choosing a small dataset may not cover an adequate range of sentiment types and semantic variations, thus limiting the model's understanding and prediction of different

sentiments and expressions. The models may overfit the training data, resulting in poor generalization and limited applicability in real-world scenarios. Nevertheless, in this paper, multi-model fusion transfer learning can be employed to overcome some of these limitations to some extent. By leveraging existing large-scale datasets and pre-trained models, this approach improves the performance and generalization ability of models on small datasets, benefiting from the pre-existing knowledge and representation capabilities of the pre-trained models. In summary, although there are still limitations in research using small datasets, the utilization of multi-model fusion transfer learning can mitigate some of these limitations and enhance model performance and generalization.

**Future prospects:** Using multi-modal fusion in research is currently a hot topic and trend. By combining image, video, and other public opinion data resources, we can further leverage the advantages and potential of multi-model transfer learning. For image data, we can employ a convolutional neural network (CNN) during the multi-modal fusion stage to extract image features. These features can then be combined with text features for joint modeling. Regarding video data, a temporal sequential convolutional neural network can be utilized to capture temporal correlations and spatial features. This approach allows the model to utilize visual information from images and videos, enabling richer and more accurate sentiment analysis. In a multi-modal context, multi-model fusion transfer learning can enhance the learning process by jointly extracting features from different modalities. This also enables the model to benefit from the features learned from existing modalities and apply them during training for the target modality. The cross-modal feature transfer achieved through the fusion of different models helps the model better understand the connections and dependencies between different modalities. Consequently, this improves the model's performance and generalization capabilities.

**Extended analysis:** In order to broaden the scope of research, it is worth exploring ways to improve the accuracy of sentiment mining across various types of opinion data. For instance, one can attempt to integrate models of different dimensions in a flexible manner during model fusion. Additionally, different transfer strategies can be explored to create diverse transfer entities. Moreover, efforts can be made to enhance the capability of handling different types of opinion data, catering to the varying requirements of information extraction, including differences in data structure, industry domains, and sentiment expressions.

Here, we present several ideas and thoughts regarding the next steps of our research work. We hope to conduct further in-depth research to propose more valuable solutions. Simultaneously, we will incorporate relevant literature and expert opinions from the field to continuously optimize and improve our research content. These contributions will serve as references and directions for future research, promoting the development and progress of this field.

## **Conclusion**

In this paper, based on the problem of lack of a large amount of labeled data for deep learning, we propose a multi-model fusion transfer learning-based method for web opinion sentiment analysis, pre-training the label-rich news data, using a joint TextCNN and BiGRU model for feature extraction, in which the TextCNN model extracts the

local feature information of the text, and the BiGRU model performs the overall feature information. The trained feature layer parameters are migrated to the opinion sentiment analysis model as the initial parameters of the web opinion sentiment analysis model, and then trained with the web opinion data. The results of eight sets of comparison experiments show that the web opinion sentiment analysis model based on multi-model fusion transfer learning achieves performance improvement on the public opinion task. In addition, with the development of the Internet, opinion data is not only text, but also pictures and videos are one of the ways people express their opinions, and it is hoped that multi-modal models can be studied to analyze the sentiment of text, picture and video data in the future.

#### **Acknowledgements**

We are grateful to the Institute of Network Technology and Information Security of Harbin Engineering University for technical support.

#### **Author contributions**

ZZ and WL wrote the manuscript. WL analyzed and organized the content. KW performed data analysis.

#### **Funding**

This work was supported by the the China Postdoctoral Science Foundation (Grant No.2019M651263).

#### **Availability of data and materials**

All the supporting data and files are shown in this manuscript.

#### **Declarations**

##### **Ethics approval and consent to participate**

Not applicable.

##### **Consent for publication**

Not applicable.

##### **Competing interests**

The authors declared no potential competing interests.

Received: 26 May 2023 Accepted: 3 October 2023

Published online: 11 October 2023

#### **References**

1. Jia F, Chen CC. Emotional characteristics and time series analysis of Internet public opinion participants based on emotional feature words. *Int J Adv Rob Syst*. 2020;17(1):1729881420904213.
2. Li Z, Li R, Jin G. Sentiment analysis of danmaku videos based on naïve bayes and sentiment dictionary. *IEEE Access*. 2020;8:75073–84.
3. Yuan S. Modeling and analyzing a public opinion influence method with K-adaboost. *Int J Mod Phys B*. 2020;34(28):2050257.
4. Szabóová M, Sarnovský M, Maslej Krešňáková V, et al. Emotion analysis in human-robot interaction. *Electronics*. 2020;9(11):1761.
5. Zhang M, Su H, Wen J. Analysis and mining of internet public opinion based on LDA subject classification. *J Web Eng*. 2021: 2457-2472-2457-2472.
6. Yang Z, Li C, Zhao Z, et al. Sentiment classification based on dependency-relationship embedding and attention mechanism. *J Intell Fuzzy Syst*. 2021;41(1):867–77.
7. Liao W, Zeng B, Yin X, et al. An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa. *Appl Intell*. 2021;51:3522–33.
8. Naz H, Ahuja S, Kumar D. DT-FNN based effective hybrid classification scheme for twitter sentiment analysis. *Multimed Tools Appl*. 2021;80:11443–58.
9. Wang W. Textual information classification of campus network public opinion based on BiLSTM and ARIMA. *Wireless Commun Mobile Comput*. 2022;2022:1.
10. Wei W, Liu YP, Wei LR. Feature-level sentiment analysis based on rules and fine-grained domain ontology. *KO Knowl Org*. 2020;47(2):105–21.
11. Li L, Yang L, Zeng Y. Improving sentiment classification of restaurant reviews with attention-based bi-GRU neural network. *Symmetry*. 2021;13(8):1517.
12. He M, Ma C, Wang R. A data-driven approach for university public opinion analysis and its applications. *Appl Sci*. 2022;12(18):9136.

13. Li X, Fu X, Xu G, et al. Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis. *IEEE Access*. 2020;8:46868–76.
14. Liu Y, Lu J, Yang J, et al. Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BIGRU-Softmax. *Math Biosci Eng*. 2020;17(6):7819–37.
15. Li M, Chen L, Zhao J, et al. Sentiment analysis of Chinese stock reviews based on BERT model. *Appl Intell*. 2021;51:5016–24.
16. Ren Z, Zeng G, Chen L, et al. A lexicon-enhanced attention network for aspect-level sentiment analysis. *IEEE Access*. 2020;8:93464–71.
17. Wei J, Liao J, Yang Z, et al. BiLSTM with multi-polarity orthogonal attention for implicit sentiment analysis. *Neurocomputing*. 2020;383:165–73.
18. Li W, Qi F, Tang M, et al. Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*. 2020;387:63–77.
19. Huang M, Xie H, Rao Y, et al. Sentiment strength detection with a context-dependent lexicon-based convolutional neural network. *Inf Sci*. 2020;520:389–99.
20. Gu T, Xu G, Luo J. Sentiment analysis via deep multichannel neural networks with variational information bottleneck. *IEEE Access*. 2020;8:121014–21.
21. Zhao N, Gao H, Wen X, et al. Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis. *IEEE Access*. 2021;9:15561–9.
22. Jain PK, Saravanan V, Pamula R. A hybrid CNN-LSTM: a deep learning approach for consumer sentiment analysis using qualitative user-generated contents. *Trans Asian Low-Resour Lang Inf Proc*. 2021;20(5):1–15.
23. Lin Z, Wang L, Cui X, et al. Fast sentiment analysis algorithm based on double model fusion. *Comput Syst Sci Eng*. 2021;36(1):175–88.
24. Lv Y, Wei F, Cao L, et al. Aspect-level sentiment analysis using context and aspect memory network. *Neurocomputing*. 2021;428:195–205.
25. Sweidan AH, El-Bendary N, Al-Feel H. Sentence-level aspect-based sentiment analysis for classifying adverse drug reactions (ADRs) using hybrid ontology-XLNet transfer learning. *IEEE Access*. 2021;9:90828–46.
26. Zhang J, Zhang T, Chen J. Sentiment analysis of chinese reviews based on BiTCN-attention model. *Int J Found Comput Sci*. 2022;33(06n07):755–70.
27. Yang H. Network public opinion risk prediction and judgment based on deep learning: a model of text sentiment analysis. *Comput Intell Neurosci*. 2022;2022:1.
28. Tao J, Fang X. Toward multi-label sentiment analysis: a transfer learning based approach. *J Big Data*. 2020;7:1–26.
29. Sanagar S, Gupta D. Unsupervised genre-based multidomain sentiment lexicon learning using corpus-generated polarity seed words. *IEEE Access*. 2020;8:118050–71.
30. Cao Z, Zhou Y, Yang A, et al. Deep transfer learning mechanism for fine-grained cross-domain sentiment classification. *Connect Sci*. 2021;33(4):911–28.
31. Chandrasekaran G, Antoanela N, Andrei G, et al. Visual sentiment analysis using deep learning models with social media data. *Appl Sci*. 2022;12(3):1030.
32. Pruttasha NJ, Sami AA, Kowshar M, et al. Transfer learning for sentiment analysis using BERT based supervised fine-tuning. *Sensors*. 2022;22(11):4157.
33. Xu G, Zhang Z, Zhang T, et al. Aspect-level sentiment classification based on attention-BiLSTM model and transfer learning. *Knowl-Based Syst*. 2022;245: 108586.
34. Hao WU, Shanliang PAN. Research on recognition of chinese illegal comments based on BERT-RCNN. *J Chin Inf Proc*. 2022;36(1):92–103.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---