

Sentiment analysis of microblog public opinion based on deep learning

Yi Xie , Kaiyu Li , Zhaomin Liu , Shicheng Li, Di Liu *

School of Information Engineering, Wuhan Business University, Wuhan 430056, China

* Corresponding author: Di Liu (Email: 63330350@qq.com)

Abstract: Public opinion sentiment analysis focuses on the public's emotional bias towards public events, and public opinion involving public health events will directly affect social stability, so it is essential for microblog sentiment analysis. In this paper, we use the RoBERTa-BiGRU-Attention model to analyze the sentiment of the text data of Weibo by taking the 2024 Paris Olympics as an example. Firstly, RoBERTa obtains the word embedding representation containing the text context information from the text data in the input layer. BiGRU obtains the character representation and sentence representation, and finally, the weight proportion of each character to the sentence in which it is located is calculated using the Attention mechanism, The text representation of the full text is obtained, and the softmax function carries out the sentiment analysis. To verify the effectiveness of the RoBERTa-BiGRU-Attention model, this paper takes Accuracy, Precision, Recall, and F1-measure as the evaluation indexes, and finds five models to compare with them. The results show that the RoBERTa-BiGRU-Attention model has good performance and advantages compared with the general model.

Keywords: microblog public opinion analysis; deep learning; BERT model; Attention mechanism; GRU

1. Introduction

With the rapid development of information technology and the popularization of social media, Weibo has become an important platform for people to obtain information, express opinions, and share emotions. The Weibo platform has accumulated a large amount of user-generated content, which reflects the development of social hot events and contains a wealth of emotional information and public opinion tendencies. Therefore, sentiment analysis of microblog public opinion has become a topic with great research value and application prospects^[2].

Traditional sentiment analysis methods are based on rules, dictionaries, or simple machine learning algorithms, which are good for small-scale, structured data, but often inadequate for large-scale, complex, and unstructured data such as microblogging. Specifically, it is difficult for traditional methods to accurately capture the complex semantic relationships, contextual information, and nuances of emotional expression in the text, resulting in limited accuracy and generalization ability of analysis results^[4].

To overcome these limitations, sentiment analysis methods based on deep learning have emerged. Deep learning can automatically learn the latent features of data by simulating the human brain's information processing mechanism to build a multi-level nonlinear network model, which can automatically learn the potential features of data and show powerful capabilities when processing large-scale and complex data. In the field of sentiment analysis, deep learning models such as recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) are widely used, which can effectively capture the temporal dependencies and contextual information in texts, to more accurately identify the sentiment tendencies of texts^[6].

Weibo public opinion sentiment analysis based on deep learning can automatically and efficiently process massive Weibo data and provide valuable public opinion monitoring

and decision-making support for governments, enterprises, and media through refined sentiment classification and tendency analysis. For example, the government can adjust the direction of the policy promptly by analyzing the public's emotional response to a policy or event; Enterprises can monitor consumers' evaluation of products or services and optimize marketing strategies; The media can more accurately grasp public sentiment and develop more compelling news stories.

In addition, with the continuous advancement of technology, sentiment analysis methods based on deep learning are also constantly optimized and innovated. For example, the combination of a sentiment dictionary for sentiment semantic enhancement, and the use of convolutional neural network (CNN) and LSTM hybrid model for feature extraction and sentiment classification are all aimed at further improving sentiment analysis's accuracy and generalization performance.

2. Related technologies

2.1. BERT

The BERT model randomly selects two sentences, and if the two sentences are adjacent, then there is a 50% probability that the two sentences are adjacent. The other 50% probability is that the two sentences are randomly selected. The BERT model mainly consists of the following three parts: TokenEmbeddings: converts the input words or subwords into vector representations; SegmentEmbeddings: Identifies the beginning and end of a sentence; PositionEmbeddings: Encodes the position information of a word into a vector representation^[4]. The BERT input feature is shown in Figure 1.

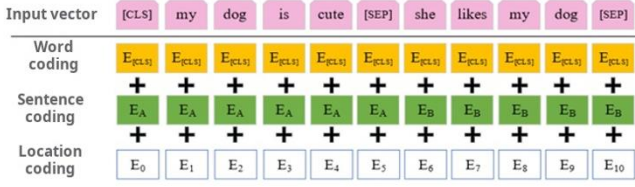


Figure 1 BERT input features

As shown in Figure 2, the structure of the BERT model trains each character $E_1, E_2, E_3, \dots, E_t$ in the input sequence through a bidirectional Transformer model to generate feature vector representations of the text, $X_1, X_2, X_3, \dots, X_t$. This model employs bidirectional Self-Attention for pre-training deep bidirectional representations, enabling the Attention calculation to take into account all input moments at each point in time^[5].

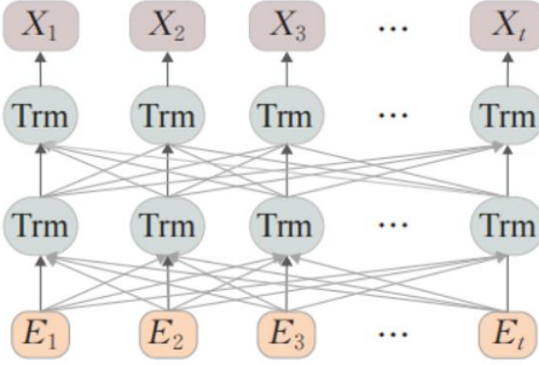


Fig.2. Structure diagram of the BERT model

RoBERTa has made three main improvements to BERT^[9]: (1) Using the dynamic masking mechanism, the BERT model will only perform a random MASK for each sample in the data preprocessing stage, and then the order of each round of word MASK is the same, which is a static MASK operation. RoBERTa, on the other hand, uses a dynamic masking mechanism that generates a new masking pattern each time a sequence is fed into the model. This way, while continuously entering a large amount of data, the model will gradually adapt to different masking strategies and learn different linguistic representations. (2) Remove the NextSentencePrediction (NSP) task and improve the efficiency of your model. (3) The corpus size of the pre-training process is increased, and the step size of the training is increased while the BatchSize is expanded. Consistent with the BERT model, the RoBERTa model also uses multiple bidirectional Transformer models to stack the encoder parts to form a main frame, which can more thoroughly capture the bidirectional relationships in the text^[5]. Figure 3 shows the logical structure of the Encoder of the Transformer model.

2.2. GRU

GRU (as shown in Figure 4) is an improved form of Recurrent Neural Network (RNN), which adopts a unique gating structure to solve the problems of gradient dispersion and gradient explosion in traditional RNNs. The reset gate controls the proportion of historical information in the current state, determining how much past information needs to be retained in the current state. The update gate controls the proportion of historical data in the current state, deciding how much new data needs to be fused into the current state. Compared to before the improvement, The training time is short, the parameters are few, and the structure is simple,

which can effectively save the training cost and is widely used in the field of deep learning^[8].

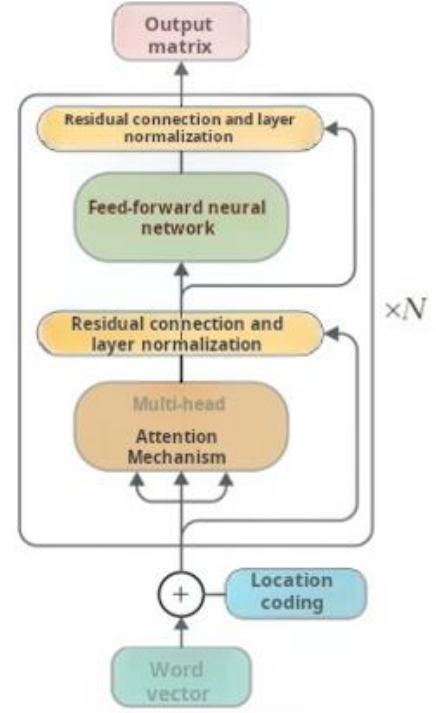


Figure 3 Transformer-encoder logical structure

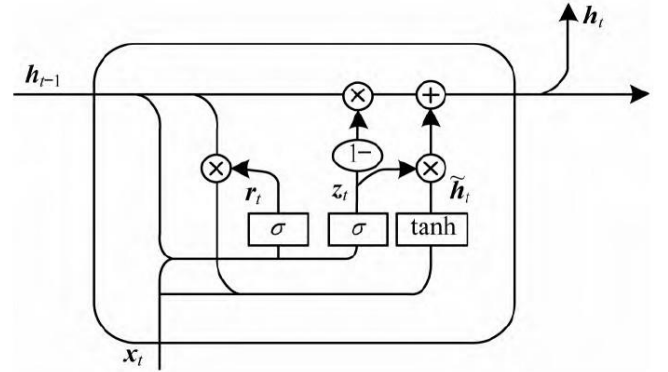


Figure 4 GRU network model structure

In the context, x_t represents the input at time t , r_t represents the reset gate at time t , z_t represents the update gate at time t , δ denotes the Sigmoid activation function, 1- indicates the gate signal weight 1, ranging from 0 to 1, h_t represents the candidate activation state at time t and h_t also denotes the activation state at time t . h_{t-1} and h_t respectively represent the output of the hidden layer from the previous moment and the state of the hidden layer at time t . In GRU networks, state information is output from front to back, allowing it to learn only information before the current moment, not after^[8]. In sentiment classification tasks, it is necessary for the state information of the current moment, the previous moment, and the subsequent moment to be interconnected. This paper adopts a Bidirectional Gated Recurrent Unit (BiGRU) instead of GRU, which can learn the contextual semantic information of a word. BiGRU is a neural network model composed of two unidirectional GRUs with opposite directions, and its output is jointly determined by the state information of these two GRUs^[7]. The network structure of BiGRU is shown in Figure 5.

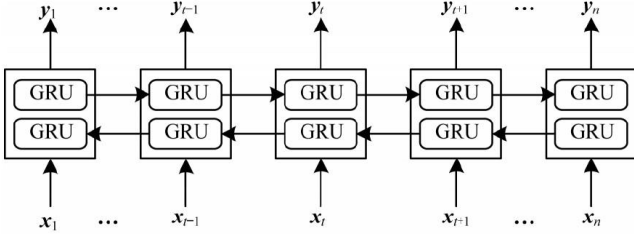


Figure 5: BiGRU network structure

As can be seen from the figure, the current hidden layer state of the BiGRU network is determined by three parts: the current input x_t , the output of the forward hidden layer state at $t-1$ time, and the output of the hidden layer state in the reverse layer at $t+1$ time. The formula for calculating the GRU network structure is as follows:

$$z_t = \delta(W_z \cdot [h_{t-1}, x_t]) \quad (1)$$

$$r_t = \delta(W_r \cdot [h_{t-1}, x_t]) \quad (2)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (3)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (4)$$

$$\vec{h}_t = GRU(x_t, \vec{h}_{t-1}) \quad (5)$$

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

$$h_t = v_t \vec{h}_t + w_t \overleftarrow{h}_t + b_t \quad (7)$$

Here, $\vec{h}_t$ and $\overleftarrow{h}_t$ represent the states of the forward and backward hidden layers at time t , respectively. v_t and w_t are weight vectors, and b_t is the bias.

In the equations, W_r and W_z are both weight matrices, and $\cdot$ denotes matrix multiplication. BiGRU comprises a forward and a backward GRU, allowing it to capture both forward and backward semantic information within the text^[1]. By incorporating context, BiGRU can deeply extract the emotional features embedded in the text. Furthermore, BiGRU utilizes a stacking approach to extract global text features. The specific model structure of Bi-GRU is illustrated in Figure 6.

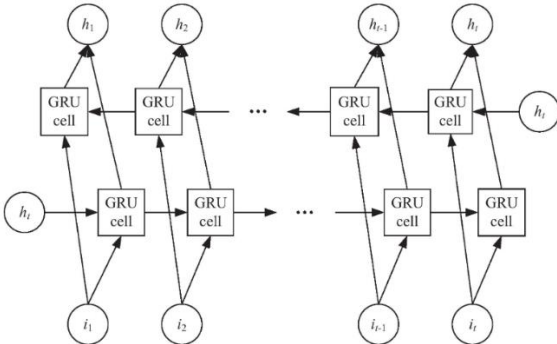


Figure 6: Structure of the Bi-GRU model

2.3. RoBERTa-BiGRU-Attention model

This study's sentiment analysis model comprises five parts: the input layer, the RoBERTa layer, the BiGRU layer, the attention layer, and the Softmax classification layer. In this model, each word in the input text is obtained using the RoBERTa pre-trained model. Then it is sent to the

downstream task BiGRU layer to extract the information between contexts and the attention mechanism to get the weighted comment text information. Finally, the layer's output is classified using the Softmax function.

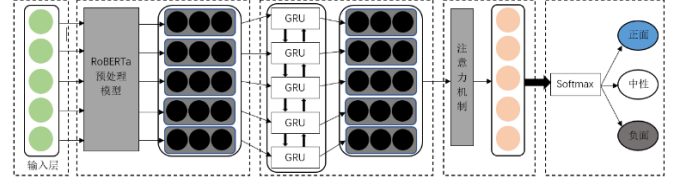


Fig.8 Structure diagram of the RoBERTa-BiGRU-Attention model

3. Experimental results and analysis

Taking the recent hot event of the 2024 Paris Olympics as an example, this study uses the Python Scrappy framework to crawl the structured data on Weibo as a dataset, with a total of 71,614 records and a period from July 27, 2024, to August 11, 2024. The sentiment judgment of the dataset contains three categories: positive, negative, and neutral.

The dataset is divided into an 8:1 training set and a test set.

3.1. Experimental environment and parameters

The experimental environment is shown in Table 1.

Table 1: Experimental environment

Development environment	Configuration & Environment
CPU	i7-13700H
GPU	RTX4060 8GB
Memory	64GB
OS	Windows 11 64-bit
Deep learning framework	Tensorflow3.19.0
Programming tools	Visual Studio Code

The model used in this article, RoBERTa-BiGRU-Attention, has many parameter combinations and settings. Google used in the RoBERTa structure: "RoBERTa-Base", where the Transform block contains 12 layers, the output dimension d is 768, the head parameter of the multi-head attention mechanism is 12, the activation function is RLU, the experimental Adam method optimizes the training model, the initial learning rate is set to 0.00001, the learning rate decay rate is set to 0.000001, and the sliding window w is set to 10. The batch batch-size=16 is set to 0.1 in the input layer and 0.5 in the hidden layer, with a maximum of $l=200$. The node of the bidirectional BiGRU is 128.

3.2. Visualization of experimental results

The Paris Olympics will be broadcast live on CCTV, Douyin, and CCTV APP, and the broadcast period will be from July 27, 2024, to August 11, 2024. As can be seen from Figure 9, the number of microblogs published every day generally decreases first and then increases. It explodes in the two days of July 31 and August 1, which is also the time when hot events such as the upset of the national table tennis king Chuqin occur.

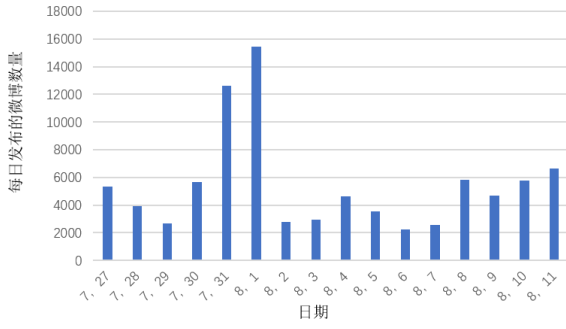


Figure 9: The number of microblogs posted per day

Figures 10, 11, and 12 show the number of daily likes, retweets, and comments on positive, negative, and neutral posts. As you can see from the graph, the number of likes and retweets of positive posts is greater than that of neutral posts per day, while the number of likes and retweets of neutral posts is greater than that of negative posts. The number of positive, negative, and neutral post comments per day is relatively close. As can be seen from the figure, when there is a hot event in the Paris Olympics, the Olympic Games' popularity will increase, and the number of Weibo posts will increase, so the number of likes, retweets, and comments on Weibo will increase.

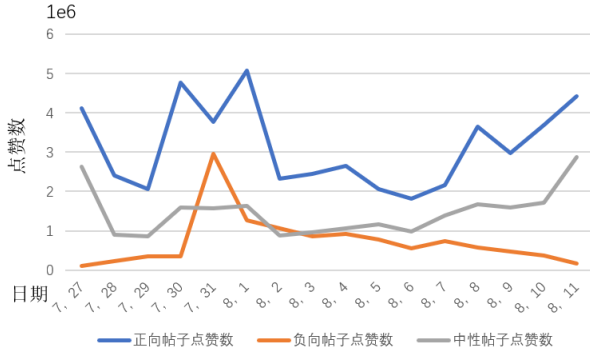


Figure 10: The number of likes on positive, negative, and neutral posts

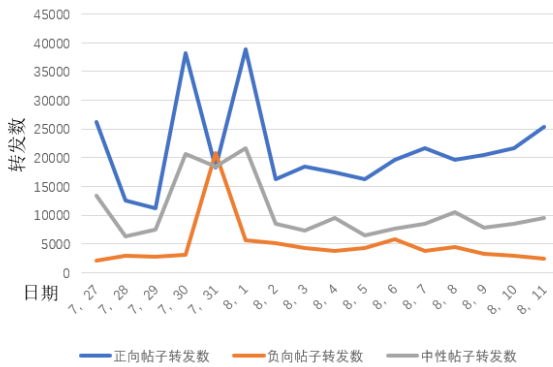


Figure 11: Number of retweets for positive, negative, and neutral posts

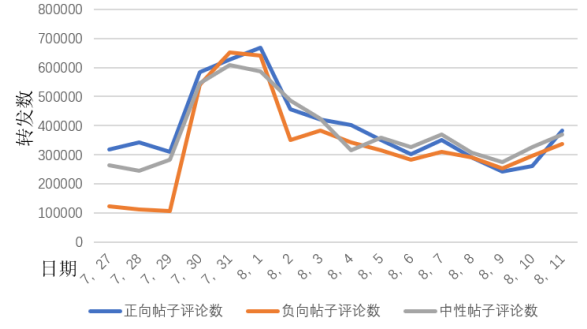


Figure 12: Number of comments on positive, negative, and neutral posts

Finally, the word cloud map is used to visualize the keywords that appear frequently in the text^[3]. In a word cloud diagram, the higher the word frequency, the larger the font size. This is shown in Figure 13. The figure shows that the highest keyword is the Chinese team, indicating that netizens have mentioned this word many times in the discussion, which shows the importance and concern of netizens about Chinese sports. Other keywords with high frequency are concentrated in athletes' names and words such as "congratulations", "victory", "come on", and "great", indicating that netizens respect and love athletes.



Fig.13 Keyword word cloud

3.3. Evaluation indicators

In this experiment, four evaluation indexes were used to evaluate the performance of the model: Accuracy (Acc), Precision (P), Recall (R), and F1-measure (F1).[6]. These standards are widely used in text classification and sentiment analysis tasks. These indicators are calculated as follows:

(1) TP: The number of genuinely positive comments correctly categorized as positive emotion.

(2) FP: The number of genuinely negative or neutral comments misclassified as positive emotions.

(3) TN: The number of comments that are genuinely negative or neutral sentiments and are correctly categorized as negative or neutral sentiments.

(4) FN: The number of genuinely positive comments misclassified as negative or neutral.

(5) Accuracy: The proportion of the number of correctly predicted reviews to the number of all reviews.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

(6) Precision: The ratio of correctly predicted positive reviews to the total number of positive reviews.

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

(7) Recall: The ratio of the number of correctly predicted positive reviews to the actual number of positive reviews.

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

(8) F1: A weighted average of precision and recall.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (11)$$

3.4. Comparative analysis of classification models

Table 2: Iterative experiment results

The number of iterations	Accuracy	Precision	Recall	F1
10	91.61%	93.69%	85.97%	92.73%
20	92.65%	94.79%	87.11%	98.72%
30	91.79%	94.16%	89.72%	95.72%
40	91.91%	91.89%	92.57%	94.75%
50	91.67%	87.99%	92.78%	90.27%
60	90.96%	91.96%	94.78%	93.72%
70	92.72%	94.17%	90.28%	95.28%

Table 3: Experimental results of model comparison

model	Accuracy	Precision	Recall	F1
RoBERTa-TextCNN	72.79%	74.15%	71.95%	78.42%
RoBERTa-LSTM	77.05%	79.25%	77.26%	86.75%
RoBERTa-GRU	83.54%	82.68%	83.71%	89.83%
RoBERTa-BiLSTM	86.52%	87.06%	86.02%	90.56%
RoBERTa-BiGRU	89.63%	88.71%	87.56%	92.57%
RoBERTa-BiGRU-Attention	91.43%	93.78%	90.96%	95.26%

The experimental results show that the RoBERTa-BiGRU-Attention model performs the best in the evaluation index, which indicates that the addition of bidirectional GRU and attention mechanism benefits the affective propensity analysis task. One of the reasons for the analysis is that the word vectors generated by RoBERTa are dynamic, taking into account the contextual semantic information of the text. The second reason is that the attention mechanism can effectively capture the context or even the overall key information sentences, and integrate them into the current word vector, which can assign different weights to different parts of the text, thereby improving the accuracy of sentiment tendency analysis.

4. Conclusion

With the continuous development and improvement of deep learning technology, its application in microblog public opinion sentiment analysis has achieved remarkable results. Through the above discussion, we can see that using deep

learning models to analyze the sentiment of microblog data improves the processing speed and analysis accuracy and provides a new perspective and method for understanding and predicting public sentiment. However, the application of deep learning in microblog sentiment analysis still faces many challenges, such as data quality control, model interpretability, algorithm fairness, and bias issues. In the future, we look forward to overcoming these challenges through more refined model design, smarter algorithm optimization, and stricter ethical standards, further unleashing the potential of deep learning in this field.

Acknowledgments

The authors of this comprehensive document extend their heartfelt gratitude and profound appreciation for the generous funding and invaluable support offered by the esteemed 2024 School-level Innovation and Entrepreneurship Training Program of Wuhan Business University. (No.202411654174). Project of Collaborative Education between Industry and Education, Ministry of Education: Project Number: 230905181154815

References

- [1] Wu Jiahui, Jia Yungang, Wang Zhixiao, et al. Sentiment analysis of Weibo epidemic public opinion text based on deep learning[J].Computer Technology and Development, 024,34(07):175-183.DOI:10.20165/j.cnki.ISSN1673-629X.2024.0087.
- [2] Xi Yuyi,Lv Guo,Bai Mohan. Research on microblog public opinion analysis based on CNN-BiLstm model[J].Journal of Hebei University of Civil Engineering and Architecture, 023,41(03):207-210+220.)
- [3] ZHANG Li,LI Ju. Research on microblog public opinion sentiment analysis based on dictionary and emoji[J].Computer & Telecommunications,2023,(07):40-44.DOI:10. 5966/j. nki. nydx.2023.07.011.
- [4] GUO Junhao. Research and application of text sentiment analysis based on deep learning[D].North University of China,2023.DOI:10.27470/d.cnki.ghbgc.2023.000118.
- [5] ZHANG Li. Research on Chinese implicit sentiment analysis for social network text[D].East China University of Technology, 023.DOI:10.27145/d.cnki.ghddc.2023.000069.
- [6] MA Yu. Research on microblog public opinion sentiment analysis and evolution trend prediction technology under major emergencies[D].Zhengzhou University of Light Industry, 023.DOI:10.27469/d.cnki.gzzqc.2023.000173.
- [7] Li Yunxiang. Research on Chinese named entity recognition based on RoBERTa-WWM[D].Nanning Normal University, 023.DOI:10.27037/d.cnki.ggxsc.2023.000246.
- [8] Wang Shuyan,Yuan Ke. Sentiment analysis model of college student forum based on RoBERTa-WWM[J].Computer Engineering, 022,48(08):292-298+ 05.DOI: 0.19678/j. ssn. 000- 428.0062008.
- [9] LIU Siyu. Research on improvement of Chinese named entity recognition method based on deep learning[D].Chengdu University of Technology, 020.DOI:10.26986/d. nki. cdlc. 020.000601.)