

# A Chinese Stock Reviews Sentiment Analysis Based on BERT Model

**Li Mingzheng** (✉ [mzli.hhu@gmail.com](mailto:mzli.hhu@gmail.com))

Huainan Normal University <https://orcid.org/0000-0002-8259-3924>

**Chen Lei**

Huainan Normal University

**Zhao Jing**

Huainan Normal University

**Li Qiang**

Huainan Normal University

---

## Research

**Keywords:** Chinese Stock Reviews, Sentiment Classification, BERT, Fine-tuning

**Posted Date:** September 8th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-69958/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

---

# A Chinese Stock Reviews Sentiment Analysis Based on BERT Model

Mingzheng Li \*<sup>1</sup> , Lei Chen<sup>1</sup>, Jing Zhao<sup>1</sup> and Qiang Li<sup>1</sup>

1 College of Computer Science, Huainan Normal University, 232038, China

Mingzheng Li is the corresponding author, his email is mzli.hhu@gmail.com

## **Abstract:**

A huge amount of stock reviews occurred on the Internet due to its rapid development, therefore, the stock reviews sentiment analysis has profound significance for the study of the financial market. Due to the lack of a large amount of labeled data, the accuracy of existing sentiment analysis of Chinese stock reviews remains to be further improved. In this paper, a sentiment analysis algorithm for Chinese stock reviews based on BERT is proposed and it improves the accuracy of sentiment classification. The algorithm uses BERT pre-training language model to perform representation of stock reviews on the sentence level, and then input the obtained feature vector into the classifier layer for classification. In the experiments, we show our method has nearly 8% and 9% improvement than TextCNN and TextRNN in F1, respectively. Our model can obtain the best results via fine-tuning which is proved to be effective in Chinese stock review sentiment analysis.

**Keywords:** Chinese Stock Reviews, Sentiment Classification, BERT , Fine-tuning

## **1.Introduction**

With the rapid development of the Internet, people's lives are deeply integrated with the Internet. The stock market has attracted many investors, and the vast majority of investors are using the Internet for stock trading. Posting stock opinions on stock message boards has become a popular way, which results in a large number of stock

---

reviews. The existing study[1] has shown that the sentiment in these messages has been related to stock prices, It can significantly improve the accuracy of stock market predictions. Therefore, this paper attempts to cope with the sentiment analysis of Chinese stock reviews. Sentiment analysis[2] is also called opinion mining and opinion analysis, which aims to find the polarity of text and classify it into positive, neutral or negative. Sentiment analysis areas include news comment analysis, product comment analysis, movie reviews analysis and other fields.

Recently, the mainstream of research on text sentiment analysis is machine learning. Some early works depend on manually extract features, eg. sentiment lexicons[3][4]. Lately, machine learning methods[5] are widely used in the sentiment classification task. The early approaches need to manually build the sentiment dictionary, and the method based on machine learning needs to manually extract features, both of which are time-consuming and labor-intensive. In recent years, assorted methods based on deep learning have become popular for sentiment analysis, as they don't require manual feature engineering. But for many specific fields, especially in the field of stock commentary, we don't have a large number of labeled data sets, which makes it difficult for us to train complex models. In this case, the sentiment analysis for stock reviews has met a bottleneck. The basic idea of this paper is to rely on BERT (Bidirectional Encoder Representations from Transformers)[6]with its pre-training on an enormous dataset and the powerful architecture to gain better results in the sentiment analysis of Chinese stock reviews.

In this paper, we designed different variants of using BERT on the sentiment analysis for Chinese stock reviews. These models used the deep learning method which could automatically extract features. Meanwhile, we used the pre-training model to fine-tune in new application areas which further boost the performance of sentiment analysis classification. Our model could solve the problem of sentiment analysis in a specific field. The experiment was performed on Chinese stock reviews data set in Guba of Eastmoney, and the result showed that the proposed model in this paper had higher accuracy than TextCNN[7]and TextRNN[8] models.

---

In summary, the main contributions of this paper are as follows:

- (1) We propose a new solution for sentiment analysis of Chinese stock reviews which can avoid building dictionaries and extracting features manually.
- (2) We designed different variants of using BERT on the Chinese stock reviews and found that the BERT+Linear model gains the best performs via fine-tuning, Experimental results indicate that the proposed method is highly effective.
- (3) The benefit of this method is that it has good generalization ability and can be widely used in some specific fields. For stock comment sentiment classification tasks, the BERT pre-training model was trained on a general text corpus, and then the model is fine-tuned on a domain-specific corpus, which effectively improves the performance of the model.

## 2.Related Work

The goal of sentiment analysis is to identify sentiment polarity. Now, abundant methods have been used to resolve sentiment analysis, including traditional methods and deep learning methods. In this section, we will briefly introduce the related work of sentiment analysis. Sentiment analysis methods are mainly divided into three types: sentiment dictionary based method, machine learning based method[9][10] and deep learning method.

The method of sentiment dictionary is based on the sentiment dictionary, which calculates the sum of positive and negative sentiment in the text, greater than zero indicates that the text tends to be positive, otherwise it is negative. The advantage of the sentiment dictionary method is simple and does not need a labeled text. However, the disadvantage is that the sentiment dictionary depends on the manual design, which leads to insufficient coverage of sentiment words. The machine learning method depends on the machine learning algorithm and has become the most popular method. This method converts a large number of features into feature vectors based on sentiment lexicons, then train a classifier. The earliest method used is Naive

---

Bayes[11], then some important methods in the development of machine learning, such as support vector machine (SVM) [12][13] and decision tree [14] have been widely used in the field of text sentiment classification. The method of text sentiment analysis based on machine learning is more and more accurate, but it relies on the quality of the feature extracted.

Nowadays, deep learning has been popularly used in the field of text sentiment classification. Convolutional Neural Network(CNN) and Recurrent Neural Network(RNN) have been broadly used to text sentiment classification. Dos et al.[15] proposed a CharSCNN which used tow convolution layers to extract features to solve sentiment analysis. Wang et al.[16] used LSTM to predict the sentiment polarities of tweets by composing word embeddings.

However, the RNN network can't be calculated in parallel, and also produce gradient explosion. Vaswani et al.[17] proposed a Transformer to solve the problem of parallel computing and also achieved the best results in many natural language processing tasks including sentiment analysis. Devlin et al. [6] proposed a pre-training model called BERT, which achieved the best results in many NLP tasks.

Due to the great progress of the pre-training model, many researchers utilize BERT to figure out the down-stream tasks. Dong et al.[18] proposed BERT-CNN model to improve the accuracy of commodity sentiment analysis. Yu et al. [19] used BERT to achieve state-of-art results for ancient Chinese sentence segmentation. Literature [20] assorted BERT to classify the Chinese short texts. Liu et al. [21] proposed BERTSUM based on BERT for extractive summarization.

### 3. Methods

This paper aims to settle the sentiment classification of Chinese stock reviews based on BERT. The model is composed of two parts: pre-training and fine-tuning layer. The structure of the model is shown in Figure 1:

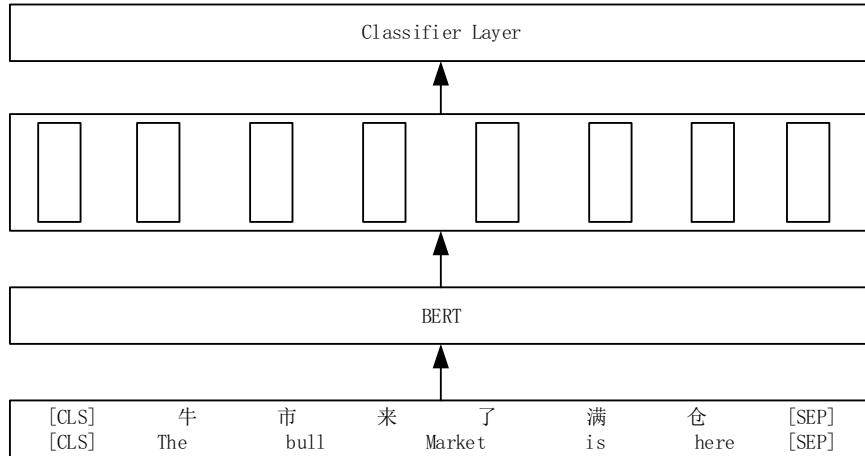


Figure 1 The structure of a sentiment analysis model

As illustrated in Figure 1, the BERT model will automatically add the symbols [CLS] and [SEP] at the beginning and the end of the sentence to indicate the beginning and the end of each sentence. Taking a Chinese stock review text as an example, the [CLS] and [SEP] symbols are added to the left and right positions of the text. In this paper, we want to get the vector of a sentence, the BERT model will output the leftmost [CLS ]vector.

After we obtain the sentence vector from the BERT, we add a classification layer stacked on the top of the BERT. The output of the BERT is the input of the classification layer, thereby capturing sentence-level features to perform sentiment classification on Chinese stock review text. Sun et al. [22] concluded that the 12th layer has the best classification capability. Therefore, this paper added several layers on the 12th layer, for example, BERT +Linear, BERT + LSTM and BERT + CNN models. These classification layers are jointly fine-tuned with BERT.

- BERT+Linear

The basic method is to add a linear classification output layer behind the BERT model. BERT output a sentence vector, which is classified through a fully connected layer. In this paper, we called this model as BERT+ Linear.

- BERT+LSTM

LSTM [23]is a special RNN network, which is good at handling sequence issues. This paper used LSTM as a classifier to access the BERT model, and the input of LSTM is the BERT output. Every BERT output is recorded and calculated as:

---


$$\begin{pmatrix} F_i \\ I_i \\ O_i \\ G_i \end{pmatrix} = \text{LN}_h(W_h h_{i-1}) + \text{LN}_x(W_x T_i) \quad (1)$$

$$C_i = \sigma(F_i) e^{-1} + \sigma(I_i) e^{-\tanh(G_i - 1)} \quad (2)$$

$$h_i = \sigma(O_i) e^{-\tanh(LN_c(C_i))} \quad (3)$$

Among them,  $F_i$ ,  $I_i$ ,  $O_i$  are the forget gate, the input gate and the output gate, respectively,  $G_i$  are the hidden layer,  $C_i$  is the memory vector, and  $h_i$  is the output vector,  $\text{LN}_h$ ,  $\text{LN}_x$ ,  $\text{LN}_c$  realize the standardization of different layers. The final output is still added to a linear layer.

#### ● BERT+CNN

A convolutional neural network has been successfully applied in computer vision, and its basic structure includes an input layer, convolutional layer, pooling layer, fully connected layer and output layer. The output of the BERT model is the input of the convolutional neural network. After the convolutional layer and the pooling layer, it is connected to the fully connected layer to finally output the sentiment classification of the Chinese stock review text. Its generation process could be described as:

$$G_i = f(G_{i-1} \otimes W_i + b_i) \quad (4)$$

$$G_i = \text{subsampling}(G_{i-1}) \quad (5)$$

$W_i$  represents the weight vector of the  $i$ -th layer convolution kernel,  $b_i$  is the offset. The operation symbol  $\otimes$  indicates the convolution kernel and the  $i-1$ -th layer convolution calculation, and sampling.

## 4. Experimental

### 4.1 Dataset

In this paper, we used the pre-trained BERT model released by Google and trained

---

on the Chinese Wikipedia corpus. There are few datasets for Chinese stock reviews with a label. The dataset used in the experiments in this paper is from the Github [24]website. There is a total of 9204 review data in this dataset, with half of the positive and negative samples. This dataset uses logistic regression for sentiment classification, and its accuracy rate reaches 88.09%, which shows that the quality of this data set is very high and can be used to train other models. This paper divided the training set, verification set and the test set according to the ratio of 6:2:2. Table 1 was some examples of stock reviews.

Table 1 Examples of data set

Review	Polarity
越发觉得现在入股市好，好多股 票都想买可惜没钱。	
I feel that it is a better time to enter the stock market now. I want to buy many stocks, unfortunately, I don't have much money.	positive
大盘进入主升浪加速运行，请系 好安全带。	
The market enters the main wave and accelerates, please fasten your seat belt.	positive
今天下跌还带量。 It fell sharply today.	
收市前仓底应该呵呵继续跌。	negative
The bottom of the warehouse should continue to fall before the market closes.	negative

## 4.2 Evaluation Indicators

This paper is designed to realize the sentiment classification of Chinese stock review text, which is a typical classification problem. The most commonly used evaluation indicator for classification problems includes accuracy rate, recall rate and F1 value. According to the changes in the two dimensions of model prediction and actual situation, the binary classification problem can be divided into four cases: See Table 2 below:

Table 2 confusion matrix

Actual prediction	Predict is positive	Predict is negative
	TP	FN
True is a positive example	TP	FN
True is a counterexample	FP	TN

The accuracy rate refers to the proportion of the samples predicted by the classifier as positive examples and the correct prediction accounts for all the samples predict positive. Through the confusion matrix, the formula can be easily obtained:

$$P = \frac{TP}{TP+FP} \quad (6)$$

The recall R refers to the proportion of the positive examples of the classifier's prediction category and the correct prediction of the samples to all positive examples.

The formula is as follows:

$$R = \frac{TP}{TP+FN} \quad (7)$$

The F1 is an indicator that combines the precision rate (P) and the recall rate (R), and its calculation formula is:

$$F_{\beta} = \frac{(\beta^2+1)*P*R}{\beta^2*P+R} \quad (8)$$

It can be seen from the formula that it is the reconciled average value of the

---

recall rate (R) and the accuracy rate (P). If the value is 1, the formula will become to:

$$F_1 = \frac{2P*R}{P+R} \quad (9)$$

The F1 score is one of the significant indicators to evaluate the performance of the classifier. The value is close to 0, indicating that the model has a bad performance. To verify the classification ability of the BERT model in Chinese stock reviews more effectively, this paper selected two indicators, which are accuracy and F1.

#### 4.3 Experimental results on different variants of models

This paper proposed different variants of the model in section 2.1.2. The BERT used in this paper is the " bert-base-uncased " version provided by Google. The network structure has a total of 12 layers, and the hidden layer has 768 dimensions. It uses a 12-head mode and has a total of 110M parameters. The dataset is small, and the BERT model requires a huge amount of data. For this reason, we used the Chinese pre-training model trained by Google[25].

To obtain a more effective classification model, this section uses experimental data based on the previous section and implements the above algorithms with Pytorch[26]. F1 value was selected as the indicator, and the experimental results are shown in Table 2:

Table 2 F1 scores of various BERT models

BERT Fine-tuning model	F1
BERT + Linear	0.9250
BERT + LSTM	0.9197
BERT + CNN	0.9197

As seen from the experimental results, the F1 scores of different models have only a very small gap, which shows that complex models have not improved the effect of sentiment classification. Therefore, this paper chose the model of

---

BERT+Linear for Chinese stock review sentiment analysis.

#### 4.4. Experimental results on the hyperparameters setting of the model

The hyperparameters in the proposed analysis sentiment for Chinese stock reviews include epochs, pad size, learning rate and so on. This section will carry out related experiments to find the best values of the parameters.

##### (1) Epoch

The hyperparameters of the model have an important impact on the performance of the model. The hyperparameters in this paper included the number of epochs, learning rate, pad size, and so on. Next, we executed several experiments to select the best hyperparameters. The data set had been prepared, and it is divided into a training set and a test set according to the ratio of 6:4.

Epoch is the number of training times, specific to the training set. Generally, with the increase of Epoch, the performance of the model is enhanced. However, if the Epoch is too large, overfitting problems may occur. Therefore, it is important to choose the correct era. Figure 2 is the relationship between epoch and F1.

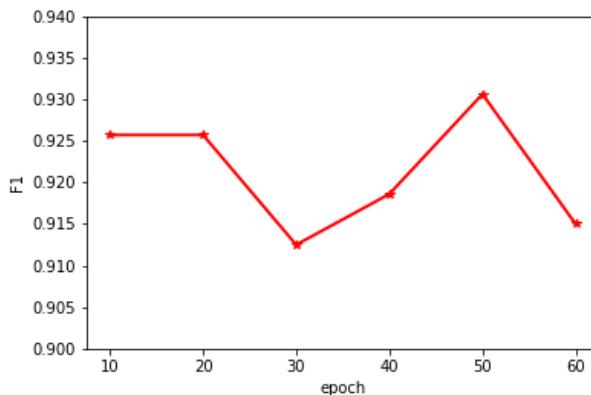


Figure 2 Relationship between Epochs and F1.

As can be seen from Figure 2, as the Epoch increases, the F1 score also increases, which means that the classification performance of the model becomes better. When Epoch is 50, the performance of the model is the best. When the epoch value is 60, the model stops training early, and the accuracy is also decreased.

##### (2) Pad size

---

The pad size is the length that the model will process each sentence. Its basic processing principle is short filling and long cutting. If a sentence is shorter than the set length, then it is filled with 0. If the length of the sentence is greater than the set value, Then the sentence shall be truncated to get the set value.

Pad size has an enormous influence on the performance of the model, hence the value should choose a suitable one, If the Pad size is too large, the zeros in the data will be filled too much. On the contrary, plenty of information will be lost. The distribution of the review length is shown in Figure 3. The stock review length is very short, and most of them are less 50.

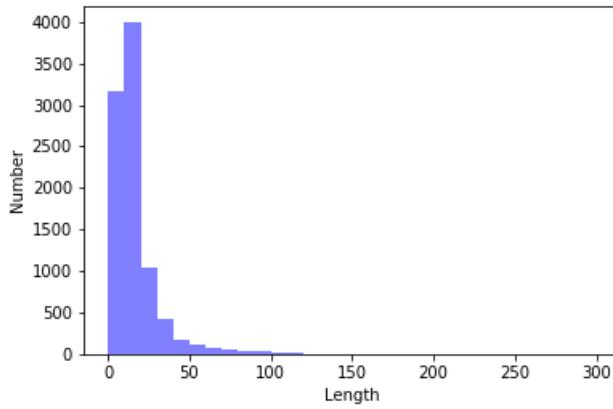


Figure 3 Text length distribution in the dataset.

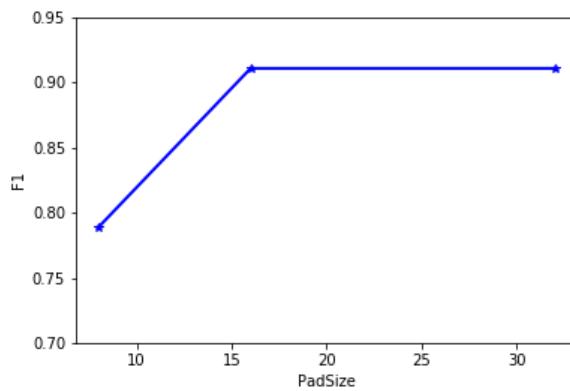


Figure 4 Relationship between PadSize and F1.

We set up different pad size to experiment, and Figure 4 is our result. When pad size increases, the F1 score will show an upward trend. When the pad size is 16, the F1 score will reach the maximum. But when pad size continues to increase, the

---

machine will have a memory leak.

It can be concluded that when pad size is small, too much information will be lost. In this experiment, the pad size is 16 which is the best.

#### 4.5 Experimental results on sentiment analysis for different methods

According to the experimental results in the previous section, this paper chose BERT + Linear model to solve the Chinese stock review sentiment classification. To verify the effectiveness of the BERT model proposed in this paper, two experiments were selected as the control group. The control group experiments used the TextCNN and the TextRNN, respectively.

Kim et al was proposed TextCNN in 2014. Since this was proposed, it's the baseline model in the field of text classification tasks because of its efficiency. TextRNN was based on a recurrent neural network and used a multi-task learning framework to train jointly. Both models achieved favorable results in text classification.

TextCNN model is one of the most commonly used methods for text classification. After many trials and comparisons, we chose the best hyperparameters as shown in Table 3:

Table 3 TextCNN model training parameters

TextCNN parameters	values
Word vector dimension	128
Number of CNN convolution kernels	100
CNN convolution kernel size	3
Dropout random inactivation rate	0.5
Model iteration round	30
Learning rate	0.001

The experiment comparison of TextRNN used in this article, of which the

---

training parameter values are shown in table 4:

Table 4 TextRNN model training parameters

TextRNN parameters values	TextRNN parameter
dimension of word vector	128
Number of neurons	64
Number of layers of lstm stack	1
Number of neurons	64
Model iteration round	30
Learning rate	0.001

To verify the BERT+Linear model can effectively perform sentiment analysis of Chinese stock review text, The experimental dataset used the Chinese stock review, indicators used accuracy and F1 value. The comparison experiment results are shown in Figure 5 and 6:

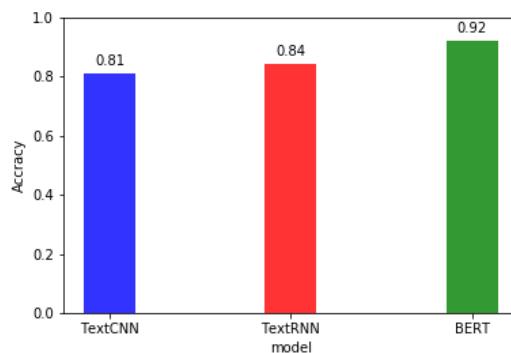


Figure 5 Accuracy experiment results

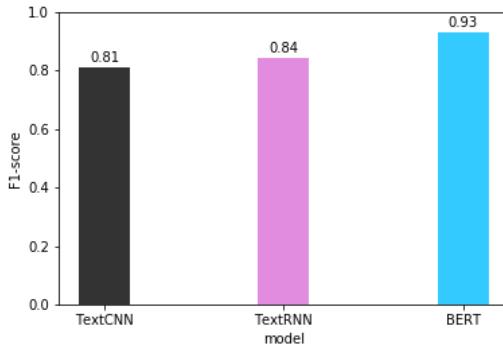


Figure 6 F1 value experiment result

Figures 5 and 6 are the results of this experiment, Figure 5 is the accuracy of the three experimental results, and Figure 6 is the F1 scores. From Figure 5, it can be seen that the BERT+Linear model has the highest accuracy of sentiment classification for review text, which has a 12% improvement over TextCNN, and 9% for TextRNN, respectively. Figure 8 shows that the BERT+Linear model also has a nearly 9% improvement compared with TextRNN in F1 scores, which demonstrate the BERT+Linear model had a better classification effect.

#### 4.6. Discussion

We tested and evaluated our proposed BERT+Linear method to address the sentiment analysis of Chinese stock review which proven the advantages of our method. The experimental results indicate that the BERT+ Linear model has achieved fine classification results in the sentiment analysis of Chinese stock review texts, which has more accurate than TextCNN and TextRNN. It shows that our model can be used as the sentiment classification algorithm for Chinese stock review text. More recently, the BERT fine-tuning model can be used as a general method for Chinese sentiment classification.

### 5. Conclusion

With the advancement of Internet technology, investors can easily post relevant tweets. Mining these data is valuable research, so we propose the BERT+Linear model for sentiment analysis of Chinese stock reviews which can avoid building dictionaries and extracting features manually. Experimental results indicate that the

---

proposed method is highly effective.

The main contribution of this paper was that we proposed a BERT+Linear model for Chinese stock review sentiment classification. The advantage of this method is that it has good generalization ability and can be widely used in some specific fields.

Although this paper innovatively applied the BERT model to the field of sentiment classification in stock reviews and eliminated the dependence on the emotion dictionary in the professional field, overcome the difficulty of manually extracting features, and improved the accuracy of sentiment classification. However, there were still some defects in this paper. The dataset was small, and the conclusion had not been verified on the larger dataset. In subsequent research, make further improvements to BERT model, and expand the sample size to carry out further experiments.

## REFERENCE

- [1] H.-J. Sheu, Y.-C. Lu, and Y.-C. Wei, “Causalities between sentiment indicators and stock market returns under different market scenarios,” *Int. J. Bus. Fin. Res.*, vol. 4, no. 1, pp. 159 – 171, 2010.
- [2] S. V. Wawre and S. N. Deshmukh, “Sentiment classification using machine learning techniques,” *International Journal of Science and Research (IJSR)*, vol. 5, no. 4, pp. 819 – 821, 2016.
- [3] R. Socher, J. Pennington, E. H. Huang, A. Y. Ng, and C. D. Manning, “Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions,” in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, Edinburgh, Scotland, UK., 2011, pp. 151 – 161, [Online]. Available: <https://www.aclweb.org/anthology/D11-1014>.
- [4] S. Feng, Y. Fu, F. Yang, D. Wang, and D. Zhang, “Blog sentiment orientation analysis based on dependency parsing,” *Journal of Computer Research and Development*, vol. 49, no. 11, pp. 2395 – 2406, 2012.
- [5] Pang B , Lee L , Vaithyanathan S . Thumbs up Sentiment Classification using Machine Learning Techniques[J]. *Empirical Methods in Natural Language Processing*, 2002:79–86.
- [6] DEVLIN J, CHANG M-W, LEE K, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[J/OL]. CoRR, 2018, abs/1810.04805.
- [7] Kim Y . Convolutional Neural Networks for Sentence Classification[J]. Eprint Arxiv, 2014.
- [8] P. Liu, X. Qiu, and X. Huang, “Recurrent Neural Network for Text Classification with Multi-Task Learning,” *CoRR*, vol. abs/1605.05101, 2016, [Online]. Available:

---

<http://arxiv.org/abs/1605.05101>.

- [9] UM, LIU B. Mining and summarizing ccustomer reviews[C]. Proceedings of the Tenth ACM SIGKDD International Conference on Konwledge Discovery and Data Mining, 2004: 168–177.
- [10] JOTHEESWARAN J, KOTEESWARAN S. DECISION TREE BASED FEATURE SELECTION AND MULTILAYER PERCEPTRON FOR SENTIMENT ANALYSIS[J]. 2015, 10(14) : 12
- [11] Pranckevičius, Tomas, and Virginijus Marcinkevičius. "Comparison of naive bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviewsclassification." Baltic Journal of Modern Computing 5.2 (2017): 221.
- [12] Zhang M, Ai X, Hu Y. Chinese Text Classification System on Regulatory Information Based on SVM[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2019, 252(2) : 022133.
- [13] Su et al., "Chinese comments sentiment classification based on word2vec and SVMperf," Expert Systems with Application, 2015.
- [14] JOTHEESWARAN J, KOTEESWARAN S. DECISION TREE BASED FEATURE SELECTION AND MULTILAYER PERCEPTRON FOR SENTIMENT ANALYSIS[J]. 2015, 10(14) : 12..
- [15] Dos Santos and M. Gatti, "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts," in Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, Dublin, Ireland, 2014, pp. 69 – 78, [Online]. Available: <https://www.aclweb.org/anthology/C14-1008>
- [16] X. Wang, Y. Liu, C. Sun, B. Wang, and X. Wang, "Predicting Polarities of Tweets by Composing Word Embeddings with Long Short-Term Memory," in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Beijing, China, 2015, pp. 1343 – 1353, doi: 10.3115/v1/P15-1130.
- [17] VASWANI A, SHAZER N, PARMAR N, . Attention is All you Need[M/OL]. GUYON I, LUXBURG U V, BENGIO S, , //Advances in Neural Information Processing Systems 30. Curran Associates, Inc. 2017: 5998 - 6008.
- [18] Dong, F. He, Y. Guo, and H. Zhang, "A Commodity Review Sentiment Analysis Based on BERT-CNN Model," in 2020 5th International Conference on Computer and Communication Systems (ICCCS), May 2020, pp. 143 – 147, doi: 10.1109/ICCCS49078.2020.9118434.
- [19] YU Jingsong, WEI Yi, ZHANG Yongwei, Automatic Ancient Chinese Texts Segmentation Based on BERT [J]. Journal of Chinese Information Processing. 2019, 33(11):57–63.
- [20] Dandan Duan Jiashan Tang Yong Wen Kehai Yuan; BERT Based Research on Classification of Short Chinese Text; Computer Engineering;<https://doi.org/10.19678/j.issn.1000-3428.0056222>
- [21] LIU Y. Fine-tune BERT for Extractive Summarization[J/OL]. CoRR, 2019, abs/1903.10318. <http://arxiv.org/abs/1903.10318>.
- [22] SUN C, QIU X, XU Y, . How to Fine-Tune BERT for Text Classification?[J/OL]. CoRR, 2019, abs/1905.05583. <http://arxiv.org/abs/1905.05583>
- [23] Hochreiter S, Schmidhuber J. Long short-term memory. [J]. Neural computation, 1997, 9(8).
- [24] [https://github.com/algosenses/Stock\\_Market\\_Sentiment\\_Analysis/tree/master/data](https://github.com/algosenses/Stock_Market_Sentiment_Analysis/tree/master/data)
- [25] <https://github.com/ymcui/Chinese-BERT-wwm>
- [26] <https://pytorch.org/>

---

## **Abbreviations**

BERT: Bidirectional Encoder Representations from Transformers; SVM: support vector machine; CNN: Convolutional Neural Networks; RNN: Recurrent Neural Network; LSTM: Long Short-Term Memory;

## **Availability of data and materials**

Experiment data URL is  
[https://github.com/algosenses/Stock\\_Market\\_Sentiment\\_Analysis/tree/master/data](https://github.com/algosenses/Stock_Market_Sentiment_Analysis/tree/master/data)

## **Competing interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## **Funding**

Industry-University Cooperation Cooperative Education Program of the Department of Higher Education, Ministry of Education of China  
NO. 201901148034

## **Authors' contributions**

Mingzheng Li: Conceptualization, Methodology, Investigation Writing - review & editing, Supervision. Lei Chen: Resources, review & editing, Supervision. Jing Zhao: Data curation, Writing – review & editing. Qiang Li: Software, Visualization, Writing - review & editing.

---

## Acknowledgements

This work is supported by the industry-University Cooperation Cooperative Education Program of the Department of Higher Education, Ministry of Education of China with Grant No. 201901148034.

## Figures:

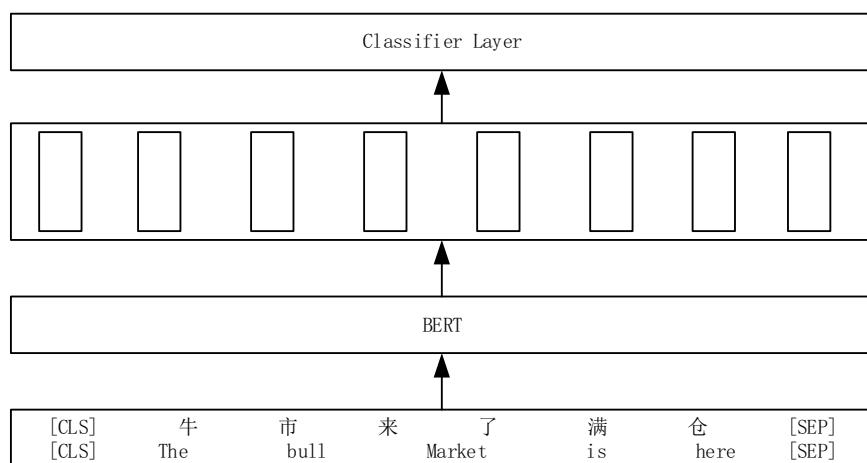


Figure 1 The structure of a sentiment analysis model

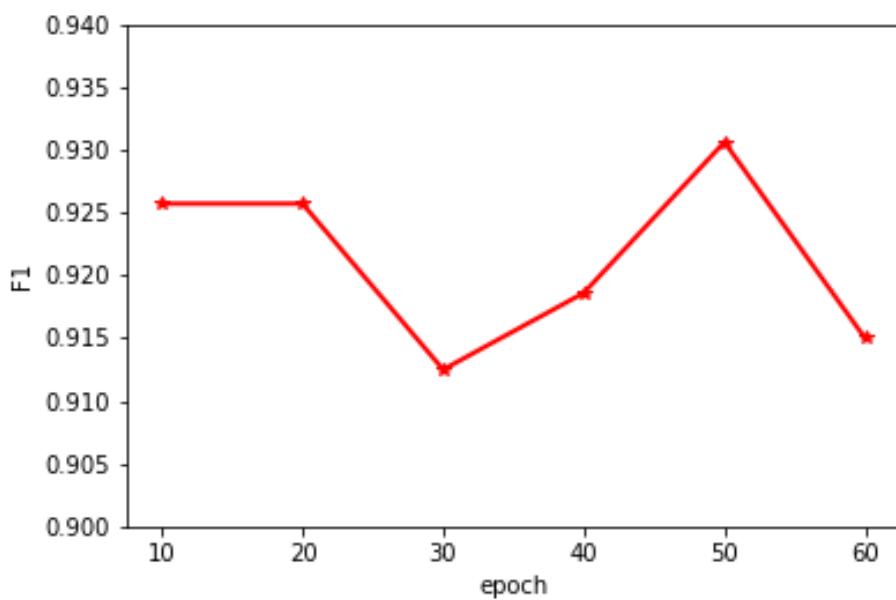


Figure 2 Relationship between Epochs and F1.

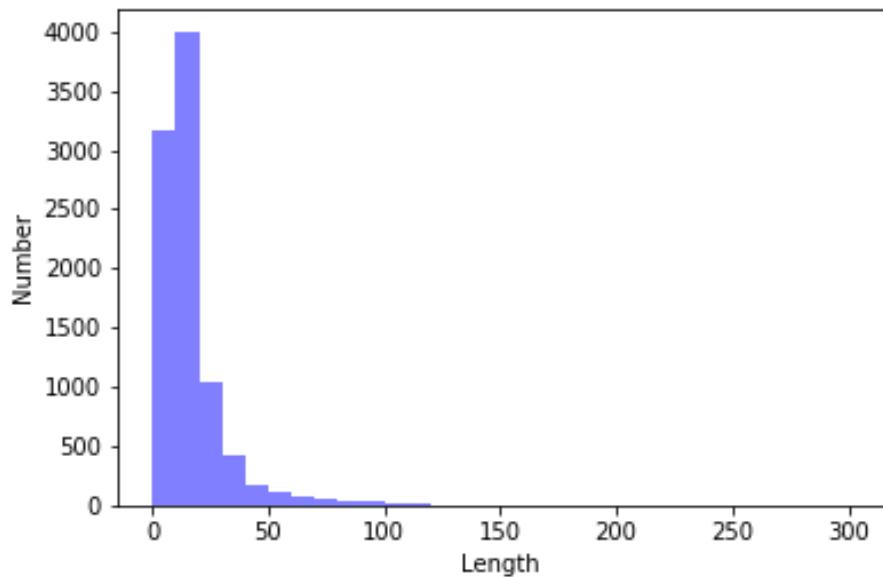


Figure 3 Text length distribution in the dataset.

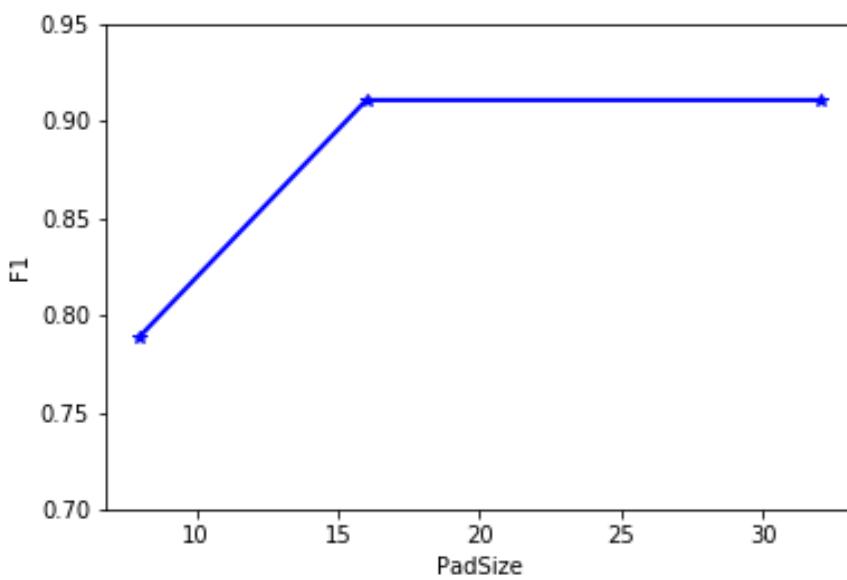


Figure 4 Relationship between PadSize and F1.

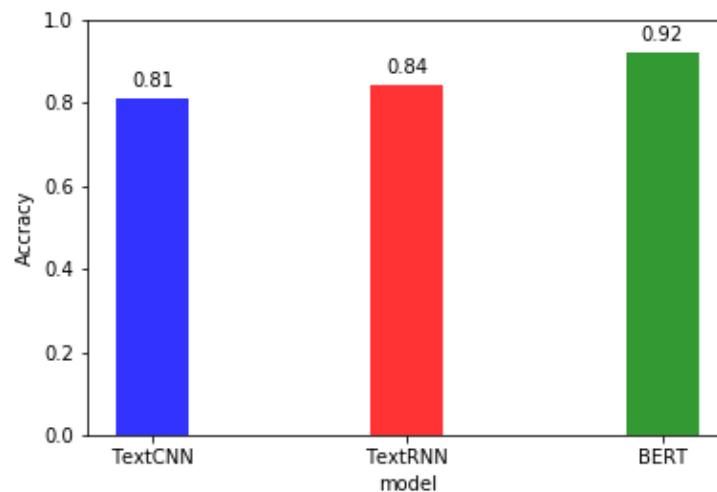


Figure 5 Accuracy experiment results

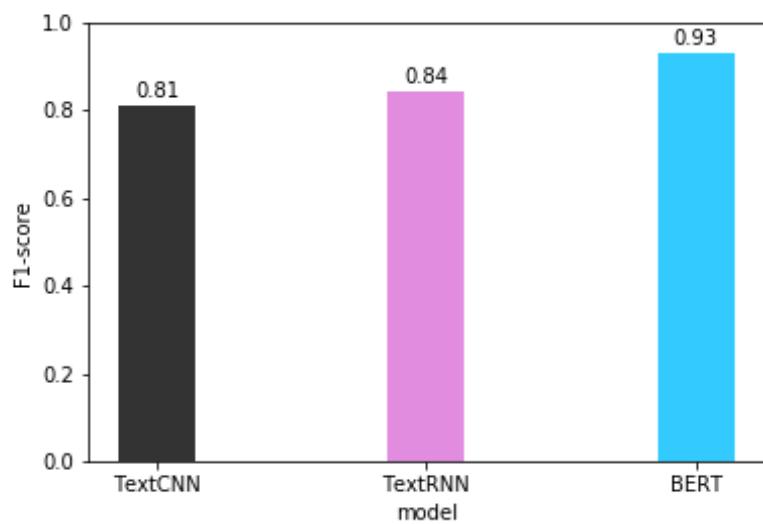


Figure 6 F1 value experiment result

# Figures

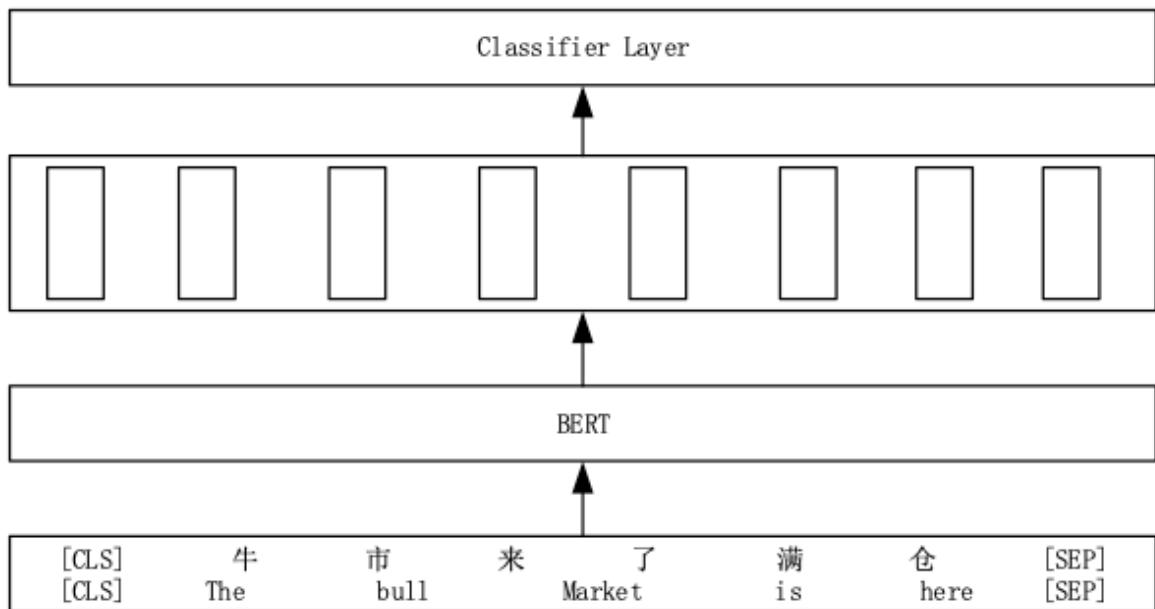


Figure 1

The structure of a sentiment analysis model

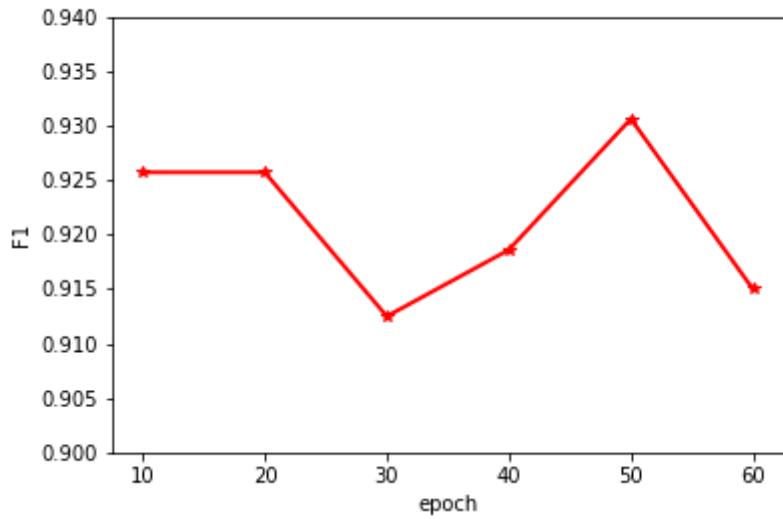
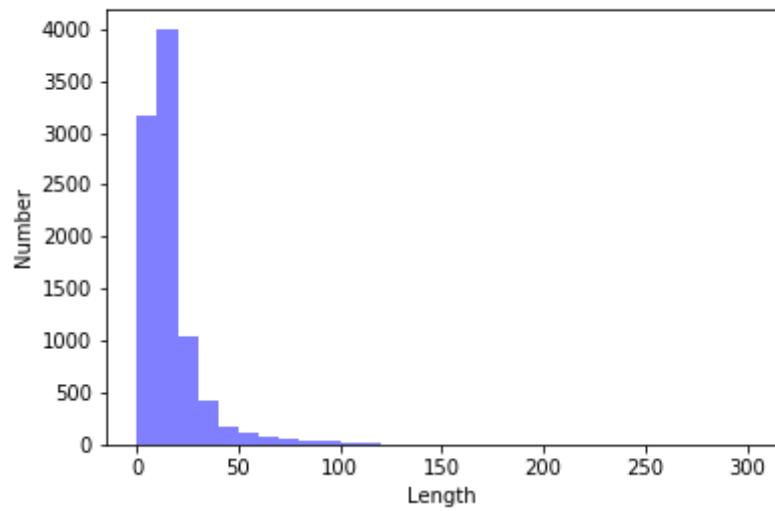


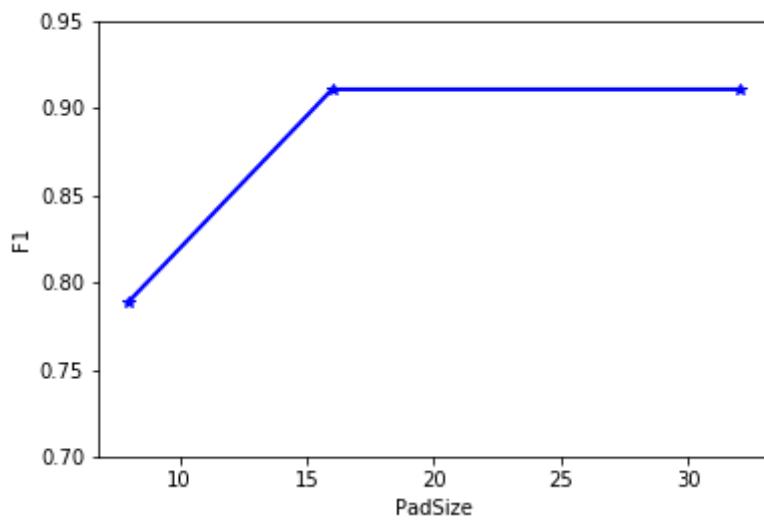
Figure 2

Relationship between Epochs and F1.



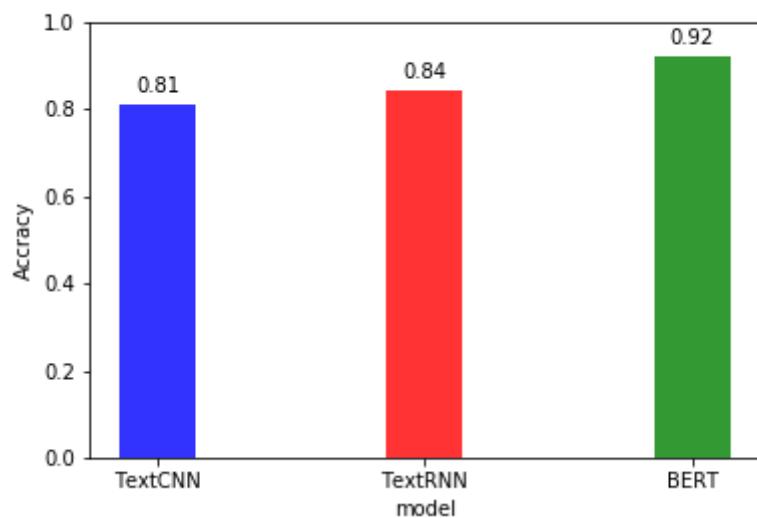
**Figure 3**

Text length distribution in the dataset.



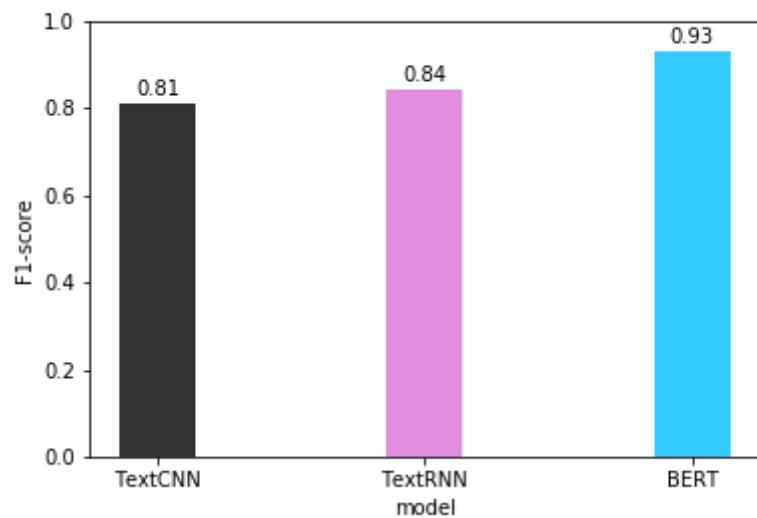
**Figure 4**

Relationship between PadSize and F1.



**Figure 5**

Accuracy experiment results



**Figure 6**

F1 value experiment result