Sentiment analysis of Chinese stock reviews based on BERT model

Mingzheng Li[1] · Lei Chen[1] · Jing Zhao[1] · Qiang Li[1]

Abstract

A large number of stock reviews are available on the Internet. Sentiment analysis of stock reviews has strong significance in research on the financial market. Due to the lack of a large amount of labeled data, it is difficult to improve the accuracy of Chinese stock sentiment classification using traditional methods. To address this challenge, in this paper, a novel sentiment analysis model for Chinese stock reviews based on BERT is proposed. This model relies on a pre-trained model to improve the accuracy of classification. The model use a BERT pre-training language model to perform representation of stock reviews on the sentence level, and subsequently feed the obtained feature vector into the classifier layer for classification. In the experiments, we demonstrate that our method has higher precision, recall, and F1 than TextCNN, TextRNN, Att-BLSTM and TextCRNN. Our model can obtain the best results which are indicated to be effective in Chinese stock review sentiment analysis. Meanwhile, Our model has powerful generalization capacity and can perform sentiment analysis in many fields.

Keywords Chinese stock reviews · Sentiment analysis · BERT · Fine-tuning

Abbreviations

BERT	Bidirectional Encoder Representations from Transformers
SVM	Support Vector Machine
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory

1 Introduction

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 method, that results in a large number of stock reviews. An existing study [1] showed that the sentiment in these messages has been related to stock prices, This 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, mainstream research on text sentiment analysis has involved machine learning. Some early works depended on manually extracting features, e.g. sentiment lexicons [3]. Lately, machine learning methods [4] have been widely used in the sentiment classification task. The early approaches needed to manually build the sentiment dictionary, and the method based on machine learning needed to manually extract features. Both approaches are time-consuming and labor-intensive. In recent years, assorted methods based on deep learning have become popular for sentiment analysis, as they do not require manual feature engineering. However, for many specific fields, especially in the field of stock reviews, the labeled data sets are usually small, which makes it difficult for us to train complex models to improve the accuracy of classification. In this case, the sentiment analysis for stock reviews has hit a bottleneck. Therefore, the accuracy of existing sentiment analysis of Chinese stock reviews can be improved.

The motivation of this paper is to rely on Bidirectional Encoder Representations from Transformers(BERT) [5] with pre-training an enormous dataset and a powerful architecture to gain better results in the sentiment analysis of Chinese stock reviews. In the case of a small amount of data, we can still train a model that works well to generate application value.

✉ Mingzheng Li
   mzli.hhu@gmail.com

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

The important significance is that our model can solve the problem of sentiment analysis in a specific field under the same case of small labeled data.

In this paper, we propose a novel sentiment analysis model for Chinese stock reviews. we design different variants of using BERT on sentiment analysis for Chinese stock reviews. These models use a deep learning method that can automatically extract features. Meanwhile, we use a pre-training model to fine-tune new areas. This further boosts the performance of sentiment analysis classification. An experiment is performed on Chinese stock reviews data set in Guba of Eastmoney, and the result shows that the proposed model in this paper has higher accuracy than TextCNN [6], TextRNN [7], Att-BLSTM [8] and TextCRNN [9]. Our model can improve the accuracy of sentiment analysis classification for Chinese stock reviews. It concludes that our model has a powerful generalization ability that could be used in other fields.

The main contributions of this paper are as follows:

(1) We propose a new solution for sentiment analysis of Chinese stock reviews that can avoid building dictionaries and extracting features manually.

(2) We design different variants using BERT on Chinese stock reviews and found that the BERT+FC 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 a powerful generalization capacity and can be widely used in specific fields. For stock comment sentiment classification tasks, the BERT pre-training model was trained on a general text corpus, and then the model was fine-tuned on a domain-specific corpus, which effectively improves the performance of the model.

The remainder of this paper is organized as follows. Section 2 provides related works. Section 3 provides details of the sentiment analysis method proposed in this paper. Section 4 presents the experimental results and discusses the results. Section 5 summarizes the proposed method and outline future work.

2 Related work

The goal of sentiment analysis is to identify sentiment polarity. Abundant methods have been utilized 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 [10]: sentiment dictionary [11], machine learning [12] and deep learning [13].

The method for the sentiment dictionary is based on the sentiment dictionary, which calculates the sum of positive and

negative sentiment in the text, A value greater than zero indicates that the text tends to be positive; otherwise, it is negative. Many techniques are used to improve the accuracy of sentiment classification, such as word sense disambiguation [14] and constructing sentiment lexicons. 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 a 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 [15]. Then some important methods in the development of machine learning, such as support vector machines (SVMs) [16] and decision trees [17] were been widely used in the field of text sentiment classification. The method of text sentiment analysis based on machine learning is more accurate, but it relies on the quality of the feature extracted.

Currently, deep learning has been popularly used in the field of text sentiment classification. The Convolutional Neural Network(CNN) and Recurrent Neural Network(RNN) have been broadly used for text sentiment classification [18]. Dos et al. [19] proposed a CharSCNN that used two convolution layers to extract features to solve sentiment analysis. Wang et al.[20] used LSTM to predict the sentiment polarities of tweets by composing word embeddings. Wu Xing et al. [21] represented the intrinsic characteristics of the stock market by GRU, which is another variant of RNN.

However, the RNN network cannot be calculated in parallel, and produces a gradient explosion. Vaswani et al. [22] proposed a transformer to solve the problem of parallel computing and achieved the best results in many natural language processing tasks including sentiment analysis. Devlin et al. [5] 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 down-stream tasks. Xu et al. [23] proposed a novel post-training approach on BERT and achieve new state-of-the-art results for aspect extraction and sentiment classification. Rosario et al. [24] used a multilingual method that has BERT perform a Named Entity Recognition task for de-identification. Yu et al. [25] used BERT to achieve state-of-the-art results for ancient Chinese sentence segmentation. Reference [26] uesd BERT to classify Chinese short texts.

Our model utilizes a deep learning method to automatically extract features from Chinese stock review texts, avoiding the time-consuming problem in dictionary construction and feature engineering in the traditional sentiment analysis model of Chinese stock reviews. This paper is inspired by the pre-training model, especially the success of the BERT model, aimed to improve the accuracy of sentiment classification for

Chinese stock reviews by BERT. At the same time, our model has strong generalization ability, which can improve the accuracy of sentiment analysis classification in different fields.

3 Methods

To improve the accuracy of Chinese stock reviews sentiment analysis, we designed a novel sentiment analysis method for Chinese stock reviews based on BERT. BERT extracts local and global features of Chinese stock reviews text vectors. A classifier layer is designed to learn high-level abstract features and to transform the final sentence representation into the appropriate feature to predict sentiment.

The proposed model is composed of two parts: BERT and the classifier layer. The architecture of this model is shown in Fig. 1.

In the input representation layer, 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, which is a vectorized representation of the entire sentence.

After obtaining the sentence vector from the BERT, we add a classification layer stacked on top of the BERT. The output of 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. [27] concluded that the 12th layer has the best classification capability. Therefore, this paper added several layers to the twelfth layer, for example, BERT+FC, BERT+LSTM and BERT+CNN models. These classification layers are jointly fine-tuned with BERT.

• BERT+FC

The basic method is to add a linear layer to the BERT outputs and use a fully connected layer to predict the sentiment

polarity. In this paper, we call this model BERT+ FC. A fully-connected layer consists of a multi-layer perceptron, and its output is calculated as shown in the formula:

$$h = relu\left(W \cdot h^f + b\right) \tag{1}$$

where W and b are parameters that are learned in the training, $W \in R^{c \times n}, b \in R^{c \times 1} h \in R^{c \times 1}.c$ is the number of the sentiment classes.

• BERT+LSTM

RNN are good at modeling sentence sequence information and capturing long-distance dependency. Although the transformer has achieved better results in some tasks, the combination of BERT and LSTM has some advantages. LSTM [28] is a special RNN network. This paper uses LSTM as a classifier and the input to the LSTM is the BERT output. The output is calculated as:

$$\begin{pmatrix} F_i \\ I_i \\ O_i \\ G_i \end{pmatrix} = \mathrm{LN_h(W_h h_{i-1})} + \mathrm{LN_x(W_x T_i)} \tag{2}$$

$$C_i = \sigma(F_i) \odot C_{i-1} + \sigma(I_i) \odot \tanh(G_i - 1) \tag{3}$$

$$h_i = \sigma(O_t) \odot \tanh(LN_c(C_t)) \tag{4}$$

where F_i, I_i and O_i are the forget gate, input gate and output gate, respectively. G_i is the hidden layer, C_i is the memory vector, and h_i is the output vector. LN_h, LN_x and LN_c realize the standardization of different layers. The final output layer is also a fully connected layer:

$$h = relu\left(W \cdot h^f + b\right) \tag{5}$$

• BERT+CNN

The convolutional neural network (CNN)is an important method and consists of a convolutional layer, pooling layer and fully connected layer. The CNN can capture key local

Fig. 1 Overview architecture of sentiment analysis model
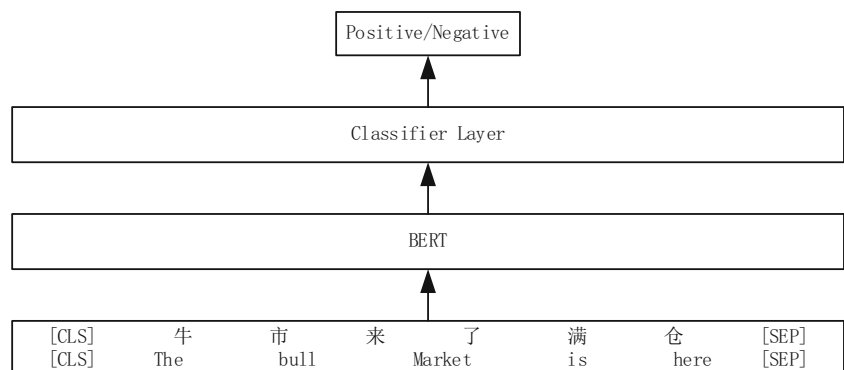
Table 1 Examples of data set

Review	Polarity
越发觉得现在入股市好，好多股票都想买可惜没钱。	positive
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.	
大盘进入主升浪加速运行，请系好安全带。	positive
The market enters the main wave and accelerates; please fasten your seat belt.	
今天下跌还带量。	negative
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.	

features which is a useful supplement to BERT. This paper uses a CNN as a classifier and the input to CNN is the BERT output. The output is calculated as:

$$G_i = f(G_{i-1} \otimes W_i + b_i) \tag{6}$$

W_i represents the weight vector of the ith layer convolution kernel, and b_i is the offset. The operation symbol $\otimes$ indicates a convolution operation. After the convolution operation, local information in the Chinese stock review is transferred to the pooling as shown in Eq. 7.

$$G_i = subsampling(G_{i-1}) \tag{7}$$

The final output layer is also a fully connected layer:

$$h = relu(W \cdot h^f + b) \tag{8}$$

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 is from the Github (https://github.com/algosenses/Stock_Market_Sentiment_Analysis/tree/master/data)website. There are a total of 9204 reviews in the dataset, half of which are positive and half are negative. This dataset uses logistic regression for sentiment classification, and its accuracy reaches 88.09%, which shows that this dataset can be used to train other models.

This paper divided the training set, verification set and test set according to a ratio of 6:2:2. Table 1 lists some examples of stock reviews.

4.2 Metrics

This paper realizes the sentiment classification of Chinese stock review text, which is a typical classification problem. The most commonly used evaluation for classification includes the accuracy, 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 the confusion matrix [29] in Table 2.

The accuracy [30] refers to the proportion of the correctly classified number of all classes in all samples of all classes. Generally, a higher accuracy shows a better classifier. Through the confusion matrix, the formula can be easily obtained:

$$P = \frac{TP}{TP + FP} \tag{9}$$

The recall [31] shows 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:

$$R = \frac{TP}{TP + FN} \tag{10}$$

F is an indicator that combines the accuracy (P) and the recall (R), and its calculation formula is:

Table 2 Confusion matrix

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

Table 3 F1 scores of various BERT models

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

Fig. 2 Relationship between
Epochs and F1
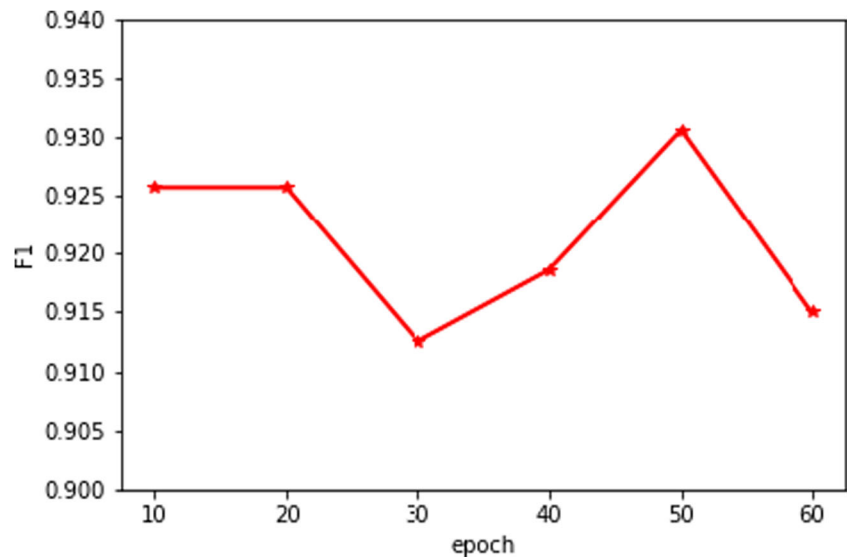
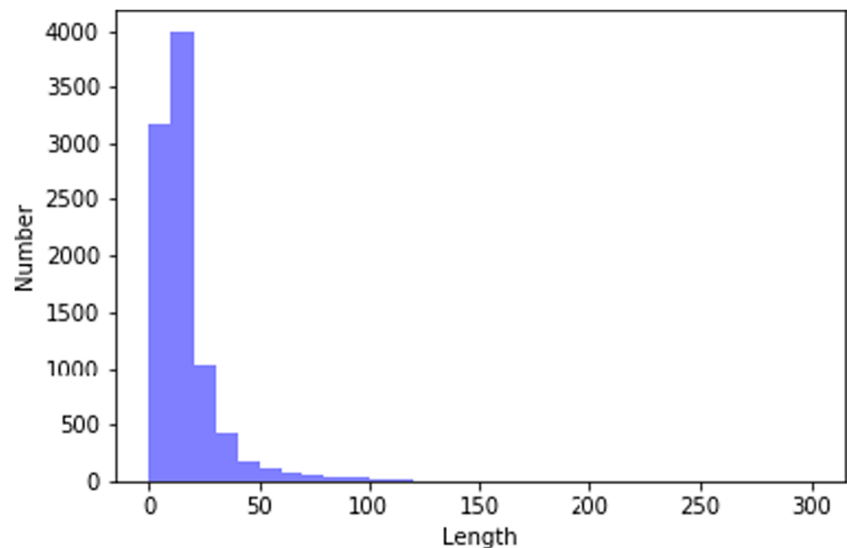
$$F_\beta = \frac{(\beta^2 + 1)*P*R}{\beta^2*P + R} \tag{11}$$

It can be seen from the formula that it is the weighted average value of the recall (R) and the accuracy (P). If the value is 1, the formula will become.

$$F_1 = \frac{2P*R}{P + R} \tag{12}$$

F1 [32] is one of the significant indicators that evaluate the performance of the classifier. The value is close to 0, indicating that the model has bad performance, When the model is close to 1, the performance of the model is great. To verify the classification ability of the BERT model in Chinese stock reviews more effectively, this paper selected two metrics: accuracy and F1.

4.3 Experimental results on different variants of models

This paper proposed different variants of the model in Section 2. The BERT used is the "bert-base-uncased" version provided by Google. The number of self-attention heads is 12 and the hidden layer has 768 dimensions. The total number of parameters for the trained model is 110 M. The dataset in this paper is small, and the BERT model requires a large amount of data. For this reason, we used the Chinese pre-training model trained by Google (https://github.com/ymcui/Chinese-BERT-wwm).

To obtain a more effective classification model, this section uses experimental data based on the previous section and implements the above algorithms with Pytorch (https://pytorch.org/). The F1 value is selected as the indicator, and the experimental results are listed in Table 3.

Fig. 3 Text length distribution in
the dataset

Fig. 4 Relationship between Pad Size and F1
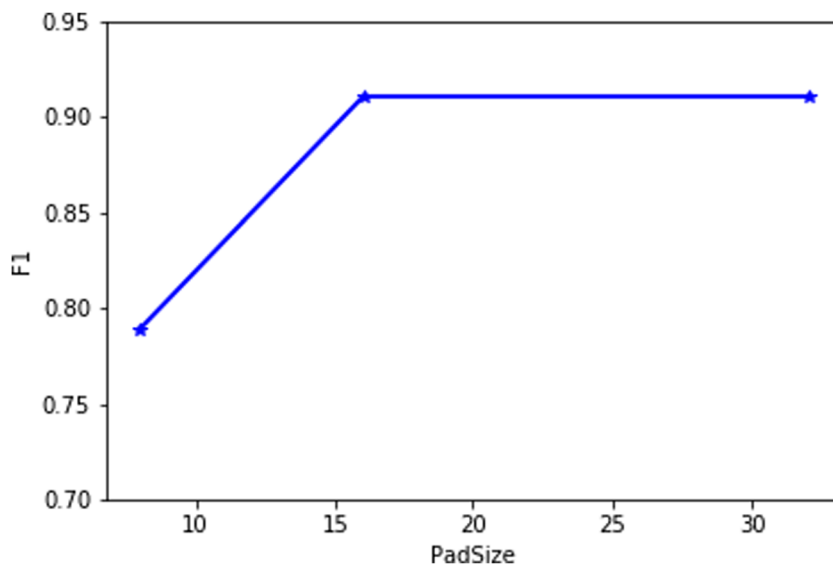
As seen from the experimental results, the F1 scores of different models have only a very small gap, which shows that complex models have not improved the effect of sentiment classification. Therefore, this paper chose the model of BERT+FC for Chinese stock review sentiment analysis.

4.4 Experimental results on hyperparameters settings of model

The hyperparameters in the proposed sentiment analysis for Chinese stock reviews include Epoch, Pad size, Batch size, Learning rate. The Learning rate is set to 5e-5. This section will carry out some experiments to find the best values of the epoch, pad size and batch size.

4.4.1 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 Epoch, Learning rate and Pad size. Next, we executed several experiments to select the best hyperparameters. The data set was divided into a training set and a test set according to a ratio of 6:4.

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

As seen from Fig. 2, as the Epoch increases, the F1 score also increases, which means that the classification performance of the model becomes better. When the 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.

4.4.2 Pad size

The Pad size is the length that the model will process for 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.

The 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 overfilled. In contrast, abundant of information will be lost. The distribution of the review length is shown in Fig. 3. The stock review length is very short, and most of them are less than 50.
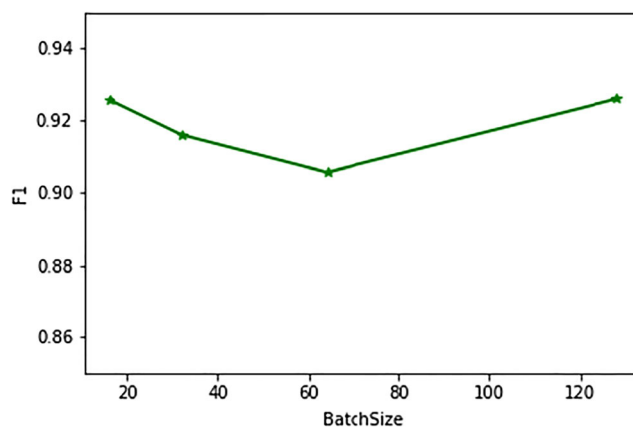
Fig. 5 Relationship between Batch Size and F1

Table 4 Experimental results of our method and other methods

Method	Accuracy	Recall	F1
Word2vec + TextCNN	90.51	90.45	90.45
Word2vec + TextRNN	89.65	89.58	89.57
Word2vec + Att-BLSTM	89.82	89.53	89.52
Word2vec + TextCRNN	90.24	90.23	90.23
Our Model	**92.65**	**92.57**	**92.57**

We set up different pad sizes to experiment, and Fig. 4 shows our results. When the Pad size increases, the F1 score will show an upward trend. When the Pad size is 16, the F1 score will reach the maximum. However when the 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.

4.4.3 Batch size

The Batch Size is the number of training samples sent to the neural network each time. The choice of batch size plays an important role in model training. The batch size will affect the efficiency of memory usage and the capacity of the memory. When the batch is larger, the memory required is larger. At the same time, the convergence time of the model may be reduced. To find a balance between efficiency and memory capacity, the Batch size should be carefully selected.

This experiment fixes the other hyperparameters and increases the Batch size from 16 to 128. The experimental results are shown in Fig. 5.

Figure 5 shows that the different batch sizes have an impact on the performance of our proposed model. When the batch size increases from small to large, the F1 decreases; and when the value of batch size is 64, the F1 is the smallest. As the batch size increases, the value of F1 gradually increases; until the batch size is 128, the model achieves the best performance. At this time, the convergence speed of the model is also the fastest.

In summary, the batch size is set to 128, and the model achieves the fastest convergence speed and the best performance.

Table 5 Results of ablation studies of our model

Model	Accuracy	Recall	F1
Word2vec + FC	**90.13**	**90.10**	**90.11**
Our Model	**92.65**	**92.57**	**92.57**

4.5 Experimental results on sentiment analysis for different methods

According to the experimental results in the previous section, this paper chose BERT + FC model to solve the Chinese stock review sentiment classification. To verify the effectiveness of the BERT model proposed in this paper, the method is strictly compared with several state-of-the-art methods: TextCNN, TextRNN, Att-BLSTM and TextCRNN.

TextCNN was proposed in 2014 by Kim et al. It uses convolutional neural networks to solve text classification problems TextRNN was based on a recurrent neural network and used a multi-task learning framework to train jointly. TextCRNN combines a convolutional recurrent neural network and recurrent neural networks and can achieve better performance. Att-BLSTM can use BLSTM with an attention mechanism that captures the most important semantic information in a sentence. The hyperparameters of TextCNN, TextRNN, Att-BLSTM and TextCRNN are the best in this experiment.

Currently, word embeddings methods have been broadly used in many NLP tasks, such as question answering [33] and sentiment analysis tasks [34]. In this paper, we use Word2vec [35] trained by Sougou (http://www.sogou.com/labs/). The metrics that were used Accuracy, Recall and F1. The comparison results are shown in Table 5.

The average accuracy, recall and F1 for different sentiment analysis methods are shown in Table 4. We can see that the accuracy of our model reaches 92.65%, recall reach 92.57% and F1 reaches 92.57%. Our model obtains the best performance across all methods. It is approximately 2.14%–3% better than the other methods in accuracy. For F1, the maximum of the other methods is 90.45%, which is lower than in our model.

This indeed shows that our model has made significant improvements. The main reason is that our model uses the BERT pre-training model, which considers more semantic information. This verifies that our model has powerful generalization ability and can be widely used in the task of sentiment classification in different fields.

4.6 Ablation study

In this section, we investigate the behaviors and effects of BERT on our model performance. Our model is consists of BERT and classifier layer. We implement a simple classifier by embedding Word2Vec. This model averages the vector of each sentence; and then feeds it to the fully connected layer for classification. The results listed in Table 5. From the table, we can see that BERT can greatly improve the results of sentiment analysis on Chinese stock reviews. The classifier layer does not have much effect on the performance of the model. This indicates that the contribution of our model is mainly dependent on BERT.

4.7 Discussion

In this paper, we propose a new model based on BERT for sentiment analysis of Chinese stock reviews. We conducted several experiments on the Chinese stock review dataset. The experimental results indicate that the BERT+FC model achieved fine classification results in the sentiment analysis of Chinese stock review texts, and is more accurate than TextCNN and TextRNN. This shows that our model can be used as a sentiment classification algorithm for Chinese stock review text.

The proposed model outperforms TextCNN, TextRNN, Att-BLSTM and TextCRNN when performing sentiment analysis of Chinese stock review text. Our main idea was to improve the accuracy of Chinese stock review sentiment analysis. Our model performs better than traditional methods is because we mainly rely on the BERT language model's finetuning, which can allow us to achieve the best performance on the Chinese stock review.

5 Conclusion

With advances in Internet technology, investors can easily post relevant tweets. Mining these comments has important value for financial markets. However, the number of labeled dataset is very small, and the accuracy of existing sentiment analysis of Chinese stock reviews needs to be further improved. In this paper, a novel sentiment analysis method of Chinese stock reviews based on BERT was proposed and applied to a sentiment analysis task. The idea of this paper was to rely on pre-training a BERT to obtain state-of-art results. The BERT pre-training language model performed representation, and then used several classifier layers for sentiment analysis. We experimented on Chinese stock reviews and found that the BERT+FC method can obtain outstanding performance.

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, overcame the difficulty of manually extracting features, and improved the accuracy of sentiment classification. However, there were still some defects. The dataset was small, and the conclusion was not verified on the larger dataset. In subsequent research, we should make further improvements to the BERT model, and expand the sample size to carry out further experiments.

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.

Compliance with ethical standards

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.

References

1. Sheu H-J, Lu Y-C, Wei Y-C (2010) Causalities between sentiment indicators and stock market returns under different market scenarios. Int. J Bus Fin Res 4(1):159–171
2. Wawre SV, Deshmukh SN (2016) Sentiment classification using machine learning techniques. Int J Sci Res (IJSR) 5(4):819–821
3. Feng S, Fu Y, Yang F, Wang D, Zhang D (2012) Blog sentiment orientation analysis based on dependency parsing. J Comput Res Dev 49(11):2395–2406
4. Pang B, Lee L, Vaithyanathan S (2002) Thumbs up sentiment classification using machine learning techniques. Empir Methods Nat Lang Process:79–86
5. J Devlin, M-W Chang, K Lee ,K Toutanova(2019) BERT: Pre-training of deep bidirectional transformers for language Understanding, Proc. NAACL-HLT, pp.4171–4186 2019.
6. Kim Y (2014) Convolutional neural networks for sentence classification. proceedings of EMNLP, Oct. 2014.
7. Liu P, Qiu X, Huang X (2016) Recurrent neural network for text classification with multi-task learning. In IJCAI, 2873–2879. AAAI Press.
8. Zhou P, Shi W, Tian J, Qi Z ,Xu B (2016) Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016), page 207.
9. Wang R, Li Z , Cao J,Chen, T,Wang L (2019). Convolutional Recurrent Neural Networks for Text Classification. 2019 International Joint Conference on Neural Networks (IJCNN) pp.1-6.
10. Abualigah L, Alfar HE, Shehab M, Hussein AMA (2020) Sentiment analysis in healthcare: a brief review. In: Abd Elaziz M, Al-qaness M, Ewees A, Dahou A (eds) Recent advances in NLP: the case of Arabic language. Studies in computational intelligence, vol 874. Springer, Cham. https://doi.org/10.1007/978-3-030-34614-0_7
11. Zubair AM, Aurangzeb K, Shakeel A et al (2017) Lexicon-enhanced sentiment analysis framework using rule-based classification scheme. PLoS One 12(2):e0171649
12. Liu Y, Bi J W, Fan Z P (2017). A method for multi-class sentiment classification based on an improved one-vs-one (OVO) strategy and the support vector machine (SVM) algorithm. Information Sciences, 394, 38–52.

13. Zhang L, Wang S, Liu B (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.

14. Wang Y, Wang M, Fujita H.(2020)Word sense disambiguation: A comprehensive knowledge exploitation framework. Knowledge-Based Systems 190 (2020): 105030.

15. Pranckevičius T, Marcinkevičius V (2017) Comparison of naive bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification. Baltic J Mod Comput 5(2):221

16. Zhang, D, Xu, H, Su, Z, Xu, Y. (2015) Chinese comments sentiment classification based on word2vec and SVMperf. Expert Systems with Applications, 42(4), 1857–1863.

17. Jeevanandam J, Koteeswaran S (2015) Decision Tree Based Feature Selection and Multilayer Perceptron for Sentiment Analysis. ARPN J Eng Appl Sci, ISSN 1819–6608 10(14):5883–5894

18. Abid F, Alam M, Yasir M, Li C (2019) Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. Futur Gener Comput Syst 95:292–308. https://doi.org/10.1016/j.future.2018.12.018

19. Santos D, Gatti M (2014) 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, pp 69–78 [Online]. Available: https://www.aclweb.org/anthology/C14-1008

20. Wang X, Liu Y, Sun C, Wang B, Wang X (2015) 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, pp. 1343–1353, doi: https://doi.org/10.3115/v1/P15-1130

21. Wu X, Chen H, Wang J, Troiano L, Loia V, Fujita H (2020) Adaptive stock trading strategies with deep reinforcement learning methods. Inf Sci 538:142–158

22. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L, Polosukhin I(2017). Attention is all you need. Advances in neural information processing systems, 30, 5998–6008.

23. Hu X, Bing L, Lei S, Philip YS (2019) BERT post-training for review reading comprehension and aspect-based sentiment analysis. Proc NAACL:2324–2335

24. Catelli R, Gargiulo F, Casola V, De Pietro G, Fujita H, Esposito M (2020). Crosslingual named entity recognition for clinical de-identification applied to a COVID-19 Italian data set. Applied Soft Computing, 97, 106779.

25. Yu J,Wei Y,Zhang Y(2019) automatic ancient Chinese texts segmentation based on BERT. J Chin Inf Process 33(11):57–63

26. Duan D, Tang J, Wen Y, Yuan K(2019) BERT Based Research on Classification of Short Chinese Text; Computer Engineering; https://doi.org/10.19678/j.issn.1000-3428.0056222

27. Sun C, Qiu X, Xu Y, Huang X(2019) How to fine-tune bert for text classification?. In China National Conference on Chinese Computational Linguistics. Springer, Cham.

28. Hochreiter S, Schmidhuber J (1997). Long short-term memory. Neural computation, 9(8), 1735–1780.

29. Li Y X, Tan C L, Ding X, Liu C(2004) Contextual post-processing based on the confusion matrix in offline handwritten Chinese script recognition. Pattern Recognition, 37(9), 1901–1912.

30. Abualigah L M, Khader A T, Hanandeh E S (2018). Hybrid clustering analysis using improved krill herd algorithm. Applied Intelligence, 48(11), 4047–4071.

31. Meng X (2019) Feature selection and enhanced krill herd algorithm for text document clustering. Comput Rev 60(8):318–318

32. Abualigah L M , Khader A T , Hanandeh E S (2017) A new feature selection method to improve the document clustering using particle swarm optimization algorithm. J Comput Sci:S1877750316305002

33. Esposito M, Damiano E, Minutolo A, De Pietro G, Fujita H (2020). Hybrid query expansion using lexical resources and word embeddings for sentence retrieval in question answering. Information Sciences, 514, 88-105.

34. Rezaeinia SM, Rahmani R, Ghodsi A, Veisi H (2019) Sentiment analysis based on improved pre-trained word embeddings. Expert Sys Appl 117:139–147. https://doi.org/10.1016/j.eswa.2018.08.044

35. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed representations of words and phrases and their compositionality, In Proc. Adv Neural Inf Process Syst Nevada, NV, USA, pp. 3111–3119

Mingzheng Li received the B.S. degree in Telecommunication Engineering from Hohai University Wentian college, Maanshan, China, in 2015 and the M.E. degree in Computer Technology from Hohai University, Nanjing, China, in 2018. He research interests include Natural Language Processing and Data Mining.