



## Research Article

# Sentiment Classification Algorithm Based on the Cascade of BERT Model and Adaptive Sentiment Dictionary

Ruixue Duan <sup>1,2</sup> Zhuofan Huang <sup>1</sup> Yangsen Zhang <sup>2,3</sup> Xiulei Liu <sup>1,4</sup>  
and Yue Dang <sup>1</sup>

<sup>1</sup>Computer School, Beijing Information Science and Technology University, Beijing 100101, China

<sup>2</sup>Beijing Laboratory of National Economic Security Early-Warning Engineering, Beijing 100044, China

<sup>3</sup>School of Information Management, Beijing Information Science and Technology University, Beijing 100101, China

<sup>4</sup>Laboratory of Data Science and Information Studies, Beijing Information Science and Technology University, Beijing 100101, China

Correspondence should be addressed to Ruixue Duan; [duanruixue@bistu.edu.cn](mailto:duanruixue@bistu.edu.cn)

Received 29 May 2021; Accepted 22 July 2021; Published 15 August 2021

Academic Editor: Jinbo Xiong

Copyright © 2021 Ruixue Duan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The mobile social network contains a large amount of information in a form of commentary. Effective analysis of the sentiment in the comments would help improve the recommendations in the mobile network. With the development of well-performing pretrained language models, the performance of sentiment classification task based on deep learning has seen new breakthroughs in the past decade. However, deep learning models suffer from poor interpretability, making it difficult to integrate sentiment knowledge into the model. This paper proposes a sentiment classification model based on the cascade of the BERT model and the adaptive sentiment dictionary. First, the pretrained BERT model is used to fine-tune with the training corpus, and the probability of sentiment classification in different categories is obtained through the softmax layer. Next, to allow a more effective comparison between the probabilities for the two classes, a nonlinearity is introduced in a form of positive-negative probability ratio, using the rule method based on sentiment dictionary to deal with the probability ratio below the threshold. This method of cascading the pretrained model and the semantic rules of the sentiment dictionary allows to utilize the advantages of both models. Different sized Chnsenticorp data sets are used to train the proposed model. Experimental results show that the Dict-BERT model is better than the BERT-only model, especially when the training set is relatively small. The improvement is obvious with the accuracy increase of 0.8%.

## 1. Introduction

Mobile social networking has emerged and spread widely with the development of Internet applications. Its main goal is to provide an online platform for sharing interests, hobbies, comments, and other information for the general population, and it contains a large amount of information in a form of business reviews. Expressing users' feelings about products (such as hotels and movies) is becoming increasingly common through sentiment comments on various e-commerce websites, forums, WeChat, and other platforms. Extracting user preferences from big data [1] and completing the sentiment analysis do not only serve as a reference for

other users but also provide valuable directions for improvement for businesses. Sentiment analysis is one of the important research tasks in the field of natural language processing, and it helps to complete recommendations in mobile social networks and has become a current research hotspot.

Early sentiment analysis was mostly based on artificially constructed sentiment dictionary design rules for sentiment discrimination [2, 3]. This method had the benefits of inherent simplicity and strong interpretability. However, since human emotional expressions are rich and diverse, and it is immensely difficult to create a complete set of rules capable of judging all complex emotional expressions. Therefore, people gradually adopted data-driven statistical machine

learning methods. Traditional classifiers, such as naive Bayes and support vector machines (SVMs), were used for sentiment analysis [4, 5], but these methods had the disadvantage of relying on hand-crafted features.

In recent years, deep learning has been successfully applied to the fields of image and speech recognition and natural language processing with its powerful representation learning ability, and it has also greatly promoted the research progress of sentiment analysis. Models such as LSTM [6] and BERT [7] were used to construct sentiment analysis algorithms, demonstrating the potential of deep learning models to improve sentiment analysis.

However, the deep learning models suffer from poor interpretability. Integration of sentiment dictionary information with better interpretability into the BERT characterization model and further performance improvement of the BERT model's sentiment analysis requires further research.

To this end, this paper proposes a sentiment analysis algorithm Dict-BERT, which is a cascade of deep learning BERT model and a sentiment dictionary. The concept of positive-negative probability ratio is proposed in this work and used alongside a threshold for deciding whether the BERT model is confident about the prediction, or the sample needs to be cascaded to the rule algorithm of the adaptive sentiment dictionary, and yields superior performance on sentiment classification. The Dict-BERT model combines the advantages of the BERT model and the sentiment dictionary and yields superior performance on sentiment classification. The Dict-BERT algorithm based on BERT and sentiment dictionary cascade performance on sentiment analysis task is evaluated on the Chnsenticorp data set. With the training corpus size of 2000, the Dict-BERT model demonstrates improved performance on the Chnsenticorp data set than just using BERT. The performance improved by 0.8 percentage points, with the achieved correct rate and F1 value reaching 0.9517 and 0.9520, respectively.

## 2. Related Works

There are three main methods of text sentiment analysis, namely, based on sentiment dictionaries, traditional machine learning, and deep learning algorithms.

Sentiment analysis based on sentiment dictionaries [8] is the most direct method. Generally, a heuristic-discriminative sentiment analysis algorithm is designed using a manually labelled sentiment dictionary, combined with adverbs and negative words. However, due to the continuous emergence of new words on the Internet, it is difficult for sentiment dictionaries to include all words referring to emotion. Moreover, human natural language is highly flexible, and it is difficult to design a discriminative sentiment analysis algorithm to determine the sentiment category of the text. In addition, the domain adaptability of the sentiment analysis algorithm based on the sentiment dictionary is very poor, and it is necessary to design a proprietary discriminant function for different scenarios, meaning this method has significant limitations.

A sentiment analysis method based on machine learning, proposed by Pang et al. [9] in 2002, used two text features of N-gram and part-of-speech and compared the effects of three

machine learning algorithms, namely, Naive Bayes, Maximum Entropy, and SVM, on sentiment analysis tasks. Kim and Hoyy [10] proposed quoting location features and evaluating word features to achieve sentiment classification. Xie et al. [11] proposed a new type of multistrategy fusion sentiment feature extraction technology, by constructing three different sentiment analysis models based on three levels of sentiment dictionary, emoticons, and SVM, and studied the fusion effects of different methods. Z. M. Liu and L. Liu [12] used the SVM algorithm, information gain, TF-IDF, and other feature weight calculation methods to improve the performance of sentiment analysis algorithm.

Deep learning models typically are a multilayer neural network, where the representation of the language model is obtained from large-scale data. The deep learning model initially used Google's Word2vec [13] to learn the representation of words, and its features were put into machine learning models, such as SVM, to perform sentiment classification. In addition to improving word representation accuracy, in order to benefit from valuable context information, a deep learning long short-term memory (LSTM) model was used to learn long-distance dependent information and enhance the semantic representation ability. Hu et al. [14] proposed building a related word lexicon on the basis of LSTM, which further improved the accuracy of text sentiment analysis. In recent years, models such as pretrained BERT and ALBERT [15] have emerged. These models are based on a multilayer Transformer model with a multilayer attention mechanism to complete semantic coding and contributed to the important breakthroughs in multiple natural language processing tasks, including sentiment analysis.

Since deep learning has gradually become a research hotspot in the field of natural language processing, technologies related to privacy protection [16] and the approach to solving sentiment analysis problems using deep learning methods of sentiment dictionary matching have also developed rapidly. Many scholars have attempted optimization of the text sentiment analysis algorithm using a sentiment dictionary [17]. Combining it with emotion distribution learning, Zhang et al. [18] proposed an end-to-end framework based on a multitask convolutional neural network, which can learn the sentiment distribution and classification simultaneously. Zhang et al. [19] proposed a Chinese microblog sentiment analysis algorithm based on sentiment dictionary, in which the sentiment value of the microblog text is obtained by the method of weight calculation, to realize the sentiment classification.

Wu et al. [20] proposed a slang sentiment word dictionary that is easy to maintain and expand, which is constructed using network resources, which demonstrated the advantages of using slang sentiment dictionary for sentiment classification. However, the sentiment dictionary-based classification algorithm is heavily related to the content of the sentiment dictionary and the weight of the sentiment word. Using only the sentiment dictionary appears to yield noticeably poorer performance, achieving 10% lower than the sentiment classification model based on deep learning. The effect of combined two approaches, sentiment dictionary, and deep learning algorithm requires further research.

### 3. Dict-BERT Model

*3.1. Dict-BERT Model Framework.* Dict-BERT is a sentiment analysis model based on cascaded BERT algorithm and adaptive sentiment dictionary. The flowchart of the Dict-BERT framework can be seen in Figure 1. The part of the algorithm containing the BERT model is mainly composed of an embedding layer, an encoder layer, two fully connected layers, and a softmax layer. The positive and negative sentiment judgments are completed through the two fully connected layers and the softmax layer. The softmax layer returns the probability of a positive and a negative sentiment. If the probability of the positive sentiment and the negative sentiment of the sample is  $[0.9, 0.1]$ , then it can be determined that the sample belongs to the positive sentiment.

In order to better quantify the model's prediction ability, this paper proposes the concept of positive-negative probability ratio.

$$\text{Positive-negative probability ratio} = \begin{cases} \frac{P_{\text{pos}}}{P_{\text{neg}}}, & \text{if } (P_{\text{pos}} > P_{\text{neg}}), \\ \frac{P_{\text{neg}}}{P_{\text{pos}}}, & \text{if } (P_{\text{neg}} > P_{\text{pos}}), \end{cases} \quad (1)$$

where  $P_{\text{pos}}$  represents the probability of a positive sentiment,  $P_{\text{neg}}$  represents the probability of a negative sentiment, and the sum of  $P_{\text{pos}}$  and  $P_{\text{neg}}$  is always 1. According to the definition, it can be seen that the positive-negative probability ratio must be greater than 1. When  $P_{\text{pos}}$  is larger than  $P_{\text{neg}}$ , the model determines that the sample belongs to positive sentiment class, and vice versa. If the positive and negative sentiment probabilities of the two samples are  $[0.9, 0.1]$  and  $[0.55, 0.45]$ , then the probability of the positive sentiment of the two documents is higher than the probability of the negative sentiment, so the model judges that these two documents are positive sentiment. However, the positive-negative probability ratios of the two articles are significantly different. The positive-negative probability ratio of the first document is 9, and the second one is 1.22. The higher the positive-negative probability ratio is, the higher the model's confidence in sentiment classification is. If the value of the positive-negative probability ratio is relatively low, it means that the model is struggling to correctly distinguish the sentiment tendency of the sample. In such cases the sentiment dictionary with the discriminant function are used to determine the sentiment of the sample.

This paper proposes a sentiment analysis model that combines a BERT model and an adaptive sentiment dictionary. The greater the value of the positive-negative probability ratio is, the greater the difference between the probabilities of the two categories of sentiment classification is, and therefore, the higher the credibility of the sentiment classification prediction is. On the contrary, when the positive and negative probability ratio is lower than the threshold, it indicates that the pretrained model cannot distinguish the sentiment categories. If the positive-negative probability ratio is higher than the predefined threshold, the output of the BERT classi-

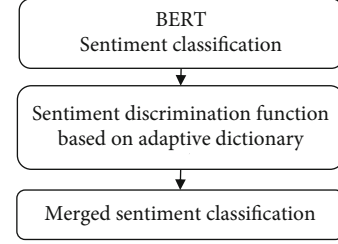


FIGURE 1: The framework of the Dict-BERT model.

fication model is directly used as the final sentiment classification. If the positive-negative probability ratio is lower than the set threshold, the discrimination function of the adaptive sentiment dictionary is used to complete the sentiment classification. In the next section, the effects of models with different thresholds will be introduced. The positive and negative probability ratio thresholds were selected as 1.2, 1.4, 1.6, 2.0, 3.0, and other values, respectively.

#### 3.2. BERT Sentiment Classification

*3.2.1. Overview of the BERT Model.* The BERT (Bidirectional Encoder Representation from Transformers) model is a pretrained model proposed by Google in 2018. The main goal of this model is to use large-scale unlabelled corpus for unsupervised training, so as to obtain a word embedding representation containing rich semantic information. The BERT model is internally composed of a two-way multilayer Transformer model. The Transformer model uses a multihead attention layer to encode contextual information. The input and output dimensions of each layer of the Transformer model remain unchanged, and the superposition of multiple layers of Transformer model can better realize the semantic encoding of the context.

*3.2.2. BERT Model for Sentiment Classification.* The flowchart of the BERT model for sentiment classification is shown in Figure 2. The embedding layer of the BERT model is composed of three parts, the word vector, the segment vector, and the position vector. Three different embedding functions are applied to transform word, segment, and position vectors. Each of the word, segment, and position vectors is in dimension 768 after the embedding processing. To represent the sentiment documents with all the word, segment, and position information, we add the values of the corresponding dimensions for all the three vectors. The word vector is obtained from the pretrained model. The segment vector is composed of 0 or 1. If the sentiment document length is shorter than max sentiment length, then the corresponding dimension is represented by 0; otherwise it is represented by 1. The position vector represents the word position in the sentiment document; it is indexed from 1 to max length. If the document is short, then 255 is filled in the position vector. The encoder layer of BERT is composed of a 12-layer Transformer model, which completes the semantic encoding of the context. Add two fully connected layers behind the encoder model to complete the conversion from the hidden layer to the hidden layer and the hidden

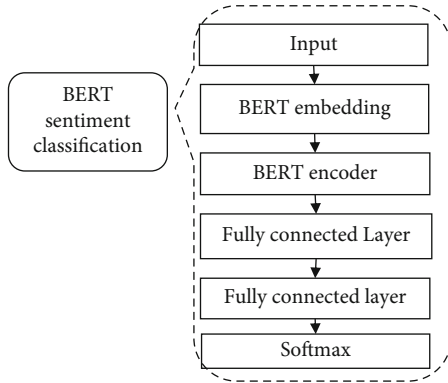


FIGURE 2: The flow chart of the BERT model for sentiment classification.

layer to the sentiment category. Finally, the softmax layer is used to calculate the probability of each sentiment category to complete the classification.

**3.3. Sentiment Discrimination Function Based on Adaptive Dictionary.** The sentiment analysis method based on sentiment dictionary generally adopts manually annotated sentiment dictionary (including positive and negative sentiment words), combined with adverbs and negative words, to design a heuristic discriminative sentiment analysis algorithm. Some sentiment dictionaries use numerical values to express the intensity of positive and negative emotions. However, the sentiment intensity of general sentiment dictionaries often cannot be consistent with the sentiment intensity of sentiment words for the test corpus. To solve this problem, this paper proposes a method of constructing a sentiment dictionary adaptively based on test corpus.

**3.3.1. Building an Adaptive Sentiment Dictionary.** In this paper, we construct a sentiment dictionary based on the HowNet sentiment dictionary and including the sentiment words frequency. The size of the sentiment dictionary is shown in Table 1. The frequency of different sentiment words in the corpus varies greatly. For example, the sentiment word “not bad” appears 211 times in the training set of the Chnsenticorp corpus, while the number of occurrences of “not occupying space” only 1 time. Obviously, the higher the frequency of sentiment words is, the stronger the emotional classification ability of sentiment words. In a corpus, the contribution of each sentiment word to the sentiment tendency is different. According to the number of occurrences of sentiment words, the sigmoid function is used to quantify the contribution of sentiment words to emotional tendency, where count represents the number of occurrences.

$$\text{Contribution} = \text{sigmoid}(\text{count}). \quad (2)$$

The calculation of contribution depends on the corpus, so the sentiment dictionary constructed in this paper is a kind of sentiment dictionary adaptive to the corpus.

TABLE 1: Number of Chinese sentiment words.

Vocabulary category	Number of word
Positive sentiment word	3638
Negative sentiment word	4401

**3.3.2. Sentiment Computing Based on Semantic Rules.** It is difficult to correctly judge the sentiment tendency of the text by relying solely on the sentiment dictionary. For example, when a negative word such as “not” or “no” accompany the sentiment word, the sentiment tendency will change. Adverbs of degree also have a great influence on the judgment of sentiment tendency. In text analysis degree, adverbs and negative words have a great influence on sentiment tendency, so this article mainly combines these two types of words and sentiment dictionary to design the discriminant function of sentiment analysis.

Next, we introduce how the judgment process of sentiment analysis is completed. First, the Peking University word segmentation tool PKUSEG is used to segment the classified text used for training and classification (in order to avoid splitting the sentiment words, you need to pass the sentiment word list to the segmentation tool) and then adjust the emotion according to the context of the emotional word, whether there are negative words or the degree adverbs. There is a contribution of words to the emotional tendency of the entire text. If the score of the positive sentiment of the text is higher than the score of the negative sentiment, it is judged that the text belongs to the positive sentiment; otherwise, it is judged to be the negative sentiment.

$$\text{Positive sentiment score} = \sum_{i=1}^N g_i \times f_i \times c_i. \quad (3)$$

$g_i$  represents the contribution of adverb words. If the adverb of degree appears in the context of positive sentiment words (the window is set to 4), then  $g_i$  is set to 2; otherwise, it is set to 1.  $f_i$  represents the contribution of negative words. If the negative words such as “no” or “not” appear in the context of positive sentiment words (the window is set to 3),  $f_i$  is set to 1; otherwise,  $f_i$  is set to -1.  $N$  indicates the number of positive sentiment words contained in the text.  $c_i$  represents the contribution of sentiment words which is defined in equation (2). Using the same method, the negative sentiment score of the text is calculated by using the context of negative sentiment words.

Using the obtained sentiment score, the emotional tendency of the text is finally determined.

## 4. Results and Discussion

**4.1. Experimental Environment and Parameter Selection.** For this study, The Pytorch was used for creating and training the classification models, using a GPU (Tesla P100) on Ubuntu16.04 system. In the experiment, the dimension of the word vector is set to 768, the maximum length of text is set to 256, and the number of Transformer layers in BERT



is set to 12. The following parameters were used to train the model: the dimension of hidden layer is 768, the size of mini-batch is set to 16, dropout is 0.1, warm up is 0.1, we use Adam optimization, epoch number is set to 5, and the learning rate is  $2 \times e - 5$ .

**4.2. Data Set.** Chnsenticorp data set [21] is a Chinese sentiment analysis data set. This corpus was compiled and published by Dr. Tan Songbo from Fudan University, which contains hotel reviews collected both from online and from review books, as well as labelled positive and negative sentiment polarities. Each sample is a short comment containing a few dozens of words. The data set consists of three parts, namely, the training set, verification set, and test set, in which the training set contains 9600 samples, the verification set contains 1200 samples, and test set contains 1200 samples. The data set details are shown in Table 2. In order to verify the effect, the sizes of the training sets are set to 2000, 4000, and 6000 samples, respectively. The results were evaluated on the test set.

#### 4.3. Evaluation Indicators

**4.3.1. Accuracy and F1.** Accuracy is used for the evaluation of the trained models, which is defined as a ratio of correctly classified samples to the total number of samples. Generally speaking, the higher the accuracy, the better. The formula is described as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (4)$$

where TN is true negative, FN is false negative, FP is true positive, and FN is false positive. In addition to accuracy, F1 value is used for the evaluation of classifier performance in this study.

**4.4. Baseline Model.** The baseline model is BERT. In the BERT model, the dimension of word vector is 768, the dimension of hidden layer is 768, the number of Transformer layers in BERT is 12, the size of minibatch is 16, the optimization function is Adam, and the drop out is 0.1. The output of the BERT model is input into two fully connected layers and a Softmax layer to complete sentiment classification. The parameters of the Dict-BERT model are the same as those of the BERT model.

**4.4.1. Performance of Sentiment Discrimination Function Based on Adaptive Dictionary.** The method of adaptive sentiment discrimination function was evaluated on the Chnsenticorp data set. The obtained results are shown in Table 3. In order to verify the effect of including adverbs and making the sentiment dictionary adaptive, the performance of the model was evaluated both with and without the adverbs and the adaptive capability of the sentiment dictionary.

From the experimental results in Table 3, it can be seen that the performance of the approach which uses only sentiment dictionary yield inferior performance with the accuracy under 80%. In the next experiment, we explain how to cascade the pretrained model with the low-accuracy method

TABLE 2: Chnsenticorp data set.

Data set	Sentiment category	Positive sentiment corpus	Negative sentiment corpus
Chnsenticorp	2	4798	4802
Chnsenticorp2000	2	996	1004
Chnsenticorp4000	2	2009	1991
Chnsenticorp6000	2	2986	3014
Chnsenticorp validation set	2	590	610
Chnsenticorp test set	2	602	598

TABLE 3: Performance analysis of sentiment discrimination module.

Model	Accuracy
Sentiment analysis model based on traditional sentiment dictionary	77.5%
Sentiment analysis model based on adaptive sentiment dictionary	81.5%
Adaptive affective analysis model-adverb of degree	81.33%

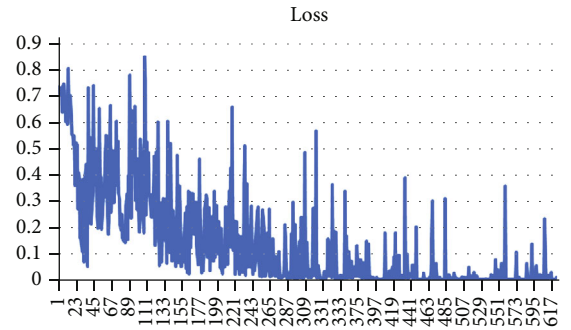


FIGURE 3: Verification of model convergence.

based on sentiment dictionary with the ultimate goal of improving the overall performance of the model.

**4.4.2. Verification of Model Convergence.** The semantic representation of each word can be obtained by pretraining the BERT model, and then, the model is fine-tuned by utilizing the training set of sentiment analysis. The cross entropy loss function is used to calculate the loss, and the parameters are updated using back propagation. Figure 3 demonstrates the convergence of the model on the Chnsenticorp data set (in 2000 samples, batch size is 16, so for each epoch, the training is done  $2000/16 = 125$  times, and epoch is 5). When the iteration times reach 625 ( $125 * 5$ ), the model converged.

**4.4.3. Comparison of Effects between the BERT Model and the Dict-BERT Model.** To verify the effectiveness of the model proposed in this paper, we conducted lots of experimental on Chnsenticorp data set.

Figure 4 and Table 4 show the accuracy of the classifier on the test set of Chnsenticorp data set with the training set size of 2000, when the positive and negative probability ratio thresholds of the BERT model are set to be 1.2, 1.4, 1.6, 1.8, 2, and 3, respectively. It can be seen from Figure 4 that when the

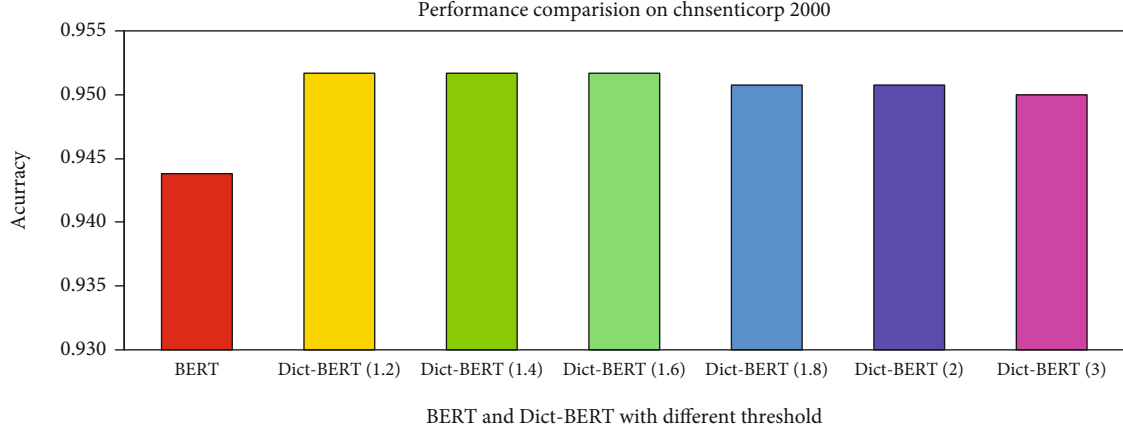


FIGURE 4: Dict-BERT and BERT models. Comparison of accuracy on Chnsenticorp2000 data set.

TABLE 4: Comparison of accuracy of each model in Chnsenticorp2000 data set and F1 score.

Model (threshold value of positive-negative probability ratio)	Accuracy	F1
BERT	0.9438	0.948
Dict-BERT(1.2)	0.9517	0.9520
Dict-BERT(1.4)	0.9517	0.9520
Dict-BERT(1.6)	0.9517	0.9520
Dict-BERT(1.8)	0.9508	0.9515
Dict-BERT(2.0)	0.9508	0.9515
Dict-BERT(3.0)	0.9500	0.9500

TABLE 5: Comparison of accuracy of each model in Chnsenticorp4000 data set and F1.

Model (threshold value of positive and negative probability ratio)	Accuracy	F1
BERT	0.9508	0.9515
Dict-BERT(1.2)	0.9558	0.956
Dict-BERT(1.4)	0.9558	0.956
Dict-BERT(1.6)	0.9542	0.954
Dict-BERT(1.8)	0.9533	0.9535
Dict-BERT(2.0)	0.9508	0.9525
Dict-BERT(3.0)	0.9500	0.956

training set size is only 2000, the accuracy of BERT model is only 0.9438. The Dict-BERT model combines the adaptive sentiment discrimination function with the BERT model. Although the accuracy rate of the model based on sentiment discrimination function is only 0.815, applying the positive and negative probability ratio method with the Dict-Bert model results in the accuracy improvement with the metric rising above 0.95. This approach makes full use of the advantages of pretraining the BERT model and the model based on emotion dictionary. When the threshold value is 1.2, 1.4, and 1.6, the accuracy rate increases to 0.9517, showing the total

accuracy increased by 0.8% and the F1 value increased by 0.4%.

Table 5 lists the accuracy and F1 value of the BERT model and Dict-BERT model with different positive and negative probability ratio thresholds with 4000 samples in the training set in Chnsenticorp data set.

Figure 5 shows the comparison between the accuracies obtained using two models. It can be seen that with the increase of training set, the accuracy of both models can be improved. The highest accuracy achieved by Dict-BERT model is 0.9558. The accuracy of Dict-BERT is better than that of the BERT model, improved by 0.5%.

In order to study the effect of training set size on the Dict-BERT model, Figure 6 compares the accuracies obtained using Dict-BERT and BERT models with varying positive and negative probability ratios with the training set size of 2000, 4000, and 9600. It can be seen that increasing the training set to 9600, the accuracy of Dict-BERT model slightly is higher than that of the BERT model, with a 0.08% difference. It is apparent that the accuracy of Dict-BERT and BERT is increasing with the increase of the training corpus. Under the condition of insufficient size of the training corpus, the Dict-BERT model has added semantic rules based on the emotion dictionary to make up for the insufficient training of the pretrained model. However, it appears that for the sufficient size of the training corpus the advantages of Dict-BERT model are reduced.

The higher the positive and negative probability ratio is, the higher the reliability of the pretrained model. When the positive-negative probability ratio is low, the credibility of the pretrained model is low, which means that the semantic information can be effectively used to improve the overall performance of the model. However, with the increase of the threshold, the amount of data used for classification using semantic rules based on sentiment dictionary increases. Since the accuracy of this method is low, the total accuracy of the model therefore decreases. It can be seen from Figure 6 that with the increase of the threshold of positive-negative probability ratio, the accuracy of the cascade model first increases, followed by a decrease. Enlarging the training set results in a steady improvement

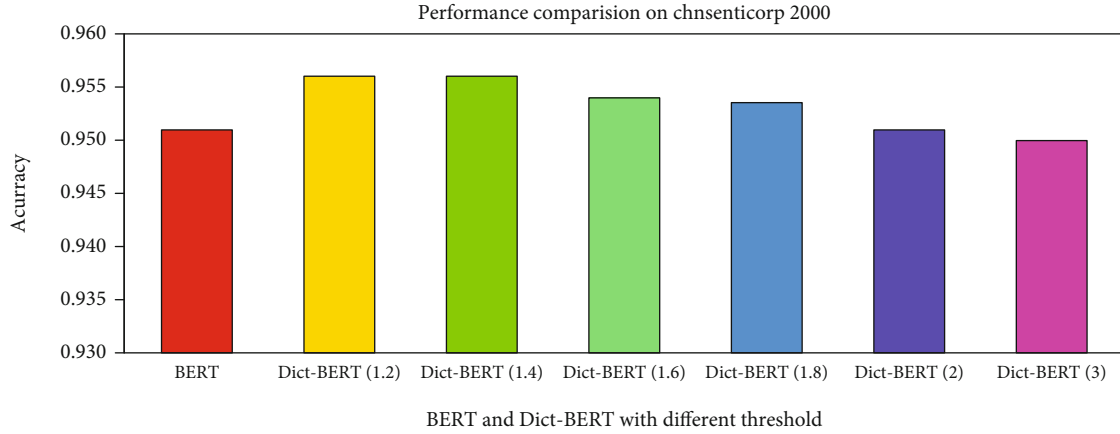


FIGURE 5: Dict-BERT and BERT model. Comparison of accuracy on Chnsenticorp4000 data set.

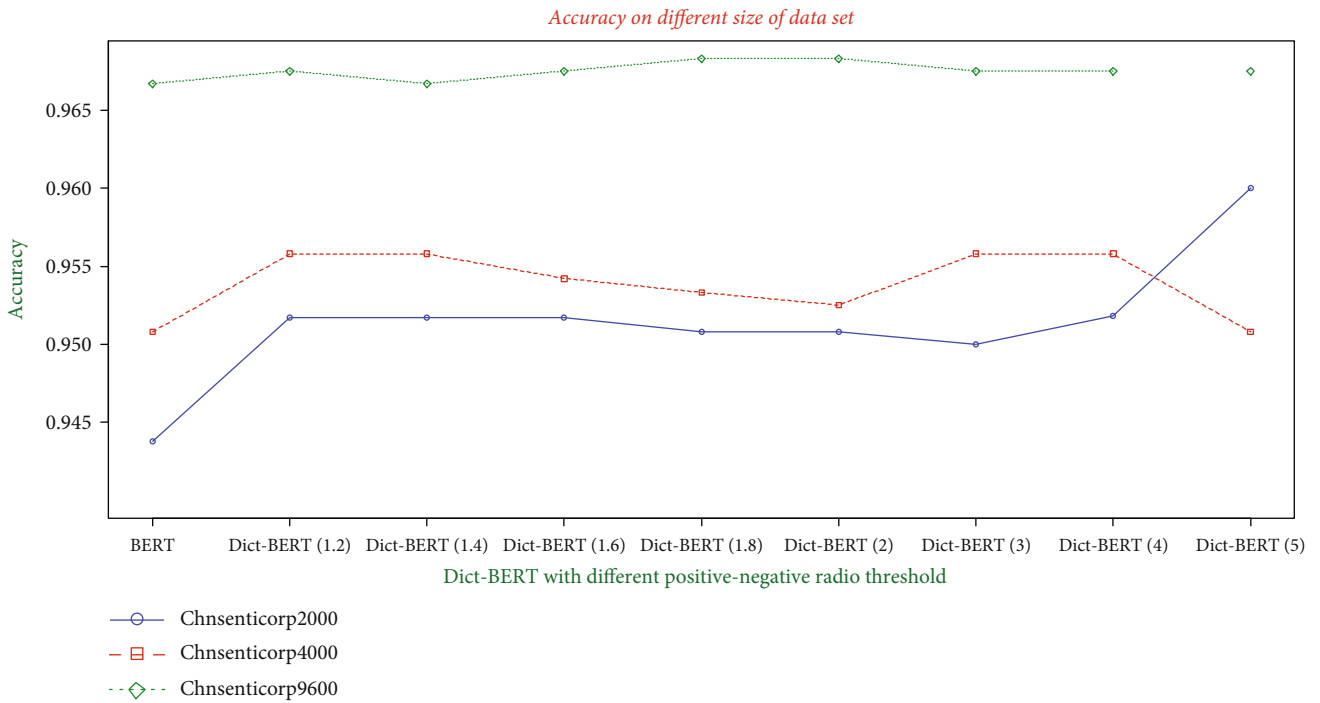


FIGURE 6: Comparison of model effects on different scales of Chnsenticorp data set.

of the accuracies of BERT and Dict-BERT models, with the accuracy of the Dict-BERT model higher than the accuracy of the BERT model. The smaller the training corpus is, the greater the improvement offered by the Dict-BERT model over the BERT model is.

### 5. Conclusion

In this paper, a sentiment analysis algorithm is proposed which is a cascade made of a pretrained deep learning BERT model and a semantic rule model. The accuracy rate of sentiment analysis based on the discriminant function of sentiment dictionary is only 81%, which is lower than the accuracy of the pretrained model. However, by cascading the pretrained model and the model based on sentiment dic-

tionary and introducing the concept of positive-negative probability ratio, the performance is further improved. The smaller the training corpus is, the more prominent the advantages of the proposed model are. If the training corpus of a task is insufficient, the cascade method proposed in this paper can be used to introduce more data knowledge to improve the performance of the system.

There are many discriminant models based on sentiment dictionary to solve the task of sentiment analysis. In the future, we would like to consider cascading these improved sentiment analysis models with improved pretrained models such as Roberta, BERT-wwm, XLNet, and ALBERT, to further improve the performance of sentiment analysis. Besides, privacy protection should also be considered in future sentiment analysis [22].

## Data Availability

The data that support the findings of this study are openly available in duanruixue/chnsenticorp at <https://github.com/duanruixue/chnsenticorp> (reference number [21]).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This paper is supported by the funding of Beijing Natural Science Foundation (Grant No. 4204100), School Foundation of Beijing Information Science and Technology University (Grant No.1825023), Subject and Graduate Education Program of Beijing Information Science and Technology University (Grant No. 5112111004), and 2021 Promote Connotation Development of Universities-Scientific Research Training Program for College Students of BISTU (Grant No. 5102110805).

## References

- [1] J. Xiong, H. Liu, B. Jin, Q. Li, and Z. Yao, "A lightweight privacy protection scheme based on user preference in mobile crowdsensing," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 5, pp. 1–16, 2021.
- [2] J. S. Wu and K. Lu, "Chinese micro-blog sentiment analysis based on multiple sentiment dictionaries and semantic rule sets," *Computer Applications and Software*, vol. 36, no. 9, pp. 93–99, 2019.
- [3] L. Yang, F. Q. Zhou, and H. F. Lin, "Sentiment analysis based on emotion commonsense knowledge," *Journal of Chinese Information Processing*, vol. 33, no. 6, pp. 94–99, 2019.
- [4] L. Ma, X. Liu, and Y. L. Gong, "Research of micro-blog short text sentiment orientation analysis based on semantic," *Application Research of Computers*, vol. 33, no. 10, pp. 2914–2918, 2016.
- [5] H. Wei and Y. H. Wang, "Sentiment analysis of Chinese micro-blog based on topic," *Computer Engineering*, vol. 41, no. 9, pp. 238–244, 2015.
- [6] T. Hu, Y. Dan, J. Hu, X. Li, and S. Li, "News named entity recognition and sentiment classification based on attention-based bi-directional long short-term memory neural network and conditional random field," *Journal of Computer Applications*, vol. 40, no. 7, pp. 1879–1883, 2020.
- [7] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," 2018, <https://arxiv.org/abs/1810.04805>.
- [8] F. Zhen, G. Yi, and Z. Zhenhao, "Sentiment analysis of movie reviews based on dictionary and weak tagging information," *Journal of Computer Applications*, vol. 38, no. 11, pp. 3084–3088, 2018.
- [9] B. Pang, L. L. Lee, and V. Shivakumar, "Thumbs up: sentiment classification using machine learning techniques," in *In the proceeding of EMNLP*, pp. 79–86, Philadelphia, USA, 2002.
- [10] S. M. Kim and E. Hooy, "Automatic identification of pro and con reasons in online reviews," in *Proceedings of the international conference on computational linguistics (COLING) and the Association for computational Linguistics (ACL): Posters*, pp. 483–490, Sydney, Australia, 2006.
- [11] L. Xie, M. Zhou, and M. Sun, "Hierarchical structure based hybrid approach to sentiment analysis of Chinese micro blog and its feature extraction," *Journal of Chinese Information Processing*, vol. 2012, no. 1, pp. 73–83, 2012.
- [12] Z. M. Liu and L. Liu, "Empirical study of sentiment classification for Chinese microblog based on machine learning," *Computer Engineering and Applications*, vol. 2012, no. 1, pp. 1–4, 2012.
- [13] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *In proceedings of International Conference on Learning Representations*, Scottsdale, Arizona, USA, 2013.
- [14] F. Hu, L. Li, Z. L. Zhang, J. Y. Wang, and X. F. Xu, "Emphasizing essential words for sentiment classification based on recurrent neural networks," *Journal of Computer Science and Technology*, vol. 32, no. 4, pp. 785–795, 2017.
- [15] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "ALBERT: a lite BERT for self-supervised learning of language representations," 2019, <https://arxiv.org/abs/1909.11942>.
- [16] J. Xiong, R. Ma, L. Chen, Y. Tian, L. Lin, and B. Jin, "Achieving incentive, security, and scalable privacy protection in mobile crowdsensing services," *Wireless Communications and Mobile Computing*, vol. 2018, 12 pages, 2018.
- [17] H. X. Wan and Y. Peng, "Topic words extraction of social media based on semantic constrained and time associated LDA," *Journal of Chinese Computer Systems*, vol. 39, no. 4, pp. 742–747, 2018.
- [18] Y. Zhang, J. Fu, D. She, Y. Zhang, S. Wang, and J. Yang, "Text emotion distribution learning via multi-task convolutional neural network," in *International Joint Conference on Artificial Intelligence*, pp. 4595–4601, Stockholm, Sweden, 2018.
- [19] S. Zhang, Z. Wei, Y. Wang, and T. Liao, "Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary," *Future Generation Computer Systems*, vol. 81, no. 8, pp. 395–403, 2018.
- [20] L. Wu, F. Morstatter, and H. Liu, "Slang SD: building and using a sentiment dictionary of slang words for short-text sentiment classification," *Language Resources & Evaluation*, vol. 2016, no. 6, pp. 1–14, 2016.
- [21] Chnsenticorp data <https://github.com/duanruixue/Chnsenticorp>.
- [22] J. Xiong, R. Ma, L. Chen et al., "A personalized privacy protection framework for mobile crowdsensing in IIoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4231–4241, 2020.