Chinese Implicit Sentiment Analysis based on Hybrid Neural Networks

Yanfang Qi¹ *, Chao Wang², Guoan Zhu³

¹School of Computer and Artificial Intelligence, Lanzhou Institute of Technology, Lanzhou, Gansu, 730030, China
²School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan, Hubei, 430000, China
³Department of Computer Science and Technology, College of Engineering, Yanbian University, Yanji, 133002, China

*Corresponding author: yanfangqi@lzit.edu.cn

Abstract. Implicit sentiment analysis is an important part of sentiment computing, especially sentiment analysis based on deep learning has become a research hotspot in recent years. In this paper, convolutional neural networks are used to extract text features, combined with long-term and short-term memory network (LSTM) structure to extract context information, and add attention mechanism to the network to construct a new hybrid neural network model to achieve implicit emotion for text analysis. The hybrid neural network model extracts more meaningful sentence semantic and structural hidden features from the word-level and sentence-level hierarchies, respectively, and pays attention to the features with a large emotional contribution rate through the attention mechanism. The classification accuracy rate of the model on the public implicit sentiment dataset has reached 77%. The research of implicit sentiment analysis can improve the effect of text sentiment analysis more comprehensively, and further promote the application of text sentiment analysis in the fields of knowledge embedding, text representation learning, user modeling and natural language.

1. Introduction

Sentiment analysis is a hot research issue in the field of natural language processing. It can be understood as the tendency to analyze emotions by recognizing text opinions. Text sentiment analysis differs from text classification and text mining in that emotions are abstract and only take literal meaning. Judging the emotion expressed by the text is one-sided, this method is only applicable to text with obvious emotion words or emoticons that can fully express emotions. [1]

Nowadays, more and more people are willing to express their life and emotional state by publishing pure text or dynamic information with pictures on social platforms such as Weibo and WeChat. A small amount of dynamic information can directly express their current emotions. However, a large number of dynamic languages are more implicit, and there are no obvious emotion words, and the implicit emotions expressed cannot be judged intuitively. As shown in the three sentences in Table 1, they express commendatory, derogatory and objective (neutral) implicit emotions [2].
The research results of explicit sentiment analysis are now abundant, and the research of implicit sentiment classification is still in its infancy. Implicit sentiment analysis plays an important role in the research of knowledge embedding, text representation learning, user modeling and natural language understanding. It can more comprehensively improve the effect of text sentiment analysis and promote the application of text sentiment analysis in related fields. [3]

In recent years, the use of machine learning methods for sentiment analysis has improved experimental performance. Some researchers solve the deficiencies in a single neural network by combining various neural networks. The current research work on sentiment analysis can be divided into two categories: sentiment analysis methods based on rules and dictionaries, and sentiment analysis methods based on machine learning. Rule-based and dictionary-based sentiment analysis methods have low classification flexibility and poor transferability. Using machine learning methods for sentiment analysis solves the problem of poor adaptability of sentiment dictionaries. The experimental performance of traditional machine learning methods needs to be further improved; the learning method can extract deeper sentence-level features in the text, thus improving the accuracy of text classification. [4]

In this paper, the attention mechanism is added on the basis of the depth model, and it shows good performance in Chinese implicit sentiment analysis. The model proposed in this paper uses double-layer convolutional neural networks to extract text-level and sentence-level features, combined with the bi-directional long-short-term memory network Bi-LSTM structure to pay attention to its context information, Bi-LSTM replaces the pooling layer in traditional convolutional neural networks. Add attention mechanism to analyze the implicit emotions contained in the Chinese text, and divide them into three categories: no emotion, praise implicit emotion and derogatory implicit emotion.

2. Proposed model framework

2.1. Model introduction

![Fig 1. Hybrid Neural Attention Network (CLA).]
The hybrid neural network used in this paper includes convolutional neural network (CNN), bidirectional long-short-term memory network (Bi-LSTM) and attention mechanism (Figure 1). The first layer of the model structure is the word embedding layer in the preprocessing process, and the second layer is a convolutional layer that includes three convolution kernels. The main purpose of using a hybrid convolution kernel is to extract as much as possible to make more meaningful to abstract the features, replace the pooling layer in the convolutional neural network with BiLSTM, and use Dropout to prevent overfitting.

The output sequence of iLSTM combined with the attention mechanism pays attention to word vectors with a large contribution rate. The third layer is our sentence embedding layer. The fourth layer has the same structure as the second layer and is used to extract abstract features of sentences. After adding attention to the sentence vector after feature extraction, it is used as the feature of the last layer of the maxmax classifier for classification. The size of the word vector in the word embedding layer is 100 dimensions, and the maximum number of sentences per document is set to 15. When using convolutional neural networks to process text information, the width of the convolution kernel is consistent with the size of the word vector, so the scope of the context covered by the convolution kernel each time is consistent with the size of the convolution kernel, so you can set the size of the convolution kernel to determine the characteristics of several consecutive words. In addition to the convolutional layer, the traditional CNN also uses the pooling layer to extract more important features to simplify the network. In this paper, we choose to use the two-way LSM instead of the pooling layer, and add an attention mechanism to the sequence of the two-way LSTM output. Further pay more attention to the words and sentences in the sequence that have a large contribution rate to the classification.

2.2. Bi-LSTM encoding model

The Bi-LSTM used in this paper can simply concatenate the obtained two directions of lt. The coding model of the sentence by BiLSTM is shown in Fig. 2.

![Fig 2. Bi-LSTM bidirectionally encodes the phrase "I love encoding".](image)

Enter "me", "love", and "encode" in sequence to get three vectors in the forward LSTM, as shown in the blue part of Figure 1, and enter "encode", "love", and "me" in sequence for the backward LSTM. The vectors are shown in the gray part of Fig. 3, and the hidden vectors obtained by forward and backward LSTM are concatenated to obtain the vectors encoded by BiLSTM. In the method of this paper, BiLSTM does not encode the vectors obtained by embedding the original words after word embedding, but encodes the vectors with abstract features of great significance after extracting their features through a convolutional neural network. The coding contains vector context information, which can make full use of the context information during classification and improve the classification performance.
### 2.3. Attention mechanism

Bi-LSTM summarizes the extracted features after convolution from two directions, fully retaining the context information contained in the features. The attention mechanism allows the model to focus more on words and sentences that have a greater contribution to sentiment classification and learn local importance.

The alignment mode adopted in the attention mechanism in this paper is the dot product mode. The vector hit output by BiLSTM is subjected to a Tanh transformation to obtain the hidden vector representation \( w_{it} \), and then the attention of each word to the emotion classification is calculated, and finally obtained by simple addition. All attention \( s_i \), \( u_w \) in a sentence represent the context vector of the word. The sentiment expressed in a certain language is always closely related to the upper and lower sentences. The method proposed in this paper can make full use of the context information. The sentiment analysis of the sentence not only encodes a single sentence, but also encodes the context of this sentence into a simple document. Put the marked sentiment sentence in the middle of the document, and fill the context sentence before and after. Therefore, the context semantic and structural information of the sentence is also included in the feature extraction, which enriches the feature information and improves the experimental performance.

\[
\begin{align*}
   w_{it} &= \tanh (W_w h_{it} + b_w) \\
   a_{ij} &= \frac{\exp (\alpha v_{it}^T u_w)}{\sum_t \exp (w_{it}^T u_w)} \\
   s_i &= \sum a_{ij} h_{it} 
\end{align*}
\]

### 3. Experimental results and analysis

#### 3.1. Data set

The data set used in the experiment is the data set provided by Shanxi University published by the National Social Media Processing Conference (SMP2010), which is mainly data from platforms such as product forums, Weibo, and travel websites. The main work of this paper is to evaluate the implicit sentiment sentences in Chinese. The texts containing the sentiment words in the data set have been filtered through a large-scale sentiment dictionary. In the processed data set, the implicit sentiment sentences are partially marked, and they are divided into three categories: praise implicit sentiment sentences (1), derogatory implicit sentiment sentences (2) and neutral sentiment sentences (0). The data is published as a sentence-segmented document, which contains the complete contextual information of the sentence.

#### 3.2. Evaluation index

The evaluation indexes used in the experiment are accuracy, accuracy (P), recall rate (R), accuracy (Acc), F1 value and hamming loss. Hamming loss (Hamming_loss) is used to calculate the accuracy of the multi-label classification model, which can measure the degree of inconsistency between the predicted markers and the actual markers of the sample.

#### 3.3. Analysis of experimental results

The convolution layer of the convolution neural network extracts meaningful abstract features from the representation of words and sentences, and carries out a comparative experiment when the size of the convolution kernel is 3, 4, 5 and the mixture of three convolution kernels, respectively. At this time, the epoch of the experiment is set to 8, and the experimental results are shown in figure 3. The (a)-(d) diagram of figure 3 represents the convolution kernel size of 3, 4, 5 and mixed convolution kernel, respectively. From figure 5, it can be found that when 8 epoch are trained, the results begin to converge,
and the training accuracy can reach 0.9, and the test accuracy fluctuates around 0.7. It is obvious that when the epoch is 3 or 4, training and testing fit best, so all epoch in later experiments are set to 3.

By comparing the four groups of diagrams from figure 3 (a) to (d), we can see that the precision and loss change in figure 3 (a) fluctuates greatly, and there is no convergence trend. This is because when the convolution kernel is 3, only the features of the adjacent three words are extracted each time, and the semantic and structural information in the sentence is not all included, so the experimental effect is relatively poor, and the curve tends to relax obviously when the convolution kernel is 4 and 5. When the mixed convolution kernel is used, the experimental effect is the best. When the convolution kernel is 3, the feature information between adjacent words can be fully extracted. When the convolution kernel is 5, the span before and after the feature extraction is increased, and the semantic and structural features in the sentence can be extracted. The combination of the features extracted by the mixed convolution kernel can ensure that the convolution neural network can fully extract meaningful feature information from words or sentences in the convolution process, thus improving the classification effect.

**Fig 3.** Convolution kernel contrast experiment.

4. Conclusion
In this paper, we replace the pooling layer in the traditional convolution neural network CNN with BiLSTM, which can save the context information in the extracted features more completely, and add the attention mechanism to the model, which enables the model to pay more attention to the words and sentences that contribute greatly to emotion classification. Learn local importance. The experimental results show that this method can significantly improve the classification effect of implicit emotional sentences. The difficulty of implicit emotion analysis lies in the implication of semantic expression. The information extraction of semantic features and context structure features of implicit emotion sentences
directly affects the final classification results. In the following implicit emotion analysis, the main work is to extract the feature information of implicit emotional sentences from a more fine-grained point of view. At present, there are some difficulties in emotion analysis, such as data scarcity, category imbalance, domain dependence and language imbalance, so the emotion analysis based on multimedia fusion, domain adaptation, deep semantics and social network will be a focus of the future research work.

Reference
[1] Wang Jin, Huang Chao, Wang Ke, et al. Pedestrian attribute classification based on attribute-sensitive convolution neural network [J]. Journal of Jiangsu University: natural Science Edition, 2019
[2] Chen Bo. Text classification method based on convolution neural network based on cyclic structure [J]. Journal of Chongqing University of Posts and Telecommunications: natural Science Edition, 2018.
[3] Luo Fan, Wang Houfeng. Emotion classification of Chinese text based on RNN and CNN hierarchical network [J]. Journal of Peking University: natural Science Edition, 2018
[4] Du Tianbao, Yu Chunhao, Wen Zhuo, et al. Psychological evaluation model based on emotional characteristics of text [J]. Journal of Jilin University: science Edition, 2019