Multifaceted sentiment analysis of public comments on the Dianping.com

Liying Sui, Lu Shang, Xiaochao Guo, and Dexue Zhang

College of Electronic Information Engineering, Shandong University of Science and Technology, Qingdao, 266590, China
1074467903@qq.com

Abstract. Emotional analysis of text has always been a hot topic in natural language research. In view of the long-term dependence of recurrent neural networks and the fact that most models do not consider the correlation between input and output, this paper proposes a bidirectional LSTM model based on attention mechanism to judge the comment emotion. This method vectorizes the semantics of comments into the LSTM model, improves the relevance of input and output through the attention mechanism, fuses the aspect category and aspect Term, and outputs the results through the classifier. The experimental results of Dianping.com review data set provided by AI Challenger competition show that the improved method adopted in this paper is better than the common deep learning method.

1. Introduction
The emotional analysis of text is a process of analyzing, processing, summarizing and reasoning subjective text with emotional color. Due to the rapid development of social media on the Internet, a large number of formal records have been generated, and thus emotional analysis has grown into an important research field of natural language processing.

The methods of emotion analysis mainly include traditional methods and deep learning methods. In recent years, affective analysis based on deep learning model has become the mainstream because deep learning can obtain more text features and improve the accuracy greatly. Muyongli et al put forward the emotion method based on the integrated convolutional neural network, which integrated several CNN models to reduce the data imbalance and omitted the process of rule making and feature extraction in the traditional method, which is of great significance for the research of emotion analysis method. However, RNN and CNN do not consider the correlation between the input and output of the model, which leads to low accuracy and poor stability of relation extraction. The long-term and short-term memory network (LSTM) [5] is an improved model of cyclic neural network proposed by Hochreiter et al. It constructs memory units to store historical information, so as to connect the current time with the previous input state, and effectively alleviates the long-distance dependence of RNN and CNN. The model has been applied to the field of natural language and achieved good results, but LSTM does not consider the impact of model input on output.

Aiming at the problems of traditional methods and common deep learning models, this paper proposes a bidirectional LSTM model based on attention mechanism to analyze emotion, which is to effectively reduce noise and improve classification efficiency.
2. Data Set
This paper obtains data from AI Challenger competition. The data includes 120000 comments in Chinese from Dianping.com's catering industry. According to the granularity, the dataset constructs a two-tier label system, which contains six first tier categories and 20 second tier fine-grained elements. For example, location, service, price, environment, dishes and other first tier labels. There are some child tags for each tag. For example, location categories include sub-tags such as convenient transportation, easy to find, and close to business districts. For each tag, there are four emotional types.

| Sentimental Label | 1 | 0 | -1 | -2 |
|-------------------|---|---|----|----|
| Meaning           | Positive | Neutral | Negative | Not mentioned |

Instance:
"Noodle restaurants with a good taste are quite cost-effective, and the portions are very large. Girls can't finish a whole portion of noodles. The environment is good, at least it looks bright and clean, and generally the small restaurant is not as good as this sanitation condition. At lunchtime, there are a lot of people, and the sidewalk is also full. It is said that the restaurant next door sometimes opens for people to sit and eat noodles."

And the label of the category service and price is as follows:

| The first layer label | The second layer label | type |
|-----------------------|-----------------------|------|
| Service               | Waiting time          | -2   |
|                       | Waiter’s attitude     | -2   |
|                       | Parking convenience   | -2   |
|                       | Serving speed         | -2   |
| Price                 | Price level           | -2   |
|                       | cost-effective        | 1    |
|                       | discount              | -1   |

The data set includes more than 120000 comments, which are divided into training, verification and test sets in the proportion of 6:2:2.

3. Data Preprocessing

3.1 Word2vec
For all the contents of the training data set, each content should be marked firstly. Here, through the use of the software package jieba to solve the problem of Chinese NLP. So what we have now is a vector of many words or phrases. Finally, we get about 55k different phrases, which can be used in vector transmission [2]. All statistical frequency word or phrase based on this, and the number of
occurrences is greater than the threshold value set 3 is replaced with ",<UNK>". Therefore, after sorting them in descending order according to the number of occurrences, a corpus dictionary is obtained. In addition, three indexes are added, namely ",< pad >", ",< start >" and ",< end >", so that the list of words can be converted into real vectors.

3.2 Input format
Now, for each training data point, this paper uses the frequency sequence list to transfer it to a vector and sets the length of the vector to the longest vector after transmission (if the length of the vector is less than the length of the longest vector, these vector will be set to zero), and records the original length of each vector so that it can be used in the neural network training process. For label format, there are 20 types of labels, each of which has four emotional types, namely - 2, - 1, 0 and 1, respectively, indicating not mentioned, negative, neutral and positive.

4. Main model
Starting from the problem of sentiment analysis, this chapter explains the idea and method of constructing the model, and puts the prepared inputs into the neural network model for model training. Embedding is used to transfer vectors before putting data into the model. It is meaningful to use word embedding, because word embedding is a representation of word "semantics", which can effectively encode the semantic information related to the current task. For emotional classification, there are always some similar annotation words, so embedding is very important for target classification. The following is a brief introduction of three comparative models:

4.1 Bidirectional LSTM model
Long short term memory (LSTM) network is a variant of cyclic neural network. Three gates (input gate, forget gate and output gate) are introduced into LSTM, and memory cells with the same shape as hidden state are used to record additional information. But sometimes the prediction may need to be determined by the previous inputs and the later inputs, which will be more accurate. Therefore, a bidirectional cyclic neural network is proposed, which is embedded in the bidirectional LSTM model, and the output is increased. Then the results are obtained in the full connection layer using softmax. Here are some equations about the model. First, for an LSTM unit:

\[ f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \]  
\[ i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \]  
\[ o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \]  
\[ c_t = f_t \odot c_{t-1} + i_t \odot \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \]  
\[ h_t = o_t \odot \sigma_o(c_t) \]

Representation: \(x_t\) is the input vector, \(h_{t-1}\) and \(h_t\) are the output of the last LSTM unit and the current unit. \(f_t\), \(i_t\) and \(o_t\) represent the activation vectors of forgetting gate, input gate and output gate respectively. \(W\), \(U\) and \(b\) represent the weight matrix and deviation vector. \(\sigma_g\) and \(\sigma_c\) represent sigmoid and tanh activation functions.

If \(h_N\) and \(g_N\) is the last output of LSTM in two directions, and the equation is expressed as follows:

\[ Label = \text{softmax}(W_{fc}(h_N + g_N)) \]

\(W_{fc}\) represents the weight matrix of the full connection layer. One thing to note is that the final dimension of the full connection layer output is the product of the number of tags and the number of emotion types of each tag, so it is necessary to adjust its size to classify each tag in the actual training.

4.2 Bidirectional GRU model
The main structure of GRU model is similar to that of LSTM model. The GRU model updates gate to replace input and forget gates of LSTM model, and the formula is as follows:

\[ z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \]  
\[ r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r) \]
\[
\begin{align*}
\hat{h}_t &= z_t \circ \hat{h}_{t-1} + (1 - z_t) \circ \sigma(g(x_t) + U \circ \hat{h}_{t-1} + b) \\
\end{align*}
\]  \hspace{1cm} (9)

\(\hat{h}_t\) and \(x_t\) represent the input and output vectors. \(z_t\) and \(r_t\) represent update and reset doors. \(W, U\) and \(b\) represent the weight matrix and deviation vector.

4.3 ATAE-LSTM model

Because the bidirectional LSTM and GUR models consider all tags at the same time, there is no real information to learn for each specific class tag. When considering a specific class of information, it is important to determine which parts of the text have a significant impact on a given aspect. Therefore, this paper attempts to use the aspect level LSTM model in [1].

![Figure 1. Aspect-level LSTM model](image)

Add a layer of interest between the hidden layer and the output layer of the LSTM. The attention mechanism [4] allows the network to consider different parts of the input sentence, and enables the model to learn the added content based on the input sentence. More specifically, this model is applicable to specific class tags, and the attention layer can evaluate the impact of various parts of the text on this aspect. The attention equation in this paper is as follows:

\[
M = \tanh(W_h H, W_v v_a \otimes e_N) \quad (10)
\]

\[
\alpha = \text{softmax}(w^T M) \quad (11)
\]

\[
r = H \alpha^T \quad (12)
\]

\(W_h, W_v\) and \(M\) is the weight projection parameter. \(\alpha\) is the vector containing the attention weight. \(r\) represents a given aspect. \(v_a\) in \(W_v v_a \otimes e_N\) is the same as the number of words in the sentence. A label result can be obtained by the following equation:

\[
L = \text{softmax}(f_c(\tanh(W_p r + W_h h_N))) \quad (13)
\]

\(F_c\) represents a fully connected layer, \(h_N\) is the last output of the hidden layer. Then the final results are obtained according to the softmax equation.
5. Experimental results and analysis

5.1 Parameter setting
Here are some parameter settings of atae-lstm model. Adam Optimizer was selected for modeling.

| Number of epochs | Batch size | Number of layers | Hidden units | Learning rate |
|------------------|------------|------------------|--------------|---------------|
| 20               | 50         | 1                | 500          | 0.002         |

5.2 experimental result
The following are the results of loss variation obtained by training three models:

The indicators of the comparative experiment only use accuracy, because this is a multi-classification problem, so the average accuracy of 20 tags is multiplied by 0-1 loss to express the final results of this paper. The benchmark model uses the SVM [3] provided by the competition website with an accuracy of about 70%. The accuracy of test set is 84.9% for GRU model, 85.2% for bidirectional LSTM model and 87.7% for atae-lstm model.

| Model         | Accuracy |
|---------------|----------|
| GRU           | 84.9%    |
| B-LSTM        | 85.2%    |
| ATAE-LSTM     | 87.7%    |

In the figure, the two losses of GRU and LSTM verification sets are similar. But there are obvious over fitting problems in LSTM model. LSTM can extract the internal relations between all tags and texts, while atae-lstm can capture the important parts and different parts of sentences when given different aspects, analyze which words determine the emotional polarity of a certain aspect of sentences, embed learning aspects, and enable aspects to participate in the calculation of attention.
weight, so the accuracy is higher and more competitive.

6. Further Work
Atae-lstm is more accurate than traditional GRU and bidirectional LSTM, but the training model needs a lot of time. How to shorten the training time is an optimization direction. At the same time, for attention mechanism, simulation of multiple aspects is also the future development direction.

Acknowledgments
Special thanks to the careful teaching of my teachers and the help of classmates.

References
[1] Wang, Yequan, Minlie Huang, and Li Zhao. (2016) Attention-based lstm for aspect-level sentiment classification. Proceedings of the 2016 conference on empirical methods in natural language processing.
[2] Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems, 26, 3111-3119.
[3] Nguyen, T., Li, M., Bass, I., & Sethi, I. K. (2005). Investigation of Combining SVM and Decision Tree for Emotion Classification. Seventh IEEE International Symposium on Multimedia (ISM 2005), 12-14 December 2005, Irvine, CA, USA. IEEE.
[4] Ying Zhou, Yue Liu, & Jun Cai. (2018). Sentiment analysis of Weibo based on attention mechanism. Information theory and practice, 41(03), 89-94.
[5] Hochreiter, Sepp, & Schmidhuber, Jürgen. Long short-term memory. Neural Computation, 9(8), 1735-1780.