Advances in Intelligent Systems Research, volume 147
International Conference on Network, Communication, Computer Engineering (NCCE 2018)

# Multi-view and Attention-Based BI-LSTM for Weibo Emotion Recognition

Xiaofeng Cai[1, a)], Zhifeng Hao [1, 2]

[1] *Faculty of Computer Science, Guangdong University of Technology, GuangZhou 510006, China*
[2] *Foshan University, Foshan 528000, China*

[a)] Corresponding author: ms.xiaofengcai@hotmail.com

**Abstract.** Weibo emotion recognition is one of the main tasks of the study of social public opinion. BI-LSTM, as a derivative model of RNN, has been widely used in the task of text emotion analysis. However, existing models do not make good use of prior information sun as emotion words and emoji, and we always capture the different keywords in order to gain a different understanding of the text. Therefore, this paper proposes a Multi-view and Attention-Based BI-LSTM method for weibo emotion recognition. Firstly, we use the emotion ontology lexicon and weibo emotions to label each word in a sentence. Secondly, we use these labels as attention information, combine the attention mechanism and BI-LSTM to get the sentiment words perspective and emoticon perspective. Finally, the output of the original semantic perspective of BI-LSTM is fused with the output of the above two perspectives to enhance the performance of the classification algorithm. Experiments show that the proposed method has a 6% increase in macro-average F1 score and an increase of 8% in micro-average F1 score compared with a AVE-BI-LSTM output in the task of weibo emotion recognition.

**Key words:** Multi-view; BI-LSTM; attention mechanism; weibo emotion recognition word embedding.

## INTRODUCTION

At present, the rapid development of the mobile Internet, users often like to use Weichat, Weibo, Taobao, etc. to leave comments with a variety of emoji to express a view of a thing or commodity. As a sharing and communication platform, Weibo has the characteristics of timeliness and randomness. A large number of users often like to express their opinions on Weibo, and these text data often contain the user's emotions. The spread of negative emotions may bring indirect economic losses to enterprises, and even affect the harmony and stability of society. The analysis of Weibo's text sentiment has great value for the monitoring of public sentiment and improvement of products by enterprises. Therefore, it is of great significance to study the text emotion recognition of Weibo, which is also one of the research focuses in the field of textual emotion mining.

There are two main research methods for text sentiment analysis: one is a combination of emotional dictionary and rules1; the other is based on machine learning2. Machine learning methods are divided into traditional machine learning methods and deep learning methods. Traditional machine learning methods mainly use Bayes, support vector machines or maximum entropy algorithms. These methods are accompanied by a large number of manual feature projects and have task specificity. The quality of feature selection directly affects the correctness of text sentiment analysis, and the characteristics of different task selections are different, and these methods are easy to lose text grammar and semantic information. In recent years, deep learning has achieved breakthrough results in the field of natural language processing such as speech recognition3, machine translation4, question and answer5, and abstract generation6. Scholars began to explore the application of deep learning to textual emotions. In the analysis, the most used one is the RNNs model. The training of RNNs relies on a large number of well-marked linguistic materials. Based on the linguistic features of Weibo's more liberal language and the special nature of the sentiment analysis task, emotional dictionaries and emoji are very helpful in improving the accuracy of Weibo's emotional analysis. However, most of the current text sentiment analysis methods based on the sequence model RNNs do not make good use of emotional dictionaries and emojis, and when reading comprehension, we also unconsciously extract the keywords in the text to analyze the text. With different types of keywords, readers will have different perspectives of text understanding. Integrating these perspectives will greatly improve the accuracy of text sentiment recognition.

In response to the above-mentioned issues, this paper proposes a multi-view and attention-based BI-LSTM weibo emotion analysis method. Firstly, we use the Chinese emotion vocabulary ontology library resources [9] collated and annotated with the information retrieval lab of Dalian University of Technology and the bidirectional LSTM based on attention mechanism to obtain the first perspective; secondly, we use weibo emoji and bidirectional LSTM based on attention mechanism to obtain the second perspective; then, using the purely semantic understanding of the bidirectional LSTM to obtain the third perspective; finally, by blending the three perspectives as new emotional characteristics, the emotion word perspective can be expressed on the expression of emotions, and the expression can be expressed on the emoji perspective. The combination of rich and complex emotions and hidden perspectives from semantic perspectives improves the accuracy of Weibo's sentiment analysis.

## MULTI-VIEW AND ATTENTION-BASED BI-LSTM FOR WEIBO EMOTION RECOGNITION

### Multi-View Learning

Multi-view learning is a machine learning algorithm that takes full advantage of multiple perspectives of the same research object to learn its intrinsic patterns. Different perspectives usually contain complementary information, and multi-view learning methods can use this information to learn to understand the data structure of objects. Weibo texts are rich in emotional dictionaries and emoji. Whether they are from emotional dictionaries, emoji, or the semantic understanding of plain text, the understanding of the same paragraph is different. This paper hopes to obtain different expressions of Weibo text sentiment through these three different perspectives, and then fuse the three perspectives to obtain the final Weibo emotion expression. Using three perspectives of complementary information and prior knowledge such as emoji and emotional lexicons to improve the accuracy of weibo emotion recognition.

### BI-LSTM Based on Attention Mechanism

RNN is an extension of traditional feedforward neural networks and is widely used to solve serial labeling problems. However, the standard RNN has the problem of the gradient disappearing. To overcome this problem, Hochreiter proposed LSTM. Different from RNN, LSTM upgrades the hidden layer, introduces a memory unit, and uses the gate control structure to avoid the long-term dependence of the RNN. The formula for LSTM is as follows:

$$\begin{cases} X = \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} \\ f_t = \sigma(W_f X + b_f) \\ i_t = \sigma(W_i X + b_i) \\ o_t = \sigma(W_o X + b_o) \\ \tilde{c}_t = \sigma(W_c X + b_c) \\ c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t = o_t \odot \tanh(c_t) \end{cases} \tag{1}$$

Where, $W_i, W_i, W_i \in R^{H\times k}$ are weight matrix, $b_f, b_i, b_o, b_c \in R^H$ are bias term. $\sigma$ is sigmoid function， $x_t$ is the input of the LSTM at time t, $\odot$ means multiplication, $f_t$,$i_t$,$o_t$,$\tilde{c}_t$ are the forgotten gate, input gate, output gate, and cell activation vector at time t. H and k denote the dimensions of the hidden layer and the dimensions of the input, respectively.

Combining past contextual information with future contextual information is very beneficial for many sequence problems. However, whether it is RNN or LSTM, only the information of the previous context at the current moment can be captured. BI-LSTM performs forward training and backward training separately for each training sequence, and then combines the results of forward training and backward training together as the output of the current time. The specific calculation is as follows:

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \tag{2}$$

Although the standard LSTM can learn the order of words, as well as lexical and syntactic information, it ignores the more important characteristics of emotional words and emoji in weibo emotion recognition. In order to solve this problem, this paper adds a layer based on attention mechanism to BI-LSTM.

As shown in Figure. 1, H $\in R^{d\times n}$ is the hidden layer output vector $[h_1, \dots, h_n]$ generated by the BI-LSTM model. V $\in R^{v\times n}$ is the attention vector corresponding to each moment$[v_1, \dots, v_n]$. Where d is the dimension of the output vector of the hidden layer and n is the length of the input sequence. The attention mechanism will produce an attentional weight α and an implicit hidden layer characteristic expression r. The specific formula is as follows:

$$\begin{cases} M = \tanh(W[H \oplus V]) \\ \alpha = softmax(\omega^T M) \\ r = H\alpha^T \end{cases} \tag{3}$$

Where, M $\in R^{(d+d_v)\times n}$,α$\in R^n$,r $\in R^d$. W $\in R^{d\times(d+d_v)}$ and ω$\in R^{d+d_v}$ are the learning parameters. $\alpha$ is a vector composed of attention weights. r is a weighted sentence representation based on the view-vector (view vector V). ⊕ is a splicing operator, which expresses the stitching of the output of BI-LSTM each time with the corresponding view-vector (view vector V), $d_v$ is the dimension of the perspective vector V.

The final output sentence is represented as:

$$h^* = \tanh(W_r r) \tag{4}$$

Where, $h^* \in R^d$ and $W_r$ are the learning parameters.

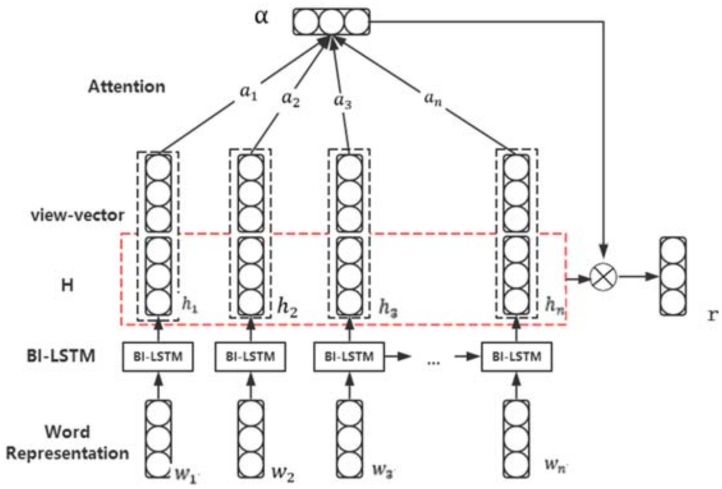


**FIGURE 1.** BI-LSTM Model Based on Attention Mechanisms

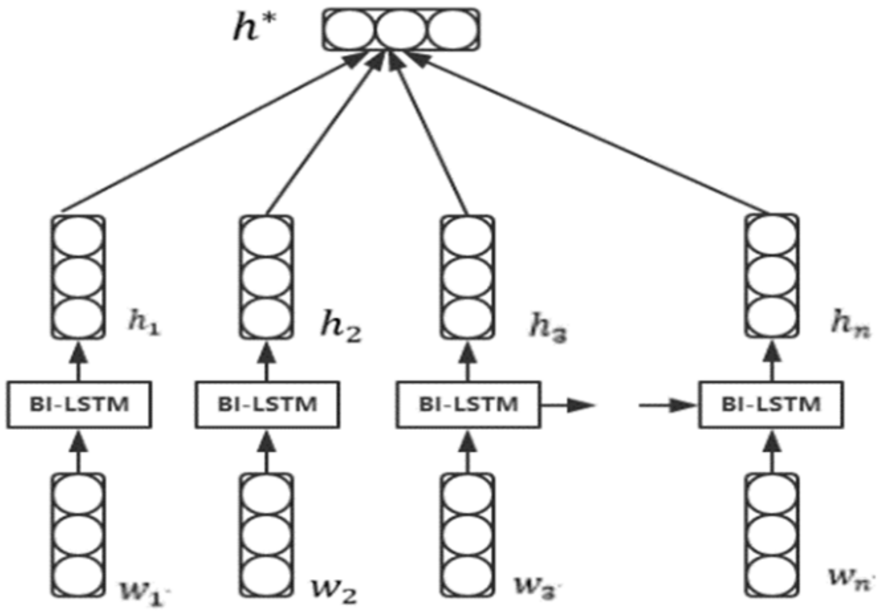


**FIGURE 2.** BI-LSTM Model Based on Pure Semantic Perspective

## Emotional Word Perspective

Emotional words are very important for emotion recognition. This paper uses the Chinese ontology emotional lexicon provided by the Information Retrieval Research Institute of Dalian University of Technology and the BI-LSTM based on attention mechanism mentioned above to obtain the weibo text representation based on emotional word perspective. The vocabulary contains 27,467 Chinese emotional words. The emotional vocabulary summary used in this paper is shown in Table.1. First, by looking up the table, each word in each weibo text is tagged with an emotion category label (PA, PE, PD, PC), and the label not found is the label null. Then use one-hot encoding as the view-vector of the emotion word view, where the view-vector is the view vector in the attention-based BI-LSTM model above. The tag categories we extracted are: PA, PE, PD, PH, PG, PB, PK, NA, NB, NJ, NH, PF, NI, NC, NG, NE, ND, NN, NK, NL, PC, null, so the dimension $d_v$ of the view vector is 22.

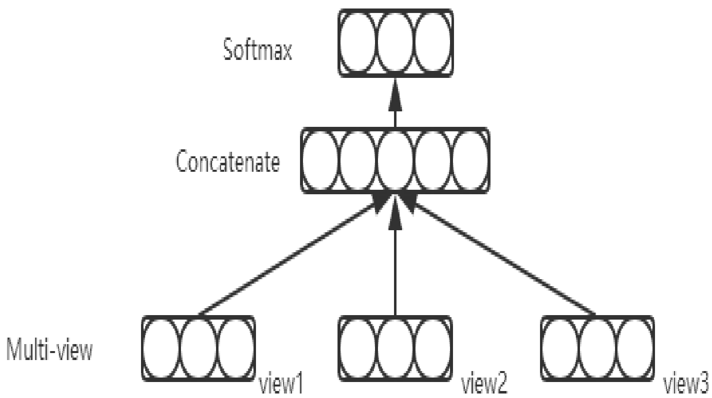


**FIGURE 3.** Multi-perspective fusion structure chart

**TABLE 1.** Chinese ontology classification of thesaurus

| Num. | Class1 | Class2 | Example |
|---|---|---|---|
| 1 | 乐 | 快乐（PA） | 喜悦、欢 |
| 2 | 好 | 尊敬（PD） | 恭敬、敬爱 |
| 3 | 怒 | 愤怒（NA） | 气愤、恼火 |
| 4 | 哀 | 悲伤（NB） | 忧伤、悲苦 |
| 5 | 惧 | 慌（NI） | 慌张、心慌 |

## Emoji Perspective

Emoji is very widely used in Weibo, so it is also very important for the expression of Weibo emotions. This paper extracts the most frequently occurring 30 emoji from many weibo emoji. By looking up the emoji table, each word in each weibo text is tagged with an emotion category tag, and the label not found is the label null. Then use the one-hot code to encode as the view-vector of the emoji view, where the view-vector is the view vector in the attention-based BI-LSTM model above. The emotional tag categories we extracted are: 抓狂]、[耶]、[鼓掌]、[委屈]、[泪]、[爱你]、[good]、[吃惊]、[偷笑]、[吐]、[哈哈]、[心]、[酷]、[奥特曼]、[大哭]、[花心]、[大笑]、[怒]、[嘻嘻]、[干杯]、[伤心]、[怒号]、[飞吻]、[鄙视]、[脸红]、[眼泪]、[害羞]、[开心]、[生病]、[礼物]、null，so the dimension $d_v$ of the view vector is 31.

## Purely Semantic Perspective

Adding a purely semantic perspective allows the expression of hidden emotions based on representational learning of non-emotional and non-emoji of text. The network structure of the pure semantic perspective is shown in Figure. 2. Compared to the attention-based BI-LSTM model, the attention layer (intentional layer) and view-vector (visual vector V) are removed and the calculation method is the BI-LSTM model. The output of each moment is averaged together as a semantic representation of the text. The output of the network is:

$$h^* = \frac{1}{n}\sum_1^n h_i \tag{5}$$

## Multiple View Fusion and Emotion Classification

This paper uses BI-LSTM based on attention mechanism and BI-LSTM from purely semantic perspective with the Chinese text sentiment dictionary ontology library of Dalian University of Technology and Weibo emoji to obtaine the expressions of the emotional word perspective（$h^*_{view1}$）, the expression of the emoji perspective（$h^*_{view2}$） and the expression semantic perspective ($h^*_{view3}$). The emotion word perspective is the expression of the emotion display, the emoji perspective can express the rich and complex emotions, and the semantic perspective is based on the hidden emotion expression of the expression learning. These perspectives are important for Weibo's sentiment analysis, so they need to be integrated. As shown in Figure 3, this paper integrates the vectors obtained from view1, view2, and view3 into one layer and connects them to a layer of softmax to obtain the probability distribution vector of the weibo emotion label. The specific formula is as follows:

$$\begin{cases} o_c = h^*_{view1} \oplus h^*_{view2} \oplus h^*_{view3} \\ o_f = softmax(W_o o_c) \end{cases} \tag{6}$$

Where, $o_c$ is the expression of the final weibo text resulting from the stitching of the three perspective vectors, of which is the final output of the model. $W_o$ is a parameter to learn. $\oplus$ represents a vector stitching operation.

### Training of Models

This model can be trained with end-to-end back propagation. The objective function (loss function) is the cross-entropy loss function. Let y denote the correct mood distribution and $\hat{y}$ denote the model's predicted value. The training goal of the model is to minimize the cross-entropy error between y and $\hat{y}$ , as shown in equation (7):

$$loss = -\sum_i \sum_j y_i^j log\hat{y}_i^j + \lambda||\theta||^2 \tag{7}$$

Where i denotes the index of the Weibo text, j denotes the index of the emotion category, λ is the parameter of the L2 canonical term, θ is the parameter set of the model, and then a random gradient descent algorithm is used to find the minimum value of the cross entropy, so that the prediction value gradually approaches actual value.

## EXPERIMENTS AND ANALYSIS

### Experimental Data

The experiment uses the NLP&CC 2013 evaluation dataset. The task of the evaluation is to determine whether the weibo contains emotions. For a weibo containing emotions, judge one of its emotional categories: anger anger, disgust, fear of fear, happiness of happiness, likeness, sadness sadness, surprise in surprise. Marked as none for weibo that do not contain emotions. The training set contains 10,000 weibo texts, the test set contains 4000 weibo texts, and each text has a basic emotion tag. Data Preprocessing

This paper uses the jieba word segmentation tool to segment the weibo text and filter out the stop words. Then, tag each word with emotional lexicons and emoji. After that, encode the tags with one-hot encoding, Finally, using the open source tool word2vec 7 to train word vector proposed by Google in 2013, the selected model is the skip-gram, the dimension is 50, the window size is 5, and the word vector is dynamically updated during the training process.

### Experimental Parameter Settings

The word vector is pre-trained with word2vec, and the view vector view-vetor is expressed with one-hot. The rest of the parameters that need to be trained are initialized with an even distribution (-0.1, 0.1). The maximum number of words is 200, and the dimension of the input word vector is 50 dimensions. In order to accelerate the training of the model, a batch training method is used, the batch size is set to 256, and the filling method is used to ensure that the length of each batch of training sentences is the same, and the learning rate is Set to 0.00125, the number of nodes in the hidden layer is set to 120, and λ is set to 0.01.

### Experimental Results and Analysis

In view of the above-mentioned method, this article conducted an experiment on the NLP&CC 2013 evaluation data set. This paper uses the NLP&CC 2013 evaluation standard, and uses Macro, Micro recall, Recall, and F(F-measure) values as evaluation indicators. The comparison methods are as follows: BI-LSTM, the output of the last moment of the BI-LSTM is text expression, then a layer of full connection, and the last layer is a SoftMax function;CNN(Kim), the first layer of the method is a sentence vector represented by n*k, and the second layer is a convolutional layer with different filter widths. The width of experiments used in this paper is [2, 3, 4], and the third layer is the max_over_time_pooling layer, then a layer of full connection, and the last layer is a SoftMax function; BI-LSTM-AVE, the results of the direct summation of the output of each time in the BI-LSTM are taken as a text representation, followed by a full-connection layer and finally a soft max function; BI-LSTM-AVE-MV, BI-LSTM based on multi-view and attention mechanism, is the method of this paper; BI-LSTM-V1, emotional word perspective; BI-LSTM-V2, emoji perspective; BI-LSTM-V1-V2, blending emotional words and emoji perspectives; BI-LSTM-AVE-V1, merging pure semantic and emotional word perspectives; BI-LSTM-AVE-V2, merging pure semantics and emoji perspectives.

It can be seen from Table.2 that:

The BI-LSTM-AVE method has a higher macro-average and micro-average FI score than the BI-LSTM method. It can be seen that taking the output of the last moment of the BI-LSTM as an expression of weibo emotionsa lot of information in front of the moment will still be missed.

The BI-LSTM-MV method increases nearly 6% in macro-average F1 score and nearly 8% in micro-average F1 score than BI-LSTM-AVE. It can be seen that the integration of pure semantic perspectives, emotional word perspectives, and emoji perspectives can understand texts from different perspectives. The fusion of multiple perspectives is very meaningful for enhancing emotional expression and further enhancing emotion classification.

FI score from the BI-LSTM-AVE-MV method are higher than the BI-LSTM-V1-V2, BI-LSTM-AVE-V1, and BI-LSTM-AVE-V2 methods. It can be seen that the pure semantic perspective, emotional word perspective, and emoji perspective are all very important for enhancing the expression of Weibo's text sentiment, and they are indispensable.

The macro-average and micro-averages of the FI score of the BI-LSTM-AVE-MV method are 7% higher than the BI-LSTM-AVE averages. And the macro-average and micro-averages of the FI score of the BI-LSTM-AVE-MV method are 17% higher than BI-LSTM-V1 and BI-LSTM-V2 averages. It can be seen that the purely semantic perspective serves as a representation of the entire microblog text and plays a role in globally understanding the text semantics. It is more important than the other two perspectives for Weibo text sentiment recognition.

In summary, the emotional perspective is the expression of emotional display, the emoticon perspective can express rich and complex complex emotions, the semantic perspective is based on the expression of the hidden emotional expression of learning, and the fusion of three perspectives enhances the expression of weibo emotions and it is of great significance to improve the accuracy of emotional classification.

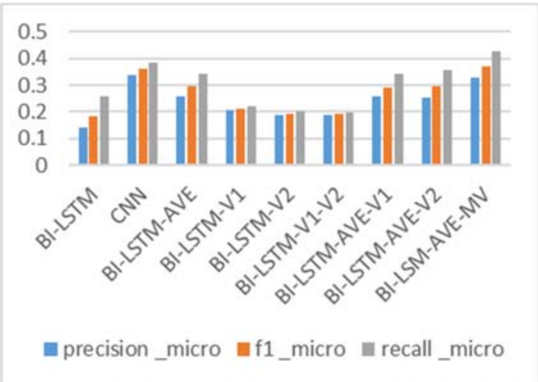


**FIGURE 4.** Weibo emotional results in the micro-average.

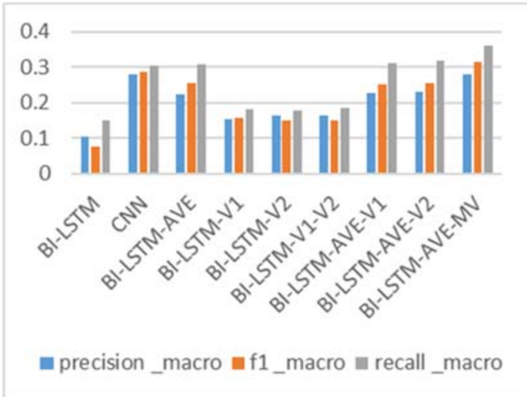


**FIGURE 5.** weibo emotional classification results

## CONCLUSION

This paper proposes a BI-LSTM weibo emotion recognition method based on multiple perspectives and attention mechanisms. Using the BI-LSTM model to make good use of the contextual information of the language while avoiding the cumbersome feature engineering and low generalization ability of using traditional machine learning methods. And Using the attention mechanism combined with emotional lexicon and emoji, the text semantics are understood from multiple perspectives and the accuracy of weibo emotion classification is improved.

**TABLE 2.** Weibo emotional classification in the macro average and micro-average evaluation results

| methods | precision _micro | f1 _micro | recall _micro | precision _macro | f1 _macro | recall _macro |
|---|---|---|---|---|---|---|
| BI-LSTM | 0.13960 | 0.18090 | 0.25691 | 0.10474 | 0.07538 | 0.14979 |
| CNN | 0.33987 | 0.36017 | 0.38306 | 0.28131 | 0.28786 | 0.30563 |
| BI-LSTM-AVE | 0.25581 | 0.29356 | 0.34438 | 0.22587 | 0.25673 | 0.30976 |
| BI-LSTM-V1 | 0.20780 | 0.21289 | 0.21823 | 0.15499 | 0.15834 | 0.18180 |
| BI-LSTM-V2 | 0.18659 | 0.19297 | 0.19982 | 0.16427 | 0.15166 | 0.17773 |
| BI-LSTM-V1-V2 | 0.18609 | 0.19141 | 0.19705 | 0.16524 | 0.15188 | 0.18522 |
| BI-LSTM-AVE-V1 | 0.25612 | 0.29292 | 0.34208 | 0.22823 | 0.25379 | 0.31062 |
| BI-LSTM-AVE-V2 | 0.25287 | 0.29535 | 0.35497 | 0.23181 | 0.25623 | 0.31867 |
| BI-LSM-AVE-MV | **0.32746** | **0.37111** | **0.42818** | **0.28057** | **0.31460** | **0.36235** |

## ACKNOWLEDGMENTS

Project fund: National Natural Science Foundation of China (61472089, 61502108); NSFC-Guangdong Joint Fund (U1501254); Guangdong Natural Science Foundation (2014A030308008, 2014A030306004).

## REFERENCES

1. LI C, CAO F Y, CAO Y D, et al. Text Sentiment Classification Based on Phrase Patterns[J]. Computer Science, 2008, 35(4): 132-134.
2. Wawre S V, Deshmukh S N. Sentiment classification using machine learning techniques[J]. International Journal of Science and Research (IJSR), 2016, 5(4): 819-821.
3. Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks[C]//Acoustics, speech and signal processing (icassp), 2013 ieee international conference on. IEEE, 2013: 6645-6649.
4. Cho K, Van Merriënboer B, Bahdanau D, et al. On the properties of neural machine translation: Encoder-decoder approaches[J]. arXiv preprint arXiv:1409.1259, 2014.
5. Golub D, He X. Character-level question answering with attention[J]. arXiv preprint arXiv:1604.00727, 2016.
6. Rush A M, Chopra S, Weston J. A neural attention model for abstractive sentence summarization[J]. arXiv preprint arXiv:1509.00685, 2015.
7. Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C]//Advances in neural information processing systems. 2013: 3111-3119.