Text sentiment analysis based on Glove model and United Network

Duojiao Li¹*, Chenwan He² and Ming Chen³

college of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan Hubei 430205, China

*Corresponding author’s e-mail: 21807010023@stu.wit.edu.cn

Abstract. Traditional convolutional neural network has been proved to have good classification effect in text sentiment analysis experiments. However, CNN ignores the context dependent information of sentences, which affects the classification effect. Therefore, this paper proposes a text sentiment analysis model based on CapsuleNets and Attention mechanism. Firstly, the word vector is trained by using Glove model; secondly, the local features are extracted by using the features of double-layer capsule network; then, attention mechanism is added to assign the corresponding weight to the extracted text information; finally, we complete classification. The experimental results show that the proposed model can effectively improve the accuracy of emotion classification.

1. Introduction
The development of social network technology helps people update information in time. In order to analysis sentiment of users, a large number of comments need to be processed, and operators formulate corresponding strategies according to the classification results, the text sentiment analysis has a wide range of value [1]. Due to the large amount of information, the traditional way cannot guarantee the timeliness. Deep learning has good results in many NLP tasks [2-3]. CNN network in the pooling layer will lead to the loss of important information, ignoring the context relevance. Therefore, this paper proposes a joint network model based on Glove model. CapsuleNets complete the vectorization of information. Attention mechanism assigns weight to features.

2. Related work
Deep learning method can automatically select text features, and has gradually become the mainstream method in the field of text sentiment analysis. Collobert et al. [4] first proposed the use of CNN (convolutional neural networks). Kim [5] used CNN for sentence modeling to solve the problem of sentiment classification. DOS Santo et al. [6] used depth CNN to analysis short text sentiment. Capsule Networks is an improvement of CNN. It uses neuron vector instead of single neuron node in CNN, which can effectively encode text information and save multi-level text features. Attention mechanism can assign different weights according to the importance of features, the model can extract features selectively. Bahdanau et al. [7] used attention mechanism and cyclic neural network for the first time in machine translation tasks. Wang et al. [8] proposed a circular neural network based on attention mechanism, which can extract more text features. These methods proved that the introduction of attention mechanism in CNN can improve the accuracy of text sentiment analysis.
In this paper, we use Glove to reduce the dimension of text features in the feature representation stage. CapsuleNets focus on the relationship between the local and the global, reduce the information loss of convolutional neural network in the pooling layer, and use the Attention mechanism to assign different weights to the semantic coding of text vector, capture the key information and improve the accuracy of sentiment analysis.

3. Our G-CN-ATT Model

This paper proposes a network model based on G-CN-ATT mechanism. The structure of the model is shown in Figure 1. The main process of the experiment as follows: firstly, the experimental data set is obtained, and the word vector is obtained by pretraining with the Glove model, which is used as the input of the capsule network; secondly, dim is obtained from the text through the double-layer capsule network. Then, attention mechanism is added, which assigns different weights to the features, so that the model can learn the importance of the features; finally, the classification is realized through the classification layer.

| Layer (type)   | Output Shape | Param # | Connected to          |
|----------------|--------------|---------|-----------------------|
| input_1 (InputLayer) | (None, 400)  | 0       | input_1[0][0]         |
| embedding_1 (Embedding)     | (None, 100, 500) | 1000000 | input_1[0][0]         |
| spatial_dropout_1 (SpatialDropout) | (None, 100, 500) | 0       | embedding_1[0][0]     |
| capsule_1 (Capsule)        | (None, 10, 10) | 8000    | spatial_dropout_1[0][0] |
| capsule_2 (Capsule)        | (None, 10, 32) | 16000   | capsule_1[0][0]       |
| concatenate_1 (Concatenate) | (None, 10, 48) | 0       | capsule_1[0][0]       |
| attention_layer_1 (Attention)   | (None, 48)   | 2332    | concatenate_1[0][0]   |
| dropout_1 (Dropout)        | (None, 48)   | 0       | attention_layer_1[0][0] |
| dense_1 (Dense)            | (None, 2)    | 98      | dropout_1[0][0]       |

Figure 1. G-CN-ATT Model

Figure 2. Structure diagram of G-CN-ATT
3.1. Word vector

Glove (Global vectors for word representation, global vector) [9] is a word representation tool based on global word frequency statistics. Its purpose is to realize the vectorization of words, so that the vectors contain semantic and grammatical information as much as possible.

The implementation of Glove is divided into three steps:
1. Construct co-occurrence matrix according to corpus.
2. The approximate relationship between word vector and co-occurrence matrix is constructed.
3. Build the loss function.

3.2. CapsuleNets

CapsuleNets[10] is a new neural network structure proposed by Geoffrey Hinton, in order to solve some defects of convolution neural network. The difference between CapsuleNets and CNN is that the neuron of capsule network is a vector rather than a scalar, and CapsuleNets neuron is called vector neuron.

1. Network connection mode. Each capsule unit in the previous layer of CapsuleNets is connected to each capsule unit in the next layer.
2. Weight update. In CapsuleNets, each capsule neuron is a vector, that is, it contains multiple values, so the weight W of each capsule neuron is also a vector. W is updated based on back propagation.
3. Network input. The input S of the capsule network is equal to the linear summation plus the coupling coefficient C. U is the output of the upper layer of the capsule network, W is the weight of each output to be multiplied, and the initial value of b is 0. In the process of seeking S by forward propagation, W is designed as a random value, B is initialized to 0 to get C, and u is the output of the upper layer of capsule network, and the input of the next layer can be obtained according to the following formulas.

\[
S_j = \sum c_{ij} \hat{u}_{ji} \quad (1)
\]

\[
\hat{u}_{ji} = W_{ij} u_i \quad (2)
\]

\[
c_{ij} = \frac{\exp(b_{ij})}{\sum \exp(b_{ik})} \quad (3)
\]

\[
b_{ij} \leftarrow b_{ij} + \hat{u}_{ji} v_j \quad (4)
\]

4. Activation function. Squash function not only preserves the direction of input vector, but also compresses the modulus of input vector between [0, 1]. It means: using the size of the vector module to measure the probability of an entity, the greater the modulus value, the greater the probability.

\[
v_j = \frac{|S_j|^2}{1+|S_j|^2} \cdot \frac{S_j}{|S_j|} \quad (5)
\]
3.3. Attention mechanism
At the beginning, attention and click attention of multi-layer perceptron were both used in the aspect of input and output sequence. In this paper, we add the attention mechanism to our model, which enables the model to allocate different weights according to the importance of feature attributes, so that the model can selectively extract features.

\[ u_i = \tanh(w_i h_i + b_w) \]  

(6)

\[ p_i = \text{softmax}(u_i) \]  

(7)

\[ H = \sum_i p_i h_i \]  

(8)

3.4. G-CN-ATT model
In this paper, the G-CN-ATT network model is proposed to complete the emotion analysis task on the IMDB movie review data set. Firstly, preprocess the data to complete the feature representation of the text. Word vectorization is carried out by using Glove model. Words are used as indexes and vectors are stored in embeddings_index as values, the feature representation of the text is calculated. Using Glove model can ensure more semantic information between word vectors. Then, the CN-ATT network model is used to extract the local features with the weight assigned, and the output of the word vector matrix is used as the input of the double-layer capsule network and the two capsule networks with capsule numbers of 16 and 32. Through the calculation of main capsule layer, squash function, dynamic routing protocol and classified capsule layer, then the local eigenvectors of the two parts are merged by concatenate function. After that, the attention mechanism is added. By giving different weights to the features, the ability of the model to focus on different positions can be extended, so that the model can pay attention to more critical semantic information. Finally, the deny layer is added, and the activation function is sigmoid function.

4. Result

4.1. Experimental environment configuration
The programming environment of this experiment is jupyter notebook, on which the third-party library is installed, including tensorflow framework and keras framework. The virtual environment is Python 3.5 version.

4.2. IMDB Data
The data set in this paper is from IMDB movie reviews. Before input, we preprocess the data, including deleting HTML tags, line breaks, adjusting data format, and completing data preprocessing. Test set and verification set are randomly assigned. In this paper, it is divided into two categories: negative is 0, positive is 1. There are 22500 training sets and 2500 verification sets, with a total of 25000.

4.3. Comparative experiment
The experiment set up the following groups of comparative experiments, the experimental data are IMDB film reviews.

(1) G-C: using Glove to represent the semantic information of each word as a vector composed of real values, using CNN network to complete the classification.

(2) G-CN: using the word representation tool of Glove model to reveal the semantic information of words by modeling the context relationship of words, and extract the local feature vector of text by double-layer CapsuleNets.

(3) G-C-ATT: the word vector is obtained by using Glove model, the local feature of text is extracted by CNN, and attention mechanism is added after CNN.
(4) G-CN-ATT: using Glove model to represent word vector, using double-layer CapsuleNets to obtain local feature matrix, then adding attention mechanism to achieve text sentiment classification.

(5) G-BiGRU: words are transformed into word vectors by Glove model, and then classified by BiGRU network.

(6) G-BiLSTM: using Glove model to train word vector as the input of BiLSTM network to complete the training task.

4.4. Analysis of experimental results

Through the experimental data, we find that the accuracy of this model is relatively high. The accuracy of G-CN is 0.0028 lower than our model, which indicates that attention can pay more attention to the more important word information in the sentences and reduce the weight of unimportant information in the process of feature extraction. The accuracy of G-C-ATT network is 0.0026 lower than our model, because CNN ignores the connection between words. Capsule can extract the text information that missed by CNN and get good experimental results. The results of G-BiLSTM model are the worst because of its complexity and many parameters. The result of G-BiGRU model is better than G-BiGRU because GRU replaces the forgetting gate and input gate in LSTM with updates, which shortens the training time. The val-acc of the model proposed in this paper reaches 0.8912 on IMDB dataset, which is 0.0028, 0.0026, 0.0324, 0.1074 and 0.1900 higher than other models. From the above analysis, it shows that the network structure proposed in this paper has a good classification effect and can realize the task of emotion analysis.

### Table 1. Experimental Results

| Model       | val-acc | time(s) | val-loss |
|-------------|---------|---------|----------|
| G-CN-ATT    | 0.8912  | 546     | 0.2787   |
| G-CN        | 0.8884  | 301     | 0.3291   |
| G-C-Att     | 0.8886  | 180     | 0.3222   |
| G-C         | 0.8588  | 236     | 0.3623   |
| G-BiGRU     | 0.7838  | 1657    | 0.5032   |
| G-BiLSTM    | 0.6978  | 5097    | 0.5188   |

Before that, we proposed a joint network model based on BiGRU-CN (i.e. the combination of BiGRU and double-layer CapsuleNets). Under the same experimental environment, BiGRU-CN network achieves good real classification results. However, due to the complexity of BiGRU network and the number of network parameters, the training time is longer and the execution efficiency is lower. In this paper, G-CN-ATT network is used to assign different weights according to the importance of words in the text. This model is simpler than BiGRU-CN network, and the training time and execution efficiency have been greatly improved.

### Table 2. BiGRU-CN and G-CN-ATT

| Model       | val-acc | time(s) | val-loss |
|-------------|---------|---------|----------|
| G-CN-ATT    | 0.8912  | 546     | 0.3291   |
| BiGRU-CN    | 0.8950  | 2930    | 0.2647   |

5. Conclusion

This paper proposes a network structure which combines CapsuleNets and Attention mechanism. Experiments show that this kind of hybrid network has good feature extraction ability and can be used for text sentiment analysis. In the future, the research work of this paper is to try to integrate the structure of other models, optimize the text sentiment analysis model, and improve the accuracy of model classification.
Acknowledgment

This work was supported by the Wuhan Institute of Technology Software engineering laboratory and the eleventh Graduate Innovation Fund of Wuhan Institute of Technology (CX2019235, CX2019236, CX2019237).

Reference

[1] Tang, H.F, Tan, S.B, Cheng, X.Q. (2007) A comparative study of Chinese emotion classification techniques based on supervised learning [J]. Chinese Journal of information technology, (06):88-94+108.

[2] Zeng K, Ding S.F, Jia W.K. (2018) Single image super resolution using a polymorphic parallel CNN[J]. Applied Intelligence.

[3] Zhang N, Ding S, Zhang J, et al. (2018) An overview on restricted boltzmann machines [J]. Neurocomputing.

[4] Collobert R, Weston J, Bottou, Léon, et al. (2011) Natural Language Processing (almost) from Scratch[J]. Journal of Machine Learning Research, 12(1):2493-2537.

[5] Kim Y. (2014) Convolutional Neural Networks for Sentence Classification [J]. Eprint Arxiv.

[6] Dos Santos C N, De Bayser M G. (2014) Deep convolutional neural networks for sentiment analysis of short texts[C] Proc of the 25th International Conference on Computational Linguistics. [S.l], Association for Computational Linguistics, 69-78.

[7] Bahdanau D, Cho K, Bengio Y, et al. (2015) Neural machine translation by jointly learning to align and translate[C] 3rd ICLR 2015, San Diego, CA, USA.

[8] Wang Y, Huang M, Zhao L, et al. (2016) Attention-based LSTM for aspect-level sentiment classification[C] EMNLP 2016, Austin, Texas, USA, 606-615.

[9] Pennington J, Socher R, Manning C. (1975) Glove: Global Vectors for Word Representation[C] Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014. Salton G, Wong A, Yang C S. A vector space model for automatic indexing[J]. Communications of the ACM, 18(11): 613-620.

[10] Sabour S, Frosst N, Hinton G E. (2017) Dynamic Routing Between Capsules[J].