Chinese text sentiment classification method based on dual-channel attention capsule network

Chunjia Liu [1], Xinyu Zhang [1,*], Juan Li [1], Jingwen Chen [1] and Chang Li [1]

[1] Jiangsu Emergency Early Warning Release Center, nanjing, China

* Corresponding Author Email: 786564119@qq.com

Abstract. In recent years, text sentiment analysis has received close attention from researchers. The CNN-based neural network is frequently used in this field. Capsule network can overcome the problem of information loss caused by pooling layer in CNN-based neural network, but the capsule network cannot selectively focus on important words in the text. To solve this problem, this paper proposes a text sentiment analysis method based on the attention mechanism of the capsule network. First, use the cw2vec model and the SAT model to train word vectors and use them as the input of the two channels of the method, then combine the attention mechanism and the convolutional neural network to extract the characteristics of the two channels, and then input the capsule network to realize the emotion classification. The static routing mechanism is also used to reduce the complexity of the model, and it has a higher accuracy than the dynamic routing mechanism. We verified the effectiveness of the model on the Tan Songbo hotel review data set and we further crawls a large amount of public opinion data related to Meteorology and early warning from Weibo to form different fields. the model is also better than many sentiment analysis methods.

Keywords: convolutional neural network; deep learning; word vector; capsule network; sentiment analysis.

1. Introduction

In recent years, deep learning methods have achieved good results in text sentiment analysis tasks. However, most traditional deep learning methods use a single word embeddingto train the model, which results in the model relying heavily on the initial value of the input vector. On the other hand, in recent years, capsule networks have also exhibited excellent performance characteristics. CNN-based neural networks usually use pooling to extract data features but cause information loss, while capsule networks use capsule vectors and routing mechanisms to overcome this problem, and can also express the positional hierarchical relationship between local features, but capsule networks can not be able to selectively focus on important words in the text. In response to the above problems, on the basis of fully considering the particularity of the Chinese language, we use two models cw2vec [1] and SAT [2] to train two different word vectors. Cw2vec can capture the morphological feature information of Chinese characters, and SAT can make full use of the existing Chinese HowNet knowledge base.Then we input these two word vectors into the capsule network which use a static routing mechanism, and attention mechanism is introduced to make the key information in the sentence easier to extract.

2. Text sentiment analysis method based on attention capsule network

Because CNN and capsule networks can successfully detect and extract local features with strong classification and representation capabilities [3]. Yang et al. [4] proposed a text classification model based on CapsNet (Capsule A, Capsule B), and prove that the classification effect is better than the CNN and LSTM-based models.

On the one hand, capsule networks use capsule vectors instead of neurons, which can indicate the hierarchical relationship of positions between local features. On the other hand, the capsule network uses a routing mechanism instead of max/average pooling, thereby reducing information loss to some extent [5]. So this article uses a capsule network to replace the pooling layer.

In the field of language, the expression of emotions in documents has great freedom. For example, in the original capsule network, learning to correctly represent the position features of the eyes, nose, and mouth when classifying faces in images is a major challenge. However, as far as documents are concerned, it cannot be said that the order of sentences in the two documents is different, that is, the two documents are absolutely different. In view of the above problems, Jaeyoung Kim et al. proposed a static routing mechanism, which effectively reduced the computational complexity of dynamic routing and improved the text classification performance [6]. Based on the above work, we propose a new emotion classification model based on the attention capsule network ATT-CapsNet. As shown in Fig.1, the model is mainly composed of five parts, as follows.

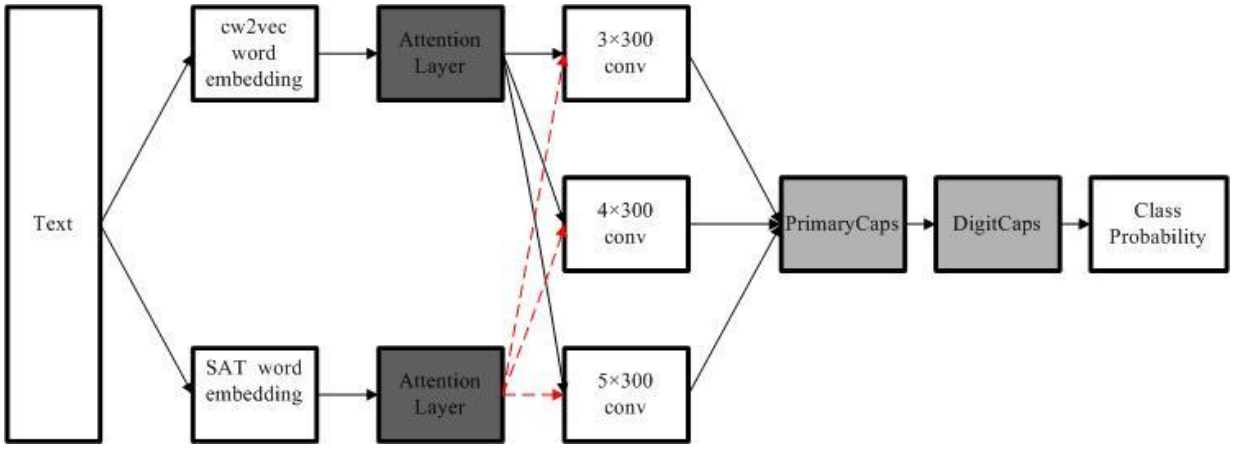

Fig.1 A method of text sentiment classification based on attention capsule network

2.1. Word embedding layer

We uses the word vectors trained by the cw2vec model and the SAT model.

2.2. Attention layer

By calculating the correlation coefficient of each word in the sentence and other words in the classification task, the context vector of the word is constructed, and the word vector corresponding to the word is connected as an input, which can make Words with a larger coefficient of relationship with other words in the sentence get more attention, and these words are often important information for classification. Fig.2 shows the core structure of the word vector attention mechanism.

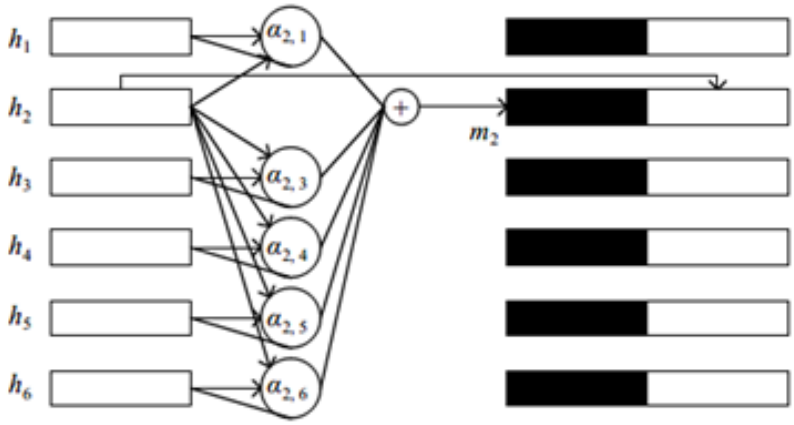

Fig. 2 The core structure of the word vector attention mechanism

2.3. Convolutional layer

In this paper, convolution kernels with heights of 3, 4, and 5 are used to convolve the inputs of two different channels, thereby extracting the local features of the inputs of different channels.

2.4. Initial capsule layer

In the initial capsule, the vector output capsule is used instead of CNN's scalar output feature detector to save the instantiated parameters.

Suppose $p_i \in R^d$ represents the instantiated parameter set of the capsule, and d is the dimension of the capsule. $M^b \epsilon W^{k_2 \times d}$ is to initialize the capsule layer filter, each N-gram vector is expressed as $S_i \in M$, the convolution kernel M^b and N-gram vector are calculated, and the window size is 1, The step size is 2, resulting in a list of capsules $p \in R^k$. Among them, each capsule $p_i \in R^k$ is obtained by equation (1).

$$p_i = g(M^b S_i + b_i) \tag{1}$$

$$P = [p_1, \ p_2 \ldots p_C] \tag{2}$$

We use two different routing methods, as shown in the following two sections.

2.5. Capsule network with dynamic routing

The capsule network updates the weight of the coupling coefficient through the iterative routing process, and determines the degree to which the initial capsule layer points to the classified capsule. The coupling coefficient represents the probability that the vector output from the capsule i of the initial capsule layer belongs to the capsule j of the classification capsule layer, which is determined by the iterative dynamic routing algorithm, that is, it is updated and adjusted in the forward training of the model.

$$C_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \tag{3}$$

Where $i \in [1, a]$, $j \in [1,k]$, and k is the number of categories. C_{ij} is the coupling coefficient and is updated in every routing iteration. The b_{ij} is determined by the similarity between the lower capsule and the higher capsule. $\hat{h}_{j|i}$ is represented by the matrix operation between the weight matrix W_{ij} and h_i.

2.6. Capsule network with static routing

For the field of images, it is reasonable to consider the spatial hierarchy of entities, and dynamic routing can identify objects similar to the way we humans recognize objects. However, in the field of language, the expression of emotions in documents has great freedom. From this perspective, we consider using the following static routing scheme:

$$s_j = \sum_i W_{ij} h_i \tag{4}$$

$$v_j = \frac{\| s_j \|^2}{1 + \| s_j \|^2} \frac{s_j}{\| s_j \|} \tag{5}$$

$W_{ij} = [M \times N]$ is a weight matrix, where n is the number of capsules in hi. W_{ij} multiplied by h_i indicates that the upper layer entity is a capsule of M-dimensional vector. v_j is the result of compressing s_j and represents the text capsule layer.

3. Experiments

3.1. Experimental data

In this paper, we used public data set which called Tan Songbo hotel review data set (ChnSentiCorp) in the experiment, And we further crawls a large amount of public opinion data related to Meteorology and early warning from Weibo to form different fields. The microblog text

public opinion data set has a total of 10,000 items and is marked with emotional polarity. The positive text has 4,785 and the negative text has 5,215. Examples are shown in Table 1.

Table.1 Sample of Weibo text public opinion dataset

Positive	Negative
#Suspected tornadoes appear in many areas of Jiangsu # Thank you for your good luck. The tornado has bypassed our village and trees have fallen around it	Just been hit by a leaf on the tree, the pain is unbearable
Hurry up and rain, so as not to wash the car	This weather is a bit scary, it changes too quickly

3.2. Parameter setting

The sliding window size of the convolution kernel is set to 3, 4, 5, the number of convolution kernels of each size is set to 100, and the Dropout parameter is set to 0.5, iteration The number of times is set to 10, and the activation function uses the ReLU function. In the training process of the convolutional capsule network model, the batch size batch_size is set to 25, and the Adam optimization algorithm with an initial value of 0.9 and a learning rate of 1e-3 is used to train the model. The capsule dimension is 16 dimensions and the number of capsules is 6.

3.3. Comparative experiment

In order to compare and verify the effects of various text sentiment classification models, we also use the following methods to conduct experiments on this data set.

1) TextCNN-multichannel: Use the TextCNN model proposed by Kim for emotion classification, and the parameter settings are consistent with the original paper.

2) AT-DouCNN [7]: A dual-path convolutional neural network sentence classification model that uses the word vector attention mechanism. This model combines the attention mechanism and the convolutional neural network to obtain words using two methods, word2vec and Glove.

3) CS-DouCNN: A two-way convolutional neural network sentence classification model that uses the word embedding attention mechanism. On the basis of AT-DouCNN, the attention mechanism is removed, and cw2vec and SAT are used.

4) CS-AT-DouCNN: On the basis of CS-DouCNN, an attention layer is added after the word embedding layer.

5) Capsule-A: Capsule A, a text classification model based on CapsNet proposed by Yang et al. [4], with parameter settings as in the original paper.

6) Capsule-B: Capsule A, same as Capsule-A.

7) CS-CapsNet dynamic routing: based on CS-DouCNN, without attention mechanism, input dual channels, extract features through convolution and enter the capsule network, between the first capsule layer and the second capsule layer The routing mechanism selects dynamic routing.

8) CS-AT-CapsNet dynamic routing: based on CS-ATT-DouCNN, using the word embedding attention mechanism, and then extracting features through convolution, then input into the capsule network, from the first capsule layer to the second capsule layer The dynamic routing mechanism chooses dynamic routing.

9) CS-AT-CapsNet static routing: The routing mechanism between the first capsule layer and the second capsule layer selects a static route.

Table.2 Sentiment classification effects of different methods

ID	Methods	Weibo text public opinion dataset	ChnSentiCorp
1	CNN-multichannel	90.37%	89.17%
2	AT-DouCNN	91.55%	91.33%
3	CS-DouCNN	91.52%	91.67%
4	CS-AT-DouCNN	92.20%	91.93%
5	Capsule-A	90.44%	90.67%
6	Capsule-B	91.62%	91.33%
7	CS-CapsNet (dynamic routing)	92.09%	91.50%
8	CS-AT-CapsNet (dynamic routing)	93.06%	92.17%
9	CS-AT-CapsNet (static routing)	93.23%	92.36%

It can be observed from the experimental results that, as shown in Table 2, compared with the baseline method, the two capsule network models based on the attention mechanism proposed in this paper have achieved good results. Among them, the CS-AT-CapsNet static routing in this paper is the most Excellent, the accuracy rate reached 93.23% on the Tan Songbo hotel review data set, and 93.02% on the Weibo text public opinion data set.

Compared with the methods 1 and 3 that only use the dual-channel model without the attention mechanism, the CS-DouCNN method in this paper is better than the CNN-multichannel method. Then compare the methods 2 and 4 that also use the dual-channel and attention mechanism. The CS-AT-DouCNN method is improved by 0.6% and 0.65% on the two data sets compared to the AT-DouCNN method. The difference between the two methods is whether to use the cw2vec and SAT models to train word vectors, while the AT-DouCNN method uses the word vectors trained by the traditional word2vec and glove models, which illustrates the superiority of the word embedding selection.

In this paper, the four methods using the attention mechanism are superior to the methods that do not use attention. Compare between CS-AT-DouCNN and CS-DouCNN, we can see that after adding the attention mechanism, The accuracy rate is improved by 0.25%, indicating that the introduction of attention mechanism makes the key information in the sentence under the sentiment classification task easier to be extracted, thereby obtaining more accurate sentence features.

Compared with CS-AT-DouCNN, CS-AT-CapsNet dynamic routing and static routing methods do not use the pooling layer, but use the capsule network, which shows that the capsule network does reduce the Information loss problem; compared with the CS-CapsNet dynamic routing without attention mechanism, the CS-AT-CapsNet dynamic routing method has improved accuracy. This is due to the fact that the capsule network cannot selectively focus on the text. The problem of important words has been improved very well when the attention mechanism is added. The static routing method has improved accuracy than the dynamic routing method, which verifies the effectiveness of the static routing method.

4. Summary

We propose a text sentiment analysis method based on the attention capsule network. A static route is also introduced. The mechanism makes the capsule network flexible and robust enough to handle subtle modifications in the text, enhancing the applicability of the capsule network to the text. Through experiments on the Tan Songbo hotel review data set andWeibo text public opinion dataset, it was verified that the use of capsule networks is indeed superior to the traditional convolution pooling layer, which can retain more features and also verify the effectiveness of the attention mechanism and proves that the use of static routing mechanisms in the capsule network is indeed superior to dynamic routing mechanisms in text sentiment analysis tasks. As a result, the training accuracy of this method reaches the best of all methods.

References

[1] Cao, S., Lu, W., Zhou, J., & Li, X cw2vec: Learning chinese word embeddings with stroke n-gram information. In Thirty-Second AAAI Conference on Artificial Intelligence. 2018.

[2] Niu, Y., Xie, R., Liu, Z., & Sun, M Improved word representation learning with sememes. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2017, Vol. 1, pp 2049-58.

[3] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules[C]//Advances in neural information processing systems. 2017: 3856-3866.

[4] YANG M , ZHAO W , YEJB , etal.Investigating capsule networks with dynamic routing for text classification [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels , Belgium: ACL ,2018 : 3110 3119.

[5] Yujia Wu, Jing Li, Jia Wu, Jun Chang, Siamese capsule networks with global and local features for text classification, Neurocomputing,2020

[6] Jaeyoung Kim, Sion Jang, Eunjeong Park, Sungchul Choi,Text classification using capsules, Neurocomputing, Volume 376, 2020, Pages 214-221

[7] GUO Bao-zhen,ZUO Wan-li,WANG Ying. Double CNN sentence classification model with attention mechanism of word embeddings [J]. Journal of Zhejiang University (Engineering Science),2018,52 (09): 1729-1737.