Advances in Artificial Intelligence, Big Data and Algorithms
G. Grigoras and P. Lorenz (Eds.)

doi:10.3233/FAIA230888

A Multi-Dimensional Text Sentiment Analysis Method Based on Joint Network

Tao CAO, Na LIU[1], Shuchen BAI
School of Information Science and Engineering, Dalian Polytechnic University, Dalian, 116034, China
ORCiD ID: Tao CAO https://orcid.org/0009-0008-7284-8875
ORCiD ID: Shuchen BAI https://orcid.org/0009-0007-5271-491X

Abstract. Text sentiment analysis in social media has problems such as irregular structure, short length and sparse features. In this paper, a text sentiment analysis method combining emotional symbols is proposed. Based on the BiGRU and capsule network joint network model, this method fully considers the influence of emotional emoji in the text to be analyzed on the sentiment analysis tendency. Secondly, the BiGRU network was used to extract the long-term dependent features of the text context, and the capsule network was used to deal with the problem of losing feature information in the CNN pooling layer to better extract the local features of the text. Finally, the Softmax classifier was used to output the sentiment tendency. Experimental results show that the proposed model is superior to the current mainstream models in accuracy, recall rate and F1 value.

Keywords. Capsule network, sentiment analysis, BiGRU, CNN

1. Related Work

With the continuous development of sentiment analysis technology methods, algorithms and resources, existing research is constantly moving closer to deep learning methods. However, the majority of scholars are more inclined to optimize and improve the model, ignoring the impact of emojis on sentiment analysis. According to the survey report, more and more post-00s express their views through social platforms, and more post-00s are satisfied with the use of emojis. At present, a large number of emojis appear in social media comments, such as😭, 😂etc.It makes the emotion expressed richer and more real. The use of emojis can reveal the emotions hidden beneath the text, for example, "This car I want to buy is worth $350,000"😭; "I slept for 20 hours😱" expresses surprise and horror. In these sentences, although the words do not explicitly convey the emotion, the emoji at the end can effectively identify the emotional state of the individual.

Capsule Network (CapsNet) model was improved on the basis of Convolutional Neural Network (CNN) model [1-2]. It replaced the pooling operation in CNN with dynamic routing, solved the problem of information loss caused by pooling operation, and could well extract the position semantic information of emotional words in the full text. Yang et al. [3] combined the BERT model with the capsule network and proposed an enhanced capsule network to accurately feedback the real word-of-mouth of the movie

[1] Corresponding author: Na LIU, School of Information Science and Engineering, Dalian Polytechnic University; e-mail: liuna@dlpu.edu.cn

through the user's comments in the sentiment analysis based on social media comments. The BiGRU model is composed of two GRU models in opposite directions superimposed up and down [4], and the single GRU model can only obtain the one-way above or below information of the text [5]. In this paper, BiGRU is used instead of GRU to better capture the bidirectional semantic dependence of the text. Basiri et al. [6] proposed an attention-based bidirectional CNN-RNN deep model (ABCDM). The model uses independent BiLSTM layer and BiGRU layer to mine context information in the time dimension [7-8], and uses the attention mechanism to act on the output of the double layer. The ABCDM model has achieved good experimental results on the sentiment analysis tendency of long and short text reviews respectively

In summary, for the existing text sentiment analysis method text information expression is not sufficient, can not extract text context information and local semantic features at the same time and other issues, this paper proposes an improved capsule network BiGUR-CapsNet hybrid model for text sentiment analysis method, the main contributions of this paper are as follows:

(1) In this model, the commonly used emojis in microblog are replaced by corresponding words, and the text and Emoji are constructed as a multi-dimensional word vector matrix by Word2Vec.

(2) Taking the obtained vector matrix as input, the BiGRU model is used to extract the text context information, and the capsule network model is used to further extract the semantic information features of the text position.

(3) The SoftMax classifier is used to classify the sentiment orientation of Chinese microblog comments.

2. Multi-dimensional Sentiment Classification Model Combining Emoji and Text

The bigRU-Capsule model consists of five layers, namely, BiGRU layer, N-gram convolutional layer, main CAPSULE layer, convolutional capsule layer, and fully connected capsule layer. The model structure is shown in Figure 1.

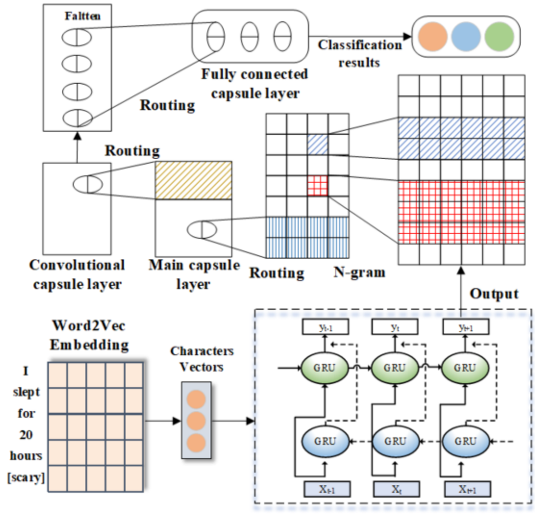


Figure 1. BiGRU-CAPSULE model structure

BiGRU layer:The BiGRU layer is used to capture bidirectional semantic dependencies in the text to enhance generalization to new datasets. The BiGRU network makes up for the inability of GRU to encode information from back to forward. In BiGRU network, there are two routes for information transmission, so it can extract richer context information and has stronger learning ability. The calculations are shown in equations (1)-(3).

$$\overrightarrow{h_t} = GRU(x_t, \vec{h}_{t-1}) \tag{1}$$

$$\overrightarrow{h_t} = GRU(x_t, \vec{h}_{t-1}) \tag{2}$$

$$h_t = f(W_{\vec{h}_t} \overrightarrow{h_t} + W_{\overleftarrow{h}_t} \overleftarrow{h_t} + b_t) \tag{3}$$

Where: $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are the states of the forward $W_{\vec{h}_t}$ and $W_{\overleftarrow{h}_t}$ backward hidden layers at time *t*, respectively; And, are the weights of the forward and backward hidden layer states at time *t*, respectively. b_t is the offset of the hidden layer state at time *t*.

N-gram convolutional layer: This layer extracts features with multiple size convolution kernels. The output $Q = [q1,q2,...,qL]$ of the BiGRU layer is used as the input, the *j* th row to *j+h-1* row of *Q* is taken as a local continuous window $q_{j:j+h-1} \in \mathrm{R}^{h \times k}$, and the convolution filter $w_{cf} \in \mathrm{R}^{h \times k}$ is applied to the window to form a new feature m_{ij}. See Equation (4) for calculation.

$$m_{ij} = f(w_{cf} \circ y_{j:j+h-1} + b) \tag{4}$$

Where $f(\cdot)$ represents the nonlinear activation function, $\circ$ which represents the multiplication of the corresponding positions of the two matrices. *b* is the bias vector. For a combination of *B* convolution kernels containing *C* kernels of different sizes, the obtained output is shown in Equation (5).

$$M = [(m_{11}, m_{22}, ..., m_{1C}), ..., (m_{b1}, m_{b2}, ..., m_{bC})] \tag{5}$$

Main Capsule Layer: The main capsule layer is used to replace the output of the convolution operation from a scalar output to a vector output. Firstly, the feature vector m_i of the sliding window of each N-gram is transformed into the appropriate feature capsule u_i, as shown in Equation (6).

$$u_i = g(W^b m_i + b_1) \tag{6}$$

Where W^b is the filter shared by all N-gram vectors, *g* is a nonlinear function, and b_1 is the capsule bias.

Convolutional capsule layer: Used to calculate the relationship between child capsules and parent capsules, then dynamic routing is used to calculate the parent capsule in the upper layer. The prediction vector is $\hat{u}_{j|i} \in R^d$ calculated as shown in Equation (7).

The parent capsule S^j is acquired by the dynamic routing algorithm, which is computed as in Equations (8)-(11). Equation (9) is the compression function that compresses the input vector module to [0,1).

$$\hat{u}_{j|i} = W^{t_1} u_i + \hat{b}_{j|i} \in R^d \tag{7}$$

$$c_{ij} = \hat{a}_{j|i} \cdot softmax(b_{j|i}) \tag{8}$$

$$S^j = \sum_i c_{ij} \hat{u}_{j|i} \tag{9}$$

$$v_j = \frac{S}{\|S\|} \frac{\|S\|^2}{1 + \|S\|^2}, a_j = |v_j| \tag{10}$$

$$b_{j|i} = b_{j|i} + \hat{u}_{j|i} \cdot v_j \tag{11}$$

Here, $W^{t1} \in \mathrm{R}^{E\times d\times d}$ represents the shared weight, E represents the count of parent capsules in the previous layer, and u_i is the child capsule in the next layer and $\hat{b}_{j|i}$ represents the bias term. Where C_{ij} is the weight of the i-th child capsule to the j-th parent capsule, which can be obtained from the given prediction vector $\hat{u}_{j|i}$ and its existence probability $\hat{a}_{j|i}$.

Fully connected capsule layer: The upper capsule flattens, becomes the capsule list V, and is entered. The capsules are multiplied with the transformation matrix to generate the final capsule $v`_j \in \mathrm{R}^d$ corresponding to each category. The probability y_j for each class is then obtained based on v_j, which is calculated as in Equations (12) and (13).

$$V' = W^V \cdot V \tag{12}$$

$$y_j = \|v'_j\| \tag{13}$$

Where $W^V \in \mathrm{R}^{e\times r}$, e is the amount of classification and r is the amount of capsules in the capsule list V.

3. Experimental Results and Analysis

3.1. Dataset

In traditional sentiment analysis experiments, emotifications will be removed as a kind of text noise. However, the length of microblog text is small, and the proportion of emotional symbols is large, so removing emotifications will directly reduce the sentiment classification effect of sentiment analysis experiments. More than 2w microblog statements with Emoji are selected as the experimental data set. Split it into 8:2 training and test sets. The detailed statistical results of the dataset can be seen in Table 1.

Table 1. Detailed statistics of the dataset

Dataset	Positivity	Neutral	Negativity
Training set	5768	5731	4500
Validation set	1443	1433	1125
Total	7211	7164	5625

Reasonable use of emotional symbols can enhance the emotional categories of the text and enrich the content of the text. In order to facilitate the use of emotifications, some emotifications are transformed into words for processing, as shown in Table 2. The description is convenient, and the sentiment is divided into positive and negative.

Table 2. Partial translation of sentiment symbols

Forward emoji translation	Negative emoji translation
[ha ha][clapping][proud][laughing][clapping] [heart][power][praise][lovely][mighty]	[bad][despised][sad][obscene][disgusting] [tears][collapse][anger][curse][fear]

3.2. Model Parameters and Evaluation Metrics

This experiment is completed based on the Keras framework and python3.6 environment. For the network model in this paper, the activation parameters are selected sigmoid function, the output layer is selected softmax for classification, dropout is used to prevent training overfitting, the loss function is binary_crossentropy, and Adam is selected as the optimizer.

In purpose of evaluating the properties of various models, this paper adopts the evaluation indicators of Precision, Recall and F1-score (F1), and the evaluation formula is (15) - (16).

$$Precision = \frac{T_p}{T_P + F_P} \tag{14}$$

$$Recall = \frac{T_P}{T_P + F_N} \tag{15}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{16}$$

3.3. Analysis of experimental results

The experiment was conducted in two parts. One part filtered out useless punctuation marks and special characters such as emoticons. Another part filters out useless punctuation marks and special characters outside emojis in the text. So as to verify the influence of emojis on sentiment classification performance.

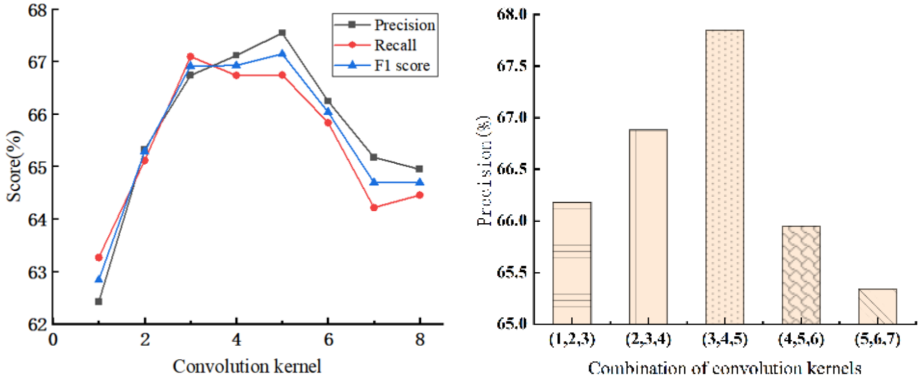
Table 3. Experimental results of different models

Discriminative model	Text			Text+Emoji		
	Precision	Recall	F1 Score	Precision	Recall	F1 Score
SVM	41.32	44.27	42.74	63.22	61.74	62.47
CNN	47.25	49.36	48.31	68.33	68.26	68.30
BiLSTM	52.36	50.23	51.27	71.44	63.84	67.43
BiGRU	55.51	54.30	54.90	74.36	75.35	74.85
BiGRU+CapsNet	63.67	68.33	65.92	85.32	84.56	84.94

From Table 3, it can be seen that in the prediction of social text containing Emoji, there are obvious advantages compared to the method using only plain text. This is because short text reviews are mostly phrases with irregular grammar and length, which cannot fully extract features in the training process, resulting in large errors. As a result, the accuracy of sentiment analysis of social text considering emojis is significantly improved.

3.4. Ablation Experiments and Analysis

In this part, the influence of the size of the convolution kernel on the results is first tested, and the convolution kernel window size is set to 1-10. Then, in order to obtain higher prediction accuracy, the influence of different combination of convolution kernels with different window sizes on the prediction results is tested. The results are s-
hown in Figs. 2(a) and 2(b).



(a)Same convolution kernel window size (b)Different combination of convolution kernel window sizes

Figure 2. The effect of convolution kernel size on Capsule

As can be seen in Fig. 2(a) that as the convolution window continues to increase, the accuracy increases first and then decreases. When the convolution kernel window size is set to 5, the prediction accuracy is the highest. Figure 2(b) shows the prediction accuracy when using different combinations of convolution kernels of different sizes, and it can be seen that when the combination of convolution kernel window sizes is (3,4,5), the capsule network obtains higher prediction accuracy.

4. Concluding Remarks

In this paper, emoji and short text information are fused in the vector representation stage, and the BiGRU model combined with neural network can effectively extract global text features. The capsule network is used to deal with the problem of losing feature information in the CNN pooling layer, so as to better extract local text features. Taking the respective advantages of the two networks can enrich the text feature information. This method can effectively improve the performance of multi-sentiment classification for text and Emoji. In the next step, the model will be further optimized to shorten the running time while ensuring the accuracy.

References

[1] Mahendhiran, PD. Subramanian, K., CLSA-CapsNet: Dependency based concept level sentiment analysis for text, *JOURNAL OF INTELLIGENT & FUZZY SYSTEMS* **43** (2022), 107–123.

[2] Gao, ZW. Li, ZY. Luo, JY.Li, XL.,Short Text Aspect-Based Sentiment Analysis Based on CNN plus BiGRU,*APPLIED SCIENCES-BASEL*,**12** (2022).

[3] Yang, KH.Liu,J.,Weibo Sentiment Analysis Based on Advanced Capsule Network,*International Symposium on Computer Science and Intelligent Controls(ISCSIC)* (2021), 211–216.

[4] Zhu, QN.Jiang, XF.Ye,RZ.,Sentiment Analysis of Review Text Based on BiGRU-Attention and Hybrid CNN,*IEEE ACCESS* **9** (2021), 149077-149088.

[5] Yu,HL.Ji,YN.Li,QL.,Student sentiment classification model based on GRU neural network and TF-IDF algorithm,*JOURNAL OF INTELLIGENT & FUZZY SYSTEMS* **40** (2021), 2301–2311.

[6] Basiri, ME.Nemati, S. Abdar, M .Cambria, E.Acharya, UR.,ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis,*FUTURE GENERATION COMPUTER SYSTEMS-THE INTERNATIONAL JOURNAL OF ESCIENCE* **115** (2020), 279-294.

[7] Xie, J.Chen, B.Gu, XL.Liang, FM.Xu, XY.,Self-Attention-Based BiLSTM Model for Short Text Fine-Grained Sentiment Classification,*IEEE ACCESS* **7** (2019), 180558-180570.

[8] Hameed, Z.Garcia-Zapirain, B.,Sentiment Classification Using a Single-Layered BiLSTM Model,*IEEE ACCESS* **8** (2020), 73992-74001.