IJLAI Transactions on Science and Engineering (ISSN: 3006-3418) 2024, vol. 2, no. 2, pp: 81–90

English Text Sentiment Analysis Based on Convolutional Neural Network and U-network

Shu Ma[1]

College of International Business, Shenyang Normal University
No. 253 Huanghe North Street, Shenyang 110034, China
mashuetd@163.com
Corresponding author: Shu Ma

Abstract. English text sentiment orientation analysis is a fundamental problem in the field of natural language processing. The traditional word segmentation method can produce ambiguity when dealing with English text. Therefore, this paper proposes a novel English text sentiment analysis based on convolutional neural network and U-network. The proposed method uses a parallel convolution layer to learn the associations and combinations between word vectors. The results are then input into the hierarchical attention network whose basic unit is U-network to determine the affective tendency. The experimental results show that the accuracy of bias classification on the English review dataset reaches 93.45%. Compared with many existing sentiment analysis models, it has more accuracy.

Keywords: English text sentiment, Convolutional neural network, U-network.

1. Introduction

Sentiment analysis, also known as opinion mining, refers to people's feelings, opinions, evaluations, attitudes and emotions about services, products, organizations, individuals, problems, events, topics and their attributes [1]. Text emotion tendency analysis is a branch of emotion analysis, whose purpose is to judge the speaker's emotion tendency to things from the original text. The knowledge-based approach is the first technique to be widely used in this field, but it requires writing very complex rules to make computers understand human language more accurately, which is difficult. Therefore, this kind of method can only achieve certain results on small-scale data [2]. With the increasing amount of text data, it is difficult to use knowledge-based methods to process text. Since the 1990s, machine learning methods have begun to emerge in the field of text sentiment analysis [3-5].

However, these methods belong to the category of shallow learning, and the function model and calculation method are relatively simple, which leads to their inability to express some complex functions under limited samples and calculation units, weak generalization ability, and the need to manually select a large number of data features. These shortcomings cause machine learning methods to hit a bottleneck on this task [6,7].

Deep learning can automatically learn important features and feature expressions from raw data to deal with a variety of complex tasks, and has obvious advantages in modeling, interpretation, expression ability and optimization. Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) are two popular models in deep learning [8,9]. The convolutional neural network can extract the localized structure information from the data, and the recurrent neural network can process the serialized structure information.

In recent years, a composite model combining the structure of the two models has also emerged, which has achieved excellent results in the field of text sentiment analysis. Attention mechanism is the latest achievement in the field of deep learning, which can capture the most representative features of the text and optimize the structure of the model. Using deep learning model to analyze text emotion is a popular research direction at present [10,11].

This paper argues that the emotional tendency of text is determined by the sentence level and the word or character level. First, text is made up of sentences, and different sentences have different levels of importance to the results of affective orientation analysis. For example, if the overall emotional tendency of the text is positive, some sentences with more negative emotional color are not important, and not all sentences will affect the judgment of the final result [12,13]. Similarly, the sentence is composed of words or characters, and different words or characters have different degrees of influence on the judgment of the emotional tendency of the sentence. The existing models seldom explore the emotion of text from this perspective, and fail to reflect the influence of the hierarchy of text structure and the context of text content on the results of tendency analysis. Therefore, this paper establishes a hierarchical affective orientation analysis model, and introduces the attention mechanism to screen out the text information that has the highest impact on the results of orientation analysis from two levels. On the other hand, the representation of word vectors is very important for text classification tasks. In recent years, the research on the granularity of word vectors has become more and more detailed, and some work based on the character level has appeared.

English text is mainly reflected at the character level. English words are made up of 26 letters, and individual letters often have no special meaning. Because of the particularity of English, most English text classification tasks use word segmentation operations. However, when the word segmentation operation is performed, the combination form between letters is fixed, which sometimes leads to ambiguity, and the correct combination form of Chinese characters cannot be divided. In order to solve this problem, this paper uses the parallel convolutional layer of convolutional neural network to learn English text word-level features, without relying on parsing trees and other syntactic analysis methods, and avoids language knowledge level analysis and complex data preprocessing. The experimental results show that for English corpus, using the trained word-level word vector as the original feature is better than using the trained word-level word vector as the original feature.

2. Related Works

The analysis of affective tendency has always been a research hotspot in the field of affective analysis. The methods used in the early days mainly include sentiment dictionary-based methods and machine learning-based methods. The method based on emotion dictionary usually matches the words in the sentence with the words that have recorded emotion tendency in the dictionary, and then obtains the final emotion tendency by aggregating the emotion tendency of the words (such as averaging or summing). Ramshankar et al. [14] used WordNet to judge the emotional tendency of words. He et al. [15] calculated the path distance between words in WordNet to obtain the emotional similarity, and then calculated the emotional tendency of words. There is a relative lack of standardized Chinese emotion dictionaries. The earliest and most widely spread one is the vocabulary set for emotion analysis provided by HowNet [16]. In fact, the real emotional judgment is not a pile of simple rules, but a complex and systematic project, and the words in the emotional dictionary need to be manually selected. Therefore, the performance of this method depends largely on prior knowledge and manual design. The problem of machine learning-based affective orientation analysis is often viewed as a supervised learning problem. Hammi et al. [17] used machine learning methods such as naive Bayes, maximum entropy and support vector machine to try to solve the sentiment analysis problem as early as 2004. However, these methods require a complex feature selection process, which also relies on manual design, resulting in poor generalization ability.

Deep learning method can automatically select features, and has gradually become the mainstream method in sentiment analysis field in recent years. Li et al. [18] first proposed the use of CNNS to solve problems in the field of NLP such as part-of-speech tagging. Zhang [19] proposed to apply CNN to sentiment analysis tasks, and Adam et al. [20] proposed wide convolution and K-max pooling methods on this basis. Zhang et al. [21] proposed the VDCNN model and adopted the deep convolutional network method. But the CNN model has its defect, that is, it can only mine the local information of the text. RNN captures long-distance dependencies between texts better than CNN. In order to model the relationship between sentences, Imron et al. [22] proposed a hierarchical RNN model to model text at the discourse level. Gaikwad et al. [23] proposed a DRNN model to fix the step size of information flow. Combining the advantages of CNN and RNN, Al-Khazaleh et al. [24] proposed the RCNN model, which first used bidirectional recurrent neural networks to obtain context representation, and then output classification results after convolution and pooling operations. Chan et al. [25] proposed a C-LSTM model, which first used convolutional neural networks to extract text features and then input recurrent neural networks to obtain classification results. The attention mechanism can capture the importance of features and is also applied in text sentiment analysis tasks. For example, Abbes et al. [26] proposed a hierarchical attention model for sentiment analysis tasks. In short, the application of deep learning method in text sentiment orientation analysis eliminates the cumbersome feature engineering steps of traditional methods, and has certain advantages.

The model structure in this paper is an improvement on the work of Ahmed et al. [27]. This traditional CNN-RNN model architecture does not fully consider the importance of different text components to the determination of emotional tendency. Based on this framework, the hierarchical attention mechanism is added to the model in this paper, which helps the model to learn the most important information for judging the outcome of emotional tendency. In addition, in the convolution part, this paper refers to the work of Das et al. [28]. The difference is that the position vector is introduced into the model, and each word is assigned with actual position coding to construct a new word vector coding, so that the model can learn more abundant word vector coding information. At the same time, this paper also draws on the work of Deepa et al. [29], which is different from it by using word level vector and word level vector respectively for experiments.

The work flow of the model in this paper is as follows: First, convolutional neural networks are used to encode word vectors into a new vector space to learn the position information and context information of words, and then hierarchical attention is used to learn the serialization information of sentences and texts and their importance to text tendency determination. All in all, the model makes comprehensive use of the advantages of several deep learning methods and considers the local information and global information of the text, which avoids the loss of information and can screen out the information that has the greatest impact on the result.

3. Proposed Model

3.1. U-network

The main contributions of U-network in this paper are as follows.

1. Network dimensionality reduction. The U-network [30] is designed for two-dimensional or three-dimensional image data by default, and seismic waveform data is one-dimensional time series, so all convolution layers, pooling layers, drop out layers, and input and output tensors need to be reduced to one dimension.
2. The loss function. The mathematical principle of deep learning algorithm classification is actually to optimize the objective function, by repeatedly comparing the difference between the predicted value of the current network and our expected target value, and constantly adjusting the weight and deviation of each layer and other hyperparameters, so as to minimize the difference. The equation that measures the difference between the predicted value and the expected value is called the loss function (or objective function). The loss function used in this study is:

$$L = -\sum_{i=1}^{3}\sum_{j=1}^{n} Y'_{ij} \times \log(Y_{ij}). \tag{1}$$

Where Y' is the label encoded by binarization. $i = 1, 2, 3$ are noise. $j = 1, \cdots, n$ is the sequence number of the sampling point, and its expression is:

$$Y'_i = \begin{cases} Y'_1 & [1,0,0] \\ Y'_2 & [0,1,0] \\ Y'_3 & [0,0,1] \end{cases} \tag{2}$$

Y is the probability value calculated by the last layer softmax function, and its expression is as follows:

$$Y_i = \frac{e^{z_i}}{\sum_{k=1}^{3} e^{z_k}}. \tag{3}$$

Where z is the output tensor of the last layer $([m, n, 3])$. m indicates the number of input data.

3.2. Convolutional Neural Network (CNN)

CNN was first used in computer vision. In recent years, it also has a superior performance in text classification [31,32]. When processing text tasks, the traditional CNN model often converts words into vector form, multiplies the convolution nuclei of different numbers and sizes with the vector by elements, and obtains the final output after a series of operations including convolution, pooling, dropout regularization, and so on. CNN can capture the local relationship between words or words in text tasks, which is a common processing method in text sentiment analysis tasks. Since the purpose of using CNN in this paper is to extract N-gram features of words in a single sentence and input them into the structure of the next layer of the model, only convolution operation is used. Its structure is shown in Figure 1.

In Figure 1, padding represents the zero-complete operation to ensure that the length of the converted sentence representation matrix is the same as the length of the word vector matrix. Suppose the current input is the $j - th$ word $x_{ij} \in R^d$ in the $i - th$ sentence, and d represents the word vector dimension. In this paper, each word is assigned a positional code $l_{ij} \in R^d$, which is independent of word vector semantics, and the specific value is learned through model training. Thus, each word has a new code a_{ij}, as shown in equation (4).

$$a_{ij} = x_{ij} + l_{ij}. \tag{4}$$

Let the convolution kernel $B^n \in R^{h \times d}$ perform convolution operations on a new word vector encoding matrix. Where h represents the length of the convolution kernel, the width of the convolution kernel is the same as the dimension of the word vector, and n represents the number of convolution kernel.

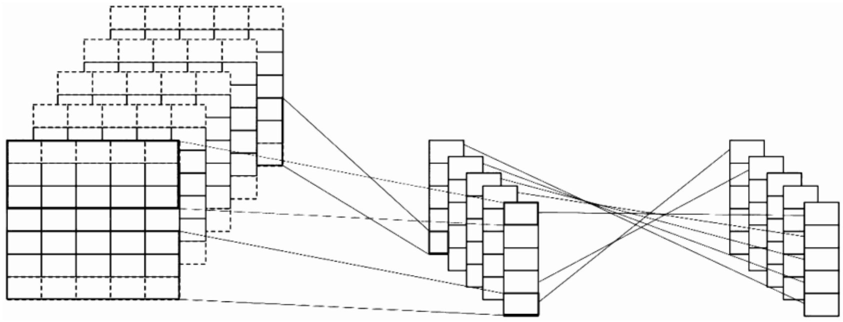

Fig. 1. CNN convolutional layer structure

Suppose that the corresponding vector obtained by each a_{ij} after convolution operation is expressed as $z_{ij} = [z_{ij}^1, z_{ij}^2, \cdots, z_{ij}^d]$, then the calculation process of each element in z_{ij} is shown in equation (5).

$$z_{ij}^n = f(\sum_{q=-(h-1)/2}^{(h-1)/2} a_{ij+q} \odot B_{(h-1)/2+q}^n). \tag{5}$$

Where $\odot$ represents the sum of vectors dotted by elements. B_m^n represents the $m-th$ row vector of the convolution kernel matrix. f stands for nonlinear activation function ReLU. Finally, the sequence of the $i-th$ sentence represents $z_{ik}, k \in [1, K]$, where K represents the total number of words in the sentence.

3.3. Attention Mechanism

The attention mechanism was applied to machine translation tasks as early as 2014, and after a period of development, many different forms of variants have been produced.

The attention model can be abstracted as Module1 and Module2 modules. Module1 is generally an encoder and transforms the input data to a certain extent. Module2 is a decoder and also outputs data after a certain transformation. The calculation process of each output value m is shown in equation (6).

$$m_i = F(C_i, m_1, m_2, \cdots, m_{i-1}. \tag{6}$$

Where, C_i is the semantic encoding corresponding to each output data, which is generated by the distribution of input data, as shown in equation (7).

$$C_i = \sum_{j=1}^{T} a_{ij} S(n_j). \tag{7}$$

Where $S(n_j)$ represents the hidden layer status of the input data obtained after Module1 processing. T represents the number of input data. a_{ij} represents the probability of attention distribution between input j and output m_i. The calculation process of a_{ij} is shown in equations (8) and (9).

$$a_{ij} = \frac{exp(e_{ij})}{\sum_{k=1}^{T} exp(e_{ik})}. \tag{8}$$

$$e_{ij} = Vtanh(Wh_j + U_{s_{i-1}} + b). \tag{9}$$

Where e_{ij} refers to the evaluation score of the influence of the $j-th$ input on the $i-th$ output. h_j is the hidden layer status of the $j-th$ input in module1. s_{i-1} is the output of module2 in the previous step. W, U and V are weight matrices. b is the bias value, which is learned during training. Attention semantic coding is used as input to module2 to generate the final deep features and obtain the most critical semantic information.

3.4. Specific Model Structure

Based on the above, this paper proposes a text emotion analysis model that integrates convolutional neural network and hierarchical attention network. The model consists of a word/word level initialization vector module operated by convolutional neural network, a bidirectional recurrent neural network, a word/word level attention module, a bidirectional recurrent neural network and a sentence level attention module.

The model first converts English characters or words into corresponding vector expressions through the operation of the CNN layer. Suppose a paragraph of English text has L sentences, denoted as $S_i, i \in [1, L]$. The sentence contains K characters or words. The vector from which the $i-th$ sentence has been convolved represents $z_{ik}, k \in [1, K]$. Therefore, after the completion of the first step, the output result of the convolutional neural network operation can be combined with its context-relevant information through the Bi-GRU network to obtain the output of the hidden layer. The specific calculation process is shown in formulas (10)-(12).

$$\overrightarrow{g}_{ik} = \overrightarrow{GRU}(z_{ik}), k \in [1, K]. \tag{10}$$

$$\overleftarrow{g}_{ik} = \overleftarrow{GRU}(z_{ik}), k \in [K, 1]. \tag{11}$$

$$g_{ik} = [\overrightarrow{g}_{ik}, \overleftarrow{g}_{ik}]. \tag{12}$$

Among them, g_{ik} is the vectorized representation obtained after bidirectional GRU.

The purpose of adding the Attention mechanism after this step is to find the word or word in a sentence that contributes the most to the meaning of the sentence. First, it inputs g_{ik} into a single-layer perceptron and the resulting u_{ik} is used as an implicit representation of g_{ik}. The importance of a word is determined by the similarity between u_{ik} and a randomly initialized context vector U_w. Then, after softmax operation, a normalized Attention weight matrix is obtained, representing the weight of the $k-th$ word in sentence i. Finally, after the attention weight matrix is obtained, the sentence vector is regarded as the weighted sum of these words or word vectors. The calculation process is shown in equations (13)-(15).

$$u_{ik} = tanh(W_w g_{ik} + b_w). \tag{13}$$

$$a_{ik} = \frac{exp(u_{ik}^T U_w)}{\sum_k exp(u_{ik}^T U_w)}. \tag{14}$$

$$S_i = \sum_k a_{ik} g_{ik}. \tag{15}$$

Where, W_w and b_w are weight matrix and bias matrix respectively. a_{ik} is the attention weight factor that measures the importance of the $k-th$ word or word in sentence i.

After obtaining the representation of S_i, we can use a similar method to process the sentence and obtain the corresponding hidden layer sentence vector G_i after bidirectional GRU, as shown in equations (16)-(18).

$$\overrightarrow{G}_i = \overrightarrow{GRU}(S_i), i \in [1, L]. \tag{16}$$

$$\overleftarrow{G}_i = \overleftarrow{GRU}(S_i), i \in [L, 1]. \tag{17}$$

$$G_i = [\overrightarrow{G}_i, \overleftarrow{G}_i]. \tag{18}$$

Then, a sentence-level context vector U_S is introduced to measure the importance of the sentence in the whole text, and the total text vector V is obtained. Finally, sentiment analysis can be performed through softmax layer. The calculation process is shown in equations (19)-(22).

$$u_i = tanh(W_S G_i + b_S). \tag{19}$$

$$a_i = \frac{exp(u_i^T U_S)}{\sum_i exp(u_i^T U_S)}. \tag{20}$$

$$V = \sum_i a_i G_i. \tag{21}$$

$$p = softmax(W_2 V + b_2). \tag{22}$$

As above, W_S, W_2 and b_S, b_2 are the weight matrix and bias matrix respectively. a_i measures the attention weight factor that measures the importance of sentence i.

In addition, the ultimate goal of this training is to minimize the loss function (negative log-likelihood function), as shown in equation (23).

$$L = -\sum_{d} \log p_{d_j}. \tag{23}$$

Where j is the emotion category label corresponding to the text d.

4. Experimental and Result Analysis

4.1. Dataset Introduction

In this paper, we use online movie review data sets (Data1) and online shopping review data sets (Data2). Since there is no emotion label directly given in the movie review data set, the emotion label is set according to the review score. The comments with rating=1 are set to label=0 and those with rating=5 are set to label=1. label=0 indicates a negative polarity comment, and label=1 indicates a positive polarity comment. The sentiment classification labels of online shopping review data sets are also two categories of positive polarity and negative polarity.

The data is preprocessed before the model training, such as removing the URL and email in the text. Then the data set is divided into three parts: training set $Train-data$, verification set $Dev-data$ and test set $Test-data$. Table 1 shows the division of the two data sets.

Table 1. Dataset partitioning

Data	English film review dataset	Dataset of online shopping reviews
$Train-data$	34000	50000
$Dev-data$	3800	3800
$Test-data$	2000	1900

4.2. Evaluation Index

In this paper, Accuracy and F1 score are used to verify the effect of the model, in which F1 score can comprehensively measure the accuracy rate and recall rate. Accuracy and F1 indexes are calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}. \tag{24}$$

$$F1 = \frac{2PR}{P+R}. \tag{25}$$

Where $P = \frac{TP}{TP+FP}$, $R = \frac{TP}{TP+FN}$. TP (True Positive) is the number of samples whose true polarity and predicted polarity are positive. FP(False Positive) is the number of samples in the test data whose true polarity is negative and predicted polarity is positive; FN (False Negative) is the number of samples in the test data whose true polarity is positive and predicted polarity is negative. TN (True Negative) is the number of samples in the test data whose true polarity is negative and whose predicted polarity is negative.

4.3. Comparison and Result Analysis

In this paper, the proposed model is compared with the popular emotion analysis models, including CRNN (Combination of Convolutional and Recurrent Neural Network) and BiGRU (Bidirectional Gated Recurrent Unit), DPCNN (Deep Pyramidal Convolutional Neural Network), AC-BiLSTM (Attention-based Bidirectional Long) Short-Term Memory with Convolution Layer) and ABCDM (An Attention-based Bidirectional CNN-RNN Deep Model). The experimental comparison results are shown in Table 2.

- BiGRU. The semantic features of text above and below are obtained by BiGRU model to classify emotion polarity.
- CRNN [33]. Combining CNN and RNN, we use RNN structure to learn the local features and order relationships retained by CNN in sentences.

- DPCNN [34]. A fixed number of equal length convolution layers are used to extract the semantic features between adjacent words in the sentence, and then the local features of the text are obtained through the pooling layer with step size 2.
- AC-BiLSTM [35]. The text representation is input to CNN, and then the output of CNN is input to BiLSTM, and the output $\overrightarrow{H}$ hidden in the forward propagation direction and the output $\overleftarrow{H}$ hidden in the reverse propagation direction are extracted respectively. Input $\overleftarrow{H}$ and $\overrightarrow{H}$ into the attention mechanism respectively, and finally fuse the output of the two attention mechanisms and classify them.
- ABCDM [36]. Two independent channels are parallel, and the first channel combines BiLSTM, attention mechanism and two parallel CNNs to carry out maximum pooling and average pooling operations for each convolutional layer output respectively. The second channel combines BiGRU, the attention mechanism, and two parallel CNNS to maximize and average the output of each convolutional layer, respectively. The output in the two channels is then fused and classified.

Table 2. Comparison results

Method	Data1	Data1	Data2	Data2
Index	Accuracy/%	F1/%	Accuracy/%	F1/%
BiGRU	89.46	89.11	92.53	92.55
CRNN	84.16	83.40	91.20	91.19
DPCNN	80.21	79.64	87.67	87.74
AC-BiLSTM	84.76	84.33	91.15	91.08
ABCDM	90.9	90.59	93.25	93.29
Proposed	93.16	93.17	94.78	94.92

As can be seen from Table 2, ABCDM uses two independent and parallel BiLSTM channels and BiGRU channels, but the features extracted by both channels are contextual semantic information of text. The feature information extracted by the two channels of the model in this paper is different. Not only the semantic information of text context is extracted, but also important local features are obtained. The experimental results show that the accuracy and F1 value of proposed model on the two data sets are superior to that of ABCDM model, the accuracy is 2.2% and 1.53% higher, and the F1 is 2.58% and 1.63% higher, respectively. The results show that the model in this paper can effectively improve the performance of the model by extracting the deep global text semantic features from different aspects through dual channels and deep fusion.

The evaluation indexes of the model in this paper are better than those of CRNN, BiGRU, DPCNN, AC-BiLSTM and ABCDM on different data sets. In summary, the comparison between the proposed model and the popular sentiment analysis model proves that the proposed model has better performance in sentiment classification.

The proposed model is compared with the deep BiGRU single-channel, DPCS single-channel and DBG-DPCS models, and the ablation experimental results are shown in Table 3.

- Deep BiGRU single channel. RoBERTa is used to obtain the text representation vector, and then BiGRU is used to extract the deep semantic information of the text context. Finally, emotion classification is carried out by the output layer.
- DPCS single channel. According to the length of the text sequence, the convolutional attention blocks are continuously added to deepen the DPCS until the sequence length is 1, and important local features of the deep text are output. Finally, the emotion polarity is classified.
- DBG-DPCS. By combining the deep BiGRU channel and DPCS channel, the dynamic attention network in the model in this paper is removed, that is, semantic features of different aspects are obtained by deep BiGRU and DPCS, and text features of different aspects are directly spliced for splicing and fusion. Finally, emotion classification is performed on the fused features.

As can be seen from Table 3, the comparison between deep BiGRU single channel, DPCS single channel and DBG-DPCS shows that the accuracy and F1 of DBG-DPCS on movie data sets are lower than those of deep BiGRU single channel and DPCS single channel. However, the accuracy of online shopping data set and F1 are higher than that of deep BiGRU single channel and DPCS single channel, indicating that the text features cannot be effectively fused by directly splicing features from different aspects.

By comparing DBG-DPCS with the model in this paper, the accuracy and F1 of DBGDCN-DPCS model on the movie review dataset and online shopping review dataset are higher than DBG-DPCS, indicating that the

Table 3. Results of ablation experiment

Method	Data1	Data1	Data2	Data2
Index	Accuracy/%	F1/%	Accuracy/%	F1/%
Deep BiGRU single channel	92.76	92.81	94.12	94.19
DPCS single channel	93.01	92.96	94.17	94.36
DBG-DPCS	92.36	92.57	94.63	94.71
Proposed	93.16	93.17	94.78	94.92

feature fusion method based on co-attention network can effectively and deeply integrate different aspects of text features and improve the classification effect of the model.

By comparing BiGRU single channel and DPCS single channel with the proposed model, the accuracy and F1 of the proposed model in movie review data reach 93.16% and 93.17% respectively, and the accuracy and F1 of online shopping review data set reach 94.78% and 94.92% respectively, both of which are superior to the two single-channel models. The results show that the extraction of different aspects of text semantic features can obtain more comprehensive and rich semantic features, achieve comprehensive and deep text semantic features extraction, and thus improve the performance of the model, to a certain extent, the classification performance of the model is better than that of the two single-channel models.

To sum up, the proposed deep BiGRU and DPCS sentiment analysis model combined with co-attention network achieves good classification performance in the task of Chinese text sentiment analysis, which proves the effectiveness of the model.

5. Conclusion

Aiming at the task of English text sentiment analysis, this paper proposes a deep BiGRU and DPCS sentiment analysis model combined with co-attention network. First, RoBERTa model is used to obtain text semantic representation vector, and then deep BiGRU and DPCS dual channels are used to extract text semantic features from different aspects, and feature fusion method based on co-attention network is used to deeply integrate text semantic features from different aspects to extract more comprehensive and deep-level global text semantic features. Finally, it classifies and outputs the emotion category. Deep BiGRU extracts the semantic feature information of text depth above and below through multi-layer BiGRU, and DPCS obtains important deep local features and text long-distance dependencies by deepening its depth. In order to prove the validity of the proposed model, comparison experiments are carried out on two datasets, movie reviews and online shopping reviews. The experimental results show that the classification effect of the proposed model is higher than other comparison models. Although the model has achieved good results, the training time is long. In future work, the training time of the model should be effectively reduced under the premise of ensuring the model effect.

6. Conflict of Interest

The authors declare that there are no conflict of interests, we do not have any possible conflicts of interest.

Acknowledgments. None.

References

1. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: A survey[J]. Ain Shams engineering journal, 2014, 5(4): 1093-1113.
2. Wankhade M, Rao A C S, Kulkarni C. A survey on sentiment analysis methods, applications, and challenges[J]. Artificial Intelligence Review, 2022, 55(7): 5731-5780.
3. Hussein D M E D M. A survey on sentiment analysis challenges[J]. Journal of King Saud University-Engineering Sciences, 2018, 30(4): 330-338.
4. Yu J, Lu Z, Yin S, et al. News recommendation model based on encoder graph neural network and bat optimization in online social multimedia art education[J]. Computer Science and Information Systems, 2024. doi: 10.2298/CSIS231225025Y.
5. Liu Z, Jiang D, Zhang C, et al. A novel fireworks algorithm for the protein-ligand docking on the autodock[J]. Mobile Networks and Applications, 26: 657-668, 2021.
6. Das R, Singh T D. Multimodal sentiment analysis: a survey of methods, trends, and challenges[J]. ACM Computing Surveys, 2023, 55(13s): 1-38.

7. S. Yin, H. Li, Y. Sun, M. Ibrar, and L. Teng. Data Visualization Analysis Based on Explainable Artificial Intelligence: A Survey[J]. IJLAI Transactionss on Science and Engineering, vol. 2, no. 2, pp. 13-20, 2024.
8. Aslan S, Kzloluk S, Sert E. TSA-CNN-AOA: Twitter sentiment analysis using CNN optimized via arithmetic optimization algorithm[J]. Neural Computing and Applications, 2023, 35(14): 10311-10328.
9. Joloudari J H, Hussain S, Nematollahi M A, et al. BERT-deep CNN: State of the art for sentiment analysis of COVID-19 tweets[J]. Social Network Analysis and Mining, 2023, 13(1): 99.
10. Aslan S. A deep learning-based sentiment analysis approach (MF-CNN-BILSTM) and topic modeling of tweets related to the UkraineCRussia conflict[J]. Applied Soft Computing, 2023, 143: 110404.
11. Baarslan M S, Kayaalp F. MBi-GRUMCONV: A novel Multi Bi-GRU and Multi CNN-Based deep learning model for social media sentiment analysis[J]. Journal of Cloud Computing, 2023, 12(1): 5.
12. Mohbey K K, Meena G, Kumar S, et al. A CNN-LSTM-based hybrid deep learning approach for sentiment analysis on Monkeypox tweets[J]. New Generation Computing, 2023: 1-19.
13. Yin S, Li H, Laghari A A, et al. An Anomaly Detection Model Based On Deep Auto-Encoder and Capsule Graph Convolution via Sparrow Search Algorithm in 6G Internet-of-Everything[J]. IEEE Internet of Things Journal, 2024. DOI: 10.1109/JIOT.2024.3353337.
14. Ramshankar N, PM J P. Automated sentimental analysis using heuristic-based CNN-BiLSTM for E-commerce dataset[J]. Data & Knowledge Engineering, 2023, 146: 102194.
15. He A, Abisado M. Text Sentiment Analysis of Douban Film Short Comments Based on BERT-CNN-BiLSTM-Att Model[J]. IEEE Access, 12: 45229-45237, 2024.
16. Shan Y. Social Network Text Sentiment Analysis Method Based on CNN-BiGRU in Big Data Environment[J]. Mobile Information Systems, 2023, 2023(1): 8920094.
17. Hammi S, Hammami S M, Belguith L H. Advancing aspect-based sentiment analysis with a novel architecture combining deep learning models CNN and bi-RNN with the machine learning model SVM[J]. Social Network Analysis and Mining, 2023, 13(1): 117.
18. Li H, Lu Y, Zhu H, et al. A Novel AB-CNN Model for Multi-Classification Sentiment Analysis of e-Commerce Comments[J]. Electronics, 2023, 12(8): 1880.
19. Zhang B. A bert-cnn based approach on movie review sentiment analysis[C]//SHS Web of Conferences. EDP Sciences, 2023, 163: 04007.
20. Adam A Z R, Setiawan E B. Social Media Sentiment Analysis using Convolutional Neural Network (CNN) dan Gated Recurrent Unit (GRU)[J]. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI), 2023, 9(1): 119-131.
21. Zhang N, Xiong J, Zhao Z, et al. Dose My Opinion Count? A CNN-LSTM Approach for Sentiment Analysis of Indian General Elections[J]. Journal of Theory and Practice of Engineering Science, 2024, 4(05): 40-50.
22. Imron S, Setiawan E I, Santoso J, et al. Aspect Based Sentiment Analysis Marketplace Product Reviews Using BERT, LSTM, and CNN[J]. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 2023, 7(3): 586-591.
23. Gaikwad R S, Gandage S C. MCNN: Visual sentiment analysis using various deep learning framework with deep CNN[J]. International Journal of Intelligent Systems and Applications in Engineering, 2023, 11(2s): 265C278-265C278.
24. Al-Khazaleh M, Alian M, Biltawi M, et al. Sentiment Analysis for People's Opinions about COVID-19 Using LSTM and CNN Models[J]. International Journal of Online & Biomedical Engineering, 2023, 19(1).
25. Chan K H, Im S K. Sentiment Analysis Using Bi-CARU with Recurrent CNN Models[C]//2023 8th International Conference on Smart and Sustainable Technologies (SpliTech). IEEE, 2023: 1-5.
26. Abbes M, Kechaou Z, Alimi A M. A Novel Hybrid Model Based on CNN and Bi-LSTM for Arabic Multi-domain Sentiment Analysis[C]//Conference on Complex, Intelligent, and Software Intensive Systems. Cham: Springer Nature Switzerland, 2023: 92-102.
27. Ahmed Z, Wang J. A fine-grained deep learning model using embedded-CNN with BiLSTM for exploiting product sentiments[J]. Alexandria Engineering Journal, 2023, 65: 731-747.
28. Das N, Sadhukhan B, Bhakta S S, et al. Integrating EEMD and ensemble CNN with X (Twitter) sentiment for enhanced stock price predictions[J]. Social Network Analysis and Mining, 2024, 14(1): 1-18.
29. Deepa N, Priya J S, Devi T. Sentimental Analysis Recognition in Customer review using Novel-CNN[C]//2023 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2023: 1-4.
30. Anand I, Negi H, Kumar D, et al. Residual u-network for breast tumor segmentation from magnetic resonance images[J]. Comput. Mater. Contin, 2021, 67: 3107-3127.
31. Jiang D, Li H, Yin S. Speech emotion recognition method based on improved long short-term memory networks[J]. International Journal of Electronics and Information Engineering, 2020, 12(4): 147-154.
32. Meng X, Wang X, Yin S, et al. Few-shot image classification algorithm based on attention mechanism and weight fusion[J]. Journal of Engineering and Applied Science, 2023, 70(1): 14.
33. Wang X, Jiang W, Luo Z. Combination of convolutional and recurrent neural network for sentiment analysis of short texts[C]//Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers. 2016: 2428-2437.
34. Johnson R, Zhang T. Deep pyramid convolutional neural networks for text categorization[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 2017: 562-570.
35. Liu G, Guo J. Bidirectional LSTM with attention mechanism and convolutional layer for text classification[J]. Neurocomputing, 2019, 337: 325-338.
36. Basiri M E, Nemati S, Abdar M, et al. ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis[J]. Future Generation Computer Systems, 2021, 115: 279-294.

Biography

Shu Ma is with the College of International Business, Shenyang Normal University. Research direction is English-Chinese translation, English linguistics and AI.