Construction of Text Emotion Classification Model Based on Convolutional Neural Network

Ruhua Lu[1], Yalan Li[2, a, *] and Yanwen Yan[1]
[1]College of Software and Communication Engineering, Xiangnan University, Chenzhou, Hunan 423000, China
[2]School of Electronic Information and Electrical Engineering, Xiangnan University, Chenzhou, Hunan 423000, China
[a]liyalan@xnu.edu.cn

Abstract—**Text emotion analysis transforms a text sequence of indefinite length into text category, which is one of the key research problems in the field of natural language processing. With the wide application of deep learning technology in natural language processing, the text emotion analysis model based on deep learning has made a new breakthrough. This paper builds a basic framework of text emotion analysis, and describes it from two aspects: data preprocessing and design of network structure of revolutionary neural network. Data preprocessing mainly involves word segmentation model and word embedding. Design of network structure of revolutionary network is the basic structure of the revolutionary network (CNN), including input layer, convolution layer, pool layer, full connection layer and output layer. Finally, the feasibility of the method is verified by experiments.**

I. INTRODUCTION

Text sentiment analysis is a common task in natural language processing. It transforms a text sequence of variable length into a text category[1], which is widely used at present. For example, through the analysis of a product review on the e-commerce platform, consumers can quickly know whether the indicators of the desired product meet their expectations, and then decide whether to buy the product; for enterprises, through the analysis of after-sales product reviews, users can know how much they like the product and what problems exist, so that they can continue to improve in the future. Through the analysis and excavation of the hot issues concerned by the people, we can always grasp the needs and thoughts of the people, which can not only control the development of public opinion, but also provide the basis for the correct guidance of public opinion, and also issue some policies to benefit the people, which can not only solve the immediate problems concerned by the people, but also provide the basis for the correct guidance of public opinion It further promoted the prosperity and stability of the society.

II. BASIC FRAMEWORK OF TEXT SENTIMENT CLASSIFICATION

Emotion recognition can be divided into three categories: text level, sentence level and phrase level [2], and then rules, dictionaries, machine learning and other methods are used for recognition. The third method is favored by scholars, and its classification process is shown in Figure 1. Firstly, the labeled text emotion information is used to train the classification model, and then the trained model is used to classify the unlabeled text emotion information. In the process of model training, the collected text emotion information is labeled first, preprocessed second, and then the text is vectorized. Finally, the machine learning classification model is built and trained. When the model parameters tend to be stable, the trained model can be used to classify the unlabeled new text emotion information. The following key problems need to be solved in the whole process: how to obtain massive text emotion information data, how to preprocess text emotion information, how to express text, how to choose text features, how to select classification model and how to train model, etc. Some of the key issues will be described in the next section.

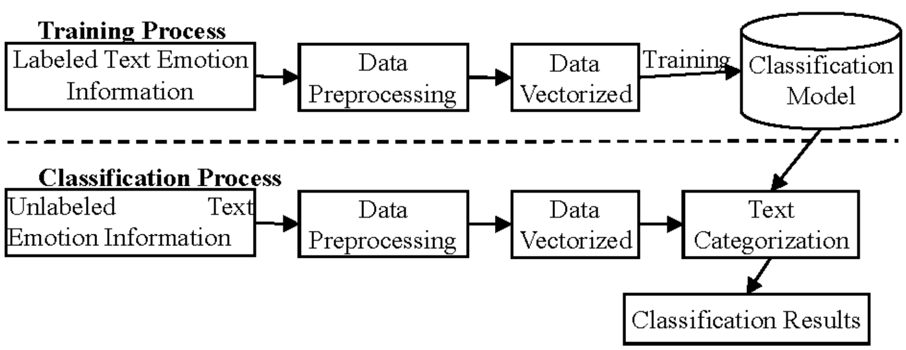

Figure 1 Basic Framework of Text Sentiment Classification

978-1-6654-3730-1/21/$31.00 ©2021 IEEE
DOI 10.1109/ICAA53760.2021.00056

III. The Main Content of the Basic Framework of Text Sentiment Classification

A. Data Preprocessing

1) Word segmentation model

To extract features from text, word segmentation is very important [3]. In English, separated by spaces, each word is a word. In Chinese, words are made up of words, which need to be divided into words according to the semantics of the text. Usually, Jieba-Participle is used to Chinese text classification. There are three patterns of Jieba-Participle. First is precise pattern, trying to cut the sentence most accurately, which is suitable for text analysis. Second is full mode, scanning all the words in the sentence that can be words. It's very fast, but can't solve the ambiguity. Three is search engine mode. The long words are segmented again on the basis of accurate mode to improve the recall rate, which is suitable for search engine segmentation. In English, separated by spaces, each word is semantic. In Chinese, words are made up of characters. We should divide text into words according to the semantics of the text. In tensorflow platform, Jieba-Participle has some methods to achieve these functions. JIEBA.CUT Method takes three input parameters. The first parameter is the string to be segmented. The second parameter is cut_all, used to control whether the full mode is used. The three parameter is HMM, used to control whether the HMM model is used. JIEBA.CUT_FOR_SEARCH method takes two parameters, the string to be segmented and the HMM used to control whether to use the HMM model. This method is suitable for word segmentation of inverted index in search engine. JIEBA.CUT and JIEBA.CUT_FOR_SEARCH structure returned by search is an iterative generator. It use the for-loop to get every word after word segmentation, or use the JIEBA.CUT method and JIEBA.CUT_FOR_SEARCH method to return a list directly.

2) Word embedding

One-Hot coding can realize the digital representation of words [4], but its matrix is sparse, the dimension is high, and it can't capture the correlation between words, which is not conducive to text sentiment analysis. Word embedding can avoid these problems. Word vector is used to express the vector or representation of a word, and can also be considered as the feature vector of a word. Embedding Words are into low dimensional continuous vector space, that is, words are represented as vectors on real number field, is called word embedding [5]. A word embedding is a dense floating-point vector (the length of the vector can be set), which can capture the relationship between words. For example, by calculating the cosine value of two words embedded, we can get the correlation degree of two words. Generally, word embedding is from 8 dimensions (for small datasets) to 1024 dimensions (for large datasets) or higher. The word embedding model transforms each vocabulary into vectors with fixed length, and makes these vectors better express the similarity and class relationship between different words, realizing the transformation from text to word vector. Common pre-training word embedding models include word2vec, glove and so on. Word2vec trains news articles with billions of words. Google provides the results of a set of word vectors, which can be downloaded from http://word2vec.googlecode.com/svn/trunk/. Word2vec is the most widely used pre-training word embedding model, and it is also used in this paper.

B. Design of Network Structure of Convolutional Neural Network

Convolutional neural network (CNN) is a kind of feed-forward neural network. At present, there are one-dimensional, two-dimensional and three-dimensional convolutional neural network. The basic structure includes input layer, convolution layer, pooling layer, full connection layer and output layer [6]. The input layer is the input of the whole neural network data. In the convolutional neural network, which deals with the emotion of text, the input layer data represents a digital matrix of text, which is embedded by the preceding words. Convolution layer is the most important part of convolution neural network. In this paper, we use two-dimensional convolution neural network to output a two-dimensional array, namely convolution layer, by cross-correlation operation of a two-dimensional input array and a two-dimensional convolution kernel array. Suppose that the shape of the two-dimensional input array is $Nh \times Nw$ and the shape of the two-dimensional convolution kernel window is $Kh \times Kw$, then the shape of the output two-dimensional array is $(Nh - Kh + 1) \times (Nw - Kw + 1)$. The convolution layer also has two super parameters: filling and stride. Filling refers to filling elements (generally element 0) on both sides of input height and width. Filling Ph on both sides of input height and Pw on both sides of input width, the shape of output two-dimensional array changes from $(Nh - Kh + 1) \times (Nw - Kw + 1)$ to $(Nh - Kh + Ph + 1) \times (Nw - Kw + Pw + 1)$. In order not to change the shape of output, that is, the shape of output and input is the same, $Ph = Kh-1$ and $Pw = Kw-1$ are defined. Stride refers to the number of rows and columns sliding each time in cross-correlation operation. When the number of rows and columns is greater than 1, the output shape will be different. The stride of height recording is Sh, the stride of width is Sw, and the output shape is $(Nh - Kh + Ph + Sh) / Sh \times (Nw - Kw + Pw + Sw) / Sw$. The pooling layer does not change the depth of the neural network. It is mainly used for feature dimensionality reduction, compressing the number of data and parameters, reducing over fitting, and improving the fault tolerance of the model. Like the convolution layer, the pooling layer computes the output for each element in a fixed shape window of the input data. According to different calculation, the most common pooling operations are average pooling and maximum pooling. Average pooling calculates the average value of the image region as the pooled value of the region, and maximum pooling selects the maximum value of the image region as the pooled value of the region. The pooling layer will continuously reduce the spatial size of the data, so the number of parameters and the amount of calculation will also decrease, which also controls the over fitting to a certain extent. After several rounds of convolution layer and pooling layer processing, the text emotional information has been abstracted into features with higher information content. Finally, the full connection layer is used to complete the classification task. The output layer uses softmax function for text sentiment classification, and gets the final analysis results.

IV. SIMULATION RESULTS AND ANALYSIS

A. Data Set

This paper uses the Internet Movie Database (IMDb), which is widely used in text sentiment analysis. IMDB is an online database about movie actors, movies, TV programs, TV stars and movie production. It includes much information about movies, such as actors, length, content introduction, rating, comments, etc. It is very suitable for text sentiment analysis. The official website of this database is https://www.imdb.com/, which can be downloaded directly from the website. There are 50000 comments in IMDB, of which 25000 are used for training (12500 positive comments and 12500 negative comments), and 25000 are used for testing (12500 positive comments and 12500 negative comments).

Some examples of positive comments and negative comments in IMDB database are shown in Table 1.

TABLE 1. EXAMPLES OF POSITIVE AND NEGATIVE COMMENTS

Positive Comments	**Negative Comments**
First of all, I would like to say that i am a fan of all of the actors that appear.	Being a long-time fan of Japanese film, I expected more than this.
Cage plays a trunk and gets high critically praise.	1980 was certainly a year for bad backwoods slasher movies.

B. Experimental Parameter Setting

Through many experiments, the parameters are constantly modified to get the following optimal settings.

The comments of IMDB data set vary in length. In order to make the converted digital list the same length, the comment length is set to 400 characters. During the process of constructing CNN model, the vocabulary size is 4000, word vector dimension is 32, input sequence length is 400 at input layer, and one-dimensional array with length of 12800 is generated in flat layer. The number of neurons is 256 and activation function is relu at network layer. In the last layer, the number of neurons is set to 2, and the activation function is softmax, then the final output layer is obtained. In the training phase, the optimization function is Adam and the loss function is categorical_crossentropy. In the process of training, the time of iterations is 25.

C. Analysis of Experiments

Experiment 1: This experiment tests that the test accuracy of the training data in the training process changes with the number of iterations. The experimental data are shown in Table 2.

TABLE 2. TEST ACCURACY OF THE TRAINING DATA WITH ITERATION

the Number of Iterations	1	5	10	15	25
Accuracy	5%	40%	78%	85%	92%

Experiment 2: This experiment tests that the accuracy of the testing data in the end of experiment changes with the number of training samples. The experimental data are shown in Table 3.

TABLE 3. THE ACCURACY OF THE TESTING DATA WITH THE NUMBER OF TRAINING SAMPLES

the Number of Training Samples	10000	15000	20000	25000
Accuracy	48%	62%	78%	84%

V. CONCLUSIONS

Text emotion analysis transforms a text sequence of indefinite length into text category, which is one of the key research problems in the field of natural language processing. The basic framework of text sentient classification is divided into two steps. The labeled text emotion information is used to train the classification model, and then the trained model is used to classify the unlabeled text emotion information. In these two steps, the key issues include Word Segmentation Model, Word Embedding and Design of Network Structure of Convolutional Neural Network. At the end of the paper, the experiment of text emotion classification is carried out, and the effectiveness of this method is verified.

Acknowledgement

This work has been supported by The Scientific Research Project of Hunan Provincial Education Department(19C1716,19C1723), Hunan Provincial College Student Innovation and Entrepreneurship Training Program (NO. S202010545013), The Scientific Research Project of Hunan Social Science Achievement Evaluation Committee(NO. XSP21YBC149),The Xiangnan University Youth Talent Support Program, Xiangnan University[2019]124, The Science and Technology Plan Projects of Chenzhou (ZDYF201810, ZDYF2020155).

References

[1] Mao Y . Long Text Emotion Classification Method based on the Attention Double-layer LSTM. Journal of Chongqing College of Electronic Engineering, 2019.

[2] Zhang C , Zhou M . A text emotion classification method based on CNN and bidirectional LSTM fusion[J]. Computer Era, 2019.

[3] Park S H , Bae B C , Cheong Y G . Emotion Recognition from Text Stories Using an Emotion Embedding Model[C]// 2020 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2020.

[4] Vks A , Nsm B , Am B , et al. Machine learning method using position-specific mutation based classification outperforms one hot coding for disease severity prediction in haemophilia 'A'[J]. Genomics, 2020, 112(6):5122-5128.

[5] Zhang L , Shen CL , Shou-Shan LI . Emotion Classification Algorithm Based on Emotion-specific Word Embedding[J]. Computer Science, 2019.

[6] Pentyala S , Dowsley R , Cock M D . Privacy-Preserving Video Classification with Convolutional Neural Networks. 2021.