

Sentiment Analysis Method based on Piecewise Convolutional Neural Network and Generative Adversarial Network

C. Du, L. Huang

Changshun Du*, Lei Huang

School of Economics and Management

Beijing Jiaotong University

Beijing 100044, China

*Corresponding author: summer2015@bjtu.edu.cn

Abstract: Text sentiment analysis is one of the most important tasks in the field of public opinion monitoring, service evaluation and satisfaction analysis under network environments. Compared with the traditional Natural Language Processing analysis tools, convolution neural networks can automatically learn useful features from sentences and improve the performance of the affective analysis model. However, the original convolution neural network model ignores sentence structure information which is very important for text sentiment analysis. In this paper, we add piecewise pooling to the convolution neural network, which allows the model to obtain the sentence structure. And the main features of different sentences are extracted to analyze the emotional tendencies of the text. At the same time, the user's feedback involves many different fields, and there is less labeled data. In order to alleviate the sparsity of the data, this paper also uses the generative adversarial network to make common feature extractions, so that the model can obtain the common features associated with emotions in different fields, and improves the model's Generalization ability with less training data. Experiments on different datasets demonstrate the effectiveness of this method.

Keywords: Sentiment analysis, Piecewise Convolution Neural Network, Generative Adversarial Network.

1 Introduction

As one of the most important tasks in the field of public opinion monitoring, service evaluation and satisfaction analysis in network environments, text sentiment analysis needs to determine the opinions and preferences of customers in the text. In this article, these subjective texts with emotional polarity mainly refer to users' comments on products or services, which can provide the decision-making bases of potential consumers when purchasing products and services. Analysis of these comments is extremely helpful in mining the potential needs of users and improving products and services, but these comments are growing in large numbers every day. The analysis by humans is not only costly, but also lagging in time. Therefore, it is necessary to analyze the emotional polarity of these texts using appropriate intelligent algorithms. Although there have been many studies on text sentiment analysis, it is still a huge challenge.

At present, text sentiment orientation analysis can be mainly divided into two methods: the first category is the sentiment orientation analysis method based on an emotion dictionary, and the second category is the statistical machine learning based analysis method. The first type of dictionary-based approach first requires the construction of an emotional dictionary. The Chinese Emotional Dictionary primarily includes two emotional dictionaries, HowNet and NTUSD. HowNet has been released by China Knowledge Network and NTUSD was released by Taiwan University. Such methods include Chen Xiaodong's [1] application of a sentiment lexicon to analyze the emotional tendencies of microblog texts; Li Chun et al [4] chose HowNet's strong tendencies as seed words, combined with contextual influence. It calculates the emotional

propensity score of a word by calculating the similarity between the common word and the emotional seed word, and then determines the emotional tendency of the sentence. This kind of sentiment analysis method is used to analyze the emotional tendency of sentences by examining the emotional polarity of words. It is a shallow analysis method that doesn't analyze or model the overall semantics of sentences. In the machine learning-based method, traditional methods usually use adjective features, word frequency features and N-Gram features as the characteristics of the sentiment orientation analysis of the text. These features only consider the meaning of the words themselves, or model adjacent words. The words are coded as vectors (word vectors) in vector space. With the development of Internet applications, the number of corresponding network data increases sharply, and text data is analyzed simply by artificial design features or traditional natural language processing grammar analysis tools. This method is not only noneffective but also inefficient. The rise of neural networks enables them to be used on emotional analysis of texts for semantic synthesis to automatically extract the features of sentences. Finally, classifiers are used to classify their emotional polarity. There are many works of this kind of methods, such as Socher et al. [5-7, 9, 28] using a cyclic neural network to extract the semantic features of sentences; Zeng Daojian et al. [12] used Convolutional neural networks to extract sentence features of specified tasks. The two neural network structures, recurrent neural networks and convolutional neural networks, are the two most effective methods for automatically learning sentence features in deep learning. This method of self-learning features has also made great progress in other fields of artificial intelligence. In terms of natural language processing tasks, it does not need to rely on tools such as traditional grammar analysis. It can automatically learn features from sentences, and thus it has received extensive attention from scholars. Note that except of text classification, in the field of pattern classification, classification and application of other methods also have a wide range of trying. [33, 34]

The research by Hang Cui et al. [2] shows that the sentiment analysis of texts has domain dependence. In different fields, traditional sentiment analysis methods are difficult to maintain at optimal levels at all times. Different methods have different adaptabilities in different fields. In a neural network approach, a convolutional neural network is capable of modeling the combination of word features with a small number of parameters. At the same time, convolutional neural networks have the following advantages over traditional natural language processing analysis tools: (1) As an effective method for automatically acquiring sentence features in deep learning, it is possible to automatically learn from the sentences the features that are most relevant to a sentiment analysis task. This improves the performance of the sentiment analysis model by extracting important features related to it in different fields. (2) Emotional analysis using convolutional neural network does not require the use of additional resources such as a sentiment lexicon, avoiding the problem of constructing an emotional lexicon and low emotional dictionary coverage. However, the original convolutional neural network model ignores the sentence structure information that is important for text sentiment analysis, and it is easy to overfit. In view of the above deficiencies, this paper adopts the piecewise pooling strategy, which enables the deep learning-based convolutional neural network model to model sentence structures and segment the main features of different structures. The Segmented Convolutional Neural Network (PCNN) analyzes the emotional propensity of the text by combining the structural information and the domain information of the text; and uses the Dropout algorithm to enhance the generic ability of the model.

At the same time, the user's feedback on the service involves many different fields, and there is less labeled data in each field. The lack of data makes the training of the convolutional neural network more difficult. The parameters cannot be fully optimized and cause the model to be under-fitting. Therefore, in the case that data volume expansion is difficult, in order to alleviate the sparseness of the data, this paper uses the generated adversarial network to extract the

common features of texts in different fields, so that the model can acquire the common features related to emotions in the feedback of different fields and enhance the generalization ability of the model in cases where there is less training data. Experiments on different data demonstrate the effectiveness of this method.

2 An emotional analysis model that combines PCNN and GAN

Convolutional neural networks are widely used in natural language processing tasks such as relation extraction, information retrieval, and sentiment analysis, and have achieved remarkable results. In this section, this paper will use a convolutional neural network to extract the features of the text, and automatically select the important features of the sentiment tendency that are suitable for the text field to be classified from the input text, and alleviate the dependence of the sentiment analysis model on the text field. Furthermore, the deficiencies of the convolutional neural network ignoring the structural information of text sentences are improved. The pooling layer is used to extract the structural features of the text and improve the performance of the sentiment analyzer. We also introduce the generative adversarial network to model, which enables the model to extract common features related to sentiment analysis between different domains of text, and improve the effectiveness of the model in the case of sparse data.

2.1 Emotional analysis model based on convolutional neural network

The model mainly consists of two parts. The first part is the feature extraction operation, which consists of two steps: convolution and pooling. The second part is the emotion classifier. The following is a description of each component separately.

Convolution operation

A standard convolutional neural network usually consists of a Convolution Layer and a Pooling Layer. Unlike the input of image processing, which is composed of image pixels, in natural language processing tasks, a matrix is usually used to represent a sentence or a paragraph as an input. Each line of the matrix represents a language token, which can be a character or a word. In the model presented in this paper, each row of the input matrix is a vectorized representation of a linguistic symbol. In image processing, the Convolution Kernel of the convolutional layer usually slides over all parts of the image, while in natural language processing the convolution kernel slides only in the direction of text expansion. That is, the width of the input matrix (the dimension of the word vector) coincides with the width of the convolution kernel. Assuming that the height of the convolution kernel is w , and the width and word vector dimensions are both d , the convolution kernel can be represented as a matrix $W \in R^{w \times d}$. Let the vectorization of the i th language symbol in the input be represented as s_i , and the input text can be represented by the matrix formula $S = (s_1^T, s_2^T, \dots, s_{|s|}^T)$. Then the convolution operation can be expressed as follows:

$$c_j = W \otimes S_{j:j+w-1} \quad (1)$$

$1 \leq j \leq |S| - w + 1$, c_j is a feature value extracted by the convolution operation between the word j and the word in the window of convolution kernel.

The effective features obtained by feature extraction of text by a convolution kernel are not comprehensive. In order to be able to extract more abundant information from text, a number of different convolution kernels are usually used, which can be expressed as a three-dimensional

tensor $\hat{W} = \{W_1, W_2, \dots, W_n\}$. The convolution operation of the convolutional layer can be expressed as follows:

$$c_{i,j} = W_i \otimes S_{j:j+w-1} \quad (2)$$

$1 \leq i \leq n$, The input text is subjected to the i -th convolution kernel convolution operation to obtain the feature vector $c_i = \{c_{i,1}, c_{i,2}, \dots, c_{i,|S|-w+1}\}$, Then all convolution kernels can get a total of n feature vectors, also known as Feature Map.

Sentiment classifier based on dropout algorithm and softmax classifier

The feature vector extracted by the convolution kernel still has many eigenvalues. If the feature analyzer is used directly to train the sentiment analyzer, the number of parameters that need to be optimized is still very large, and the model training is difficult and easily susceptible to overfitting. The pooling operation can further select features, effectively reduce parameters, and select some features that are most suitable for sentiment analysis as the final features of the sentiment analyzer.

The pooling layer compresses the input feature map, which can simplify the network computation complexity; on the other hand, the pooling operation can extract the main features. The essence of the pooling operation is sampling; which extracts features of the text from a sentence-level or higher-level text representation consisting of a vector representation of words. For different input texts, even when some of the expressions associated with the sentiment analysis of the current domain change, the output after the pooling operation is constant, which can enhance the robustness of the convolutional neural network. There is a certain anti-disturbance effect. Because the model uses a maximum pooling operation in the local neighborhood, the model can obtain the maximum degree of translational invariance of the text features. This feature is very important for sentiment analysis models because the model can effectively extract strong emotional features from the text. These features are in different positions in the text, and the pooling layer can select the most relevant features from the feature map through the pooling operation [10]. For the feature vector c_i obtained by the input text S through the i th convolution kernel, the maximum pooling operation can be expressed by the following formula:

$$p_i = \max(c_i) \quad (3)$$

Before the pooling layer, the feature vector obtained by the convolution kernel is connected to the pooling layer through a nonlinear mapping layer. Then Formula 3 can be rewritten as follows:

$$\begin{aligned} \hat{c}_i &= f(c_i) \\ p_i &= \max(\hat{c}_i) \end{aligned} \quad (4)$$

f represents a nonlinear mapping function.

At this time, the neural network of the feature extraction operation section can be represented as shown by the convolution and pooling layers in Figure 1. The resulting feature vector is $p_S = [p_1, p_2, \dots, p_n]$.

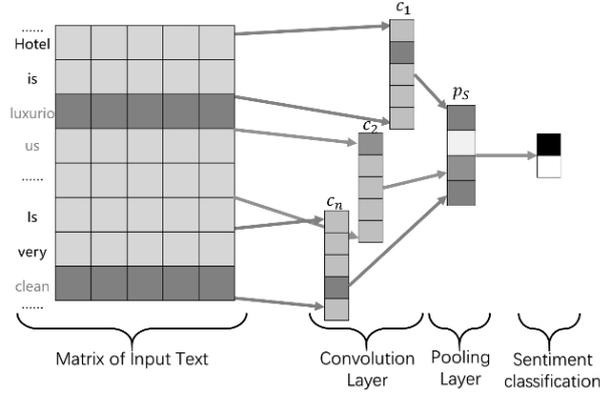


Figure 1: Sentiment analysis model based on convolution neural network

Pooling operations

The output obtained by the convolution operation and the pooling operation represents the advanced features of the original input text. In order to get the final sentiment analysis result, these features need to be input into the classifier through the fully connected layer for emotional tendency classification.

As shown in Figure 1, after the features of the text are extracted by the convolutional neural network, the feature vectors are input to the softmax classifier for classification. This article uses the dropout method to connect feature vectors with the softmax classifier. The input data of the neural network is randomly connected to the neurons of the next layer according to a preset proportional parameter ρ . In the convolutional neural network of this paper, the pooled feature vector is set to 0. Therefore, in the subsequent calculation process, only other elements that are not set to 0 participate in the operation to obtain the output of the network. The specific process of parameter update is as follows: First, the pooled feature vector is set to 0 according to the ratio ρ , and the element without 0 participates in the operation of the softmax classifier and obtains the gradient-optimized neural network outputting the neural network parameters. Then, the input vectors of all the samples in the sample set are sequentially accepted, and the connection of the input elements is randomly set to 0 in the same manner, and the elements participating in the training are selected until all the samples are used by the neural network to train the model parameters. Each time a sample is entered, the probability that the input element is set to 0 is ρ , so each time the input element of the participating operation is connected differently, the updated network weight parameters are also different. When using neural networks for prediction, it is necessary to multiply the parameters of the entire network by $1 - \rho$ to obtain the final classifier network parameters. Assuming that the eigenvector obtained by the convolution and pooling operations of the input text S is p_s , the way the dropout algorithm sets its element to 0 can be represented by the Bernoulli distribution B . First, use the Bernoulli distribution to generate a binary vector of the same dimension as c (the element takes only 0 or 1) r , ie $r \sim B(\rho)$, and then do the Hadamard product with the eigenvector. The vector that is ultimately entered into the classifier can be expressed as:

$$c_d = p_s \times r \quad (5)$$

Set the network parameter W_e of the softmax classification layer, and the bias term is b_e , then the neural network output can be expressed as:

$$o = f(W_e c_d + b_e) \quad (6)$$

Where f is the activation function. Then the probability that the entered text sentiment tends to i is:

$$p(i|\theta) = \frac{e^{o_i}}{\sum_{j=1}^N e^{o_j}} \quad (7)$$

where θ represents all parameters of the neural network, o_i represents the value of the i -th item of the output vector, and N represents the number of categories of the text. Let the sample set be expressed as Ω , then the model's optimization objective function can be calculated by the following formula.

$$L_{sen} = \sum_{i=1}^{|\Omega|} -\log p(y_i | S_i, \theta) + \lambda \|\theta\|_2^2 \quad (8)$$

Where λ is the parameter of the regularization term. In the actual experiment, the traditional convolutional neural network uses the stochastic gradient descent method to optimize the objective function. The update method of the parameter θ is:

$$\theta = \theta - \alpha \frac{\partial L}{\partial \theta} \quad (9)$$

α is the learning rate.

2.2 Piecewise pooling based convolutional neural network sentiment classification model

There is a deficiency in the process of extracting text features using the sentiment analysis model based on a traditional convolutional neural network. Regardless of whether or not the texts or Chinese, English or other natural language texts, the sentences have a certain structure, and traditional convolutional neural networks ignore the structural features of these sentences. As shown in Figure 2, both Chinese and English sentences can contain grammatical structures such as a subject, predicate and object.

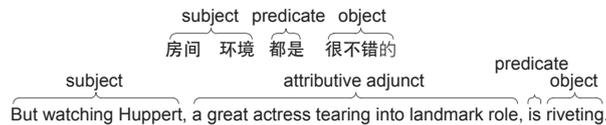


Figure 2: Structures in sentences

While existing deep learning method has difficulty parsing sentences, modeling the sentence structures will make the calculation process more complicated. However, if the simulation of grammatical structures can be added to network structures, the learning of sentence features will be significantly improved. Traditional maximum pooling extracts a maximum value from the features of a convolution kernel convolution, and does not make any distinguish between the grammatical elements of sentences. In response to this problem, this paper proposes the use of

Piecewise Pooling strategy for sentiment analysis. The piecewise pooling strategy divides the feature vector of a sentence into segments and performs maximum pooling operations on each segment. By segmenting the different grammatical structural components in a sentence, the segmented pooled feature vector can extract the features of the corresponding components in the sentence. The difference between segmented pooling and non-segmented pooling operations can be seen from the comparison in Figure 3 .

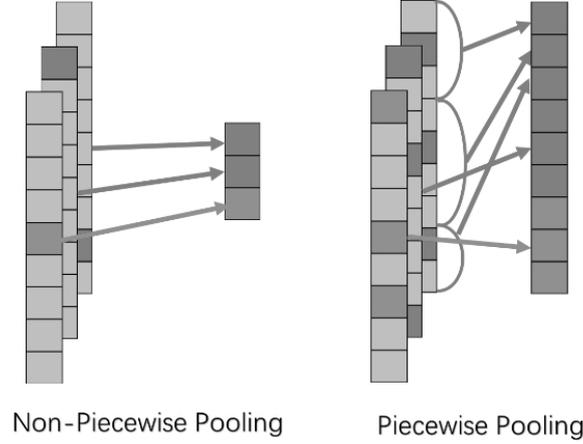


Figure 3: Piecewise pooling and non-piecewise pooling

The rest of the model is identical to the traditional convolutional neural network model described above, with the main difference being the pooling part. The feature vector c_i obtained after the convolution kernel convolution operation is different from the original pooled layer pair by Single Max-Pooling. The model in this section uses Piecewise Max-Pooling. The convoluted feature vector is divided into m segments, that is, $c_i = (c_{i,1}, c_{i,2}, \dots, c_{i,m})$. For different segments, the maximum pooling operation is performed separately, and the piecewise pooling operation may be expressed as Equation 10 .

$$p_i = [\max(c_{i,1}), \max(c_{i,2}), \dots, \max(c_{i,m})] \quad (10)$$

The piecewise pooling operation is performed on the feature vectors obtained by all convolution kernels, and then the feature vectors are nonlinearly transformed by the fully connected layer to obtain the feature vector of the input text, namely:

$$p_S = [f(p_1), f(p_2), \dots, f(p_i)]$$

$1 \leq i \leq n$, f represents a nonlinear activation function.

2.3 Multi-domain common sentiment feature extraction based on generative adversarial networks

Although the piecewise convolution neural network can effectively analyze the emotional tendency of the text, in order to make the neural network fully trained, the generalization of the network parameters is good enough, and more annotation data is needed. However, in each segment, there is relatively little data to be labeled, and data sparseness makes network training more difficult. Therefore, this paper proposes to use the generated adversarial network to extract common features related to emotions in different fields to alleviate the above problems.

The shared private model separates the feature space into shared and private spaces, but there is no guarantee that shared features cannot exist in the private feature space, and vice

versa. Therefore, some useful sharing features can be ignored in the shared private model, and the shared feature space is also vulnerable to certain task-specific information. Therefore, a simple principle can be applied to multitasking learning. A good shared feature space should contain more public information and no task-specific information. To solve this problem, we introduce adversarial training into the multitasking framework.

Introduction for generative adversarial network

First, We give a brief introduction to the generative adversarial network (GAN), generated by the Goodfellow et al. [3] was proposed in 2014, and quickly received extensive attention from the academic and industrial circles. The goal of GAN is to learn a generator distribution $P_G(x)$ that is fully matched to the real data distribution $P_{data}(X)$. Specifically, the GAN learns $P_G(x)$ by training a generation network G and a discrimination network D, wherein G is a sample generated from the distribution $P_G(x)$ to generate forged samples, and attempts to make the discrimination network unable to distinguish the forged samples from the real samples. ;D determines whether the sample is from $P_G(x)$ or $P_{data}(X)$, trying to distinguish the forged sample from the real sample. The two games are mutually optimized and alternately optimized. When the discriminator reaches the Nash equilibrium, that is, the discriminator cannot distinguish the forged sample from the real sample, the obtained $P_G(x)$ is consistent with $P_{data}(X)$. This minimum-maximum game can be optimized by the following risks:

$$\max_D \min_G V(D, G) = E_{x \sim P_{data}}[\log D(x)] + E_{z \sim p(z)}[\log(1 - D(G(x)))] \quad (11)$$

Due to the role of the discriminator, the sample generated by the generator is forced to approach the true distribution of the data progressively and unbiased. If the generator produces a sample distribution that perfectly matches the real data distribution, the discriminator will not be able to tell whether the input is from real data or a sample generated by the generator network, giving a probability value of 0.5 for all inputs, which is the Nash equilibrium. Although GAN was originally proposed for the generation of random samples, GAN can be used as a general tool for measuring the equivalence between distributions [11], which can be used for feature extraction.

Multi-domain common sentiment feature extraction GAN

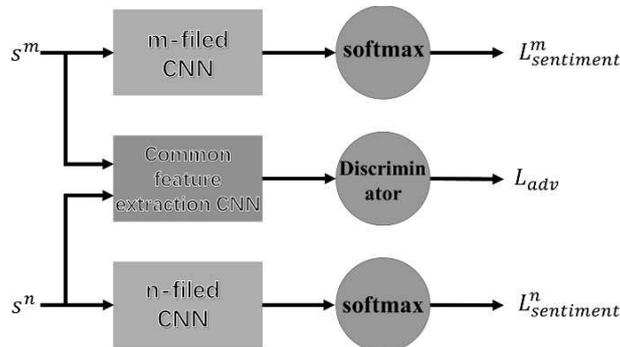


Figure 4: Multi-domain common sentiment feature extraction with generative adversarial network

Inspired by GAN, we propose a adversarial-sharing-private model for multi-domain common sentiment feature extraction. The shared convolutional neural layer first, and then the private

convolutional neural network extract features from the input text. The features extracted by the shared convolutional neural network are input into the discriminator network, trying to make the discriminator unable to distinguish which subdivision from which the feature comes from, and the discriminator tries to distinguish the features of different domains. Then, the extracted common features and the private features extracted by the piecewise domain are spliced as a feature to the accurate prediction of the text sentiment tendency in the sentiment orientation analyzer. This kind of adversarial training encourages the sharing space to be purer, ensures that the shared representation is not polluted, and can well preserve the common emotional features of different fields and eliminate the private emotional characteristics of the segment.

The domain discriminator network is used to map the shared representation of the sentence into a probability distribution and estimate which fields the encoded sentence comes from.

$$D(p_s, \theta_D) = \text{Softmax}(b + Up_s) \quad (12)$$

U is a learnable parameter and b is the bias.

In order to train the discriminator network, the parametric loss L_{adv} is used to optimize the parameters of the network to prevent the private sentiment features related to the emotional tendency of the specific domain from spreading to the shared sentiment orientation feature space. The counter-loss is used to train the model to extract common sentiment features in different domain texts so that the discriminator network cannot reliably predict which category of the currently entered text belongs to base on these features. The original form discriminator of GAN can only be used in the case of two classifications. In order to overcome this problem, this paper extends its s to multiple types of forms, so that the model of this paper can train sentiment analyzers in multiple fields together. The objective function against the network can be expressed as follows:

$$L_{adv} = \min_{\theta_s} (\lambda \max_{\theta_D} (\sum_{k=1}^K \sum_{i=1}^N d_i^k \log[D(E(S^k))])) \quad (13)$$

d_i^k indicates the domain tag to which the currently input text belongs. The basic idea is that, given a sentence, the shared convolutional neural network extracts a representation of an eigenvector misleading discriminator network. At the same time, the discriminator tries to classify the field of the input text correctly. After training, the shared sentiment feature extractor network and the discriminator network reach the Nash balance, and the discriminator cannot distinguish which domain these common features come from. The common feature extractor model can represent Figure 4.

Finally, the loss function of the sentiment analysis method model of PCNN and GAN proposed in this paper can be expressed as follows:

$$L = L_{sen} + \lambda L_{adv} \quad (14)$$

λ is a hyperparameter for balance sentiment analysis loss and adversarial loss.

3 Experiment settings

3.1 Data sets

Three data sets were used primarily in the experiment. The first is the Chinese Hotel Data Collection (Ctrip Hotel for short), which was compiled by Tan Songbo from the Insti-

tute of Computing Technology of the Chinese Academy of Sciences and is a commentary on the service of large-scale customers. The corpus was automatically collected from the Ctrip (<http://hotels.ctrip.com/>) and then organized. The corpus size is 10,000, including 7000 positive evaluation samples and 3,000 negative evaluation samples. The second data set is the English data set, from Stanford University’s emotional tree library, which labels the emotional polarity of each phrase in the sentence and the entire sentence. In this paper, only the emotional polarity of the sentence is extracted and classified. The data set contains a total of 11,855 sentences, including training sets, validation sets, and test sets containing 8544, 1101, and 2210 samples, respectively. The emotional polarity of each sentence is in the range of [0,1]. The smaller the score, the more the emotion tends to be negative. Otherwise, the emotion tends to be positive. The emotional scores of all sentences in the data set are manually labeled and then averaged. It has a good reliability. According to the distribution of the emotional scores of the dataset and its data description, the sample with the sample score between [0, 0.5) is first divided into negative samples, and the sample score is divided into positive emotions between [0.5, 1]. This is a coarser-grained sentiment analysis. Further divided into five levels of sentiment, including negative [0, 0.2), negative [0.2, 0.4), neutral [0.4, 0.6), positive [0.6, 0.8), positive [0.8, 1]. In order to make the experiment closer to a real production environment, and to verify the effectiveness of the sentiment analysis method of PCNN and GAN proposed in this paper, this paper uses the third data set. It is the review text (Dianping for short) from different fields that we crawled from the public service website (<http://www.dianping.com/>). Its data includes six segments of food, hotels, movies, entertainment, marriage, home improvement, and the comments are divided into different emotional tendencies based on the scoring information in the comments. This paper selects 30,000 reviews as the training set and 10,000 as the test set.

3.2 Data pre-processing

For the Chinese data set, the Chinese word piecewise package NLPIR developed by the Chinese Academy of Sciences is first used for Chinese word piecewise. Since the experiment in this article has a Chinese data set, you need to call the package participle. The English data itself contains independent words, so there is no need for word piecewise. Since the minibatch training model is used during training (multiple samples are learned at a time, and the text length of multiple samples may be different), so the length of the text needs to be fixed-length. Since the length of the natural language text is inconsistent, the longest sentence length l_max is first calculated. For sentences with a sentence length less than l_max , the text is uniformly filled with the $\langle \backslash s \rangle$ symbol to the length l_max (the vector of $\langle \backslash s \rangle$ is always set to 0), this will unify the text length. The purpose of this approach is to improve computational efficiency, and when the length of the data is uniform, the computational time overhead can be effectively reduced. At the same time, in order to ensure the features are extracted at the beginning and the end of the text during the convolution process, a certain number of $\langle \backslash s \rangle$ corresponding to the convolution kernel should be added at the beginning and end of the longest text as Padding.

3.3 Pre-training of word embedding

Word embedding is required before training of the model. Word embedding acts as a distributed representation of words as an input suitable for neural networks. Many current studies have shown that executing word embedding pre-training on a large-scale corpus, and then applying the obtained word embedding to subsequent training, can speed up the convergence of neural network models and achieve a better local optimal solution. In this paper, the word2vec algorithm is used to pre-train word embedding. The word embedding of this algorithm shows better performance in many natural language processing tasks, and has higher efficiency. This

paper chooses the Skip-gram model and the Negative Sampling model to pre-train the word embedding of Chinese and English words. The pre-training of Chinese word embedding uses the text content crawled on Baidu Encyclopedia. The English word vector is pre-trained using New York Times.

3.4 Setting of experimental parameters

In the training optimization process of the model, we use the Adam optimizer to train and optimize the model. The parameter settings of the optimizer adopt the recommended parameter values. In this paper, the model mainly has the following hyperparameters: the dimension d of the word vector, the number n of convolution kernels, the ratio ρ of the Dropout algorithm, and the piecewise number t of the piecewise pooling in the piecewise convolutional neural network. In order to obtain the optimal hyperparameter setting, this paper uses grid-search to determine the value of some hyperparameters. The word embedding dimension N selects a value from $\{50, 100, 200, 300\}$; the number n of convolution kernels takes a value in $\{100, 150, 200\}$. The ratio ρ of the Dropout algorithm is 0.5 based on experience. In the experiments in this paper, multiple experiments were performed using these parameters, and then the average of the results was obtained. The number of segments of the piecewise pooling was performed in this paper. The results were given in the analysis.

4 Results and analysis

In order to verify the effectiveness of the sentiment analysis method of the fusion piecewise convolution neural network and the generated adversarial network, the method of this paper is compared with some mainstream emotional analysis baseline methods. The parameter sensitivity experiment is carried out on the piecewise pooling segment number hyperparameters which are important to the performance of the model in the hyperparameters proposed in this paper, so as to find the more suitable parameter settings in the application scenario.

4.1 Comparison method

In order to verify the validity and correctness of the proposed model, this paper selects a model based on traditional methods and a neural network method such as RNTN proposed by Richard Socher et al. as a baseline method. The first method of comparison is the naive Bayesian method (NB for short) using the word bag feature. The second method is to use the word bag feature as input to perform emotion classification using a support vector machine (SVM for short) classifier. The third method is the naive Bayesian method (BiNB for short) of the bag feature obtained using the binary grammar language model. The fourth method is to use the average word vector of the sentence as the input feature and use the fully connected network as the classifier (VecAvg for short). The fifth method is the Recurrent Neural Network (RNN). The sixth method is a recurrent neural network (MV-RNN) with a semantic transformation matrix [7]. The seventh method is based on a tensor-based cyclic neural network (RNTN). The last comparison method is the traditional convolutional neural network.

4.2 Analysis of experimental results

First, because there is no multi-domain information on the Ctrip Hotel dataset and the Stanford English Emotional Tree Database dataset, the performance of the segmented convolutional neural network model and the baseline model is first compared. It can be seen from the results in Table 1 that the neural network method generally has higher performance than

Table 1: Text classification precision on different data sets(%)

Model	Stanford		Ctrip
	Two levels sentiment	Five levels sentiment	
NB	81.8	41.0	80.2
SVM	79.4	40.7	86.7
BiNB	83.1	41.9	85.9
VecAvg	80.1	32.7	82.1
RNN	82.4	43.2	87.8
MV-RNN	82.9	44.4	87.6
RNTN	85.4	45.7	89.3
CNN	81.9	45.6	88.5
PCNN	85.4	45.9	89.7

Table 2: Text classification recall on different data sets(%)

Model	Two levels sentiment	Five levels sentiment	Two levels sentiment(Reduced)
NB	72.1	35.7	66.1
SVM	72.6	33.2	61.2
BiNB	75.2	36.6	62.4
VecAvg	74.1	36.3	62.3
RNN	77.3	40.2	69.7
MV-RNN	78.0	42.2	73.1
RNTN	80.3	42.6	75.7
CNN	79.1	41.5	73.5
PCNN	80.8	43.2	74.9
PCNN+GAN	83.3	47.9	80.1

the conventional method. Especially in the five-level sentiment analysis with finer granularity, the neural network method can obtain the key features of the text well. An important reason for the BiNB method to achieve better results is that the binary grammar model considers a certain combination of semantics, but with it is a large computational overhead. At the same time, unlike the method of cyclic neural networks, the use of maximum pooling in convolutional neural networks can automatically extract the features most relevant to sentiment analysis tasks, and has positive significance for text sentiment analysis tasks, so it has achieved good results. Traditional methods such as BiNB can only extract combined features from adjacent words, and traditional convolutional neural networks cannot model grammatical structures. The segmented convolutional neural network method has the best effect on both datasets because it simulates the grammatical structure information of the text, supplements the original emotional words, and extracts the combined semantic features of different positions.

In the Dianping data, the data comes from different fields. We compare the performance of different methods in this data to fully demonstrate the effectiveness of the proposed PCNN and GAN methods. It can be seen from Table 2 that the fusion segmented convolutional neural network and the sentiment analysis method of the generated adversarial network proposed in this paper have achieved the best results. The comments crawled from the Dianping contain different subdivisions. The comment text in each field is relatively small, and the comment text itself is relatively short. Therefore, samples in different fields have a certain degree of sample sparseness.

This paper also reduces the number of samples in each field, making sparse data a more serious side-effect. It can be seen that all methods have performance degradation problems to varying degrees, but the method proposed in this paper still achieves good results. The stability of the proposed method in the case of multi-domain data sparseness is fully explained.

In the piecewise convolutional neural network, a very important parameter is the piecewise pooling segment number. Since the piecewise is a simulation of the text’s grammatical structure, the piecewise number determines the validity of the text structure information extraction. In this paper, the parameter sensitivity test is carried out on the three data sets for the number of segments. The results are shown in Figure 5. It can be seen that a reasonable piecewise can better obtain the grammatical structure information of the text, and an excessively large number of segments destroys the structure of the sentence itself. The pooling operation cannot effectively extract the features most relevant to sentiment analysis, making the sentiment analysis effect worse. Therefore, it is very important to select the appropriate segment number to simulate the grammatical structure of the sentence.

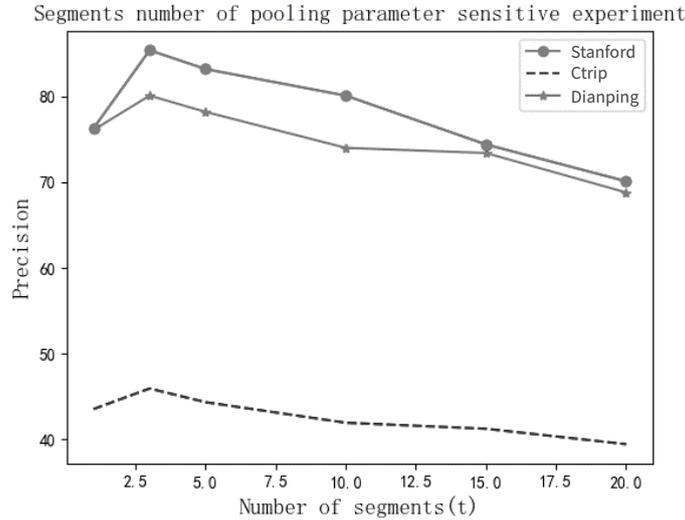


Figure 5: Segments number of pooling parameter sensitive experiment

5 Conclusion

In this paper, we show the difficulty of maintaining optimality in different fields for traditional text sentiment analysis methods. Different methods have different applicability problems in different fields. The original convolutional neural network model ignores the sentence structure information that is very important for text sentiment analysis, and it is susceptible to over-fitting. In this paper, the piecewise pooling strategy is adopted to enable the deep learning-based convolutional neural network model to model the sentence structure and segment the main features of different structures. It combines the structural information and domain information of the text to analyze the emotional tendency of the text; and uses the Dropout algorithm to enhance the generic ability of the model.

At the same time, user feedback on services involves many different areas, with less data in each field. Less data makes the training of convolutional neural networks more difficult, and the parameters cannot be fully optimized, resulting in under-fitting of the model. Therefore, in the case that data volume expansion is more difficult, in order to alleviate the sparseness of data, this paper uses the generated adversarial network to extract common features of texts in different

fields. This enables the model to acquire common features related to emotions in feedback from different fields, and to enhance the generalization ability of the model with less training data. Experiments on different data demonstrate the effectiveness of this method.

Bibliography

- [1] Chen, X. (2012); *Research on Sentiment Dictionary based Emotional Tendency Analysis of Chinese MicroBlog*, Huazhong University of Science & Technology, 2012.(in Chinese)
- [2] Cui, H.; Mittal, V.; Datar, M. (2006); Comparative experiments on sentiment classification for online product reviews, *AAAI*, 6, 1265-1270, 2006.
- [3] Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M. et al. (2014); Generative adversarial nets, *Intl Conf. on Neural Information Processing Systems*, MIT Press, 2672-2680, 2014.
- [4] Li, D.; Qiao, B.; Yuanda, Y.; et al. (2008); Word Orientation Recognition Based on Semantic Analysis, *Pattern Recognition & Artificial Intelligence*, 2008.(in Chinese)
- [5] Luong, T.; Socher, R.; Manning, C.(2013); Better word representations with recursive neural networks for morphology, *Proc. of the Seventeenth Conf. on Computational Natural Language Learning*, 104-113,2013.
- [6] Socher, R. (2014); Recursive deep learning for natural language processing and computer vision, *Stanford University*, 2014.
- [7] Socher, R.; Huval, B.; Manning, C.D. et al.(2012); Semantic compositionality through recursive matrix-vector spaces, *Proc. of the 2012 joint conf. on empirical methods in natural language processing and computational natural language learning*, Association for Computational Linguistics, 1201-1211, 2012.
- [8] Socher, R.; Chen, D.; Manning, C.D. et al (2013); Reasoning with neural tensor networks for knowledge base completion, *Advances in neural information processing systems*, 926-934, 2013.
- [9] Socher, R.; Perelygin, A.; Wu, J. et al.(2013); Recursive deep models for semantic compositionality over a sentiment treebank, *Proc. of the 2013 conf. on empirical methods in natural language processing*, 1631-1642, 2013.
- [10] Srinivas, S.; Sarvadevabhatla, R.K.; Mopuri, K.R. et al. (2016); A taxonomy of deep convolutional neural nets for computer vision, *Frontiers in Robotics and AI*, 2:36, 2016.
- [11] Taigman, Y.; Polyak, A.; Wolf, L. (2016); Unsupervised Cross-Domain Image Generation, *arXiv preprint arXiv*, 1611.02200, 2016.
- [12] Zeng, D.; Liu, K.; Lai, S. et al. (2014); Relation classification via convolutional deep neural network, *Proc. of COLING 2014, the 25th Intl Conf. on Computational Linguistics: Technical Papers*, 2335-2344, 2014.
- [13] Zhang, W., Zhang, Z., Chao, H.C. et al. (2018); Kernel mixture model for probability density estimation in Bayesian classifiers, *Data Mining and Knowledge Discovery*, 32(3), 675-707, 2018.
- [14] Zhang, W.; Zhang, Z.; Qi, D. et al. (2014); Automatic crack detection and classification method for subway tunnel safety monitoring, *Sensors*, 14(10), 19307-19328, 2018.