Computer Modeling in Engineering & Sciences
DOI:10.32604/cmes.2020.010948
Article

An Emotion Analysis Method Using Multi-Channel Convolution Neural Network in Social Networks

Xinxin Lu[1,*] and Hong Zhang[2]

[1]School of Computer Science and Technology, Zhoukou Normal University, Zhoukou, 466001, China
[2]School of Network Engineering, Zhoukou Normal University, Zhoukou, 466001, China
[*]Corresponding Author: Xinxin Lu. Email: zkxxl2016@126.com
Received: 09 April 2020; Accepted: 05 June 2020

Abstract: As an interdisciplinary comprehensive subject involving multi-disciplinary knowledge, emotional analysis has become a hot topic in psychology, health medicine and computer science. It has a high comprehensive and practical application value. Emotion research based on the social network is a relatively new topic in the field of psychology and medical health research. The text emotion analysis of college students also has an important research significance for the emotional state of students at a certain time or a certain period, so as to understand their normal state, abnormal state and the reason of state change from the information they wrote. In view of the fact that convolutional neural network cannot make full use of the unique emotional information in sentences, and the need to label a large number of high-quality training sets for emotional analysis to improve the accuracy of the model, an emotional analysis model using the emotional dictionary and multi-channel convolutional neural network is proposed in this paper. Firstly, the input matrix of emotion dictionary is constructed according to the emotion information, and the different feature information of sentences is combined to form different network input channels, so that the model can learn the emotion information of input sentences from various feature representations in the training process. Then, the loss function is reconstructed to realize the semi supervised learning of the network. Finally, experiments are carried on COAE 2014 and self-built data sets. The proposed model can not only extract more semantic information in emotional text, but also learn the hidden emotional information in emotional text. The experimental results show that the proposed emotion analysis model can achieve a better classification performance. Compared with the best benchmark model gram-CNN, the F1 value can be increased by 0.026 in the self-built data set, and it can be increased by 0.032 in the COAE 2014 data set.

Keywords: Emotion analysis model; emotion dictionary; convolution neural network; semi supervised learning; deep learning; pooling feature; feature mapping

1 Introduction

As an interdisciplinary comprehensive subject involving multi-disciplinary knowledge, emotional analysis has become a hot topic in psychology, health medicine and computer science. It has a high comprehensive and practical application value [1–5]. Emotion research based on the social network is a relatively new topic in the field of psychology and medical health research. At present, with the new era of user self-media data driven by Web 2.0 becoming more and more popular, there are more and more online comments from forums, blogs, microblogs, Tencent chat or other social networks, and millions of college students express their feelings, feelings, experiences and opinions. Their motivation is to share information, get feedback from friends and classmates, keep in touch with each other, and pay attention to relevant news. Therefore, college students have created a rich information flow including text, multimedia files, emoticons and other contents. This kind of unbiased, real-time and emotional user self-Media data comes from students' experience of teaching and research services, life experience and other aspects provided by the University, which directly affects the future decision-making of students. In analyzing the information released by these complex college students, the recognition of attitude and topic is very important. For example, information published on social media will affect the decision of potential college students whether to apply for the school. In addition, students can use this information to understand the advantages and disadvantages of teachers, and to understand what useful courses are. This information is also very effective for university teachers and administrative departments to master students' emotions and opinions. Therefore, when university administrators pay attention to all aspects of students' emotions, they are easy to make correct decisions to deal with students' problems. In addition, it allows teachers to learn more about students without paying for the survey. University management should understand current trends and improve services by analyzing students' opinions. Therefore, it is always considered important for decision-making to understand students' ideas correctly [6].

Emotional analysis is a comprehensive application of natural language processing, viewpoint extraction and text analysis, which can identify and retrieve emotional polarity from natural language texts [7–9]. This early approach to emotional analysis focuses on determining the overall emotional orientation (i.e., positive, neutral, or negative) or emotional polarity (i.e., one to five stars) of a comment [10]. Emotional polarity can be formed at different levels: Document level [11], sentence level and vocabulary level [12]. The document level is based on the whole document. Sentence level extracts and determines the emotion of each subjective sentence in the text. At the level of words, each word in the text is analyzed and classified. In this paper, sentence level is considered, because microblogs and other social media are short texts.

However, as far as we know, no researchers analyze and study the social media data of Chinese college students. In the field of education, researchers focus on e-learning by exploring students' feelings and feelings, but e-learning data is lack of interactivity, which cannot reflect the whole college life of students. This gap in research limits our ability to analyze students' emotions, which may affect their personal and intellectual growth. In addition, in previous studies, emotional polarity was limited to the judgment level of positive, negative or neural emotions, and the emotional intensity of these emotions and opinions has also been lack of research. At the same time, unstructured text brings difficulties to automatic emotion analysis, which makes the development of this technology challenging [13].

Finally, emotional expression is usually indirectly or implicitly related to the emotional expression of text information released by students. In this perspective, our research focus is to analyze the emotional data about college students in Tencent's circle of friends by constructing college students' emotional vocabulary. Tencent talks about data similar to Weibo. People talk about their feelings in life and record them bit by bit. Because of its group, spontaneity and naturalness in Chinese young students, it was selected as the source of our experimental data set. In Tencent, the "wall" is a service for students to release their own information, content, status, etc. There are several categories of SMS: Status information (each student writes on his or her own wall), posts on other people's walls, and comments on others. Usually, a student's wall is visible to his/her friends, who can comment on specific information or express their feelings.

In particular, we will focus on the emotional analysis of text. In order to get more hidden information and express the importance of each word in a sentence, a multi-channel semi supervised convolutional neural network model is proposed. By combining the features of emotional words and the position of words in sentences in college students' emotional dictionaries, the emotional text is vectorized in two different ways, which is used as the input of convolution neural network, and construct convolution neural network to realize semi supervised convolution neural network.

2 Related Research

Emotion is a complex word, which is related to personal judgment. Traditionally, emotion research has been a field of philosophy or psychology. At present, emotion analysis has been studied in the field of computer and social science, and a software system for automatically dealing with emotion problems has been established. This part discusses our research from two aspects: Emotional Analysis and college students' emotion.

Emotional analysis has been applied in many research fields, such as reviewing customer products, monitoring reputation in social networks, tracking people's feelings towards politicians, promoting marketing activities, etc. [14,15]. For example, we can find the emo value header analyzer, which analyzes the suggested headers entered by users to assign them emo value scores [16]. In [17], a semi supervised learning algorithm is proposed, which is called deep trust network with feature selection (DBNFS). The use of chi square based feature selection reduces the complexity of vocabulary input, because some irrelevant features are filtered, which makes the learning stage of DBN more effective. Literature [18] proposed emotion analysis technology to identify and monitor potential emotions in texts written by developers in GitHub project problems.

There are many methods to calculate emotion analysis from text [19], among which machine learning method [20] and dictionary-based method [21] are the main ones. Many machine learning methods were originally developed for data mining. They are used to learn the underlying patterns from the sample text data, and try to classify the new unmarked data for emotional classification [22]. The most commonly used machine learning methods are: Bayesian method [23], maximum entropy, support vector machine and k-nearest neighbor. These methods are the most widely used in the field of NLP emotion analysis because they are easy to adapt across domains and languages [24].

Many affective analysis tools rely on manual construction or automatic derivation of polar dictionaries, especially in English. Affective dictionary is defined as a list of positive and negative opinion words or affective words in English. Suppose that if we use it to deal with Chinese, such a dictionary can also be used in any other language. As recent research has revealed, students' emotions have an impact on their performance in college learning and life. In the field of student curriculum evaluation feedback. They suggest classifying students' views on the course and comparing the accuracy of different classifiers in their online forum posts. In this paper, a multi-channel semi supervised convolution neural network model combining multi features is proposed, which is used as the input of convolution neural network, and the method in [25] is used to construct convolution neural network to realize semi supervised convolution neural network.

3 Emotional Analysis Model

3.1 Task Definition

Emotional analysis is mainly based on the given sentence $s = (w_1, w_2, \ldots, w_n)$ (w_i is the i-th word in s, n is the length of s), using the hidden feature information of the word sequence in the sentence to determine the emotional tendency of the sentence. In this paper, a semi supervised convolution neural network model is constructed by using the method proposed in [25]. By using word vectors and different features as the input of convolutional neural network, the structure is spliced. This multi feature input can enable the network model to learn different feature information and optimize the super parameters of the model during the training process, so as to effectively obtain more hidden information in sentences [26,27].

3.2 Part of Speech Tagging in Emotional Dictionary

In this paper, based on the existing emotion vocabulary ontology, this paper constructs a dictionary for students' emotion analysis. There are 56610 emotional words (15093 commendatory words, 17847 derogatory words and 21670 objective words). Because it is necessary to use appropriate Chinese dictionaries and methods to carry out emotional analysis of their text information, this section will introduce the construction steps of College Students' dictionaries.

Research shows that emotion analysis usually depends on characteristic domain, because the emotional polarity of related domain terms may vary according to the context in which they are used. Therefore, sentics is an explicit and implicit emotional vocabulary, which aims to detect the emotion of college students in the text, calculate their polarity intensity, and predict their personality characteristics. Many polar words of this word are domain and context specific. This will significantly improve the accuracy of emotional analysis of college students. For example, from "going to class," we construct the following vocabulary, such as "teacher, classmate, homework, etc." refers to the relationship between people and events, and other words and phrases, such as "interesting, boring, severe" refers to the positive and negative emotions of classroom teaching.

Combined with corpus filtering, web search and dictionary extension technology, this paper constructs sentics manually and computer aided. Based on the experience of experts and word frequency, a dictionary is built. Delete words with a total frequency of less than 5 or more than 30. The reason is that when the frequency is greater than 30, they are too common, and when the frequency is less than 5, they are too unique. In this paper, affective words are defined as a five tuple containing word names. The affective categories of words can be happiness, love, disgust, sadness, etc., the part of speech includes nouns, verbs, adjectives, etc., and the emotional intensity and polarity include positive, negative, neutral, etc. Therefore, a word is formally defined as a 5-tuple: (w, wsc, wc, ws, wp). Among them, w is the word name, wsc is the word emotion classification, wc is the part of speech, ws is the word emotion intensity, wp is the word polarity.

The related emotional words are mainly defined as shown in Tab. 1. These classifications basically cover all the main emotions of students in school life. Parts of speech can be divided into nouns, verbs, adjectives, adverbs, pronouns, etc. Traditionally, adjectives, verbs and specific nouns express emotions in texts. However, other features of vocabulary are also used to determine the emotional polarity in the text.

Table 1: Emotional classification of emotional words

Serial number	Sentiment classification	Examples of words	Polarity	Describe	Example	Part of speech
1	Happiness	Happy, excited,	Pos	The name of the thing	Mobile phone, textbook	Noun
2	Feel at ease	Calm down, relax,	Pos	Actions of people or things	Sleep, play	Verb
3	Admire	Respect, adore, admire, support	Pos	The state or quality of a person or thing, used to modify words.	Diligent and friendly	Adj
4	Praise	Praise	Pos	A word used to modify the degree of expression of a verb or adjective.	Very much. Very, a little	Adv
5	Believe	Believe, trust, confidence	Pos	High frequency words appear in the network	Decline embarrassed, loser	Nw
6	Love	Love, like,adore	Pos	Negative words	No	No
7	Wish	Wish, bless, celebrate	Pos	Words in place of other words	They, we	Pronoun
8	Miss	Miss, care	Pos	Words indicating sentence and mood in writing	!, ○, ?	Punc
9	Surprise	Surprise, amazed, curiosity	Pos	Words expressing mood	Ah, wow, oh	Tone
10	Sympathy	Sympathy, pity, poor	Pos	The name of a place	Beijing, Harbin	Addr
11	Anger	Angry, ignition	Neg	Human name	Zhang San, Li Si	Pername
12	Sadness	Sadness, grieved, grief,	Neg	Quantitative words	5, 4, 31	Numb
13	Cold	Apathy, indifference, coldness	Neg	Represents the relationship between sentences	But, and	Turn
14	Bashful	Ashamed, shy shame	Neg	Name of organization or group	Zhoukou Normal University	Agent
15	Doubt	Doubt, suspicion,	Neg	To express an emotion with symbols	^_^	Emotion
16	Fear	Terrified, fear	Neg	Time words	February 2018	Time

Other parts of speech (parts of speech) listed in Tab. 1 are also for improving emotional classification, such as adverbs, negatives, punctuation and conjunctions. In this study, adverbs are used to increase or decrease the intensity of emotion as an emotional booster. Degree adverbs can be used to calculate the emotional value according to the hierarchy. The first level maximizes and enhances emotions, such as “extraordinary,” “absolute” and “unbelievable,” while the second level is weaker, such as “very,” “true” and “certain.” The third kind of emotional enhancement is the weakest, such as “a little,” “a little,” “a little.” When students only use adjectives or nouns to express emotional intensity, the calculation of intensity is not accurate. Therefore, emotional intensity should increase or decrease with the appearance of degree adverbs. Negative words, such as “no,” change the direction of emotional words. For example, “I don’t like him.” The emotional value of “like” is positive, but the combination of “not” and “like” is a derogatory emotional word. This matching affects the polarity of the whole sentence, so the polarity of the sentence is negative.

This paper also considers connection rules. Generally speaking, when there are some conjunctions in a sentence, such as “although,” “but,” “and” and “then,” the sentence can express multiple polarity. Through the use of connection rules, accurate semantics is easier to be accurate, and emotional analysis is more accurate. There are two main conjunction rules, one is the progressive relationship “a sentence, and B sentence,” the other is the turning relationship “a sentence, but B sentence.” The first conjunction rule is that if the emotional polarity of a sentence is positive, it will strengthen the positive emotion of B sentence. The second conjunction rule is that if sentence a is positive, then sentence B should be negative. For example, progressive relationship: “he is not only handsome, but also kind-hearted.” The turning word “and” in this sentence strengthens the positive emotion of “he is not only handsome.” For example, “he is handsome, but he is evil inside.” The conjunction “Dan” negates the positive emotion of “he is very handsome,” so the emotional polarity of the whole sentence should be negative emotion. The conjunction rules not only ensure that the relationship between words is considered, but also the relationship between sentences, and also enhance the accuracy of emotional analysis.

Similarly, the combination of exclamation and punctuation is also an emotional booster. For example, “it’s finally time off!.” Exclamation “La” and punctuation “!” Both increase the positive emotional intensity of sentences. If the part of speech of a sentence is recognized, then we can use the following model to calculate the polarity intensity to get the final score. The model is as follows:

$$sentiscore = \sum_{i=1}^{n} \sigma * wv_i * wp_i * ws_i + \alpha + \delta + \beta \tag{1}$$

Here, wv_i is the strength of adverbs, wp_i is the polarity of emotions, the value is 1 (positive), -1 (negative), 0 (neutral). The ws_i is the emotional intensity of emotional words, the value is 0 to 9. The σ is the value of negative words, and the value of -1. The α is the value represented by conjunctions. If the conjunctions are progressive, the value is the same as the emotional intensity of the emotional words before the conjunctions. If it is a turning relationship, the value is opposite to the emotional intensity of the emotional words before the conjunctions. And δ is the emotional intensity of emoticons, β is the emotional intensity of punctuation and interjection.

3.3 Construct Word Vector

In this paper, different input channels are used to receive the combination of different feature information, so that the model can learn more emotional feature information in the training process, effectively identify the emotional polarity of emotional text sentences, and improve the emotional classification performance of the network model. For each word in the sentence, this paper maps it to a multi-dimensional continuous value vector, and each word w_i in sentence $s = (w_1, w_2, \ldots, w_n)$ can be mapped to an m-dimensional vector, namely $V_i \in R^m$. The term vector matrix of the dataset is defined as $E = R^{m\times|V|}$, where $|V|$ is the size of the term set of the dataset.

This paper uses the emotional words database in Section 3.2 to mark the part of speech of the special words in the input sentence, as shown in Tab. 1. By tagging the special words in sentences, and paying attention to learning the characteristic information of special words in the training process of network model, we can make full use of these special words in emotion classification. In this paper, through the vectorization operation of different part of speech tagging, part of speech tagging mapping is defined as $tag_i \in R^k, tag_i$ is the i-th part of speech vector, k is the part of speech vector dimension. Convolutional neural network can learn more advanced natural language feature information by learning different part of speech tagging to fine tune the component parameters of part of speech vector.

Because the emotional information contained in the sentence is limited, and the same word in the sentence may appear in different positions, which may contain different information, so the position of the entry in the sentence is also an important feature of the short text. The position value of the i-th entry w_i in sentence s is

$$p(w_i) = i - len(s) + maxlen \quad (2)$$

Here, $p(w_i)$ is the position value of entry (w_i) in sentence s; $len(s)$ is the length of sentence s; maxlen is the maximum length of input sentence.

3.4 Semi Supervised Convolutional Neural Network

Convolution neural network is a kind of feedforward neural network including convolution calculation and with depth structure. It is one of the representative algorithms of depth learning [28,29]. Convolutional neural network has good classification performance in emotion analysis task, but it depends on the availability of a large number of training data, which limits its applicability. In this paper, we use a semi supervised learning inversion strategy proposed in reference [25], which is suitable for a variety of depth neural networks. This method does not add extra super parameters, and the stability and robustness of input noise and gradient update in semi supervised tasks are good.

Let x be the input of the network, the output of each layer of the network is $z^{(1)}(x)$, the activation function is Relu, and the transformation of activation function is $f^{(l)}_{\theta(l)}$. Therefore, the output of the network is

$$z^{(1)} = f^{(l)}_{\theta(l)}\left(z^{(l-1)}(x)\right) \quad (3)$$

Here, $\theta^{(l)}$ represents the parameters of the l-th layer in the network, which are obtained through network learning. The variance loss function is reconstructed, and the loss function after reconstruction is

$$L_R(x) = \left\| x \frac{dz^{(l)}(x)^T}{dx} z^{(l)}(x) \right\|^2 \quad (4)$$

The reconstructed loss function acts as a data-driven network regularize, which makes the information of unlabeled samples be taken into account in the input of the network, so that the square difference between unlabeled samples and the output of the network modeling for the samples is the smallest, which is the opposite of the standard regularization. In addition, the cross-entropy loss function of unlabeled samples is

$$L_E\left(\hat{y}(x)\right)-\sum_{c=1}^{C}y(x)_c\log\hat{y}(x)_c \tag{5}$$

Here, C is the number of classification categories; $y(x)_c$ is the real sample label; $\hat{y}(x)_c$ is the network prediction label.

For labeled supervised learning, minimizing entropy is a part of the optimal output distribution, which makes the output distribution in the correct classification category. For unsupervised learning, although the category is unknown, the optimal value is to minimize the entropy distribution. Given a set of labeled data set $D_s=\{(x_n,y_n),n=1,2,\ldots\}$, the internal parameters of the network are learned by cross entropy $L_{CE}(x,y)$ and unlabeled data set U_s to get the unsupervised examples of known labeled clustering. Therefore, the semi supervised loss function is

$$L(x,y)=\alpha 1_{\{y\neq\varnothing\}}L_{CE}(x,y)+(1-\alpha)\beta 1_{\{y\neq\varnothing\}}L_E(x)+(1-\alpha)(1-\beta)L_R(x) \tag{6}$$

Here, α and $\beta\in[0,1]^2$ are weighting coefficients of supervision loss and unsupervised loss respectively.

Each loss is re normalized to ensure that their impact is evenly distributed with reference to the overall loss. In this case, the semi supervised loss function is

$$L(x,y)=\frac{1}{\log C}\left(1_{\{y\neq\varnothing\}}L_{CE}(x,y)+1_{\{y\neq\varnothing\}}L_E(x)\right)+\frac{1}{D}L_R \tag{7}$$

Here, D is the dimension of input x. The loss functions $L_E(x)$ and $L_{CE}(x,y)$ have the same order of magnitude. In this case, $L_{CE}(x,y)\approx\log C$ and $L_E(x)\approx\log C$ can be initialized to $\hat{y}(x)\sim\pi(C)$, in which $\pi(C)$ Dirichlet's parameter distribution is uniform, and the range of $L_R(x)$ depends on the infinite norm of the input x under consideration.

Since Eq. (7) is not stable for input noise and gradient update, in order to improve its stability, the loss of each layer is defined as Eq. (2), let $\frac{1}{L}\sum_{l=0}^{L-1}\frac{1}{D^t}{L^l}_R(x)$, then the semi supervised loss function is

$$L(x,y)=\frac{1}{\log C}\left(1_{\{y\neq\varnothing\}}L_{CE}(x,y)+1_{\{y\neq\varnothing\}}L_E(x)\right)+\frac{1}{L}\sum_{l=0}^{L-1}\frac{1}{D^t}{L^l}_R(x) \tag{8}$$

3.5 Multichannel Semi Supervised Convolutional Neural Network

The multi-channel semi supervised convolution neural network model combined with multi features is shown in Fig. 1. For the embedded layer of the network, this paper adds word features, part of speech features and location features to the input layer of the network. These three features form four different channels through simple splicing in the input layer of the network. The convolution layer mainly obtains the local features of different channel inputs to form the feature information graph. The convolution operation is carried out in the way of multi windows and multi convolution kernels for each different channel, and mean pooling is used to obtain the important feature information in each channel. The local features obtained from different channels are merged into a feature vector by using a merging layer, and then the feature vector is input to the next layer, that is, the hidden layer. In this paper, the local feature vector is extracted in the hidden layer to get the relationship between the local features of different channels, and the relationship between different channels is learned through the weight matrix. In order to get the classification results of the sentences to be classified, the output layer uses the softmax function to output the classification results. In this paper, the input layer of the network is added with emotional dictionary word features, part of speech features and location features. These three features are simply spliced in the input layer of the network to form four different channels. In order to simplify the network model, this paper uses a simple splicing operation to realize the feature combination in the experiment, as follows:

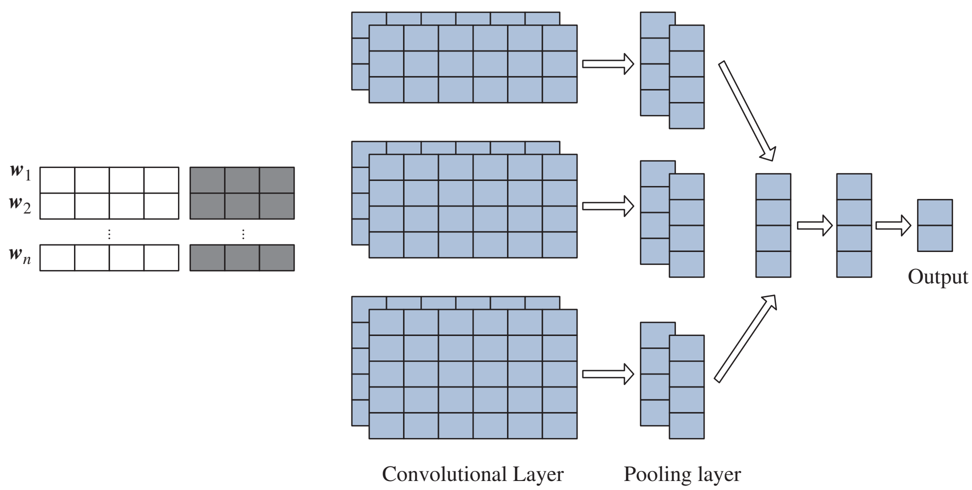

Figure 1: Results of multi-channel semi supervised convolution neural network

$$\boldsymbol{V}_1 = \boldsymbol{tag} \oplus \boldsymbol{position} \tag{9}$$

$$\boldsymbol{V}_2 = \boldsymbol{w} \oplus \boldsymbol{position} \tag{10}$$

$$\boldsymbol{V}_3 = \boldsymbol{w} \oplus \boldsymbol{tag} \oplus \boldsymbol{position} \tag{11}$$

$$\boldsymbol{V}_4 = \boldsymbol{w} \oplus \boldsymbol{tag} \tag{12}$$

Here, $\oplus$ is the join operator; ***tag*** is the part of speech feature; ***position*** is the position feature; ***w*** is the common word vector.

For different channels, this paper uses different convolutions to extract the feature information of sentences. Then the most important feature information is obtained through the pooling operation of the pooling layer. The eigenvector graph and pooled eigenvector obtained by convolution operation are as follows:

$$c_j = \mathrm{relu}\left(\boldsymbol{W} \cdot \boldsymbol{V}_i + \boldsymbol{b}\right); \quad \hat{\boldsymbol{C}}_j = \left[\hat{\boldsymbol{c}}_1, \hat{\boldsymbol{c}}_2, \ldots, \hat{\boldsymbol{c}}_d\right] \tag{13}$$

Here, $j = 1, 2, 3, 4$ is the channel subscript; $\boldsymbol{W}$ is the weight matrix of convolution kernel; $\boldsymbol{V}_i$ is the characteristic matrix of splicing; $i = 1, 2, 3, 4$; $\boldsymbol{b}$ is the offset matrix; d is the number of convolution kernel. By merging the eigenvector graphs of different channels from the merging layer, the eigenvector is obtained.

$$\hat{\boldsymbol{c}} = \left[\hat{\boldsymbol{c}}_1, \hat{\boldsymbol{c}}_2, \ldots, \hat{\boldsymbol{c}}_n\right] \tag{14}$$

In addition, a hidden layer is added behind the pooling layer of the network model to get more important feature information and the relationship between different channel information. It is expressed as

$$\boldsymbol{R} = relu\left(\boldsymbol{W}_h \hat{\boldsymbol{c}}_h + \boldsymbol{b}_h\right) \tag{15}$$

Here, $\boldsymbol{R} \in R^q$ is hidden layer output, q is hidden layer output dimension; $\boldsymbol{W}_h \in R^{q \times d}$ is hidden layer weight matrix; $\boldsymbol{b}_h \in R^q$ is offset matrix. In the full connection layer, we use softmax function to map the output eigenvectors of hidden layer to interval $[0, 1]$, so as to output the classification results of sentences to be classified.

4 Experiment

4.1 Data Set and Parameter Setting

QQ said that it is a popular microblog like service for college students, which was officially launched in China on July 15, 2010. It usually lists learning related, life-related or health-related problems, promotes the sharing of personal experience, and reveals the emotional state of students. Their friends and classmates uploaded their comments in response to the author's comments. Like other media such as microblog, it uses the web to publish short updates. This paper collects QQ data of Qiqihar University students. These students are undergraduates. We selected data of 60 students, including 7051 pieces of information (3703 positive information, 2255 negative information and 1093 objective information). Tab. 2 shows the polarity of some student comments used in our experiment. In the experiment, 90% of the texts are used as training sets and 10% as test sets. MATLAB is used to simulate the experiment. In the experiment, the data set is called self-built data set temporarily.

Table 2: Polarity of some student' speaking data

ID	Student name	Number of reviews	Positive	Negative	Objective
1	S1	271	161	93	17
2	S2	126	91	28	7
3	S3	370	154	138	78
4	S4	448	190	109	149
5	S5	86	31	51	4
…	…	…	…	…	…

Experiments are also conducted on COAE 2014 dataset [30], which is a balanced dataset. For the experiment of two corpus sets, 90% of the texts are used as training set and 10% as testing set. In the experiment, the input vectors are convoluted by using the multi window convolution kernel. The corrected linear units are functions of the convolution kernel. During the network training process, Adam is used to update and optimize the parameters of the network model. According to repeated experimental tests, the parameter settings of the model are shown in Tab. 3.

Table 3: Parameter settings of the model

Parameter name	Parameter symbol	Value
Window size	h	3,4,5
Feature mapping	m	100
Learning rate	b	0.001
Mini-batch size	r	50
Dropout rate	p	0.1

4.2 Evaluation Criteria

Text emotion classification can be regarded as a task of text classification. In this paper, precision, recall and F1 score are used as the evaluation criteria. Accuracy is defined as the ratio of the number of texts correctly classified as a specified category to the number of all texts classified into a specified category. Recall rate is defined as the ratio of the number of texts correctly classified as a specified category to the number of texts in all specified categories. F1 value is the harmonic mean value of accuracy and recall, which can be used for comprehensive evaluation of accuracy and recall. If TP is the number of documents correctly classified to the positive category, TN is the number of documents correctly classified to the negative category, FP is the number of negative texts wrongly classified to the positive category, and FN is the number of positive texts wrongly classified to the negative category, then for the positive category, each evaluation standard can be expressed as:

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

$$Recall = \frac{TP}{TP + FN} \tag{17}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{18}$$

The accuracy and recall rates used in this paper are the mean of accuracy and recall rates of positive and negative categories respectively, while the F1 value is calculated by Eq. (18) using the mean of accuracy and recall rates.

4.3 Comparison of Emotional Analysis Results

In order to compare the performance of various models on two sets of data to verify the effectiveness of the model proposed in this paper in the emotional analysis task, this paper selects a number of traditional machine learning and benchmark text convolution neural network models to compare with the model proposed in this paper. For traditional machine learning, support vector machine (SVM) and logistic regression (LR) are used to classify emotion. According to the different initialization methods of word vectors, a variety of convolution neural networks for reference texts are selected, which are Rand CNN, DCNN and skip gram CNN, which are randomly initialized and dynamically adjusted by back propagation during the network training.

Tabs. 4 and 5 show the classification effect obtained by fitting each model on the data set. It can be seen from this that the RAND CNN with random initialization input has no significant improvement compared with the traditional machine learning method, and its effect on COAE 2014 data set is even worse than SVM, with F1 value only 0.790; Rand CNN needs to adjust the original word vector while learning the classification network parameters, and the quality of the training word vector is poor when the training data is insufficient; by contrast The DCNN and skip gram CNN models which use the pre training word vectors to initialize the input have obvious improvement compared with the RAND CNN and, which proves that the classification effect of the text convolution neural network is directly related to the input expression, and the network is quite sensitive to the input noise.

Table 4: Emotional analysis results of self-built dataset

Model	Accuracy	Recall ratio	F1 value
LR [31]	0.848	0.862	0.853
SVM [32]	0.831	0.842	0.824
Rand-CNN [33]	0.909	0.906	0.909
DCNN [34]	0.916	0.917	0.918
Skip gram-CNN [35]	0.922	0.923	0.921
Proposed	0.937	0.935	0.934

Table 5: Emotional analysis results of COAE 2014 data set

Model	Accuracy	Recall ratio	F1 value
LR [31]	0.781	0.783	0.787
SVM [32]	0.806	0.800	0.807
Rand-CNN [33]	0.790	0.799	0.796
DCNN [34]	0.821	0.816	0.809
Skip gram-CNN [35]	0.867	0.865	0.861
Proposed	0.909	0.898	0.903

Through comparison, it is found that the model proposed in this paper is superior to other models in terms of evaluation indexes. In the self-built data set, compared with the best benchmark model skip Gram-CNN, the F1 value increased by 0.015; in the COAE 2014 data set, the F1 value increased by 0.042. Experimental results show that word sense disambiguation can improve the input expression of text convolution neural network, and then get better classification ability.

In order to further analyze the impact of part of speech features of different dimensions on the classification effect, we use different dimensions of part of speech features on the self-built data set and COAE 2014 data set for comparative analysis. The experimental results are shown in Fig. 2, where the part of speech vector dimension of 0 means no part of speech vector is used.

As can be seen from Fig. 2, when the part of speech feature dimension is less than 50, the four models in the two datasets show an upward trend, among which the RAND CNN model with randomly initialized feature vector is the most obvious one, and the classification accuracy fluctuates after the part of speech feature dimension is more than 50; however, when the part of speech feature dimension is more than 50, the DCNN and sk of initialization input of pre trained word vector With the increase of dimensions, the classification accuracy of IP gram CNN model shows a downward trend. However, the classification accuracy of the model proposed by using word vectors fluctuates after the part of speech feature dimension exceeds 50. With the increase of part of speech feature dimension, more weight and vector parameters need to be adjusted in one iteration of the model. When the feature vector is initialized randomly, it is possible to give the part of speech a feature vector which is quite different from the real value, making it difficult for the model to approach the real feature vector through parameter adjustment in the training process. So when the feature dimension exceeds a certain threshold, the model classification accuracy of randomly initialized feature vector will decrease with the increase of part of speech feature dimension. When the part of speech feature dimension is more than 50, the classification accuracy of the proposed model is not significantly increased, and the larger the part of speech feature vector dimension is, the longer the training time of the model is. So the bigger the dimension of part of speech vector, the better.

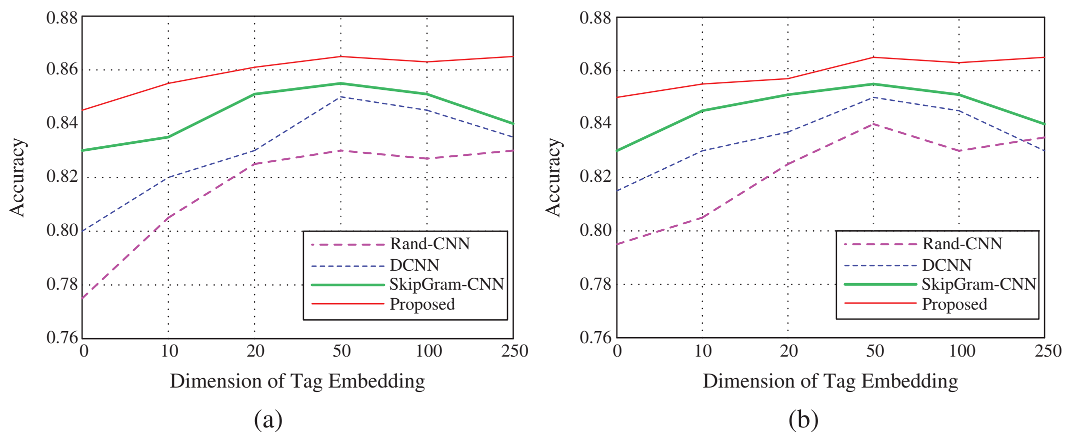

Figure 2: Comparison of tag embedding in different dimensions. (a) Self-built dataset. (b) COAE 2014 dataset

In order to analyze the influence of word vectors of different dimensions on the classification performance of the model, the comparison experiment is carried out on two datasets by taking word vectors of different dimensions from several models, and the experimental results are shown in Fig. 3.

As can be seen from the results in Fig. 3, on the two datasets, when the word vector dimension is less than 100, the accuracy of the four models has an obvious upward trend. When the word vector dimension is larger than 50 dimensions, the classification accuracy of the model proposed by using the construction word vector shows a gentle rise, which indicates that the classification accuracy of the model proposed by using the construction word vector can increase with the increase of the word vector dimension. However, when the word vector dimension of Rand CNN model is more than 50, the classification accuracy fluctuates. Because of the dependence on the initial value of vector, with the increase of word vector dimension, the model given initial value by random initialization can not learn the characteristic information of vector well, so the classification accuracy fluctuates. And with the increase of vector dimension, the training time of the model will also increase. Therefore, in the experiment, the word vector dimension is set as 100 dimensions, the part of speech vector dimension is 50 dimensions, and the location feature vector as auxiliary network model training is set as 10 dimensions.

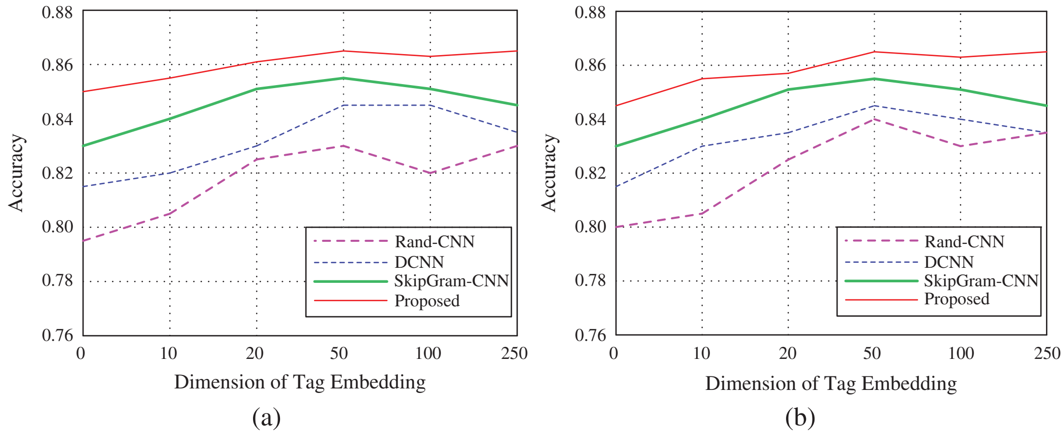

Figure 3: Comparison of word embedding in different dimensions. (a) Self-built dataset. (b) COAE 2014 dataset

Figs. 4 and 5 respectively show the classification accuracy of the first 50 iterations of each network in two emotional analysis experiments. It can be seen that the proposed model is superior to each reference text convolution neural network in each iteration, and its classification accuracy is significantly higher than other network models from the fifth iteration; in two groups of experiments, Rand CNN needs to dynamically adjust the word vector, and the classification accuracy of each iteration is the lowest.

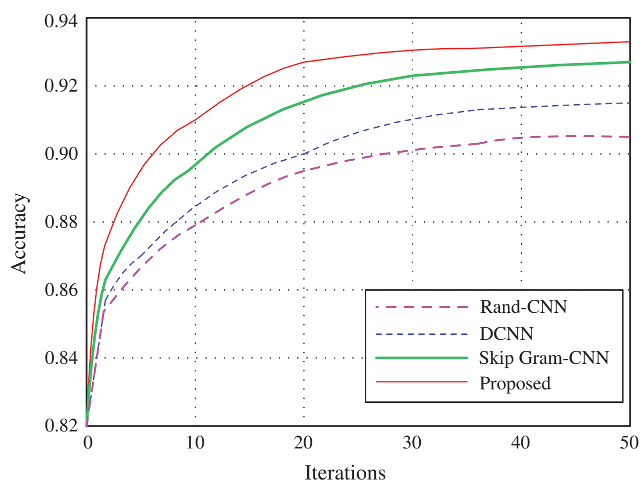

Figure 4: Results of the first 50 iterations of the self-built dataset experiment

Through observation, it can be found that the number of iterations required for the convergence of the model proposed in this paper is basically consistent with DCNN and Skip Gram-CNN. With the increase of iterations, it can show better classification effect than other models. This trend also proves that multi-channel input strategy can bring more space for network training, and the network will not be prematurely affected by the input noise of a single channel. Subsidence is over fitted.

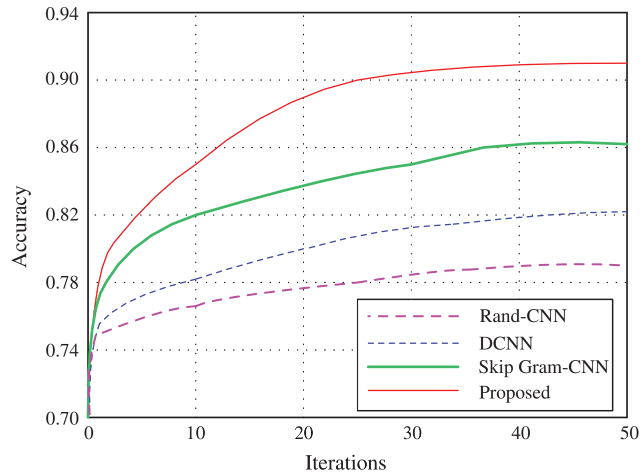

Figure 5: Results of 50 iterations before COAE 2014 dataset experiment

In conclusion, the proposed model can not only extract more sentence semantic information, but also learn the hidden emotional information in sentences. With the increase of the number of iterations, this model can show better classification effect than other models.

5 Conclusion

In this paper, a text emotion analysis model based on emotion dictionary and multi-channel convolution neural network is proposed. The model realizes semi supervised learning by reconstructing the loss function of convolutional neural network and not changing the structure of the network. This model can not only extract more sentence semantic information, but also learn the hidden emotional information. With the increase of the number of iterations, the model can show better classification effect than other models. This trend also proves that the multi-channel input strategy can bring more space for network training, and the network will not fall into over fitting prematurely due to the input noise of a single channel.

In the next step, we can use different feature combination methods for different channels and different activation functions for different channels, so that the model can learn more feature information, and improve the multi-channel convolution neural network model proposed in this paper for sentence patterns with satire emotion.

Funding Statement: This paper was supported by the 2018 Science and Technology Breakthrough Project of Henan Provincial Science and Technology Department (No. 182102310694).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Wang, S. H., Zhang, Y., Li, Y. J., Jia, W. J., Liu, F. Y. et al. (2018). Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. *Multimedia Tools and Applications, 77(9),* 10393–10417. DOI 10.1007/s11042-016-4222-4.
2. Wang, S., Sun, J., Mehmood, I., Pan, C., Chen, Y. et al. (2020). Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling. *Concurrency and Computation: Practice and Experience, 32(1),* e5130.
3. Peng, G., Xu, D. (2019). Weakly supervised learning of image emotion analysis based on cross-spatial pooling. *Chinese Conference on Biometric Recognition*. pp. 116–125. Cham: Springer.
4. Zhang, Y., Wang, S., Sui, Y., Yang, M., Liu, B. et al. (2018). Multivariate approach for Alzheimer's disease detection using stationary wavelet entropy and predator-prey particle swarm optimization. *Journal of Alzheimer's Disease, 65(3),* 855–869. DOI 10.3233/JAD-170069.
5. Kang, C., Yu, X., Wang, S. H., Guttery, D., Pandey, H. et al. (2020). A heuristic neural network structure relying on fuzzy logic for images scoring. *IEEE Transactions on Fuzzy Systems*. DOI 10.1109/TFUZZ.91.
6. Wang, S. H., Zhang, Y. D., Yang, M., Liu, B., Ramirez, J. et al. (2019). Unilateral sensorineural hearing loss identification based on double-density dual-tree complex wavelet transform and multinomial logistic regression. *Integrated Computer-Aided Engineering, 26(4),* 411–426. DOI 10.3233/ICA-190605.
7. Choi, K., Joo, M. I., Aich, S., Kim, H. C. (2018). Design of regional-based emotion analysis system using twitter feeds. *Advanced Science Letters, 24(3),* 2054–2057. DOI 10.1166/asl.2018.11846.
8. Sinha, S., Mishra, S. K., Bilgaiyan, S. (2020). Emotion analysis to provide counseling to students fighting from depression and anxiety by using CCTV surveillance. *Machine Learning and Information Processing*. pp. 81–94. Singapore: Springer.
9. Wang, S. H., Xie, S., Chen, X., Guttery, D. S., Tang, C. et al. (2019). Alcoholism identification based on an AlexNet transfer learning model. *Frontiers in Psychiatry, 10,* 205. DOI 10.3389/fpsyt.2019.00205.
10. Lubis, N., Sakti, S., Yoshino, K., Nakamura, S. (2018). Emotional triggers and responses in spontaneous affective interaction: recognition, prediction, and analysis. *European Journal of Clinical Nutrition, 33(1),* 36–40.
11. McDonald, S. E., Ma, L., Green, K. E., Hitti, S. A., Cody, A. M. et al. (2018). Evaluation of the parent-report inventory of callous-unemotional traits in a sample of children recruited from intimate partner violence services: a multidimensional Rasch analysis. *Journal of Clinical Psychology, 74(3),* 418–441. DOI 10.1002/jclp.22497.
12. Ellis, M., Kleinplatz, P. J. (2018). How contingencies of self-worth influence reactions to emotional and sexual infidelity. *Canadian Journal of Human Sexuality, 27(1),* 43–54. DOI 10.3138/cjhs.2017-0023.
13. Ma, L., Ju, F., Wan, J., Shen, X. (2019). Emotional computing based on cross-modal fusion and edge network data incentive. *Personal and Ubiquitous Computing, 23(3–4),* 363–372. DOI 10.1007/s00779-019-01232-1.
14. Wang, Y. (2019). Multimodal emotion recognition algorithm based on edge network emotion element compensation and data fusion. *Personal and Ubiquitous Computing, 23(3–4),* 383–392. DOI 10.1007/s00779-018-01195-9.
15. Sangvikar, P. C. (2018). A survey on sentiment analysis and opinion mining. *Multimedia Tools & Applications, 5(4),* 1–29.
16. Baali, M., Ghneim, N. (2019). Emotion analysis of Arabic tweets using deep learning approach. *Journal of Big Data, 6(1),* 89. DOI 10.1186/s40537-019-0252-x.
17. Zhang, Y. D., Govindaraj, V. V., Tang, C., Zhu, W., Sun, J. (2019). High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. *Journal of Medical Imaging and Health Informatics, 9(9),* 2012–2021. DOI 10.1166/jmihi.2019.2692.
18. Jurado, F., Rodriguez, P. (2015). Sentiment analysis in monitoring software development processes: an exploratory case study on GitHub's project issues. *Journal of Systems and Software, 104,* 82–89. DOI 10.1016/j.jss.2015.02.055.

19. Huang, Y. A., Jastorff, J., Van den Stock, J., Van de Vliet, L., Dupont, P. et al. (2018). Studying emotion theories through connectivity analysis: evidence from generalized psychophysiological interactions and graph theory. *NeuroImage, 172,* 250–262. DOI 10.1016/j.neuroimage.2018.01.026.
20. Schouten, K., Van Der Weijde, O., Frasincar, F., Dekker, R. (2017). Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data. *IEEE Transactions on Cybernetics, 48(4),* 1263–1275. DOI 10.1109/TCYB.2017.2688801.
21. Vandin, A. (2018). Replicated computations results (RCR) report for a holistic approach for collaborative workload execution in volunteer clouds. *ACM Transactions on Modeling & Computer Simulation, 28(2),* 1–3. DOI 10.1145/3182167.
22. Tubishat, M., Idris, N., Abushariah, M. A. (2018). Implicit aspect extraction in sentiment analysis: review, taxonomy, oppportunities, and open challenges. *Information Processing & Management, 54(4),* 545–563. DOI 10.1016/j.ipm.2018.03.008.
23. Kang, H., Yoo, S. J., Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications, 39(5),* 6000–6010. DOI 10.1016/j.eswa.2011.11.107.
24. Zhu, T., Xia, Z., Dong, J., Zhao, Q. (2019). A sociable human–robot interaction scheme based on body emotion analysis. *International Journal of Control, Automation and Systems, 17(2),* 474–485. DOI 10.1007/s12555-017-0423-5.
25. Thulasidasan, S., Bilmes, J. (2017). Acoustic classification using semi-supervised deep neural networks and stochastic entropy-regularization over nearest-neighbor graphs. *IEEE International Conference on Acoustics, Speech and Signal Processing*. pp. 2731–2735. New Orleans, LA: IEEE.
26. Rawat, S., Chopra, A., Singh, S., Sinha, S. (2019). Mid roll advertisement placement using multi modal emotion analysis. *International Conference on Artificial Neural Networks*. pp. 159–171. Cham: Springer, DOI 10.1007/978-3-030-30490-4_14.
27. Yang, G., Jin, J., Kim, D., Joo, H. J. (2019). Multi-modal emotion analysis for chatbots, high-performance computing and big data analysis. TopHPC 2019. In: Grandinetti, L., Mirtaheri, S., Shahbazian, R. (Eds.). *Communications in Computer and Information Science*. vol. 891. pp. 331–338. Cham: Springer, DOI 10.1007/978-3-030-33495-6_25.
28. Chen, L., Bo, K., Lee, F., Chen, Q. (2020). Advanced feature fusion algorithm based on multiple convolutional neural network for scene recognition. *Computer Modeling in Engineering & Sciences, 122(2),* 505–523. DOI 10.32604/cmes.2020.08425.
29. Wang, S. H., Sun, J., Phillips, P., Zhao, G., Zhang, Y. D. (2018). Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units. *Journal of Real-Time Image Processing, 15(3),* 631–642. DOI 10.1007/s11554-017-0717-0.
30. Chen, K., Liang, B., Ke, W., Xu, B., Zeng, G. C. (2018). Chinese micro-blog sentiment analysis based on muti-channels convolutional neural networks. *Journal of Computer Research and Development, 55(5),* 945–957. DOI 10.7544/issn1000-1239.2018.20170049.
31. Jacob, A. (2017). Modelling speech emotion recognition using logistic regression and decision trees. *International Journal of Speech Technology, 20(3),* 1–9. DOI 10.1007/s10772-017-9457-6.
32. Wang, S., Tang, C., Sun, J., Zhang, Y. (2019). Cerebral micro-bleeding detection based on densely connected neural network. *Frontiers in Neuroscience, 13,* 422. DOI 10.3389/fnins.2019.00422.
33. Latha, G. C. P., Priya, M. M. (2017). A novel and enhanced facial electromyogram based human emotion recognition using convolution neural network model with multirate signal processing features. *Journal of Computational and Theoretical Nanoscience, 14(3),* 1572–1580. DOI 10.1166/jctn.2017.6480.
34. Sun, L., Chen, J., Xie, K., Gu, T. (2018). Deep and shallow features fusion based on deep convolutional neural network for speech emotion recognition. *International Journal of Speech Technology, 21(4),* 931–940. DOI 10.1007/s10772-018-9551-4.
35. Zhang, C., Gao, S., Zhao, B., Yang, L., Xi, X. et al. (2017). An adaptive sentence representation learning model based on multi-gram CNN. *International Conference on Intelligent Environments*. pp. 130–135. Seoul, South Korea. DOI 10.1109/IE.2017.18.