Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

Emotion recognition based on deep learning to predict potential depression patients

Li Ding, Feng Li*

School of management science and Engineering, Anhui University of Finance and Economics,

Bengbu 233000, China

*Corresponding author Email: feli@ustc.edu.cn

Abstract: In society, the incidence of depression shows a trend of rapid growth, and the early treatment of depression has been using medical and psychological methods. Aiming at the existence of these methods in a series of faults, this paper puts forward a kind of emotion recognition algorithm based on depth study to predict the potential of depression patients, namely through the TCNN - GRU helped model for text analysis, and thus for the text feature extraction, and finally through the model training, depression index formula in reference to give users depressive state are obtained. Through experimental comparison, the algorithm can effectively improve the accuracy of identifying potential depressed patients.
Keywords: Deep learning; emotion recognition; TCNN-GRU; people with depression.

---

1. INTRODUCTION

Because China does not pay enough attention to mental health problems, people generally lack a correct understanding of mental health diseases, and there are also a lack of relevant professionals and rescue centers in mental health. The traditional diagnosis of mental health diseases mainly depends on the knowledge of psychology. Depression detection scale combined with manual interview is usually used to judge whether users have depression tendency. Therefore, there are some disadvantages, such as slow diagnosis, high cost of manual detection of depression, passivity and invasiveness.

The incidence rate of depression is high, the diagnosis rate is low, and the etiology is complex. It is not only affected by genetic factors, but also largely affected by environmental factors. It is difficult to show detectable substantive pathological characteristics, resulting in researchers' inability to accurately understand its pathogenesis, and then affect the research and development of antidepressants. Therefore, the early research on the pathogenesis and related symptoms of depression mainly focused on the fields of medicine and psychology. Therefore, researchers have proposed many evaluation scales. Although these scales and their extended versions are still widely used, their accuracy is low. In recent years, due to the rapid development of network information technology and the wide application of mobile social Internet, more patients with depression prefer to

use the Internet to vent their personal emotions and express their needs online. The psychological status of users analyzed through social networks has also been paid attention to by scholars in the field of social psychology and computer. However, due to its characteristics, the traditional character emotion recognition methods are easily constrained by the quality of dictionaries, and it is easy to ignore the specific semantics of sentences. At the same time, the character characteristics of machine learning methods divided according to the characteristics often need the in-depth research of domain experts, and the conversion efficiency between various fields is not good. Therefore, the traditional emotion recognition methods have some bottlenecks.

In recent years, the rise of deep learning methods can better make up for these defects. For example, Wang Peng uses convolutional neural network (CNN) model and word embedding clustering to improve short article classification, and ER Meng Joo adds attention mechanism to CNN for text classification. Experiments show that the classification method is efficient on some data sets. Tung tran analyzed the case description of patients by using the circular neural network (RNN) model. The experiment confirmed that the efficiency of classification methods has improved in common mental diseases. Later, Hassan abdalraouf adopted the RNN model to replace the pooling layer of CNN. The experiment also confirmed that the classification efficiency of this model was also improved on the mainstream film evaluation data set of Stanford. It can be seen that deep learning technology has laid a solid technical foundation for accurately identifying the characteristics of patients with depression in community websites, and can judge the user's emotional tendency of depression through the deep learning model, so as to evaluate the user's emotional and depressive status of depression.

Based on tcnn-gru deep learning model, this paper identifies and detects the emotion of microblog comment users. Firstly, the evaluation criteria of depression degree and depression index are introduced to evaluate the emotional state of microblog users. Then, the text classification model based on tcnn-gru is introduced to obtain the emotion label of microblog content. Finally, the experimental data, algorithm model and experimental results are analyzed, come to conclusion.

2. CRITERIA FOR USERS WITH DEPRESSION UNDER MICROBLOG COMMENTS

2.1 Depression index.

Depression refers to a person's obvious and lasting depression. In order to characterize this situation, Tao Jiong evaluated the mental health status of cancer patients through SDS depression index in scientific research. Although the text classification model of depression can judge whether it is a single line example text, However, we cannot judge that the user who issued this statement must be depressed. Therefore, for those patients who have made sentences such as "it's raining heavily today, the plan is ruined, and they really want to cry", it is very likely that the patients have an impulsive response to some stimulation in a short time, and the emotional patients themselves can quickly calm down in a short time, so as to return to a non depressed state. In view of this, when Shi Zhiwei assessed the depression tendency of Internet users from the perspective of text analysis, he used the percentage of depression microblog in the total number of microblogs to measure the depression index of users, and thus to judge the degree of depression tendency of users in a long time. Although from the perspective of text analysis, depression bias does not fully take into account the characteristics such as the number of fans of microblog users, the text according to the data can be

divided into two types, namely, the comments made by users under others' microblogs and the views expressed by users' own microblogs. Therefore, the depression index specified in this paper refers to the number of blog posts and comments in a specific time period, and puts forward the calculation formula of depression index based on microblog as follows:

$$DI = \frac{N_{\mathrm{cd}}}{2N_{ct}} + \frac{N_{md}}{2N_{mt}}$$

Among them, it refers to the number of microblogs with depression tendency published by users under other people's accounts, the number of microblogs published by users under other people's accounts, the number of microblogs with depression tendency published and forwarded by users under their own accounts, and the number of microblogs published and forwarded by users under their own accounts, which is the depression index.

2.2 Degree of depression

Zung's self rating Depression Scale (SDS) is a self rating scale, which can effectively distinguish the degree of depression. It can be divided into no depression, mild depression, moderate depression and severe depression with the thresholds of 0.5, 0.7, 0.85 and 1.00. SDS has high reliability. In this paper, the user's comments and published content are collected and marked, so as to calculate the user's depression index. Through the data analysis, the user's SDS score and the index are tested by Pearson correlation coefficient. The results show that the di index and SDS score are significantly correlated at the level of 0.01 (bilateral), and R = 0.5564, indicating that there is a strong correlation between them. Through the above analysis, it is proposed that the relationship between depression index and depression degree is as follows:

$$S(DI) = \begin{cases} \text{正常} & DI \in [0, 0.50] \\ \text{轻度抑郁} & DI \in [0.50, 0.70] \\ \text{中度抑郁} & DI \in [0.70, 0.85] \\ \text{重度抑郁} & DI \in [0.85, 1.00] \end{cases}$$

Fig.1 Relational formula

Among them, s (DI) refers to the depression status of microblog users. According to the distribution of depression index, it can be divided into normal, mild depression, moderate depression and severe depression.

3. EMOTION CLASSIFICATION ALGORITHM BASED ON TCNN-GRU MODEL

3.1 Basic structure and idea of TCNN model.

Construct a depression emotion classification model based on tcnn-gru. Firstly, classify and label the collected data set, and then preprocess each corpus, such as word segmentation and de stop words, and then train and optimize the model until the final model can accurately judge whether each social evaluation contains depression tendency. Tcnn-gru model combines textcnn model and Gru model in order to give full play to their respective advantages for model recognition. The discrimination of depressive tendency is analyzed according to the relationship between the user and the depression index and degree proposed above.

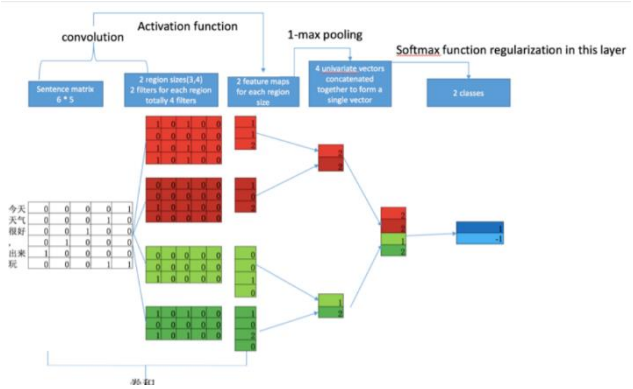

Fig.2 Architecture of TCNN

In 2014, Yoon Kim changed the CNN model and proposed the text classification model textcnn, which has no change compared with the traditional structure in the network structure. It can be seen from Figure 1 that textcnn has one layer of Max pooling and one layer of convolution. Finally, the output is connected with softmax for classification. In fact, compared with the traditional structure, the model only changes when inputting data. For example, when textcnn deals with the sentence "today's weather is good, go out to play", it first divides it into "today / weather / good /, / come out / play". Each word is mapped into a five-dimensional word vector through embedding methods such as word2vec or Glov, such as "today" - "[0,0,0,0,1]," weather "-" [0,0,1,0], "good" - "[0,0,1,0], etc. The advantage of this is to convert the text into numerical value, which makes the subsequent processing more convenient. After the word vector is constructed, it is spliced into a 6 * 5 two-dimensional matrix as the initial input.

3.2 Gru model basic structure and idea

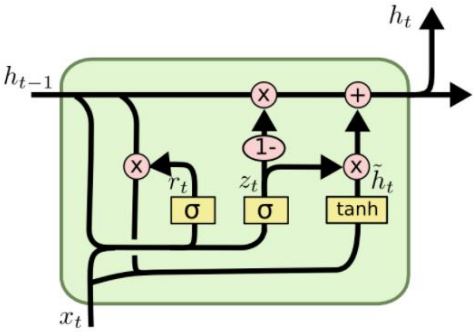

Fig.3 Basic structure of GRU

Gru is not like the three control gate of LSTM. Although it also has gates, it has only two gates, which are called reset gate and update gate respectively. As the name suggests, the reset gate controls whether to reset, that is, to what extent the previous state state is erased; The update gate indicates the extent to which the current hidden layer should be updated with candidate.

$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Fig.4 Formula

From the formula of this figure, R and Z are two gates, representing reset and update respectively. First, use X (T) and H (t-1) to generate two gates, and then use the reset gate to multiply the state of the previous time to see whether to reset or to what extent. Then, splice with the newly input x, pass through the network and activate with tanh to form the implicit variable \ hat {h_t} of candidate. Then, make a linear combination of h of the previous time and h of candidate, and the sum of the weights of the two is 1, The weight of candidate is the output of update gate, which indicates the update intensity. However, since h is a variable, it needs to be updated with the previous h and the current alternative answers at every moment, including the final linear combination.

3.3 Text classification model based on tcnn-gru

In the experiment, convolution kernels of various sizes are used for convolution, and then the tcnn-gru structure of gated loop unit is used by combining the text local feature extraction ability of textcnn model and the text sequence information learning ability of gated loop unit model, Thus, the practical problems of text local characteristic data information granularity solidification and rigidity caused by the solidification of convolution calculation core scale in convolution neural network are solved, and the long-term dependence on cyclic neural network is overcome.

Words are transformed into word vectors by word2vec method. Textcnn layer performs convolution operation through word convolution kernels of different sizes. RNN model captures text sequence information through Gru. The general process of text content processing in deep learning is as follows:

Input layer: the text of microblog information content is cut into word units, and then the word vector is obtained by word mapping, so as to obtain the text matrix, so as to realize the transformation from microblog information content to input information matrix.

Hidden layer: textcnn is a deep neural network module invented by Kim in 2014. It extends the theory of processing picture messages in convolutional neural network to the field of text analysis. When processing text messages, it has the characteristics of multi local perception and shared parameters, so it can better capture local messages. However, the local information captured by this method still has the problem of granularity solidification, Therefore, textcnn will efficiently process the local feature information in the corpus sentence in this model, and the output result is the feature vector, which is used as the input of Gru model at the next time, and then process the sequence information through the update gate and reset gate.

Hidden layer and output layer: the results after different convolution kernels and Gru models are spliced. In order to prevent over fitting, the dropout layer will be added. Finally, the full connection operation is carried out, and the softmax classifier is used to output the tcnn-gru model to predict the emotion category probability model of microblog content.

At the same time, avoid gradient explosion or gradient disappearance, and better cooperate with the softmax activation function of the full connection layer.

4. EXPERIMENT

4.1 Experimental data set.

It can be said that the research on depression has been at the forefront of the medical and psychological circles. However, due to the lack of obvious patient characteristics of patients with

depression and the estrangement of manual communication itself, the research has been stagnant. The novel coronavirus pneumonia epidemic has been spreading and spreading in recent two years. Many people's psychological condition has changed. Many people have been letting their feelings out in micro-blog's social media. Sina Weibo is currently the most widely used social platform for user communication in China. Its publishing, forwarding, following and commenting functions make it easy for Chinese users to convey emotion and exchange information. The experimental data of this study are extracted from the microblog comments during the epidemic, including the user's comments on others' microblog, the user's personal original microblog and the forwarded content on the microblog. The obtained text content is emotionally annotated, and then the labeling results are filtered to filter out the image information and unqualified information. A total of 10606 comments were collected, including user ID, microblog content, region and corresponding tags. There are 8606 pieces of data in the training set and 2000 pieces of data in the test set. The microblog content of the data set is divided into six tags according to emotion: neural: no emotion, happy: positive, anger: anger, sad: sadness, fear: fear, surprise: surprise. Among them, positive emotions are positive emotions, no emotion and surprise are neutral emotions, and anger, sadness and fear are negative emotions.

4.2 Experimental algorithm model.

In this experiment, tcnn-gru model is used for emotion recognition of depressed users. The structure diagram of this model is as follows:

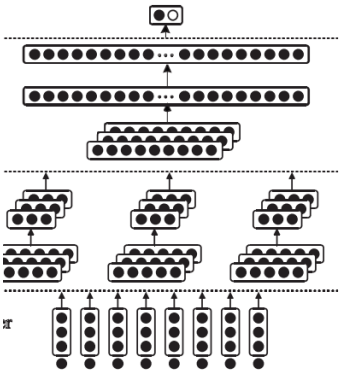

Fig.5 Internal structure of model TCNN-GRU

Where VI represents the word mapping result of the ith word after word segmentation of the original microblog content.

The textcnn feature extraction layer mainly extracts the in-depth features of the input text matrix. The text constructs a convolution structure composed of three convolution cores of different sizes and parallel to each other to obtain the abstract feature information of microblog content text with different granularity. According to the characteristics of convolution neural network for text classification, the convolution mode of each parallel convolution channel is set as one-dimensional convolution. And the relu activation function is used for activation. After the convolution layer processing, the original microblog content data will be mapped to the hidden layer and abstract feature space, build a parallel convolution structure, convert it to output, and extract the feature vector through the activation function.

$$C_1 = f(\omega_1 \otimes M + b_1) = \mathrm{Re}lu(\omega_1 \otimes M + b_1)$$
$$C_2 = f(\omega_2 \otimes M + b_2) = \mathrm{Re}lu(\omega_2 \otimes M + b_2)$$
$$C_3 = f(\omega_3 \otimes M + b_3) = \mathrm{Re}lu(\omega_3 \otimes M + b_3)$$

Fig.6 Feature vector

Where C1 is the result of convolution layer 1, C2 is the result of convolution layer 2, C3 is the convolution result of convolution layer 3, W1, W2 and W3 are the weight matrix of corresponding convolution layer, and B1, B2 and B3 are the deviation of corresponding convolution layer.

In the actual template design, there may be many convolution kernels of the same size to obtain various characteristic information, and then the results are spliced. In the traditional textcnn prediction model, the feature vector information obtained through convolution calculation can realize the pooling process of maximum pooling, mean pooling and other methods, so as to reduce the total amount of parameters Select the key feature information representing the text features. Therefore, in tcnn-gru prediction model, Gru mode is used instead of pooling layer to further obtain characteristic sequence signals.

Gru sequence information extraction layer learns and extracts the sequence information of the convolution result of the previous layer, and Gru sets reset gate and update gate.

The hidden layer and the output layer first splice three kinds of convolution and Gru processed information, and then set the dropout layer to prevent training over fitting. Finally, the full connection layer is used to map the learned "distributed feature representation" vector to the sample tag space to distinguish the tag y of microblog content text, that is, positive, neutral and negative probability distribution. Its calculation formula is:

$$p(y \mid s) = soft\max(w \bullet v^* + b^*)$$
$$\tilde{y} = \arg\max_y p(y \mid s)$$

Fig.7 Calculation formula

Among them, y is the real emotion tag of microblog content, which is expressed by independent heat coding, y ~ is the emotion tag vector of microblog content obtained through training. For the predicted probability of each emotion category, the tag with the highest probability is selected as the output. In addition, the minimization loss function of training goal uses the cross entropy loss function.

4.3 Modeling.

The storage format of the original data set is TXT, in which a py list object is stored. Each element contained in the list is a microblog content and its corresponding label.

The visual display of its original data is:

{"Id": 26, "content": "# national confirmed cases of new pneumonia# http://t.cn/RXnNTiO ?? Fuzhou "," label ":" neural "}.Then, the data is cleaned and segmented, and the data after cleaning and word segmentation are stored for subsequent loading. The style of the data set after cleaning is:

{"content": ["[heart]", "health", "peace", "[forward]", "tribute", "epidemic", "front", "medical", "personnel", "wish", "all", "all", "health", "peace", "angel in white", "tribute", "[heart]"], "label": "happy"}.After cleaning and preprocessing the data set, analyze the distribution of data of various labels in the training set and test set as follows:

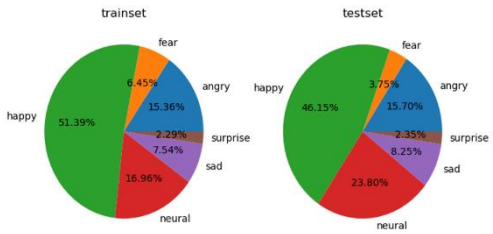

Fig.8 The distribution of data for various labels in training and test sets

Tcnn-gru model is adopted in this experiment, and the relevant model parameters are set as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} * 100\%$$

Among them, is the accuracy of the model, which refers to the real case (both the model prediction result and the actual label are positive), is the false positive case (the model prediction result is positive, but the actual label is negative), refers to the true negative case (both the model prediction result and the actual label are negative), and refers to the false negative case (the model prediction result is negative, but the actual label is positive). A balanced data set was used in the experimental process. Therefore, the above four indicators are used as the standard to measure the quality of model training.

In the process of training, the change images of loss value and accuracy rate are as follows:

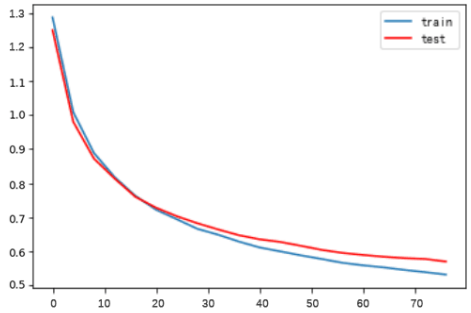

Fig.9 Variation in loss values for training and test sets

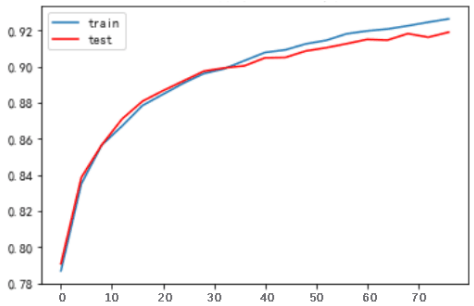

Fig.10 Variation in accuracy values for training and test sets

5. EXPERIMENTAL RESULT

As can be seen from the change trend of accuracy in the figure above, the accuracy of the training set has reached 93.96%, and the accuracy of the test set has also reached 92.78%. In order to compare the experimental results, the analysis results of other models currently used for emotion recognition are selected, as shown in the following table:

Tab.1 Comparative experiment

Comparative experiment	Accuracy	Model
machine study	0.7962	MLP_TF_IDF
	0.8605	RF_TF_IDF
	0.835 0.8556	SVM_TF_IDF CNN_rand
TextCNN	0.8772 0.8924	CNN_static CNN_nonstatic
	0.8912 0.8674	CNN_multichanne LSTM
RNN	0.8605 0.8706	GRU EBiLSTM
blend Model	0.9123	CNN-BiLSTM
	0.9035 0.9268	HMAN C-LSTM
this paper Model	0.9396	TCNN-GRU

From the previous models used for training, it can be seen that the accuracy of general hybrid models is greater than that of single models. These models are optimized step by step. The hybrid models extract polar text features through CNN model and extract text context information by LSTM model, Note that the network model also avoids the complexity of machine learning and manual design features. Its accuracy has reached quite high, but the classification effect is still lower than the model in this paper. In this paper, textcnn model and Gru model are mixed to give full play to the advantages of their respective models, which are well connected, and good results can be seen from the accuracy. This paper proposes emotion recognition based on deep learning to predict potential depression patients. The deep learning model uses tcnn-gru model. The deep learning model based on tcnn-gru can accurately identify users' depression from the perspective of text analysis, and analyze it according to the proposed depression index and depression degree standard, so as to determine users' depression level. And then make timely diagnosis and treatment for patients in need. The experimental results show that the recognition accuracy of this method is very high, but some aspects still need to be improved, and the shortcomings will continue to be improved in the follow-up research. However, this experiment provides a new way for the diagnosis and treatment of patients with potential depression.

ACKNOWLEDGMENTS

This work was financially supported by Preferential subsidy scheme for overseas students' innovation project (Project No.: 2021lcx032); Innovation and entrepreneurship training program for college students (Project No.: 202110378168)

REFERENCES

[1] World Health Oganization Depression and other common mental disorders: global health estimates[J]. 2017.

[2] Yifan Ye,Jing Chen. Multi-modal Speech Emotion Recognition Based on TCN and

Attention[C]//Proceedings of the 11th International Conference on Computer Engineering and Networks(CENet2021)Part I.,2021:342-349.DOI:10.26914/c.cnkihy.2021.045034.
[3] Araño Keith April et al. Multimodal sentiment and emotion recognition in hyperbolic space[J]. Expert Systems With Applications, 2021, 184
[4] Lee Shih-Chieh et al. Using machine learning to improve the discriminative power of the FERD screener in classifying patients with schizophrenia and healthy adults[J]. Journal of Affective Disorders, 2021, 292 : 102-107.
[5] Badal Varsha D. et al. Computational methods for integrative evaluation of confidence, accuracy, and reaction time in facial affect recognition in schizophrenia[J]. Schizophrenia Research: Cognition, 2021, 25 : 100196-100196.
[6] Ramzan Munaza and Dawn Suma. Fused CNN-LSTM Deep learning emotion recognition model using Electroencephalography signals.[J]. The International journal of neuroscience, 2021,: 11-10.
[7] Huang Junlian and Zhao Dongyan and Lv Boxue. Application of Improved Neural Network Strategy in Emotion Information Discrimination[J]. Journal of Physics: Conference Series, 2021, 1952(4).
[8] Liu Haiou, Yao Sumei, he xutao, Su Yanyuan Research on depression user portrait in online health community based on deep learning [J] Small microcomputer system, 2021,42 (03): 572-577.
[9] Ji Jianlin, Zhang Hong. Analysis of somatic 9 symptoms and related factors of depression [J] Chinese Journal of mental health, 2002,16(9):605-608.
[10] Zhao Yunlong, Li Yanbing Research on personality prediction and relationship strength of social network users [C] / / collected works of business intelligence branch of the seventh China management academic year conference 2012,10-19.
[11] Dong Li. Inference of social network user attributes based on semi supervised learning [D] Harbin University of technology, 2011.
[12] Research on network public opinion recognition based on deep learning [J] Xu Feng [D] Beijing University of Posts and telecommunications, 2019.
[13] Xu Feng, Zhang Xuefen, Xin Zhanhong. Chinese word segmentation scheme based on deep neural network model [J] Journal of Harbin Engineering University, employed (EI source journal).
[14] Ji Jianlin, Zhang Hong Analysis of somatic symptoms and related factors of depression [J] Chinese Journal of mental health, 2002, 16 (9): 605-608
[15] Zhao Yunlong, Li Yanbing Research on personality prediction and relationship strength of social network users [C] / / collected works of business intelligence branch of the seventh China management academic year conference 2012, 10-19.