MENTAL HEALTH ASSESSMENT OF SOCIAL NETWORK USERS BASED ON CONVOLUTIONAL NEURAL NETWORK

Sheng Bin*, Gengxin Sun

Abstract
In the age of the Internet, more and more users mention their mental health problems anonymously on social network sites. These users should be identified and intervened timely to prevent their mental health from deteriorating. This paper constructs a mental health assessment model based on convolutional neural network (CNN), which automatically evaluates the mental health of social network users, and the urgency of their intervention demand. Under the guidance of psychological knowledge, this model relies on the CNN to mine the statistical features of word frequency in different categories of posts, and then pinpoints the posts that need to be intervened. Experimental results show that, compared with other methods, our model can achieve a desirable recognition effect for social network users who need urgent psychological intervention. The research results reflect the value of psychological knowledge in feature extraction through deep learning.

Key words: Mental Health Assessment, Psychological Intervention, Convolutional Neural Network (CNN), Social Network.

Received: 24-02-19 | Accepted: 22-07-19

INTRODUCTION
With the rapid development of mobile Internet, social network such as twitter, Facebook, Sina Weibo, have become an indispensable part of people’s daily life. Social networks are often used to express the real feelings and psychological state of social network users, which makes social networks become an important source of research on depression, self-harm and other mental health problems.

The study of social network users found that through analyzing the user’s self-reported content and behavior on social network, it can help to find the user’s mental health status in time (Ferrari, Norman, Freedman et al., 2014; Ciobanu, Ferrari, Erskine et al., 2018). More and more researchers from computer science and psychology field are attracted by using the data of social network to automatically evaluate users’ mental health. Although abundant research results had been achieved, the accuracy of automatic mental health assessment still has a lot of room to improve, especially the characteristics of the automatic assessment model. Based on text data mining and deep learning technology, we constructed an automatic classification model to effectively identify posts reflecting different mental health status in social network.

Our study can alleviate the pressure of resource shortage of psychiatrists. At present, mental health assessment mainly relies on manual detection, but the professionals in this field are limited, and its invasive characteristics will cause users to deliberately cover up their real ideas because of shame. With the help of computer technology, users with mental health problems can be identified automatically, which can help mental health experts diagnose quickly and improve the efficiency of mental health counseling. In addition, the use of social network...
data to analyze mental health problems is not limited by region, it does not need the participation of psychiatrists in the whole process, and it can be effectively identified in the early stage of the disease.

RELATED WORKS
Mental health, refers to the state of mental well-being and tranquility, or the state without psychological illness (Marquez, Bustamante, Kozy-Mayle et al., 2012). From the perspective of positive psychology, mental health can also refer to a person’s ability to enjoy life, to achieve balance and happiness in various activities and work.

The definition standard of mental health is abstract, but there is not a unified definition at present. According to the definition of the World Health Organization (WHO), mental health includes "subjective well-being, the sense of personal efficacy, autonomy, interaction with other people, the realization of personal potential in intelligence and emotion, etc. (Idris & Dollard, 2014)".

At present, there are many researches on the automatic assessment of mental health, which can be summarized as depression detection and analysis, suicide intention detection and analysis, and Big Five personality analysis and emotional change (Rantanen, Metsäpelto, Feldt et al., 2007).

There are two kinds of methods to evaluate mental health automatically based on text data mining: unsupervised method based on dictionary judgment rules and topic clustering (Xu & Oard, 2011), supervised method based on text mining (Krogel & Scheffer, 2004). In the early stage of using computers to study the methods of automatic evaluation of users' mental health, most of them used unsupervised methods. Nerbonne (2014) constructed a dictionary (LIWC) containing psychological knowledge. Later, many researchers used LIWC dictionary to carry out mental health assessment research. Simmons, Gordon, & Chambless (2005) found that the first person singular pronouns were used more frequently in texts written by users with mental health disorders. Sun & Bin (2018) found the statistical rule of word frequency of big five personality in post texts of social networks. Roy's (1999) research on depression patients and individuals who have already committed suicide shows that most of the content expressed in their writing texts is self-centered, more use of singular first person and combination with negative emotional words. Stirman & Pennebaker (2001) found that compared with the normal control group, the first person singular and death related words were used more frequently in poems published by poets who had committed suicide.

The algorithm based on dictionary judgment is simple and effective, but the result of evaluation depends on the quality of dictionary. Automatic dictionary construction can overcome the shortcomings of manual dictionary. The conventional automatic construction method of emotion dictionary is to expand the seed emotion dictionary by using the data of different relations between seed words and other words. Kong & Zhang (2010) first annotated some suicide words manually as a seed dictionary. Based on word2vec model, the similarity between words was found to expand the suicide dictionary. McNamara, Crossley, Roscoe et al. (2015) used the semantic information in HowNet to guide deep network learning deeper semantic information of words to expand LIWC dictionary.

For supervised method based on text mining, the adopted method is mainly based on classification. Firstly, psychological experts directly analyzed social network data, construct training set of automatic mental health assessment. Then candidate features are extracted, classification model is trained, and automatic mental health assessment and detection of large-scale social network users is realized. Sun & Bin (2017) confirmed the correlation between personality and mental health in traditional theories through the online data of microblog users, and predicted the mental health status of microblog users by multi task regression learning method. Jung, Park, & Song (2017) regard each user as a node, and calculate the connection weight between nodes through information such as common friends, common labels, emotional tendencies in interaction. By adding the feature of connection weight between user nodes, the accuracy of depressive user detection results is increased by 15%. Scherer, Baysinger, Zolynsky et al. (2013) constructed a model through deep learning method to predict the personality of social network users.
MENTAL HEALTH ASSESSMENT MODEL BASED ON CNN

For social network set \( U = \langle P, B, R \rangle \), \( P = \{p_1, p_2, ..., p_N\} \) represents \( N \) posts, which included author of post, published time and other information. \( B = \{b_1, b_2, ..., b_M\} \) represents \( M \) topic blocks, each topic block contains several posts. \( R \) is the relationship between posts and topic blocks, \( r_{ij} \) represents post \( p_i \) belongs to topic block \( b_j \). The content and topic of the post can reflect the mental health of the author, concerns of different emergency levels are needed to social network users according to mental health status. There are four emergency levels \( E = \{C: \text{very urgent}, U: \text{urgent}, A: \text{not urgent}, G: \text{not require any psychological intervention}\}\).

In this paper, the definition of social network users' mental health automatic assessment is: for post \( p_i \) in \( U \), a mapping function or classification model \( m \) and a set of features \( F \) are found, the results of \( m(F(p_i, B, R, O)) \in E \) can correctly reflect the urgency that the post \( p_i \) needs to be concerned by mental health experts, it is the mental health of the author of the post. The feature \( F \) can come from the post \( p_i \) itself, the topic block \( b \) where the \( p_i \) is belong to, or other external data resource \( O \).

In this paper, we used the statistical characteristics of each word in different types of posts in LIWC dictionary to judge the ability of these words to distinguish users' mental health. And it guided the convolution neural network to extract the semantic information related to mental health of posts, which can be used for users' mental health detection.

Word weight calculation in LIWC

Linguistic inquiry and word count (LIWC) is a text analysis and measurement tool for the measurement of psychological words. The original intention of LIWC is to replace human professionals in the form of computer program automation to conduct statistical analysis of various texts. LIWC dictionary contains a variety of psychological process words, social process words and language process words, which make it widely used in the field of psychological science and computer science. Researchers had found that using LIWC dictionary to analyze text can effectively identify the emotional tendency and even mental health of the speaker. Therefore, LIWC is widely used in the analysis of depression, big five personality, suicidal tendency and other mental health problems.

In order to better mine and represent the psychological feature information in the text, a convolutional neural network model based on LIWC is proposed. Based on the LIWC dictionary, the ability of each word to distinguish the users’ mental health is calculated, and then the convolution neural network is guided to extract more effective psychological features from the posts, and the mental health automatic assessment model is trained.

There are 80 categories of words in the LIWC dictionary. Because some of them are general descriptive categories, there is no actual semantics. Therefore, we only used 64 categories of valid words with semantics. Based on standard deviation analysis, we calculated the weight of every word in LIWC dictionary, which mainly includes the following three steps:

Step 1: Calculating word frequency. The frequency of words in four kinds of posts in LIWC dictionary is counted.

Step 2: Calculating the standard deviation of word frequency. The standard deviation of each word frequency in the four kinds of posts obtained in Step 1 is normalized with the maximum value. The standard deviation can be used to measure the difference of LIWC words in different mental health categories. The larger the standard deviation is, the better it is to distinguish the mental health categories to which the samples belong.

Step 3: Adjusting the weight of words in the post. For each word in the post, its weight is adjusted according to its own LIWC words, so that words with strong ability to distinguish each post category have higher weight.

CNN model based on LIWC

Figure 1. The structure of CNN model based on LIWC

In this paper, a convolution neural network...
model based on psychological knowledge is proposed. The model is based on CNN model combined with LIWC dictionary, including input layer, attention layer, convolution layer, pooling layer and full connection layer. Its structure is shown as Figure 1.

Input layer: Firstly, the post is processed as a sequence of \( n \) words, then, Word2Vec (Zhang, Xu, Su et al., 2015) model is used to train the word vector model on the full data set, and each post is represented as the word vector matrix \( X = [\vec{x}_1, \vec{x}_2, ..., \vec{x}_n] \), where \( \vec{x}_i \in R^{1 \times d} \), \( d \) represents the dimension of word vector, \( \vec{x}_i \) represents vector of the \( i \)-th word.

Attention layer: In LIWC, the standard deviation analysis shows that the standard deviation of word frequency of different words in four types of posts is different. It means that their ability to distinguish four types of samples is also different. Words with large standard deviation are more helpful to enhance the impact of the model on post recognition of different mental health conditions. Therefore, in CNN, the attention layer is used to introduce the ability of LIWC to distinguish posts and strengthen the identification of posts with mental health problems.

Convolutional layer: The layer is used for feature detection. The convolution kernels of different size windows are used to obtain the local semantic information in the text sequence, so as to mine the features in the text data. In this paper, the convolution kernel with the sliding window size of \( h \) is used to select the local word vector information of the sentence in the sample post, and the new word vector matrix is obtained. For a sentence of length \( n \), it is divided into \( < x_{0:h-1}, x_{1:h-1}, ..., x_{n-h+1:n} > \), then the convolution of each component \( x \) is performed according to Eq. (1):

\[
\hat{g}_j = \delta(f(c \cdot x_{j:h-1} + a))
\]

where, \( f(\cdot) \) represents convolution function, \( c \in R^{h \times k} \) represents weight parameters of convolution kernel, \( b \in R \) represents offset value.

And through the \( \delta(x) = 1/1 + e^{-x} \) activation function to carry on the linear activation processing. Finally, the feature matrix \( G \) with global semantic information is obtained by splicing the feature vectors. Feature matrix \( G \) can be calculated as follows:

\[
G = \hat{g}_1 \oplus \hat{g}_2 \oplus ... \oplus \hat{g}_j \oplus ... \oplus \hat{g}_{n-h+1}
\]

where, \( g_j \) represents eigenvalue after convolution of component \( x_{j:h-1} \), \( \oplus \) represents vector splicing operation.

Pooling layer: The most effective eigenvalue for the target task is selected in the global semantic information by the pooling function. The \( \text{max}() \) function is used to complete this process, as shown in Eq. (3):

\[
\bar{G} = \{\text{max}(\hat{g}_1), \text{max}(\hat{g}_2), ..., \text{max}(\hat{g}_{n-h+1})\}
\]

where, \( \text{max}(\hat{g}_j) \) represents the maximum value selected from all elements of the eigenvector \( G \).

Full connection layer: In this layer, the feature vectors obtained by convolution operation and pooling operation of different convolution windows are used to represent the semantic feature information of the original text, it is denoted as \( \bar{F} \), as shown in Eq. (4).

\[
\bar{F} = \sum_{i=1}^{n-h+1}(c \cdot G_i + a_i)
\]

Finally, we use the function of \( \text{Con} \_\text{max}() \) to calculate the post classification result \( y \), shown as Eq. (5):

\[
y = \text{Con} \_\text{max} (V \cdot \bar{F} + a)
\]

where, \( V \in R^{1 \times d} \) represents weight coefficient.

**EXPERIMENT AND ANALYSIS**

In the experiment, the trained word vector was used to initialize the post representation, back propagation is used to minimize cross entropy loss function for learning model parameters. In order to avoid over fitting, the weight coefficient \( w \) in the network is regularized when the loss function is optimized.

The experiment data adopts posts of Tianya forum, which contains posts such as personal identification, post user name, forum block where the post is, number of times the post is viewed, number of feedback of the post, text content of the post and other information. The dataset includes posts published between January 2015 and June 2019, which are divided into training sets and test sets. There are 65680 posts in the training set, including 1205 manually annotated posts and 64475 not annotated. There are 92315 posts in the test set, including 550 manually annotated
posts for evaluation, and 91765 other posts related to manually annotated posts (the same author or with reply relationship).

Table 1 shows the distribution of all kinds of samples. It can be seen that the distribution of all kinds of samples is extremely uneven, and the number of people with mental health problems is very small, which is consistent with the actual situation.

Table 1. Emergency levels distribution of samples

| Emergency levels | Training data | Test data |
|------------------|---------------|-----------|
| C                | 38 (3.15%)    | 55 (11%)  |
| U                | 143 (11.86%)  | 62 (12.4%)|
| A                | 305 (25.32%)  | 121 (24.2%)|
| G                | 719 (59.67%)  | 262 (52.4%)|
| Total            | 1205          | 500       |
| Not annotated    | 64475         | 91765     |

Parameters of the proposed model is set as Table 2, a large number of verification experiments on training data is performed.

Table 2. Model parameter setting

| Parameters | Description                     | Value |
|------------|---------------------------------|-------|
| D          | Word vector dimension           | 128   |
| m          | Number of convolution kernels   | 60    |
| w          | Convolution kernel sliding window size | 7    |
| ε          | Learning rate                   | 0.02  |
| γ          | Regularization coefficient      | 0.1   |
| p          | Dropout proportion              | 0.6   |

Due to the small number of manually annotated training samples, more iterations may lead to model over fitting, while less iterations may make it difficult for model to learn effective features. In order to investigate the influence of the number of iterations on classification performance, the original training set is divided into training set and verification set according to 9:1 for cross validation. The experimental results are shown in Figure 2.

Figure 2. The influence of the number of iterations on the model

As shown in Figure 2, with the increasing number of iterations, the accuracy of training set is increasing, even approaching to 1.0. In the validation set, the accuracy rate increases rapidly first and then decreases gradually with the increasing number of iterations. At 600 iterations, the accuracy of the validation set reaches the maximum, so we choose the model at 600 iterations as the final test model.

In order to verify the effect of introducing psychological knowledge into CNN model, we compared the proposed model in this paper with traditional CNN and other prediction models (Bank, MacNeill, & Lichtenberg, 2000; Xu, 2011). Table 3 shows the experimental results of the proposed model and other models.

The experimental results of various models in Table 3 are analyzed as follows:

FastText model is ineffective in identifying posts in urgent need of psychological intervention. FastText has the highest accuracy of 0.89 in all models on the A class. However, the F1 value of G class is only 0.16, and the performance of other indicators is significantly lower than other models.

Using Word2Vec as the input of CNN model can improve the performance of CNN. It shows

Table 3. Comparison of experimental results of various models

| Model            | G     | A     | U     | C     |
|------------------|-------|-------|-------|-------|
| FastText         | 0.16  | 0.82  | 0.89  | 0.01  | 0.02  | 0.35 | 0.61 |
| CNN              | 0.35  | 0.76  | 0.73  | 0.45  | 0.38  | 0.44 | 0.63 |
| CNN+Word2Vec     | 0.38  | 0.78  | 0.83  | 0.51  | 0.67  | 0.49 | 0.64 |
| Our Model        | 0.41  | 0.80  | 0.81  | 0.53  | 0.72  | 0.51 | 0.66 |
that the word vector features of Word2Vec are able to describe the semantics of short text than directly using words, which is conducive to identifying the mental health status of short text users.

The proposed model in this paper is more accurate in identifying posts in urgent need of psychological intervention. It is because the proposed model weights the word vectors in the input layer, and then uses the convolution kernel to extract the deeper features. Therefore, after updating each back propagation model parameter, the weight of the word vectors of each input word will be adjusted as a whole, which can force the model to better learn the more relevant psychological features.

Through the above comparison, it shows that under the guidance of psychological knowledge, the proposed model in this paper can more effectively extract the information with mental health characteristics, and has more advantages in identifying the mental health status reflected by posts, especially for the identification of posts in urgent need of psychological intervention. It reflects the guiding role of psychological knowledge in the process of deep learning feature extraction.

**CONCLUSIONS**

The automatic evaluation of users’ mental health by post content needs a good fusion and processing of sentence or word context information, while convolutional neural network can fully consider the semantic information of words in post and the semantic information of the whole text. In addition, different words in the post play different roles in detecting mental health. In this paper, a convolutional neural network model based on psychological knowledge is proposed, through LIWC dictionary, the distribution difference of words in different kinds of posts is used to measure the role of each word in identifying different mental health samples, so as to guide CNN to extract more effective semantic information of posts and evaluate social network users’ mental health more effectively. Compared with other models, the proposed model is more effective in identifying social network users with mental health problems.

**Acknowledgments**

This research was funded by Shandong Provincial Natural Science Foundation, China, grant number ZR2017MG011. This research is also supported by the Humanity and Social Science Youth foundation of Ministry of Education of China (grant no. 15YJC860001), National Statistical Science Research Project (grant no. 2017LZ38).

**REFERENCES**

Bank, A. L., MacNeill, S. E., & Lichtenberg, P. A. (2000). Cross validation of the MacNeill-Lichtenberg Decision Tree: Triaging mental health problems in geriatric rehabilitation patients. *Rehabilitation Psychology, 45*(2), 193.

Ciobanu, L. G., Ferrari, A. J., Erskine, H. E., Santomauro, D. F., Charlson, F. J., Leung, J., & Baune, B. T. (2018). The prevalence and burden of mental and substance use disorders in Australia: Findings from the Global Burden of Disease Study 2015. *Australian & New Zealand Journal of Psychiatry, 52*(5), 483-490.

Ferrari, A. J., Norman, R. E., Freedman, G., Baxter, A. J., Pirkis, J. E., Harris, M. G., & Whiteford, H. A. (2014). The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the Global Burden of Disease Study 2010. *PloS one, 9*(4), e91936.

Idris, M. A., & Dollard, M. F. (2014). Psychosocial safety climate, emotional demands, burnout, and depression: A longitudinal multilevel study in the Malaysian private sector. *Journal of occupational health psychology, 19*(3), 291.

Jung, H., Park, H. A., & Song, T. M. (2017). Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals. *Journal of medical internet research, 19*(7), e259.

Kong, Y., & Zhang, J. (2010). Access to farming pesticides and risk for suicide in Chinese rural young people. *Psychiatry research, 179*(2), 217-221.

Krogel, M. A., & Scheffer, T. (2004). Multi-relational learning, text mining, and semi-supervised learning for functional genomics. *Machine Learning, 57*(1-2), 61-81.

Marquez, D. X., Bustamante, E. E., Kozey-Keadle,
S., Kraemer, J., & Carrion, I. (2012). Physical activity and psychosocial and mental health of older caregivers and non-caregivers. *Geriatric Nursing, 33*(5), 358-365.

McNamara, D. S., Crossley, S. A., Roscoe, R. D., Allen, L. K., & Dai, J. (2015). A hierarchical classification approach to automated essay scoring. *Assessing Writing, 23*, 35-59.

Nerbonne, J. (2014). The secret life of pronouns. What our words say about us. *Literary and Linguistic Computing, 29*(1), 139-142.

Rantanen, J., Metsäpelto, R. L., Feldt, T., Pulkkinen, L. E. A., & Kokko, K. (2007). Long-term stability in the Big Five personality traits in adulthood. *Scandinavian Journal of Psychology, 48*(6), 511-518.

Roy, A. (1999). Depressed patients who suicide at their first attempt have had few admissions. *Depression and anxiety, 9*(2), 75-77.

Scherer, K. T., Baysinger, M., Zolynsky, D., & LeBreton, J. M. (2013). Predicting counterproductive work behaviors with subclinical psychopathy: Beyond the Five Factor Model of personality. *Personality and Individual Differences, 55*(3), 300-305.

Simmons, R. A., Gordon, P. C., & Chambless, D. L. (2005). Pronouns in marital interaction: What do “you” and “I” say about marital health. *Psychological science, 16*(12), 932-936.

Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic medicine, 63*(4), 517-522.

Sun, G., & Bin, S. (2017). Router-level internet topology evolution model based on multi-subnet composited complex network model. *Journal of Internet Technology, 18*(6), 1275-1283.

Sun, G., & Bin, S. (2018). A new opinion leaders detecting algorithm in multi-relationship online social networks. *Multimedia Tools and Applications, 77*(4), 4295-4307.

Xu, J. (2011). An extended one-versus-rest support vector machine for multi-label classification. *Neurocomputing, 74*(17), 3114-3124.

Xu, T., & Oard, D. W. (2011). Wikipedia-based topic clustering for microblogs. *Proceedings of the American Society for Information Science and Technology, 48*(1), 1-10.

Zhang, D., Xu, H., Su, Z., & Xu, Y. (2015). Chinese comments sentiment classification based on word2vec and SVMperf. *Expert Systems with Applications, 42*(4), 1857-1863.