Psychological crisis warning of international students based on deep learning and computational mathematics

Le Liu[1], Hao Zhang[1,2*], Chengzhu Li[1], Hao Cui[1], Wenwen Zhang[1]

[1]Institute of Education, Guizhou Normal University, Guiyang, China

[2]Education Policy and Law Research Center, Guizhou Normal University, Guiyang, China

*Corresponding author: zh@gznu.edu.cn

Keywords: Graph CNN; Psychological crisis; International students; Deep learning

Abstract: With the rapid development of education for overseas students in China, psychological problems caused by cross-cultural factors have become more and more prominent. Based on the psychological questionnaire data, in order to make full use of the relationship information between students, this paper proposes a psychological crisis individual identification model based on bipartite graph convolution network model (B-GCN) to make up for the shortcomings of traditional identification methods. Because the GCN model can not realize inductive learning, this paper improves the GCN model based on the structural characteristics of psychological tests, and proposes a bipartite graph neural network model. The model can classify the students who have never appeared in the graph structure, so as to realize the early warning of a psychological crisis. Experiments demonstrate that the proposed B-GCN model has a good performance.

1. Introduction

In recent years, the scale of international students in China has been expanding, and the countries of origin and types of students is further increased year by year. According to statistics, in 2018, a total of 492,200 students from 196 countries and regions chose to study in China, among which 260,600 students from 64 countries along the "Belt and Road" route, accounting for 52.95 percent of the total number. The number of international students in China is increasing, and international educational exchange activities of Chinese universities are gradually frequent. Because of this, in recent years, the management of international students in China has triggered heated social discussion, and the emergence of various emergencies has also prompted the collective attention to quickly shift to the mental health of international students. International students are at various stages of life[1-2]. At the same time, the relevant research is slightly weak.

In the process of education and teaching, colleges and universities should not only give attention to the cultivation of college students' professional learning skills and the improvement of ideological and moral quality, but also pay more attention to the psychological health of college students. In the four-level work network system of mental health education, psychological crisis intervention is the most important link, timely and accurate crisis intervention is very likely to save more lives, and how to accurately identify the crisis, timely warning, effective measures are the key to achieve psychological crisis intervention.

This paper studies the psychological crisis of international students based on graph convolutional neural network model. The second section introduces the related work of psychological crisis research, the third section describes the model proposed in this paper, the fourth section is the experiment and analysis, and the conclusion is in the fifth section.

The study of psychological crisis first began with the study of bereavement, and Lindeman[3] established the theory of grief crisis based on it. In 1964, American psychologist Gaplan published his theory for the first time and put forward the concept of psychological crisis. He believed that psychological crisis was a serious psychological imbalance in the face of sudden and major life adversity [4].

Classification is an important branch of data mining, which can be used to extract models

DOI: 10.25236/iemetc.2023.026

describing important data categories or predict future data trends. Through classification and prediction, it can provide good decision support for the future development of an industry. In fact, data mining technology has played a great role in many problems of early warning and decision making at home and abroad [5]. Noh[6] proposed the construction of context-aware library based on data mining technology, which provides context-aware access services, reference services, security services, comfort services and personalized recommendation services by using user behavior data, mobile path data and environmental data. Judith and other scholars obtained the performance information of undergraduates and postgraduates from the graduate exemption plan of ETH Zurich, applied data mining methods to predict the results, and analyzed how to predict GGPA with the data on undergraduate transcripts, so as to infer the performance of students during the postgraduate period [7]. Based on the training system of California Institute of Technology, Ravikumar et al. [8] designed an academic early warning system to predict students' academic risks by monitoring their scores of each course and assigning appropriate study partners to help them. Yang et al. realized a prediction model with high accuracy for the final scores of Java programming courses by collecting students' past learning data and using data mining algorithm. Huang et al. [9] used Weibo, forums and other public platforms to collect keywords, analyze their frequency, and judge whether the text has suicidal tendency. Chattopadhyay et al. [10] used the method of data mining to verify whether there was difference between suicide note and ordinary text or simulated art, so as to judge whether the author had suicidal tendency.

Some studies have used data mining techniques for the detection of specific mental illnesses, such as depression. Depression is affected by a variety of factors, such as the gene expression profile of patients [11], whether they have long-term depression [12], psychiatric and general health comorbidities [13], and intolerable side effects of drugs. Patients often need more than one course of treatment to obtain treatment, and even those who have responded to the treatment stage are at potential risk of relapse. Zhi Nie et al. proposed a regression method to predict the likelihood of relapse in patients with depression after initial remission from antidepressant therapy [14].

2. The proposed model

2.1 Graph Convolutional Network

Graph Convolutional Neural Network (GCN) belongs to the extended Network of Graph Neural Network (GNN). It is a Graph Neural Network that adopts convolution operation and supports the application of Neural Network structure in traditional topological graphs.

In this kind of feature graph, the graph convolution process can be abstractly understood as that each node in the graph will be affected by neighboring nodes or indirectly adjacent nodes at any time, and then continuously change its own attribute state until it reaches a relative balance. Among them, neighbor nodes with smaller shortest path length will have greater influence. However, any graph convolution layer can be expressed as a corresponding nonlinear function, and its formal general formula is as follows.

$$H^{l+1} = f(H^l, A) \tag{1}$$

Among them, the $\mathrm{H}^0 = \mathrm{X}$ for the first layer of input, $\mathrm{X} \in \mathrm{R}^{\mathrm{N}*\mathrm{D}}$, number of nodes for the Fig.1 dimensions of feature vector for each node as the adjacency matrix, different models of differences is the realization of the function different.

Common implementation methods include the following three:

(1) $\mathrm{H}^{\mathrm{l}+1} = \sigma(\mathrm{A}\mathrm{H}^{\mathrm{l}}\mathrm{W}^{\mathrm{l}})$. The W^{l} matrix for the weight of the first layer parameters, is a nonlinear activation function, such as ReLU, Sigmoid, etc.

(2) $\mathrm{H}^{\mathrm{l}+1} = \sigma(\mathrm{L}\mathrm{H}^{\mathrm{l}}\mathrm{W}^{\mathrm{l}})$. The Laplacian matrix is introduced to solve the problem of not considering the self-transmission of node information.

(3) $\mathrm{H}^{\mathrm{l}+1} = \sigma\left(\mathrm{D}^{-\frac{1}{2}}\widehat{\mathrm{A}}\mathrm{D}^{-\frac{1}{2}}\mathrm{H}^{\mathrm{l}}\mathrm{W}^{\mathrm{l}}\right)$. This method improves the original Laplacian matrix operator to a symmetric normalized Laplacian operator, which can not only solve the self-transfer problem, but

also normalize the adjacency matrix. At the same time, the multi-label classification model proposed in this paper also adopts this implementation scheme.

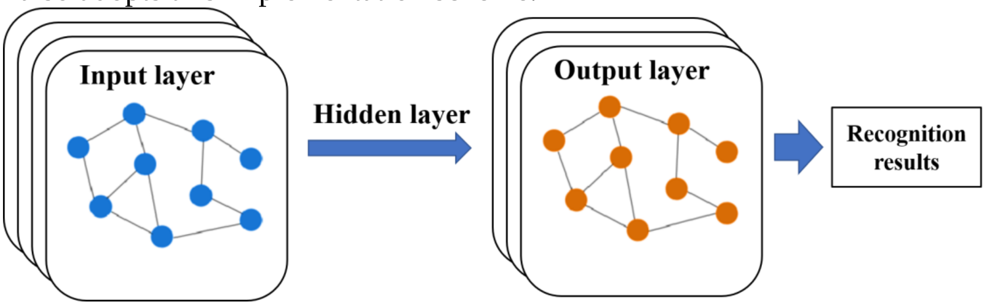

Fig.1. Schematic depiction of model

In this paper, the multi-layer graph convolutional neural network GCN is trained to perform graph convolution operation on students' representation vectors, so as to identify individuals with psychological crisis.

2.2 The construction of Graph

In this study, students and the questions in the psychological test questionnaire are taken as nodes at the same time, and two types of nodes with different properties are connected to form a bisection Graph. A Bipartite Graph Convolutional Network (B-GCN) is proposed to identify individuals in psychological crisis. Because students and questionnaires can be separated and reorganized, it is expected to improve the generalization and practicability of the model by reducing the coupling degree between student nodes and problem nodes in the figure. The overall framework of individual identification of psychological crisis based on B-GCN is shown in Figure 2.

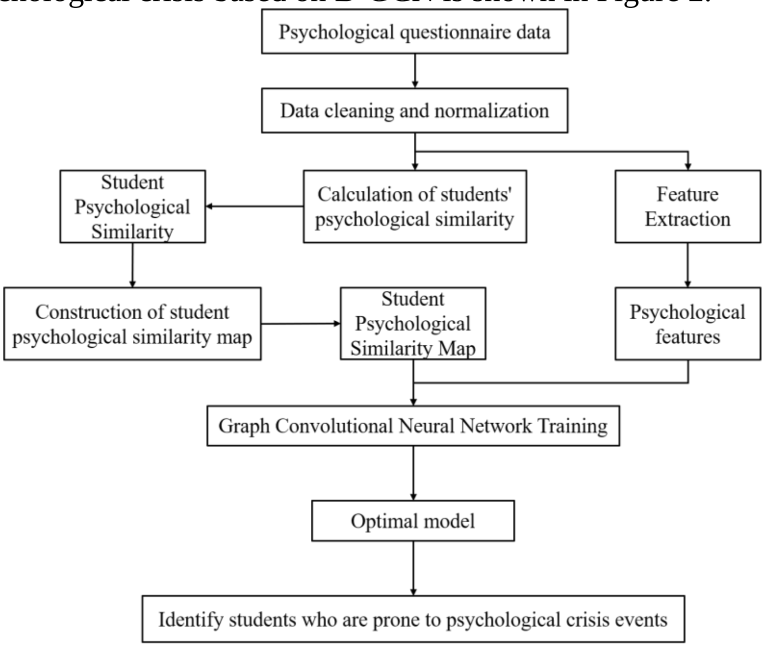

Fig.2. Overall framework

Next, it explains how to construct a dichotomous graph based on the student's psychological test questionnaire. The complete bisection graph structure as shown in fig. 3 is formed:

(1)Calculation of weight matrix

As shown in figure 3, the student node S_i and each psychological test questionnaire an edge node Q_j connections, while right do set according to the psychological test questionnaire answers. There are two types of answer options for the psychometric questionnaire. There were yes and no questions, such as "Have you ever had counseling or therapy?" and "Do you think people waste too much time scrimping to secure their future?" One is multiple choice, where the answer options are used to indicate how much you feel about a situation. Most of these psychometric questions are graded on a five-point scale, such as "I tend to study hard but hate studying", "I feel lonely, The options for questions like "No one really understands me" were set to "totally agree", "mostly agree", "not sure", "mostly don't agree", and "not at all agree".

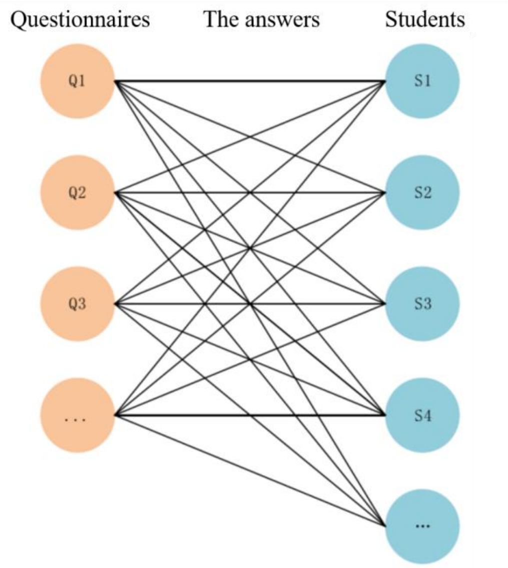

Fig.3. Complete bipartite graph of students and questionnaires

3. Experiment and Analysis

3.1 Experimental setup

In this study, 515 psychological test data questionnaires were distributed from 10 colleges and universities by stratified sampling method, and 502 valid questionnaires were recovered, with an effective rate of 97%. Among them, 314 were male students and 188 were female students; 415 were undergraduate students and 87 were master and doctoral students; the average age was 27.8 years old. The psychological test data questionnaire consisted of 40 items, including 9 dimensions of parental marital status, attachment relationship, family function, childhood trauma, coping style, past experience, social support, life meaning, learning motivation, and students' personal information. The entire dataset, containing 102 positive and 400 negative samples. Among them, the test set contains 32 positive samples and 120 negative samples, and the rest of the data are used as the training set for learning the parameters of the B-GCN model. In the model testing phase, the students in the test set are used to construct into a new bipartite graph, but the parameters learned by the B-GCN model remain unchanged. In this experiment, two cases of directed and undirected graphs are tested separately. The B-GCN model in the experiment uses three layers of graph convolution layers. The dropout layer is added between the convolutional layers, the discard probability is set to 0.5, and the number of hidden features is set to 80 and 30, respectively. the learning rate is set to 1e-2 using the Adam optimization function. the negative log-likelihood function is used as the loss function.

3.2 Evaluation Metrics

Commonly used evaluation metrics in classification tasks include precision, accuracy and F1 score.

Experimental results and analysis

In the case of directed graphs, the influence of student nodes by the question nodes is related to the answers to the questions of the student nodes, and the influence of the question nodes by each student node is of the same size, and the characteristics of each student node are considered in equilibrium. Under this assumption, the model achieves an accuracy of 82.24%, an precision of 55.55%, a recall of 78.12%, and an F1 score of 64.93%.

In the case of undirected graphs, the student node is influenced by the same amount of influence from the problem node as the problem node is influenced by the student node. This means that if the features of a problem have a large influence on the features of the student node, then the features of the student node also have a large influence on the features of that problem.

In the case of undirected graph, the model achieves 89.47% accuracy, 72.22% precision, 81.25% recall, and 76.47% F1 score, which is a higher model performance compared to the model with directed graph.

Table 1 Comparison of experimental results

Method	Accuracy	Precision	Recall	F1
GCN	73.68%	42.3%	68.75%	52.38%
B-GCN-directed	82.24%	55.55%	78.12%	64.93%
B-GCN-undirected	89.47%	72.22%	81.25%	76.47%

Table 1 compares the model performance for the three cases of GCN, directed graphs and undirected graphs. Table 1 shows that the GCN model, which uses only student mental similarity graphs constructed with student nodes, has significantly lower model performance than the two types of bipartite graph convolutional networks, which use more information-rich student-problem bipartite graphs. Specifically, using the bipartite graph convolutional network improved the accuracy rate by more than ten percentage points, with a significant improvement of 0.1255 and 0.2409 in F1, respectively. In particular, the accuracy of the B-GCN based on undirected graph is nearly twice that of the GCN model. Not only that, the GCN model cannot classify student nodes that have never been seen before, so it is not suitable for use in practical scenarios. In contrast, the bifurcated graph convolutional network does not have this problem. For new student nodes, it is only necessary to convert their options into bipartite graphs and plug them into the trained network model to recognize them.

Among the two types of models of bipartite graph convolutional networks, the performance of the undirected graph-based model is better than that of the directed graph-based model. This may be due to the fact that in the undirected graph setting, the interaction between a particular problem node and a particular student node is rapidly enhanced, thus favoring student classification. Although directed graphs may also learn this influence after enough layers of network superposition, graph neural networks are generally not superimposed with too many layers to avoid over-smoothing effects. With fewer layers, the undirected graph setup achieved better performance compared to the directed graph. It is worth noting that some students were flagged by the model as individuals in psychological crisis in this study, but it was not possible to determine whether the students were suffering from psychological disorders for the time being. In fact, these students identified as psychological crisis individuals by the model may be at some risk, or may even be already ill but undiagnosed, and schools may give some procedural care and attention to such students.

4. Conclusion

Strengthening the psychological health education for international students and guiding their healthy growth and success is an important research issue in the education of international students. This study uses psychological questionnaire data for research and analysis. In this paper, a dichotomous graph convolutional network model (B-GCN) is firstly proposed to make up for the drawback that the GCN model cannot achieve the problem of inductive learning. Experiments show that the model possesses a good performance with Accuracy reaching 89.47%, Precision reaching 72.22%, Recall 81.25%, and F1 score 76.47%.

References

[1] JONES R, FINLAY F.Medical students'experiences and perception of support following the death of a patient in the UK, and while overseas during their elective period[J].Postgraduate Medical Journal, 2014, 90(1060):69-74.

[2] WANG C C, ANDRE K, GREENWOOD K M.Chinese students studying at australian universities with specific reference to nursing students:A narrative literature review[J].Nurse

Education Today, 2015, 35(4):609-619.

[3] Lindemann E. Symptomatology and management of acute grief[J]. Pastoral Psychology, 1944, 151(6): 155-160.

[4] Matei Zaharia, Mosharaf Chowdhury, Tathagata Das, et al. Resilient Distributed Datasets: A Fault_Tolerant Abstraction for In-Memory Cluster Computing[C]. Presented as part of the 9th USENIX Symposium on Networked Systems Design and Implementation. 2012.

[5] Wong T. Alternative prior assumption for improving the performance of naive Bayesian classifiers[J]. Data Mining & Knowledge Discovery, 2009, 18(2): 183-213.

[6] Noh Younghee. A study on next-generation digital library using context-awareness technology[J]. Library Hi Tech, 2013, 31(2).

[7] Saarela M, Karkkainen T. Analysing Student Performance Using Sparse Data of Core Bachelor Courses[J]. Journal of Educational Data Mining, 2015, 7(6): 393-401.

[8] Ravikumar R, Aljanahi F, Rajan A, et al. Early Alert System for Detection of At-Risk Students[C]. 2018 Fifth HCT Information Technology Trends(ITT). IEEE, 2018. 138-142.

[9] Yang H, Diao J, Zhou T,et al. Research and Implement of Course Early Warning System Based on Teaching Behaviour Data[C]. Journal of Physics: Conference Series. IOP Publishing, 2019. 42-52.

[10] Chattopadhyay S, Ray P, Chen H S, et al. Suicidal Risk Evaluation Using a Similarity-Based Classifier[J]. Advanced Data Mining and Applications, 2008, 51-61.

[11] Eyre H A , Eskin A , Nelson S F , et al. Genomic predictors of remission to antidepressant treatment in geriatric depression using genome-wide expression analyses: a pilot study[J]. International Journal of Geriatric Psychiatry, 2016, 31(5):510-517.

[12] J. Alpert and M. Fava. Handbook of Chronic Depression: Diagnosis and Therapeutic Management.Medical Psychiatry[M]. Marcel Dekker Incorporated,2014.

[13] Otte C . Incomplete remission in depression: Role of psychiatric and somatic comorbidity[J]. Dialogues in Clinical Neuroscience, 2008, 10(4):453-460.

[14] Zhi Nie, Pinghua Gong, Jieping Ye. Predict Risk of Relapse for Patients with Multiple Stages of Treatment of Depression[C]// Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM, 2016.