International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5219-5228

https://doi.org/10.70517/ijhsa464362

Construction of an Intelligent Monitoring and Risk Warning Algorithm Model for Student Mental Health

Min Qin[1], Yang Li[2] and Jihua Cao[3,*]
[1] School of Design and Creativity, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China
[2] Basic Teaching Department, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China
[3] Student Work Office, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China
Corresponding authors: (e-mail: caojihua2024@126.com).

Abstract Traditional mental health management mainly relies on manual assessment and regular screening, which is characterized by inefficiency, limited coverage and lagging early warning. The rapid development of artificial intelligence technology provides new solutions for mental health management, and the early warning of students' mental health risks based on intelligent algorithms is of great practical significance. In this study, the Trans-LSTM neural network model was used to integrate the SCL-90 psychological assessment data and students' daily behavioral data to construct a multivariate time series classification mental health risk early warning system. The 8766 valid psychological assessment data were privacy-protected by k-anonymization technique, the SENet attention mechanism was applied to enhance the feature extraction capability, and the Transformer positional coding and multi-head attention mechanism were combined to optimize the time-series feature learning. The experimental results show that the proposed Trans-LSTM model performs well in the mental health prediction task, with an accuracy of 83.36%, a precision of 87.64%, a recall of 78.27%, and an F1 value of 80.14%, which are all significantly better than the comparative models such as GCN, SVM, and GraphSAGE. The study found that the detection rates of the three factors of anxiety, relationship sensitivity, and depression were 22.57%, 16.34%, and 14.69%, respectively, providing an important basis for mental health risk identification. The study shows that the model can effectively integrate heterogeneous data from multiple sources, realize accurate prediction and timely warning of students' mental health status, and provide an intelligent solution for mental health management in colleges and universities.

Index Terms intelligent algorithm, mental health, risk warning, Trans-LSTM, time series classification, multi-head attention mechanism

I. Introduction

In today's society, students' mental health problems are receiving increasing attention [1]. With the rapid development of science and technology, Artificial Intelligence (AI) technology is gradually applied to the field of students' mental health management, providing new possibilities for timely detection and intervention of students' psychological problems [2], [3]. Students face many pressures and challenges during their growth, such as academic competition, changes in the family environment, and social distress, which may have a negative impact on their mental health [4]-[6]. Traditional mental health monitoring methods often rely on questionnaires and interviews, but these methods have certain limitations, such as high subjectivity and difficulty in real-time monitoring [7], [8]. And the introduction of intelligent algorithms brings new ideas to solve these problems.

The application of intelligent algorithms in mental health management is of great significance [9]. Using intelligent algorithms, a more convenient and efficient management of students' mental health can be realized from the realization of psychological counseling, psychological assessment and psychological treatment [10], [11]. At the same time, intelligent algorithms can also realize the early warning of students' mental health risks through the analysis of students' emotions and behaviors, which is conducive to understanding the potential problems of students' psychology and providing timely help [12]-[14]. However, the application of intelligent algorithms in students' mental health management is not always smooth, and faces some challenges, especially data privacy issues [15], [16]. Students' personal data involves privacy, and how to use these data for monitoring while ensuring the security and legal use of the data is an urgent issue [17], [18].

In this study, we designed a Trans-LSTM-based student mental health risk early warning model, which firstly ensures data privacy and security through k-anonymization technique, then adopts multivariate time series classification method to process SCL-90 scale data and daily behavior data, constructs LSTM network architecture integrating Transformer attention mechanism, and enhances the temporal feature learning ability through positional coding and multi-head attention mechanism. Enhance the temporal feature learning ability, and finally verify the effectiveness and superiority of the model through comparative experiments, which provides intelligent technical support for mental health management in colleges and universities.

II. Student Mental Health Data Processing

II. A. Data sources

The data used in this study came from the results of mental health assessment and daily behaviors of students in a school in Shandong Province, and the mental health assessment was based on the SCL-90 scale, which was used by the assessors, and 8842 SCL-90 scales were recovered by the assessors using the mental health assessment system, of which 8766 were valid data. The assessment data mainly included students' basic information after desensitization (including name, age, hukou type, ethnicity, whether they were only children, etc.), as well as the answers to the 90-item scale and the scores of each dimension.

II. B. Data Privacy Calculations

k-anonymization is a processing method used to protect sensitive data; however, as the level of protection increases, the actual availability of the data decreases. In order to achieve k-anonymization, generalization and hiding techniques are usually used.Generalization refers to a more abstract and general description of the data, where the values are abstracted so that no specific data can be identified, and this is done in order to reduce the level of detail and precision of the data.The partially anonymized data information after k-anonymization privacy calculations is shown specifically in Table 1. It can be seen that the student's psychological data information was replaced with *, and the mental health assessment part of the data information obtained after the privacy calculation, the sensitive information has been anonymized to ensure the security of the mental health assessment data.

Table 1: Partial anonymous data information after k-anonymous privacy calculation

Grade	Type of account	Evaluation time	Student number	Contact Us
2022	node type_2	2024/11/8	2022040****	156****0503
2022	node type_1	2024/11/8	2022040****	184****5201
2022	node type_2	2024/11/8	2022040****	186****0448
2022	node type_2	2024/11/8	2022040****	138****3124
2022	node type_1	2024/11/8	2022040****	198****2132
2022	node type_1	2024/11/8	2022040****	132****7705
2022	node type_2	2024/11/8	2022040****	186****3548

II. C. Data pre-processing

The data used in this paper has some incomplete and invalid noise data, and the direct use of this data for experiments will have a certain impact on the experimental results, and the complex data will also have an impact on the efficiency of the implementation of the algorithm to a certain extent.

The data obtained in this paper is of high quality, and the number of data samples after processing is still 8422 data. Secondly, in this paper, the training data and test data are divided according to the ratio of 7:3, and there are 8422 sample data, so there are 5895 sample data in the training data set and 2527 sample data in the test data set.

II. C. 1) Analysis of basic user attributes

In this paper, the features that do not change over time in the short term are called static features, and the discussed static features of basic user information data analysis mainly include gender, age, type of household registration, and whether or not it is an only child. The statistics of the number of students' basic attributes are specifically shown in Table 2. It can be seen that the proportion of male and female students is higher in the proportion of women, accounting for 63.04%. Most of the students are concentrated between the ages of 14 and 25, and there are 6,286 students between the ages of 17 and 22, accounting for about 71.71% of the total, which can be deduced that students at this stage may be in the rebellious stage, and the incidence of psychological abnormality is relatively high. In terms of hukou type, the proportion of students with rural hukou was higher, totaling 6,314 students, accounting for 72.03% of the total. The number of only children is slightly higher than the number of non-only children, and the numbers of the two groups are quite close.

Table 2: Statistical analysis of the number of students with basic attributes

Static features	Attributes	Number of students	Proportion
Gender	Male	3240	36.96%
	Women	5526	63.04%
Age distribution	< 14	98	1.12%
	14-16	642	7.32%
	17-18	2099	23.94%
	19-22	4187	47.76%
	23-25	1366	15.58%
	> 25	374	4.27%
Distribution of household registration types	Rural	6314	72.03%
	Towns	2452	27.97%
Whether the only child	Yes	4905	55.95%
	No	3861	44.05%

II. C. 2) Analysis of SCL-90 scale data

The dataset of this paper is from the symptom self-assessment scale SCL-90.Given that the scale involves several measurement dimensions, this paper will focus on statistically analyzing several important dimensions, and will not describe in detail the statistical analyses of the other dimensions.The statistical results of the factor scores of the SCL-90 scale are specifically shown in Table 3. A total of 1,375 students scored more than 160 points on the scale, and its positive detection rate was 15.69%. Among the factors, the factors with the highest number of detections with factor scores greater than or equal to 2 were anxiety, interpersonal sensitivity, and depression, with the percentages of detections being 22.57%, 16.34%, and 14.69%, respectively. In the following section, the three factors of anxiety, interpersonal sensitivity, and depression will be analyzed separately.

Table 3: SCL-90 scale evaluation factor score statistics

Evaluation factor	number of students	Proportion	Number of positive cases (scale score > 160)	Proportion
Somatization	487	5.55%	1375	15.69%
Anxiety	1978	22.57%		
Interpersonal sensitivity	1432	16.34%		
Depression	1288	14.69%		
Obsessive-compulsive symptoms	974	11.11%		
Hostility	1044	11.91%		
Terror	1012	11.54%		
Biased	951	10.85%		
Psychiatric	713	8.13%		
Other	941	10.73%		

The statistical distribution of the number of anxiety scores is shown specifically in Figure 1. The higher the anxiety score, the stronger the anxiety level of the subjects. When the anxiety score is more than 30, the anxiety level of the subjects is stronger; when the anxiety score is less than 20, the anxiety level of the subjects is weaker. According to the results of this batch of data, the number of people with severe anxiety symptoms can be up to 1978, and the number of people with severe anxiety symptoms is 488, and the number of people suffering from severe anxiety symptoms amounts to 5.57%.

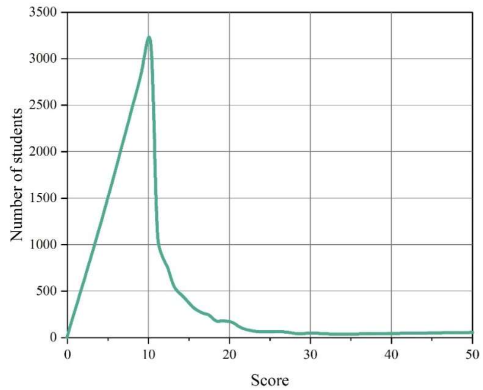

Figure 1: Statistical distribution of the number of anxiety scores

The statistical distribution of the number of depression scores is shown in Figure 2. The range of depression scores in the scale is between 13 and 65, and when the depression score exceeds 26, it indicates that the subject is strongly depressed, while exceeding 39 indicates a more severe level of depression. Looking at the distribution of the scores, it can be observed that the percentage of people with high depression scores is relatively high. In fact, those with depression scores over 26 occupied 10.12% of the total number of people. This indicates that depression is still a serious problem in the mental health of college students.

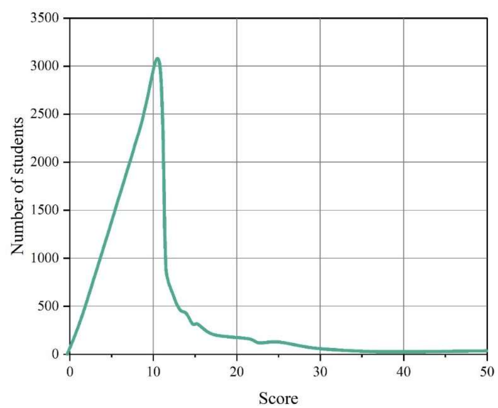

Figure 2: Statistical distribution of the number of depression scores

The score scores for interpersonal sensitivity as measured by the SCL-90 scale are shown specifically in Figure 3. Higher scores indicate greater sensitivity in these relationships, with scores above 8 indicating that individuals are sensitive or highly sensitive in their interactions with others. The figure shows that approximately one-quarter of students exhibit sensitivity in relationships that is characterized by introversion and difficulty interacting with others. Over time, this tendency to withdraw can lead to the development of a withdrawn mindset, which in turn increases the risk of other psychological disorders, such as depression and anxiety.

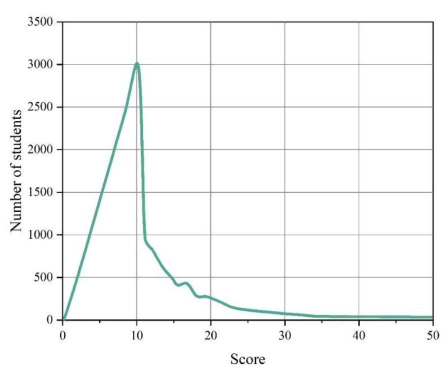

Figure 3: Statistical distribution of interpersonal sensitivity scores

II. C. 3) Visual analysis of daily behaviors

In this section, the collected data on students' daily behaviors are visualized and analyzed to determine the behavioral variability among students with different mental health conditions.

(1) Early rising data analysis

Students' daily habits can reflect their mental health status, specifically whether they go to bed early and get up early, eat on time, etc. Most of the students with good living habits have normal mental status. On the contrary, students with poor daily living habits often have certain mental health problems. The number of breakfasts and the number of early risings of the students are as specific as shown in Table 4. As shown in Table 4, there are some differences in the number of breakfasts and early risings among students with different psychological conditions, with normal students having a higher number of breakfasts and a higher degree of self-discipline.

Table 4: Number of times students had breakfast and got up early

Types of students	Average number of breakfasts per week	Average number of early waking hours per week
Normal students	4.4	2.4
Abnormal students	2.8	1.1

(2) Consumption data analysis

a) Consumption time

The consumption time of students' three meals is shown in Table 5. The consumption time of the two types of students is similar, breakfast is around 8:30-9:00, lunch is around 11:30-12:30, and dinner is around 17:30-18:00. The reason may be that the school cafeteria is open at regular intervals, so if you want to eat fresh food, you can only go to the cafeteria at a fixed time to consume it. So as long as you go to the cafeteria, the time of consumption is about the same.

Table 5: Student meal consumption time figure

Types of students	Breakfast time point	Lunch time point	Dinner time point
Normal students	8:30	11:45	17:35
Abnormal students	8:50	12:10	17:40

b) Consumption amount

The weekly amount of students' cafeteria consumption was counted and the obtained statistics are shown in Table 6. There is no obvious difference between the two types of students in the amount of weekly consumption in the cafeteria. The reason may be that there is not much difference in the consumption level of the cafeteria, and the difference in the amount of money spent by each student in the cafeteria is not very big.

Table 6: Student meal consumption amount figure

Types of students	Breakfast consumption amount	Lunch consumption amount	Dinner consumption amount
Normal students	3.8	12	12.6
Abnormal students	3.5	11.5	12.8

(3) Analysis of the dependence on the time spent on the Internet

By analyzing the number of times students go online and the number of times they go to bed late, we can find out how often students stay up late to go online, and at the same time, by analyzing the length of time students spend online and the length of time they spend resting in the dormitory, we can find out the degree of students' dependence on the Internet. Students' dependence on the Internet is shown in Table 7. Normal students have lower frequency of surfing the Internet and sleeping late, and there is a small difference in the duration of surfing the Internet and resting in the dormitory, while the duration of surfing the Internet and resting in the dormitory of abnormal students is slightly higher than that of normal students, which indicates that compared with normal students, abnormal students prefer staying in the dormitory to surf the Internet.

Table 7: Student Internet dependence figure

Internet dependence	Normal students	Abnormal students
Average number of Internet access per week	4.2	5.5
Average number of night sleeps per week	3.6	5.4
The average length of Internet access per day	3.5	4.8

The average rest time per day	10.2	11.5

III. Early warning model of students' mental health risk based on Trans-LSTM

Through the processing of students' psychological data, it is known that most students have different degrees of mental health disorders. In order to detect students' mental health problems in time and provide early warning, so that students can develop physically and mentally in all aspects and reduce the occurrence of vicious events, this chapter will construct a Trans-LSTM-based early warning model of students' mental health risks.

III. A. Model selection

In this paper, Support Vector Machines and TabNet algorithm are selected [19]. Support vector machine can be used to solve linear indivisible problems, by mapping the features to the high-dimensional space and solving the sum function, to find the classification of the super half-plane, in line with the needs of this task, and simple, efficient, easy to implement, high interpretability; and TabNet was born for the table data, especially for the table has to do too much processing, you can quickly learn the data characterization, and combined with the classification of the tree model, through the feature value of the classification, with high interpretability, and has the self-supervised learning ability to automatically fill in the missing values of the data by learning TabNet has high interpretability and self-supervised learning capability, which can automatically fill in the missing values of the data through learning.

On the other hand, the class algorithm KNN is selected when the time series classification model is chosen. the In addition, LSTM-FCN is selected as the main classifier for this experiment, because the algorithm is able to integrate the image processing model FCN and the sequence model LSTM to learn the data features from different perspectives, the model performance is excellent, the structure is simple, and there is a large space of modification, and it is able to derive the MLSTM-FCN suitable for multivariate time series classification. MLSTM-FCN algorithm [20]. Finally, as one of the best-performing models in time series classification research in the past two years, TapNet is mainly used for experimental comparison with the improved algorithm proposed in this paper, to verify that the optimization strategy is practicable and to enhance the persuasive power.

III. B. One-dimensional time series classification

It can be found through the characterization that there are obvious differences in the distribution of the number of times consumed per week by students with different psychological states. In order to test the correlation between the two, in this paper, the sequence of weekly consumption counts is input into SVM, TabNet, KNN-DTW and LSTM-FCN networks for one-dimensional time series classification preexperimentation, respectively [21]. Among them, the non-time series classifiers SVM and TabNet process the input sequences as static features, while the temporal models KNN-DTW and LSTM-FCN take into account the data temporal order. Finally, the importance of data temporal order for psychological abnormality detection is verified by comparing the experimental results of different models.

III. C. Multivariate time series classification

Multivariate Long Short-Term Memory Fully Convolutional Network (MLSTM-FCN) introduces the Seize-and-Excitation (SENet) module on top of LSTM-FCN and places it after the temporal convolutional block to further improve the model classification effect.SENet learns the feature weights of each channel of the input tensor through the fully-connected layer according to the loss training SENet learns the feature weights of each channel of the input tensor according to the loss training, which makes the strong correlation feature weights increase and the weak correlation feature weights decrease.

By analyzing the network structure, it can be seen that before the features are input into the SENet network, they need to undergo a convolutional transformation, which can be defined as follows: $F_a : X \rightarrow D, X \in \mathbb{R}^{L' \times M' \times N'}$, $D \in \mathbb{R}^{L \times M \times N}$, and the related solution procedure is shown in equation (1):

$$d_n = v_n * X = \sum_{s=1}^{N'} v_n^s * x^s \tag{1}$$

where v_n denotes the nth convolution kernel, * denotes the convolution operation, X^s denotes the sth input, and d_n denotes the nth featuremap, which is converted into tensor D after processing tensorX. Subsequently, the compression operation is executed and the related calculation formula is shown in equation (2):

$$z_n = F_b(d_n) = \frac{1}{M \times L}\sum_{i=1}^{M}\sum_{j=1}^{L} d_n(i,j) \tag{2}$$

The action is similar to global average pooling, which yields $1*1*N$ -dimensional global information. Immediately after the activation operation, in this step, the first input needs to go through the fully connected layer W_1, and then into the ReLU layer activation, and then into the fully connected layer W_2, and finally through the sigmoid output feature weight representation s. The whole computation is shown in Equation (3):

$$s = F_n(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \tag{3}$$

where W_1 and W_2 have the same dimension of $\frac{N}{r} * N$, and r is a scaling parameter which serves to reduce the number of channels and thus the computational effort. Finally, the output s is channel multiplied with the original input.

The feature vector representation with variable weights is obtained and the related expression is shown in equation (4):

$$x_n^{'} = F_d(d_n, s_n) = s_n \cdot d_n \tag{4}$$

III. D. Algorithm Optimization

The network operation process is shown in equation (5):

$$E^{(l)} = \max_{pooling}(f(W * E^{(l-1)} + b)) \tag{5}$$

$E^{(l)}$ represents the lth layer network in the encoder, which mainly consists of a temporal convolutional layer, a linear activation layer, and a maximal pooling layer, $E^{(l)} \in \mathbb{R}^{F_l \times T_l}$, L number of layers in the encoder network, F_l represents the number of filters in the lth layer of the network, T_l represents the number of time steps in the corresponding layer, w is the set of parameter weight matrices, and b is the bias. So the sensory field is calculated as shown in equation (6):

$$r(d, L) = d(2^L - 1) + 1 \tag{6}$$

From the above equation, it can be seen that the size of the receptive field of the TCN network is jointly determined by the number of network layers, the expansion factor and the size of the convolution kernel.

In order to solve the problem that the sensory field of FCN branch in MLSTM-FCN network is small and can only be carried out unidirectionally, this paper introduces the attention mechanism which is commonly used in Transfomer and replaces the TCN module in order to compose a new Trans-LSTM network.

For location coding, the mathematical expression is as follows:

$$PE_{(pos,2i)} = \sin(pos / 10000^{2i/d_{model}}) \tag{7}$$

$$PE_{(pos,2i+1)} = \cos(pos / 10000^{2i/d_{model}}) \tag{8}$$

where pos represents the position where the current encoded token is located, i denotes the dimension, and d_{model} is the dimension of the processing sequence. In addition, the multi-head attention is represented as follows:

$$MultiHead(Q, K, V) = Concat(head_1, \ldots, head_h)W^o \tag{9}$$

where W^o represents the projection matrix, which is used to solve the cross-layer dimension inconsistency problem, and $head_h$ represents the hth head, which is computed as follows:

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \tag{10}$$

The W^Q, W^K, W^V denote the query (used to match other sequence fragments), key (used to be matched) and value (information that will be extracted), respectively. The feedforward network mainly consists of two linear transformation functions and a ReLU activation function, and the computational procedure is shown in equation (11):

$$FFN(x) = max(0, xW_1 + b_1)W_2 + b_2 \tag{11}$$

IV. Experiment on Predicting Risks to Students' Mental Health

In this chapter, a student mental health risk prediction experiment will be conducted to test the validity of the student mental health risk early warning model proposed in this paper. The dataset used in this experiment comes from an extensive student mental health survey conducted by a team from Peking University in China, which interviewed a population of approximately 20,842 students aged 17-25 years old from 28 Chinese provinces, and includes data from interviews conducted from 2013 to 2023.

IV. A. Data pre-processing

There are a large number of missing values, outliers, etc. in the dataset, for which data cleaning is performed to remove outliers and fill in missing values, and data merging and normalization is performed at the end. Taking the 2023 data as an example, the sample size of the dataset is shown in Table 8. For the irregular data in the original dataset is mapped, and finally 5548 normalized sample datasets are obtained. Among them, there are 3936 psychologically unhealthy samples and 1612 psychologically healthy samples.

Table 8: Sample size of the dataset

Datasets	Sample size
Total	5548
Negative	3936
Positive	1612

IV. B. Comparative analysis of experimental results

In this paper, the low-dimensional vector representation of each interviewed middle school student node is used as the input of the downstream task, and the mental health prediction problem is transformed into a node classification task, and the changes of the accuracy and loss value of the model in this paper are shown in Fig. 4. As can be seen from the above figure, the accuracy of this paper's model gradually increases with the increase of training rounds and stabilizes after about 110 rounds of training. The accuracy of the final model is maintained at about 83.3% and the loss stabilizes after about 135 rounds of training.

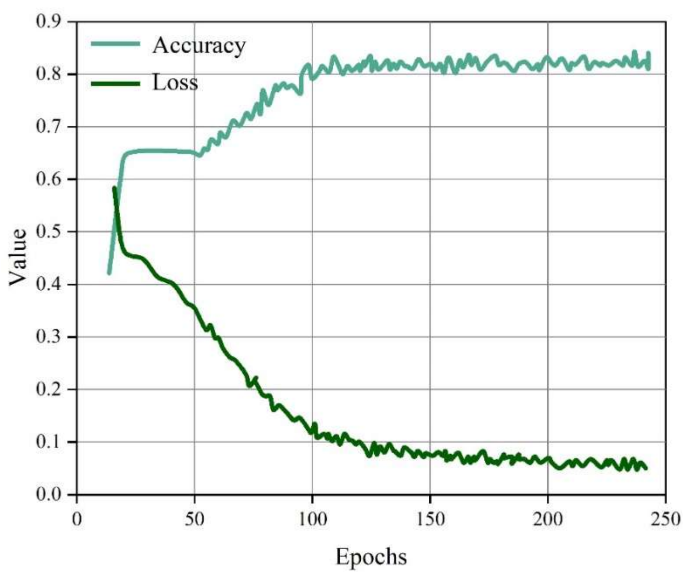

Figure 4: Changes in accuracy and loss

In order to further validate the effect of the students' mental health risk warning model proposed in this paper, five models, GCN, SVM, GraphSAGE, DGNN and MLSTM-FCN, are selected as the comparison object, and the quantitative results of the mental health prediction task comparing all models including the model in this paper are specifically shown in Table 9. It can be seen that this paper's model has a more excellent performance on the mental health prediction task than the other comparison models in terms of accuracy, precision, recall and F1 value indicators on the dataset, with the accuracy, precision, recall and F1 value reaching 83.36%, 87.64%, 78.27% and 80.14%, respectively.

Table 9: Quantitative results of mental health prediction tasks

Model	Accuracy rate	Precision rate	Recall rate	F1 value
GCN	0.7277	0.7735	0.7174	0.7364
SVM	0.7594	0.7889	0.7267	0.7526
GraphSAGE	0.7737	0.817	0.7151	0.7433
DGNN	0.7971	0.8531	0.7228	0.7855
MLSTM-FCN	0.8154	0.8537	0.7677	0.7856
The model in this paper	0.8336	0.8764	0.7827	0.8014

V. Conclusion

In this study, by constructing a Trans-LSTM-based mental health risk early warning model, we successfully realized the intelligent identification and precise prediction of students' psychological state. The experimental results show that the model has excellent comprehensive performance on 5548 sample datasets, with accuracy, precision, recall, and F1 value reaching 83.36%, 87.64%, 78.27%, and 80.14%, respectively, which is an improvement of 10.59 percentage points compared with the accuracy of the traditional GCN model. Data analysis revealed that 71.71% of the total number of students were in the age group of 17 to 22 years old, and the incidence of psychological abnormalities in this group was relatively high, requiring focused attention. Analysis of the SCL-90 scale showed that the anxiety factor had the highest detection rate of 22.57%, followed by interpersonal sensitivity and depression, with detection rates of 16.34% and 14.69%, respectively. Daily behavioral data showed that psychologically abnormal students had a higher degree of Internet dependence, with an average of 5.5 times of Internet use per week and 5.4 times of late sleeping, both significantly higher than normal students. The application of k-anonymization privacy-preserving technology ensures data security while maintaining data availability. The Trans-LSTM model effectively improves the temporal feature extraction capability by introducing the attention mechanism and location coding, which provides a reliable technical support for mental health risk early warning. The study provides an important reference for the intelligent construction of mental health management system in colleges and universities, and has good application prospects and promotion value.

References

[1] Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College students: mental health problems and treatment considerations. Academic psychiatry, 39(5), 503-511.

[2] Wang, C. (2023). Application of MPP database and artificial intelligence system in online evaluation of college students' mental health. Preventive Medicine, 173, 107580.

[3] Tian, Z., & Yi, D. (2024). Application of artificial intelligence based on sensor networks in student mental health support system and crisis prediction. Measurement: Sensors, 32, 101056.

[4] Markoulakis, R., & Kirsh, B. (2013). Difficulties for university students with mental health problems: A critical interpretive synthesis. The Review of Higher Education, 37(1), 77-100.

[5] Allen, H. K., Lilly, F., Green, K. M., Zanjani, F., Vincent, K. B., & Arria, A. M. (2022). Substance use and mental health problems among graduate students: Individual and program-level correlates. Journal of American college health, 70(1), 65-73.

[6] Nagy, G. A., Fang, C. M., Hish, A. J., Kelly, L., Nicchitta, C. V., Dzirasa, K., & Rosenthal, M. Z. (2019). Burnout and mental health problems in biomedical doctoral students. CBE—Life Sciences Education, 18(2), ar27.

[7] Fischbein, R., & Bonfine, N. (2019). Pharmacy and medical students' mental health symptoms, experiences, attitudes and help-seeking behaviors. American journal of pharmaceutical education, 83(10), 7558.

[8] Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. Research policy, 46(4), 868-879.

[9] Chen, J., Yuan, D., Dong, R., Cai, J., Ai, Z., & Zhou, S. (2024). Artificial intelligence significantly facilitates development in the mental health of college students: a bibliometric analysis. Frontiers in Psychology, 15, 1375294.

[10] Gao, Y. (2025). The impact and application of artificial intelligence technology on mental health counseling services for college students. Journal of Computational Methods in Sciences and Engineering, 25(2), 1686-1701.

[11] Aljarallah, N. A., Dutta, A. K., Alsanea, M., & Sait, A. R. W. (2023). Intelligent Student Mental Health Assessment Model on Learning Management System. Comput. Syst. Sci. Eng., 44(2), 1853-1868.

[12] Rajkishan, S. S., Meitei, A. J., & Singh, A. (2024). Role of AI/ML in the study of mental health problems of the students: a bibliometric study. International Journal of System Assurance Engineering and Management, 15(5), 1615-1637.

[13] Fu, Y., Ren, F., & Lin, J. (2025). Apriori algorithm based prediction of students' mental health risks in the context of artificial intelligence. Frontiers in Public Health, 13, 1533934.

[14] Lee, J. (2025). Smart Solutions for Student Well-being: AI-driven Mental Health Support in a Polytechnic. In The Rise of Intelligent Machines (pp. 211-242). Chapman and Hall/CRC.

[15] Nag, A., Das, A., Sil, R., Kar, A., Mandal, D., & Das, B. (2022, December). Application of artificial intelligence in mental health. In International conference on intelligent systems design and applications (pp. 128-141). Cham: Springer Nature Switzerland.

[16] Sun, S., Dong, Y., Li, Y., & Liu, H. (2024). A mechanism for analyzing and managing undergraduates' mental health based on large-scale behavior data: AI-based approach. Internet Technology Letters, 7(6), e467.

[17] Liu, P., & Yu, H. (2025). The Exploration of Integrated Learning Algorithms in the Comprehensive Evaluation of College Students' Mental Health and the Optimization of Education Management. International Journal of High Speed Electronics and Systems, 2540525.

[18] Zhang, J. (2023). A study on the application of neural networks to determine students' mental health status in school education. Revista de Psicología del Deporte (Journal of Sport Psychology), 32(2), 207-218.

[19] Asencios Rodrigo,Asencios Christian & Ramos Efrain. (2023). Profit scoring for credit unions using the multilayer perceptron, XGBoost and TabNet algorithms: Evidence from Peru. Expert Systems With Applications,213(PC).

[20] Idin Noorani & Farshid Mehrdoust. (2025). Forecasting Bitcoin Price by a Hybrid Structure Based on ARIMA, SVM and LSSVM Models. Annals of Financial Economics,19(04).

[21] Patxi Ortego,Alberto Diez-Olivan,Javier Del Ser,Fernando Veiga,Mariluz Penalva & Basilio Sierra. (2020). Evolutionary LSTM-FCN networks for pattern classification in industrial processes. Swarm and Evolutionary Computation,54,100650-100650.