International Journal of Information Technologies and Systems Approach
Volume 16 • Issue 2

Early Warning Model of College Students' Psychological Crises Based on Big Data Mining and SEM

Rui Liu, Guizhou Police College, China*
https://orcid.org/0000-0003-4369-7138

ABSTRACT

In recent years, the psychological problems of college students could not be ignored, as they have seriously affected the growth of students and the normal teaching order of colleges and universities. However, there exists a strong noise in college students' psychological sample data set and a strong correlation between its data. Aiming to solve this problem, this paper proposes a psychological crisis warning method for college students based on big data mining and structural equation model (SEM). This method is oriented to massive user data in social networks. Particle swarm optimization is introduced to improve the random forest algorithm, and the original data is labeled to alleviate the impact of data noise on the recognition accuracy. The simulation example comes from an efficient actual data set in the southern China. The experimental results show that the proposed method can achieve an efficient analysis of actual complex data, and can provide reliable psychological auxiliary diagnosis for practitioners.

KEYWORDS

Mental Health, Crisis Warning, SEM, Random Forest, Intermediary Effect, Big Data Mining

1. INTRODUCTION

Due to the increasingly fierce social competition, college students inevitably face and bear various pressures (Shervington et al, 2020). College students are in the transition stage of school and social roles, and are often faced with the challenges of learning, employment, love, family, and other aspects. The internal transformation and external development are not balanced, making students' psychological adjustment and self-healing issues call for more attention (Yang ZL et al, 2022).

In recent years, malignant events caused by psychological problems have occurred frequently (Ben-Yehuda et al, 2022). As far as depression is concerned, the report on the development of China's national mental health (2019~2020) published by the team of literature (Wang, 2019) that 13.6% of people with undergraduate degrees and above suffer from depression, and 4.2% suffer from

DOI: 10.4018/IJITSA.316164

*Corresponding Author

severe depression. Therefore, it is the key to put forward and formulate psychological early-warning methods for colleges and universities, and to increase the attention to the psychological health of college students.

Timely and accurate early crisis warning is likely to save more lives. How to accurately identify the crisis, receive the early warning in time, and take effective measures is the key to achieve an early warning of psychological crisis (Guan et al, 2020).

In most colleges and universities, the way to study and judge the psychological state of college students is still relatively traditional, lacking efficiency and accuracy (Antunes-Alves et al, 2021). The traditional psychological intervention methods mostly adopt a psychological questionnaire screening, lectures and post psychological treatments. However, these practices cannot effectively and dynamically grasp students' psychological conditions in a timely manner, so as to timely intervene in possible crises (Yang & Fu, 2022).

With the rise of social networks, researchers have conducted data mining and analysis on user data and behavior on social networks, which is helpful to understand and predict users' psychological state (Tadesse et al, 2019). Reference (Mitja, 2010) shows that personal information of social networks can reflect the real personality of individuals. Research in reference (Youyou & Kosinski, 2015) shows that computers can make more accurate judgments about people's personalities based on Facebook data than people do. Therefore, big data analysis technology is applied to the analysis of mental state, and a new solution is proposed for college students' psychological crisis intervention research (Jia et al, 2021).

Through the in-depth mining of the data characteristics of the massive collection of samples, a complete and reliable decision-making database is built to achieve an accurate analysis of the complex psychology of college students (Li et al, 2021). Reference (Wang, 2021) established a statistical analysis model for the psychological rescue function of music education through the method of fuzzy feature extraction. Reference (Sang, 2021) collects students' psychological state data and analyzes the psychological crisis state based on support vector machine. Reference (Pang, 2021) used the big data analysis technology and the decision tree algorithm to establish a psychological assessment model to achieve accurate psychological assessment and early warning. The emergence of machine learning provides more powerful help for the processing and analysis of psychological data sets. Reference (Low DM et al, 2020) identified suicide-related forum speeches based on machine learning technology during COVID-19, and verified that patients with borderline personality disorder and post-traumatic stress disorder had higher psychological crisis. Reference (Walsh CG et al, 2018) analyzed the clinical data with machine learning algorithm, and improved the accuracy of suicide risk prediction for adolescents by 9 times. Reference (Jia, 2022) carried out research on the stressors of psychological crisis and realized the analysis and early warning of college students' psychology based on artificial neural network (ANN).

However, it should be noted that the data set of college students' psychological samples has a strong noise. If not processed, it will have a greater impact on the analysis accuracy of early warning methods. There is also a strong correlation between psychological sample data. When conducting the early warning analysis of psychological crisis on them, we should optimize the early warning model in a fine granularity.

To solve the above problems, this paper combines structural equation model (SEM) with stochastic forest algorithm, and proposes a method of psychological state analysis and early warning for college students.

1. The particle swarm optimization algorithm is used to reorganize the random forest algorithm, solve the problem of over fitting caused by excessive data noise in big data mining, realize the labeling processing of the original sample data set, and reduce the amount of calculation when constructing the early warning model of psychological crisis.

2. The intermediary theory is introduced into the process of SEM to optimize the analysis model of college students' psychological state, improve the fine-grained model, better fit the actual situation, and achieve a more accurate and reliable psychological state analysis, benefitting the crisis early warning research.

2. USER DATA PREPROCESSING BASED ON IMPROVED RANDOM FOREST ALGORITHM

Social media has become the main way for college students to express their feelings, which can reflect their real state.

We obtain the corresponding user identity information and text information through the web crawler, and achieve massive user data labeling based on the improved random forest algorithm, solving the problem of large noise of original user data and providing reliable data support.

Random forest (RF) algorithm can achieve the most accurate data labeling processing (Bangpan et al, 2019) and is applicable to the processing of big data mining. However, the crawler data itself has a large noise; thus, the traditional RF algorithm needs to be optimized to solve the problem of over fitting.

2.1 Data Acquisition

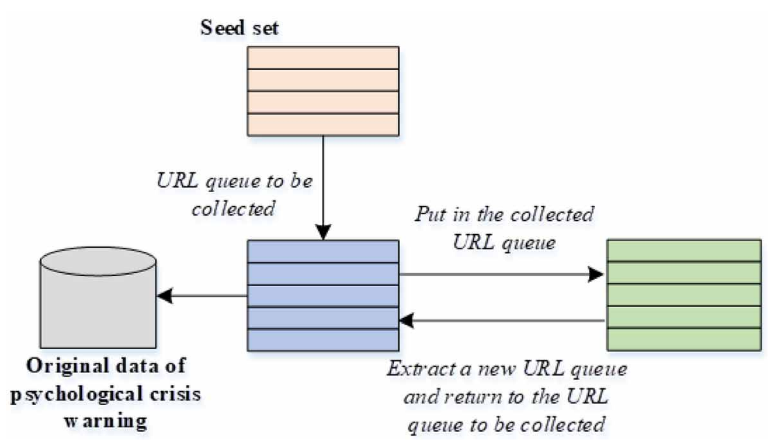
The identity information and comments of college students' users are indispensable data bases, and an effective raw data acquisition is an important prerequisite for crisis early warning analysis.

This paper uses the web crawler to achieve information acquisition for social platforms.

The principle of web crawler is shown in Figure 1.

1. Determine the seed set, that is, determine the data source of the web crawler. As a web page is to be crawled, the university BBS can use the web page link as the seed set URL (unified resource locator), and it is stored in a URL queue to be collected.
2. Read the URLs in the URL queue to be collected in turn, convert the link to the corresponding IP address of the website server through DNS resolution, analyze the page content, download and store valid information.
3. Extract a new URL and store it in the queue to be crawled, and repeat step 1 until the URL queue is empty, so as to traverse the web page. This process is also called web crawling.

Figure 1.
Acquisition of data related to students' psychological state



2.2 Random Forest Algorithm

Many studies have shown that RF algorithm has a strong anti-noise and over-fitting ability. It can handle high-dimensional complex data, and achieve an orderly and accurate classification of data. Therefore, based on RF algorithm, the collected data can be labeled (Xia S et al, 2021).

Set the crawler to get the original data set to (Rao M V et al, 2022):

$$T = \{(a_i, b_i)\}(|D| = n, a_i \in R^m, b_i \in R^m) \tag{1}$$

where, n is the number of data set samples collected by the crawler; m is the number of characteristics of each psychological data set sample. Among them, the training data set is:

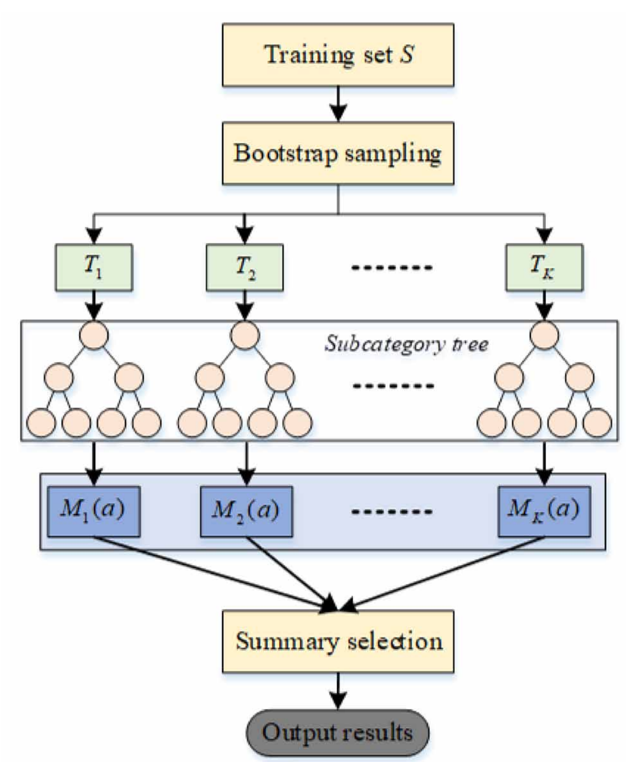
$$S = \{(a_1, b_1), (a_2, b_2), \&, (a_l, b_l)\} \tag{2}$$

where, l represents the number of training sets. The specific RF algorithm is shown in Figure 2.

The steps are as follows:

1. Bootstrap method is used to randomly sample the crawler acquisition data training set, and K sub training sample set $\{T_1, T_2, \&, T_K\}$ and K test sample sets are obtained.
2. Use the new training set obtained in step (1) to establish multiple regression model $M_1, M_2, \&, M_K$.

Figure 2.
Random forest algorithm framework



3. The test set is brought into the trained regression tree model, and the predicted value is $M_1(a), M_2(a),\&, M_K(a)$.
4. The average of the predicted values of all regression trees is taken as the prediction result of the RF model.

In the RF algorithm, because the Bootstrap method is used in this paper, the new training set will not contain all the data of the original training set. Data sets that are not included are called Out of bag (OOB) samples.

Generally, two-thirds of the original training set samples are selected to construct the regression function. The rest of the data constitutes an OOB sample, that is, a new test set. A new training set is used to build a regression tree once a time.

The calculation formula of OOB error is:

$$MSE \approx MSE^{OOB} = \frac{\sum [\hat{b}(a_i) - b_i]^2}{s} \tag{3}$$

where, $\hat{b}(a_i)$ is the predicted value of sample a_i; b_i is the true value of sample a_i; s is the number of OOB samples.

2.3 Tagging of User Data Based on PSO-RF Algorithm

In the traditional random forest model, the voting weight of each decision tree is equal. This resuls in the fact that the voting of some decision trees or leaf nodes with poor performance will affect the results of the whole random forest classifier's labeling classification of data sets (Zhao MM et al, 2021).

This paper proposes a leaf weighted random forest algorithm based on particle swarm optimization (PSO), which applies different weights to each leaf node in the random forest, and uses PSO algorithm for adaptive optimization.

Due to the increased time and user span of data collection, students sometimes express some abnormal phenomena in social media, such as watering and garbled code, making a large noise interference in the original data set, which requires the introduction of noise reduction processing before data set analysis.

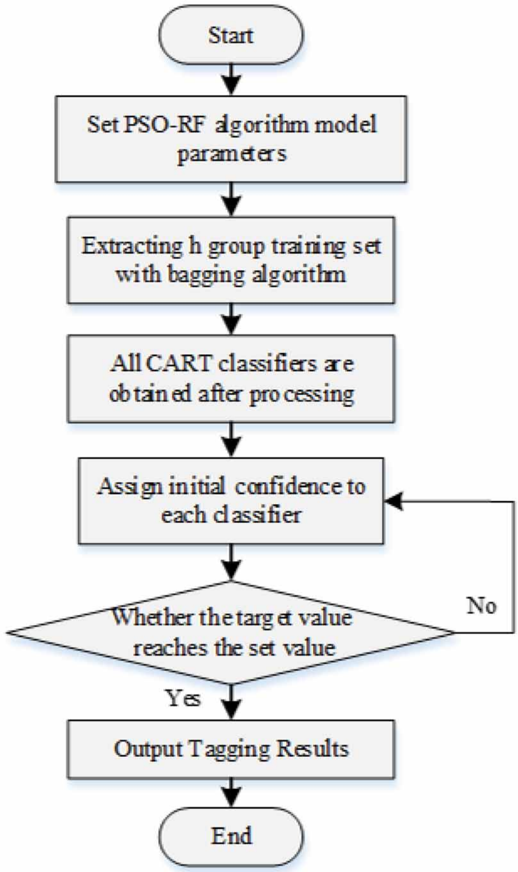
In view of the over-fitting problem of the random forest classifier in some noisy classification or regression problems, the bagging algorithm is used to select features from the feature set and training set to train the classifier and to improve the accuracy of the algorithm in feature selection.

After feature selection, initial confidence is randomly allocated to each sub leaf to form a confidence vector, and PSO algorithm is used to optimize the confidence vector until its performance reaches the preset completion threshold. Figure 3 is the flow chart of PSO-RF algorithm.

The steps of labeling college students' mental state data based on PSO-RF algorithm are as follows:

Input: Crawler collection data sample set $S = \{(a_1, b_1), (a_2, b_2), \&, (a_l, b_l)\}$, forest size k, iteration depth e, optimization threshold d.
Output: The results of processing labeling of students' mental state data.
(1) The bagging algorithm is used to extract h training sets from the crawler data sample set S.
(2) FOR (i from 1 to h)
(3) Use the i training set to get the i CART classifier;
(4) ENDFOR
(5) Each leaf node is assigned an initial confidence level, which is preset as 0.6 to form the confidence vector $P = [p_1^1, p_2^1, \&, p_n^h]$. In

Figure 3.
Tagging processing of college students' mental state data based on PSO-RF algorithm



p_j^i, i represents the decision tree to which the corresponding leaf node belongs, and j represents the ranking in the sequence count. The preset value does not affect the final result;
(6) FOR (j from 0 to e)
(7) Use PSO algorithm to optimize P. If the target value is greater than d, it ends;
(8) ENDFOR
(9) PSO optimized leaf weighted random forest is obtained. The function of the judgment sample is

$$F = \arg\max \sum \{I[(f_n(r) = i) \times p_j^i]\} \tag{4}$$

where, r is the test set, $f_n(r)$ is the classification result of the n-th decision tree, and $I()$ is a judgment function. When the output result of the base classifier meets the conditions, it is 1, and when it does not, it is 0.

3. THE EARLY-WARNING MODEL BASED ON IMPROVED SEM

PSO-RF algorithm provides reliable data preprocessing operation for this study. However, it is still noted that although the user data has been labeled accordingly, there is still some correlation between the data, causing considerable computational and analytical pressure on the modeling.

In this paper, Logistic regression analysis is used to screen the data. Based on this, the SEM is introduced to form a mental state model to realize the research on the analysis of psychological crisis.

3.1 Data Screening

Because the test data sets are assigned by category and belong to the discrete numerical type, there are different correlation degrees between the data.

Therefore, we choose to filter and preprocess the data based on Pearson test and logistic regression analysis to ensure that the samples are second-order independent and uncorrelated, to avoid invalid features and the over-fitting; thus, improving the model training effect.

Pearson test is used to show the significant correlation among variables with an easy calculation. The correlation coefficient is calculated by:

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{5}$$

where, $\rho_{X,Y}$ represents the correlation coefficient, $\operatorname{cov}()$ represents the covariance, and σ represents the standard deviation. If $\rho_{X,Y} = 1$, it indicates a full positive correlation; if $\rho_{X,Y} = -1$, it indicates full negative correlation, and if $\rho_{X,Y} = 0$, it indicates no correlation.

Then binary logistic regression analysis (Lyu et al, 2018) is conducted on the test data set processed by Pearson test, which can exclude the redundant attributes with low correlation that do not enter the regression model, and can determine the significant characteristic attributes of psychological model.

The analysis tool SPSS software is used to input the whitened independent variable and the predictive value data matrix. We select the binary logistic regression analysis, and screen out the characteristic attributes that enter the regression equation.

The maximum likelihood estimation used by forward stepwise regression method is adopted for regression analysis. $\alpha_{input} = 0.06$ and $\alpha_{output} = 0.15$ are set. Some data processing results are shown in Table 1.

According to Table 1, "personality characteristics", "economic situation", "emotional relationship", and "learning pressure", are all less than 0.01, and are included in the regression equation

Table 1.
Partial Data Processing Results of Logistic Regression Analysis

Influencing factor	P value
Personality characteristics	0.001
Economic situation	0.004
Emotional relationship	0.001
Only child or not	0.015
Failure to pass the examination	0.002
Personal achievements	0.011

model. The P value of "only child or not" and "personal achievement" is greater than 0.05, so they are excluded from the equation. Therefore, the four characteristic attributes entering the equation are determined as high correlation attributes to participate in modeling.

3.2 Construction of Model of Psychological Crisis

Traditional multiple regression statistical models cannot explain complex things, which requires simultaneous discussions and predictions of the relationship between multiple variables and the path analysis of causal models among variables -- SEM (Dube et al, 2019).

This paper uses SEM to describe the influence of subjective and objective factors on college students' psychological changes. SEM can show the complex relationship between variables and use multivariate functions and confirmatory coefficient analysis to comprehensively build a college students' psycho-analysis model (Dube et al, 2019).

The latent variables of SEM include subjective and objective factors, such as personal information, economic situation, etc. The calculation results are obtained by collecting measurement variables, testing and identifying the model, so as to build a mental state evaluation model for college students.

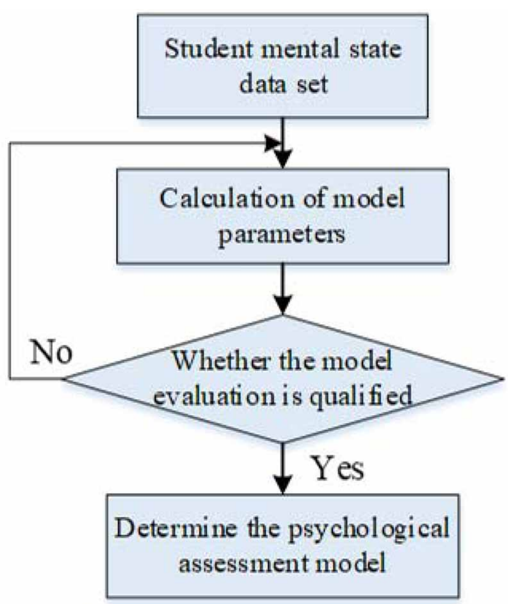
The SEM is characterized by a high accuracy, a strong correlation and a flexible application. The specific model formula is as follows:

$$X = A_X \xi + \tau \tag{6}$$

$$Y = A_Y \zeta + \upsilon \tag{7}$$

In Formula (6), X is a vector composed of exogenous observation variables, which is the mental health of college students; A_X is the factor load matrix of exogenous observation variables; A_Y is the factor load matrix of endogenous observation variables; ξ is a vector composed of exogenous latent variables, including "personality characteristics", "economic situation", "emotional relationship" and other factors; τ is the measurement error of X .

Figure 4.
Flow chart of college students' mental state assessment model



Structural equation:

$$\zeta = B\zeta + C\xi + \eta \tag{8}$$

In Formula (8), η represents the residual term of the structural equation, reflecting the unexplained part of ζ in the structural equation.

At the same time, aiming to further refine the model and better match with the actual situation, the intermediary effect theory is selected to analyze the path of the early warning model.

Intermediate variables are variables between independent variables and dependent variables (Kang et al, 2021). For example, personality factors are fundamental in the formation of college students' sad thoughts, and sudden negative life events have played a role in boosting the role of sad thoughts. Among them, negative life events first need to play a role through personality factors, especially values, which is an obvious intermediary variable.

After the implicit variable is centralized, the following equation can be used to describe the relationship between independent variable, intermediate variable and dependent variable. Figure 5 is the diagram of mediation variables.

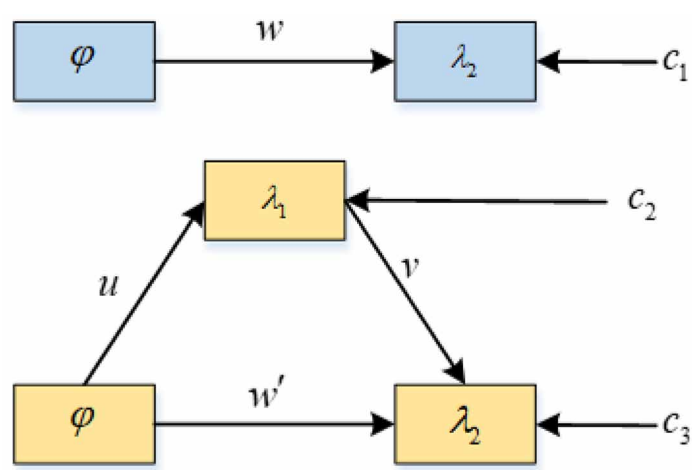
When studying the influence of the variable φ on the variable λ_2, if the independent variable φ is changed through the third influencing variable λ_2, then the variable λ_1 is called the intermediary variable.

For example, sudden negative life events should first play a role in grief thoughts through values:

$$\lambda_2 = w\varphi + c_1 \tag{9}$$
$$\lambda_1 = u\varphi + c_2 \tag{10}$$
$$\lambda_2 = w'\varphi + v\lambda_1 + c_3 \tag{11}$$

In equation (9), w is the total effect parameter. In equation (11), w' is called the direct effect; u and v represent the intermediary effect of the intermediary variable λ_1. Substituting equation (6) into equation (11), we can get:

Figure 5.
Schematic Diagram of Mediation Effect



$$\begin{aligned}\lambda_2 &= w'\varphi + v(u\varphi + c_2) + c_3 \\ &= (w' + uv)\varphi + vc_2 + c_3\end{aligned} \tag{12}$$

Comparing equation (12) with equation (11), we can get:

$$w = w' + uv \tag{13}$$

Thus, the total effect can be divided into two parts, namely, uv is the size of the intermediary effect and w' is the size of the direct effect.

Based on the above analysis, the psychological crisis warning model based on PSO-RF and SEM is shown in Table 2.

4. SIMULATION EXPERIMENT ANALYSIS

The research on psychological warning analysis relies on the programming language Python 3.8 to verify the feasibility of the proposed detection method. Meanwhile, other experimental environments are built according to the implementation shown in Table 3.

The experimental data set comes from the BBS forum of a university in the south of China, and is composed of student user comments and user information delivered from March to June 2022.

Table 2.
Assessment and Analysis Model of College Students' Psychological Crisis

Crisis degree	Interventions
Lower	Mental health care
Alert	Mental health care, Real-time monitoring
Higher	Peer counseling, psychotherapy, real-time monitoring, timely prevention and control
Very high	Psychotherapy, real-time monitoring, timely referral, security serious

Table 3.
Simulation experiment environment

Project	Parameter
CPU processor	Intel(R) Core(TM) i5-6300
Memory	16GB
Video memory	GTX 1070ti*1
Disk type	SSD
Operating system	Windows 10
Development language	Python 3.8
Deep learning framework	Python 0.2 - GPU
Image processing library	OpenCV 4.2.0
CUDA Version	10.2
CUDNN Version	7.4.1

4.1 Evaluation Index

The experimental evaluation index is mainly composed of two parts: model fitting consideration and performance consideration.

The model fitting consideration is the goodness of fit index (GFI).As the sample size has a great impact on the test, the GFI index is introduced to refer to the ratio of the square sum of the difference between the observation matrix AA and the covariance replication matrix of the sample data and the observed variance.

The formula expression of GFI is as follows:

$$\begin{aligned} GFI &= 1-\frac{tr((\sum(\hat{\theta})^{-1}(S-\sum(\hat{\theta})))^2)}{tr((\sum(\hat{\theta})^{-1}S)^2)} \\ &= 1-\frac{tr((\sum(\hat{\theta})^{-1}S-I)^2)}{tr((\sum(\hat{\theta})^{-1}S))} \end{aligned} \tag{14}$$

where, $\sum(\hat{\theta})$ is the covariance matrix implied in the college students' psychological assessment model; χ^2 is the chi square value of the model, but it should be pointed out that $tr((A)^2) = tr(AA')$.

Performance evaluation mainly consists of square root of mean square error (RMSE) and absolute mean error (MAE).

1. $RMSE$: $RMSE$ can indicate the accuracy of prediction. The smaller the $RMSE$ value is, the more accurate it is, and the stronger the effect is:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \hat{x}_i\right)^2} \tag{15}$$

In the formula, x_i is the actual psychological state of college students; $\hat{x}_i$ is the psychological prediction state of college students; N is the number of samples.

2. MAE: MAE represents the deviation of the regression model. The meaning of numerical value change is the same as that of RMSE:

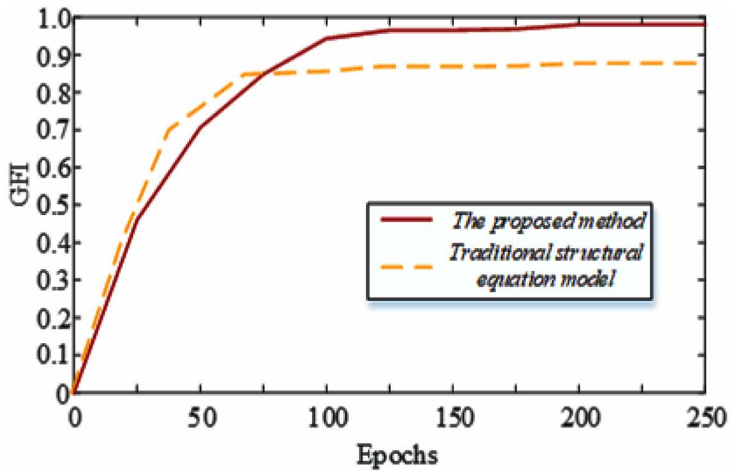
$$MAE = \frac{1}{N}\sum_{i=1}^{N}| x_i - \hat{x}_i | \tag{16}$$

4.2 Convergence and Divergence Analysis of Experimental Model

This paper first optimizes and evaluates the mathematical model based on the GFI index. The convergence of the GFI index is shown in Figure 6.

As shown in Figure 6, the proposed modeling method has a better fitting effect than that based on traditional SEM. In 125 rounds of operation, the model can achieve the effective convergence, and the GFI index can be practically close to 1, and keep a certain constant; While the modeling method based on the traditional SEM is faster than the proposed method in obtaining the final college students' mental health model in 60 calculations, its model fitting ability is far less than the proposed method.

Figure 6.
Changes of GFI Index



At the same time, based on RMSE and MAE, this paper also analyzes the precision of mental state pre-warning analysis. Figure 7 and Figure 8 show the changes in $RMSE$ and MAE of the method proposed in this paper.

Figures 7 and 8 show the fitting ability and the calculation convergence ability of the method proposed in this paper. As shown in the figures, the algorithm proposed can realize the fitting and prediction of college students' psychology after 50 rounds of calculation, with indicator $RMSE$ of 37.32 and indicator MAE of 0.33.

4.3 Experimental Optimization Analysis

Aiming to verify the performance superiority of the early warning method of student psychological crisis proposed in this paper, this paper uses reference (Wang, 2021) and reference (Sang, 2021) as

Figure 7.
Changes of RMSE Indicators

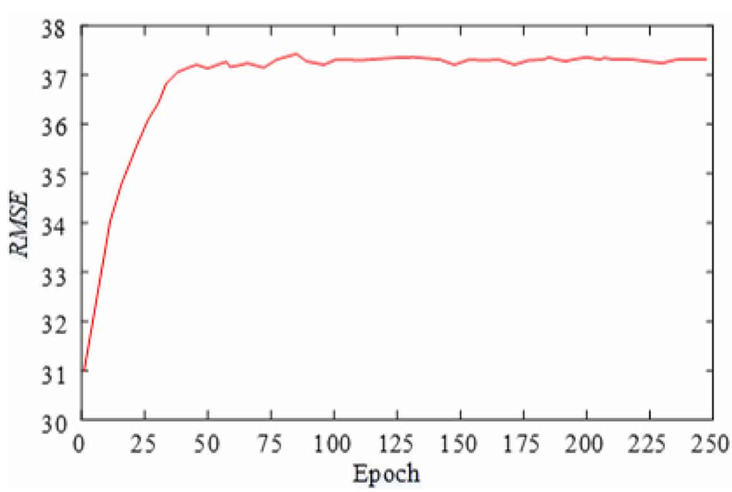
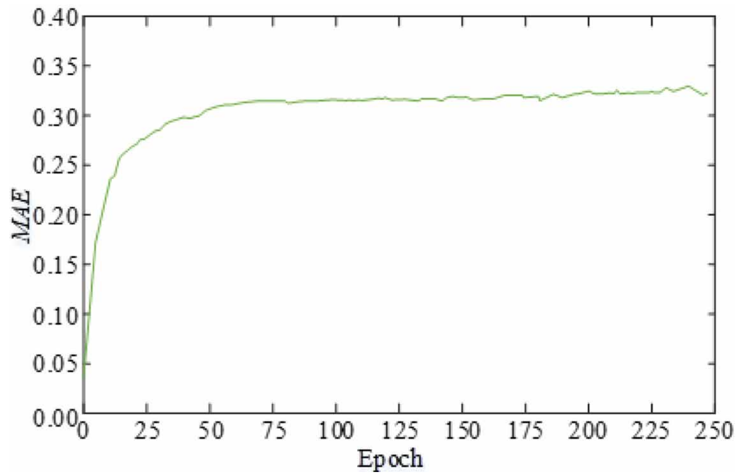


Figure 8.
MAE Index Change



the comparison method to achieve the optimal verification of simulation experiments. It should be noted that all experimental methods run in the same physical environment.

Figure 9 shows the RMSE indicators under different methods.

Furthermore, this paper also conducts a quantitative analysis on MAE indicator variables. Figure 10 shows the MAE evaluation indicators of different methods.

As shown in Figure 9 and Figure 10, the index RMSE of the psychological fitting prediction method for college students proposed in this paper is 37.21, which is 1.91 and 1.19 lower than that in reference (Wang, 2021) and reference (Sang, 2021). For the indicator MAE, the index value of the proposed method is 0.33, while that of reference (Wang, 2021) and reference (Sang, 2021) is both 0.399.

The reason is that when modeling the students' psychological state, the method introduces the intermediary effect theory to optimize the early warning model, improve the model's granularity,

Figure 9.
Evaluation Index RMSE under Different Methods

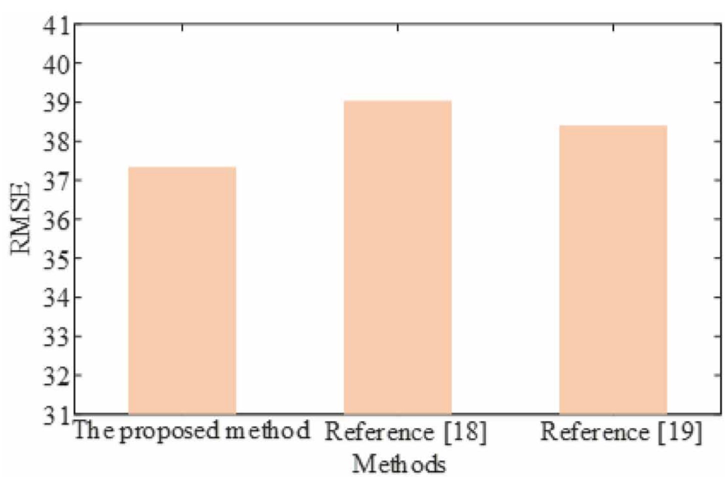
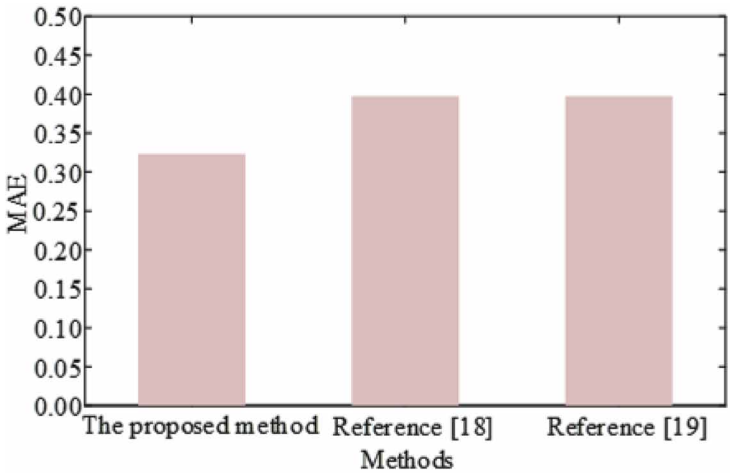


Figure 10.
Evaluation Index MAE under Different Methods



and make the model closer to the actual situation. However, the comparison method neglects the consideration of intermediary variables, making it difficult to achieve the efficient fitting.

In contrast to the comparison method, this paper also optimizes the random forest algorithm based on the particle swarm optimization algorithm, realizes data labeling, and solves the problem of the local solution results' undesirable effect on the global optimization. This further improves the fitting ability of the early warning method proposed in this paper, and supports the method to achieve an accurate analysis of actual complex data sets.

Furthermore, this paper also analyzes different methods based on their computational efficiency. The calculation time cost of different methods for psychological prediction is shown in Table 4.

As shown in Table 4, the analysis time of the proposed method for psychological crisis early warning is 1.109s, which is 0.348s shorter than that in reference (Wang, 2021). The calculation time in reference (Sang, 2021) is 1.111 s, which is close to the calculation time of the proposed method. However, due to the above analysis of RMSE and MAE evaluation indicators, the prediction performance of the proposed method is superior than that in reference (Sang, 2021). Therefore, it can be concluded that the method proposed in this paper can achieve a faster and more efficient analysis while ensuring excellent prediction performance.

5. CONCLUSION

This paper provides an early warning method of psychological crisis for college students using the SEM and the random forest algorithm, which provides a reliable guarantee for their mental health.

Table 4.
Calculation Time Cost of Different Analytical Methods

Method	Time (s)
Proposed method	1.109 s
Reference (Wang J, 2021)	1.357 s
Reference (Sang HY, 2021)	1.111 s

The proposed method introduces PSO algorithm to improve the RF algorithm to achieve data labeling; the SEM is optimized by using the intermediary effect theory. The simulation results prove that the method can effectively analyze the actual complex data sets, and its evaluation indicators $RMSE$ and MAE are only 37.32 and 0.33, which provides certain decision support and reference for college students' mental health defense work.

However, it should be noted that although the proposed method has achieved good validation results in the experimental analysis, the experimental data group only comes from one college student. Therefore, in future work, the team will expand the data set on a larger scale to make the model more universal and further improve the early warning performance.

ACKNOWLEDGMENT

This work was supported by the Natural Science Research Project of Guizhou Provincial Department of Education (the Qian Jiao He KY Zi [2021] 287).

REFERENCES

Antunes-Alves, S., & Langmuir, T. (2021). Evaluating a combined intervention targeting at-risk post-secondary students: When it comes to graduating, mental health matters *Counselling & Psychotherapy Research*, *21*(2), 313–323. doi:10.1002/capr.12390

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook Profiles Reflect Actual Personality, Not Self-Idealization *Psychological Science*, *21*(3), 1–8. doi:10.1177/0956797609360756 PMID:20424071

Bangpan, M., Felix, L., & Dickson, K. (2019). Mental health and psychosocial support programmes for adults in humanitarian emergencies: A systematic review and meta-analysis in low and middle-income countries *BMJ Global Health*, *4*(5), 1–11. doi:10.1136/bmjgh-2019-001484 PMID:31646006

Ben-Yehuda, O., Dreazen, E., Koren, D., & Peleg, M. (2022). Participatory Design of a Mobile App to Safeguard Mental Resilience in the Context of Drug Use in Young Adults: Multi-Method Study *JMIR Formative Research*, *6*(2), 1–11. doi:10.2196/34477 PMID:35212631

Brito, M., Silva, J., & Costa, A. (2020). Suicidal behavior and prevention strategies from teachers' perspective *Escola Anna Nery*, *24*(4), 1–7.

Cai, L. H. (2022). A Novel Recognition Model of University Students' Psychological Crisis Based on DM *Scientific Programming*, *2022*(1), 1–10. doi:10.1155/2022/8908848

Chen, X. J. (2022). Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristic Attribute *Scientific Programming*, *2022*(1), 1–11. doi:10.1155/2022/6026248

Davis, J. P., Rao, P., Dilkina, B., Prindle, J., Eddie, D., Christie, N. C., DiGuiseppi, G., Saba, S., Ring, C., & Dennis, M. (2022). Identifying individual and environmental predictors of opioid and psychostimulant use among adolescents and young adults following outpatient treatment *Drug and Alcohol Dependence*, *233*(1), 1–10. doi:10.1016/j.drugalcdep.2022.109359 PMID:35219997

Dube, A., & Kamath, R. (2019). Microfinance Group Processes and Crises: Responses to Economic and Psychological Threats *The Journal of Development Studies*, *55*(10), 2273–2285. doi:10.1080/00220388.2018.1539475

Eunok, K., & Hyeseong, J. (2021). A Mediating Effect of Self-Reflection and a Moderating Effect of Marital Emotional Interaction in the Influence of Psychological Crisis in Married Middle-Aged Men Upon Psychological Well-Being *Family and Family Therapy*, *29*(2), 385–401.

Guan, R. X., & Zhang, J. B. (2020). Methods and Approaches of College Counselors to Deal with Students' Psychological Crisis *Revista de Psicología del Deporte*, *29*(4), 103–114.

Jia, X. J., Jin, T., & Wang, Q. (2021). Application of Machine Learning in the Field of Mental Health [J] [Education Science Edition]. *Journal of Ningbo University*, *43*(4), 117–122.

Jia, Z.P. (2022). Prediction of College Students' Psychological Crisis with a Neural Network Optimized by Harmony Search Algorithm *International Journal Of Emerging Technologies In Learning*, *17*(2), 59-75.

Li, T. T., Li, T., & Guo, X. N. (2021). An Early Warning Method for Psychological Crisis of College Students Based on Social Media Big Data *Modern Distance Education Research*, *33*(04), 92–103.

Lin, C. X., & Lin, K. N. (2022). Exploration on the collaborative innovation path of college students' ideological education and psychological education *Frontiers in Psychology*, *13*(1), 1–10. doi:10.3389/fpsyg.2022.969508 PMID:36092039

Liu, J., Wang, M., & Sun, Q. W. (2022). An Investigation on the Administrations of Counseling and Psychological Crisis Intervention for College Students in China *Chinese Journal of Clinical Psychology*, *30*(2), 477–482.

Low, D. M., Rumker, L., Talkar, T., Torous, J., Cecchi, G., & Ghosh, S. S. (2020). Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on Reddit during COVID-19 *Journal of Medical Internet Research*, *22*(10), 1–11. doi:10.2196/22635 PMID:32936777

Lyu, J. C., Wang, Y. Y., Shi, H., & Zhang, J. (2018). Early warnings for suicide attempt among Chinese rural population *Journal of Affective Disorders*, *238*(1), 353–358. doi:10.1016/j.jad.2018.06.009 PMID:29908474

Pang, Y. J. (2021). Psychological Crisis Intervention of College Sports Majors Based on Big Data Analysis *Revista de Psicología del Deporte*, *30*(3), 124–132.

Rao, M. V., Midhunchakkaravarthy, D., & Dandu, S. (2022). Deep Learning CNN Framework for Detection and Classification of Internet Worms *Journal of Interconnection Networks*, *22*(3), 1–10.

Sang, H. Y. (2021). Analysis and Research of Psychological Education Based on Data Mining Technology *Security and Communication Networks*, *2021*(1), 1–8. doi:10.1155/2021/8979507

Shervington D. & Richardson L. (2020). Mental health framework: coronavirus pandemic in post-Katrina New Orleans. *J Inj Violence Res, 12*(2), 1-9.

Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2019). Detection of Depression-Related Posts in Reddit Social Media Forum *IEEE Access: Practical Innovations, Open Solutions*, *7*(1), 44883–44893. doi:10.1109/ACCESS.2019.2909180

Walsh, C. G., Ribeiro, J. D., & Franklin, J. C. (2018). Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *59*(12), 1261–1270. doi:10.1111/jcpp.12916 PMID:29709069

Wang, C. X. (2019). “Big data” Mode and Non-inductive Intervention of“New Media”—The New Ways of the Prevention of College Students’ Psychological Crisis. *Journal of Shenyang University*, *21*(03), 328–332.

Wang, J. (2021). Music Education to Rescue Psychological Stress in Social Crisis Based on Fuzzy Prediction Algorithm *Scientific Programming*, *2021*(1), 1–6. doi:10.1155/2021/2039235

Wu, Y. (2015). Kosinski Michal, Stillwell David. Computer-based personality judgments are more accurate than those made by humans *Proceedings of the National Academy of Sciences of the United States of America*, *112*(1), 1036–1040. PMID:25583507

Xia, S., Chen, B., Wang, G., Zheng, Y., Gao, X., Giem, E., & Chen, Z. (2021). mCRF and mRD: Two Classification Methods Based on a Novel Multiclass Label Noise Filtering Learning Framework *IEEE Transactions on Neural Networks and Learning Systems*, *33*(7), 2916–2930. doi:10.1109/TNNLS.2020.3047046 PMID:33428577

Yang, Z. L., & Fu, X. M. (2022). Effect of fitness exercise on risk assessment of mental and emotional health *Revista de Psicología del Deporte*, *31*(1), 210–216. RPD.

Zhao, J. P., & Song, H. T. (2021). Identification of Psychological Crisis Signals of College Students Based on the Dufferin Equation *Advances in Mathematical Physics*, *2021*(1), 1–12. doi:10.1155/2021/4057328

Zhao, M. M., Jin, Y. Y., & Wang, Y. J. (2021). Application Research of Random Forest Algorithm in Decision