

Fuzzy Cluster Analysis and Prediction of Psychiatric Health Data Based on BPNN

Hong Xiang^{1*}, Anrong Wang¹, Guoqun Fu², Xue Luo³, Xudong Pan¹

¹School of Electronic and information Engineering, Chongqing Three Gorges University, Wanzhou, 404100, China

²Institute of Public Administration, Chongqing Three Gorges University, Wanzhou, 404100, China

³Students Affairs Division, Chongqing Three Gorges University, Wanzhou, 404100, China

Received: August 3, 2021. Revised: December 31, 2021. Accepted: January 12, 2022. Published: January 13, 2022.

Abstract—PMH (psychiatry/mental health) is affected by many factors, among which there are numerous connections, so the prediction of PMH is a nonlinear problem. In this paper, BPNN (Back Propagation Neural Network) is applied to fuzzy clustering analysis and prediction of PMH data, and the rules and characteristics of PMH and behavioral characteristics of people with mental disorders are analyzed, and various internal relations among psychological test data are mined, thus providing scientific basis for establishing and perfecting early prevention and intervention of mental disorders in colleges and universities. Artificial neural network is a mathematical model of information processing, which is composed of synapses similar to the structure of brain neurons. The fuzzy clustering analysis and data prediction ability of optimized PMH data are obviously improved. Applying BPNN to the fuzzy clustering analysis and prediction of PMH data, analyzing the rules and characteristics of PMH and the behavioral characteristics of patients with mental disorders, can explore various internal relations among psychological test data, and provide scientific basis for establishing early prevention and intervention of mental disorders.

Keywords—BPNN, PMH, Fuzzy clustering analysis, Predict

I. INTRODUCTION

With the fierce competition in today's society and the increasing pressure of life, the PMH (Psychiatric/Mental Health) problems are becoming more and more prominent, and the PMH status is worrying [1-3]. From the current actual situation, although some relevant monitoring modes are established in the early warning work psychological crisis, the monitoring effect on specific problems arising from psychological health is still not ideal [4-5]. PMH education has a very positive significance for cultivating healthy psychology,

developing psychological potential, enhancing adaptability and improving their personality [6]. PMH is a comprehensive reflection of many factors. The common psychological test data is relatively single, resulting in a certain subjectivity in the test results [7]. Psychological problems are private personal problems, so most people are unwilling or afraid to consult. Because the factors causing psychological stress to people are becoming more and more complex, and psychological problems are becoming more and more complex and hidden, some people are unaware of the hidden dangers in their mental health, so the existing psychological crisis early warning mechanism lacks initiative. Whether PMH is affected by many factors, and there are countless connections between various factors. Therefore, the prediction of PMH is a nonlinear problem [8-9]. Cluster analysis method is an unsupervised learning process. It aggregates things into classes according to some attributes of things, so as to minimize the similarity between different classes and maximize the similarity between the same classes, so as to realize the classification of data [10]. This paper applies BPNN to the fuzzy cluster analysis and prediction of PMH data, analyzes the laws and features of PMH and the behavioral features with mental disorders, excavates various intrinsic relationships between psychological test data, and provides a scientific basis to establish the early prevention and intervention of mental disorders. With the increasing progress of society and the continuous improvement of people's material living standards, it is regrettable that people pay little attention to PMH, resulting in the increasingly serious psychological problems of people [11]. The prediction of PMH based on BPNN is a new artificial neural network technology of intelligent information system. In practical application, it solves many problems with complex causality. Its function not only has associative memory, nonlinear mapping, classification and recognition, but also has intelligent features such as optimization calculation and knowledge processing. In this competitive and increasingly interdependent world, what society needs is comprehensive senior talents who can adapt to the future world. Builders must have good psychological

qualities such as indomitable will, hard-working spirit, positive optimism and forge ahead.

II. RELATED WORK

Many scholars have done relevant research on PMH prediction. Literature [15] uses multiple stepwise regression analysis to establish a prediction model of retirement psychological adaptation period, which can make a good prediction. Deng Xiaowei et al.] used multiple linear regression and path analysis to discuss the influencing factors and inhibition mechanism [16]. Gu Youzhong and others established the mathematical model of PMH prediction system based on neural network, which provided the basis for PMH prediction to a certain extent [17]. Yang Huanhuan Zhang Yanxin established a mathematical model to evaluate the status of PMH by using neural network and fuzzy mathematics, and the verification result is good [18]. The amount of data stored in the database is increasing, and these data contain a lot of information, which is beneficial to prevent the problems of PMH as early as possible [19]. It is necessary to analyze and predict these data timely and effectively for improving and optimizing decision-making and improving the effectiveness of psychological counseling. And there is a fundamental difference between clustering and classification. The effective application of clustering analysis technology can provide powerful support for practical decision-making [20]. Qin Lina research shows that integrating cluster analysis into PMH analysis can help people better understand all aspects, and lay a good foundation for providing different psychotherapy methods for different groups by analyzing and customizing different types [21]. The PMH status evaluation model established by Karimu R Y et al. can evaluate the status of PMH well based on its fuzziness and highly nonlinear characteristics [22].

This study is based on BPNN model. Through the optimization of the algorithm, the fuzzy clustering analysis and prediction system of PMH data based on multiple factors is established, and the model system is simulated according to the situation, and compared with other models, which shows that the model has certain theoretical level and practical value.

III. MATERIALS AND METHODS

A. Cluster Analysis

Cluster analysis refers to the analysis process of dividing a group of physical or abstract objects into multiple sub-categories composed of similar objects. This process should group the data according to the distance or similarity of the data itself [23]. From the entire sample n , arbitrarily select k objects as the initial cluster center $m_i (i=1,2,\dots,k)$. Find the distance $d(p,m)$ from each p to k cluster centers,

$$d(i,j) = \sqrt{(x_{i1} - y_{j1})^2 + (x_{i2} - y_{j2})^2 + \dots + (x_{in} - y_{jn})^2} \quad (1)$$

In the formula, $i = (x_{i1}, x_{i2}, \dots, x_{in})$ and $j = (y_{j1}, y_{j2}, \dots, y_{jn})$

respectively represent two n -dimensional data objects. Find the smallest distance $d(p,m)$ of each object p , and put p into the same cluster as m_i . After traversing all objects, use the following formula to recalculate the value of m_i as the new cluster center,

$$m_k = \sum_{i=1}^N x_i / N \quad (2)$$

Where m_k represents the center of the k -th cluster, and N represents the number of data objects in the k -th cluster. Re-assign the objects in the entire data set to the most similar cluster. Repeat the above steps until the following square error E is the smallest

$$E = \sum_{i=1}^k \sum_{p \in c_i} |p - m_i|^2 \quad (3)$$

Where E represents the sum of the squared errors of all objects, p represents the objects in the space, and m_i represents the average value of c_i . Euclidean distance is usually used to determine the distance between each cluster and the centroid of the initial cluster. Then divide the sample into K classes according to the distance and calculate the new class mean avg. Use avg as the new cluster centroid, and then use Euclidean distance to cluster again until the new cluster centroid does not change.

Set $\{X(t), t=1,2,\dots,n\}$ to represent the psychological time series set of university students at different stages. As university students' antagonistic psychology has obvious differences in manifestations, it is necessary to reconstruct the phase space of the psychological formation process to obtain the matrix

$$\begin{bmatrix} x(1) & x(1+\tau) & \dots & x(1+(m-1)\tau) \\ x(j) & x(2+\tau) & \dots & x(2+(m-1)\tau) \\ \dots & \dots & \dots & \dots \\ x(K) & x(K+\tau) & \dots & x(K+(m-1)\tau) \end{bmatrix} \quad (4)$$

$j = 1, 2, \dots, K$

In the formula, m represents the complexity of university students' rebellious psychology, and τ represents the psychological duration, which satisfies

$$K = n - (m-1)\tau \quad (5)$$

The clustering analysis algorithm flow is shown in Figure 1.

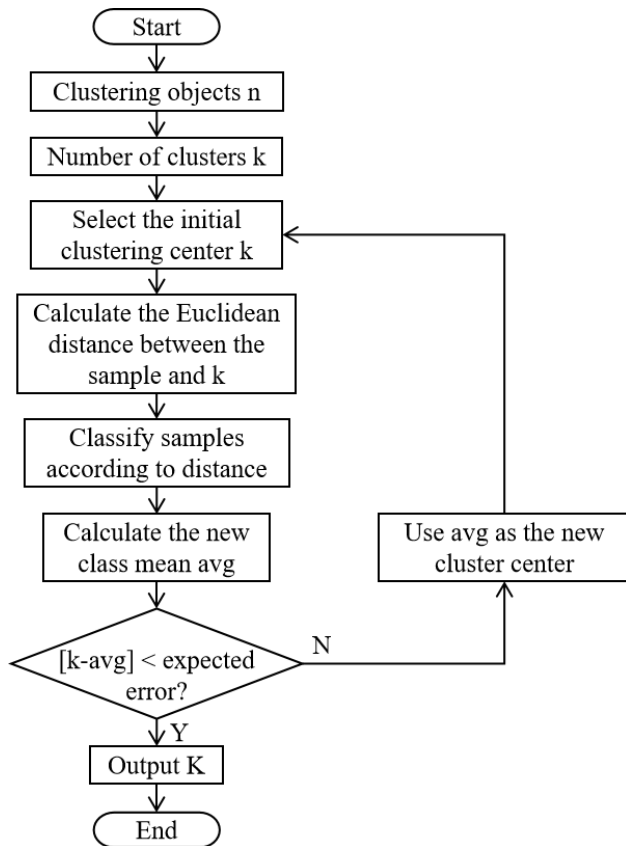


Figure 1. Cluster analysis algorithm flow

Each row in the matrix is regarded as a variety of influencing factors of inverse psychological changes, and the total number is K . The features of each influencing factor are sorted according to the degree of psychological resistance, and the

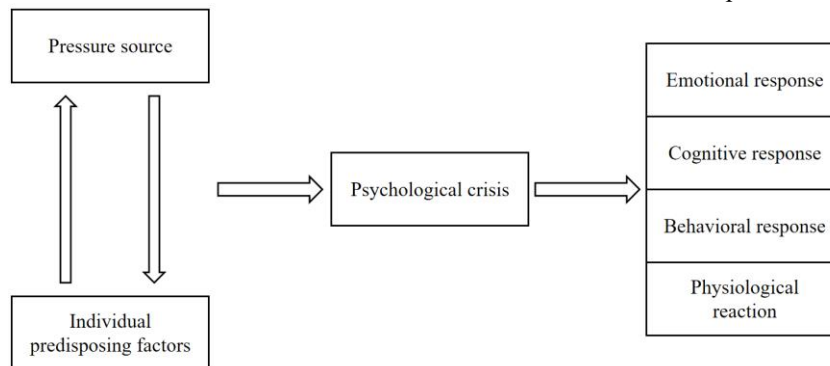


Figure 2. Evolutionary structure of mental health problems

For the two different stages of subjective will and objective events to form the rebellious mentality, the behavioral tendency time series X and Y are based on the performance of rebellious mentality. From two different stages, the average amount of information that is hindered by psychological behaviors caused by subjective factors and objective factors is screened out, and formulas (8) and (9) are used to calculate the thresholds of behavior tendency for different reasons to form rebellious psychology.

index of the categories affected by each influencing factor in the initial state of antagonistic mental performance is obtained, and the antagonistic mental performance status of different influencing categories is obtained. For m -dimensional university students' psychological complexity, the total number of different mental performance state sequences is m , assuming that the probability of k different mental performance state sequences is P_1, P_2, \dots, P_k , and the probabilities of reverse mental state sequences are ranked. Define permutation entropy.

$$H_{PE}(m) = -j = 1 \sum_{j=1}^k P_j \ln P_j \quad (6)$$

Integrate the state sequence given by formula (6), and use formula (7) to express the emotional features of behavior corresponding to the reverse mental state at this stage

$$0 \leq H_{PE} = H_{PE} / \ln(m) \leq 1 \quad (7)$$

Suppose $\{x(t), t=1, 2, \dots, N\}$ represents the time series set of behavioral emotional features of university students' rebellious mental state. Due to the complexity of the behavioral tendency of university students dominated by rebellious psychology, it is necessary to establish the relationship with the rebellious psychology duration τ , and choose the duration sequence $x(t + \tau)$ to form a new continuous stage point sequence $y(t)$. Determine the psychological duration τ based on the correlation between $x(t)$ calculation and $y(t)$.

The outbreak of mental health problems is characterized by suddenness and urgency. Figure 2 shows the evolutionary structure of mental health problems.

$$H(X) = - \sum_{i=1}^n P_x(x_i) \log_2 P_x(x_i) \quad (8)$$

$$H(Y) = - \sum_{i=1}^n P_y(y_i) \log_2 P_y(y_i) \quad (9)$$

Among them, $P_x(x_i)$ represents the time series set X of behavior tendencies governed by subjective will to form rebellious psychology. x_i represents the extracted subjective

factor feature index set. $P_y(y_i)$ represents the objective factor index feature set in the time series set Y of the behavior tendency dominated by the psychology formed by objective events.

B. BPNN Learning Process

People's psychology will change constantly with their own age, experience or other subjective and objective factors. The causes of psychological crisis are complex and difficult to control. Traditional early warning indicators mainly come from the judgment criteria in psychological theory. Although these indicators have certain credibility, they are also subjective and exist relatively independently. The psychological test database established by psychological counseling service and regular psychological questionnaire survey only records the psychological state when they are surveyed or consulted by teachers, but cannot dynamically monitor the psychological changes of college students. BPNN is a multi-layer perceptron structure model, which is composed of several layers of neurons. In the network, there are not only input layer nodes and output layer nodes, but also one or more hidden layer nodes [25]. The structure of BPNN is shown in Figure 3.

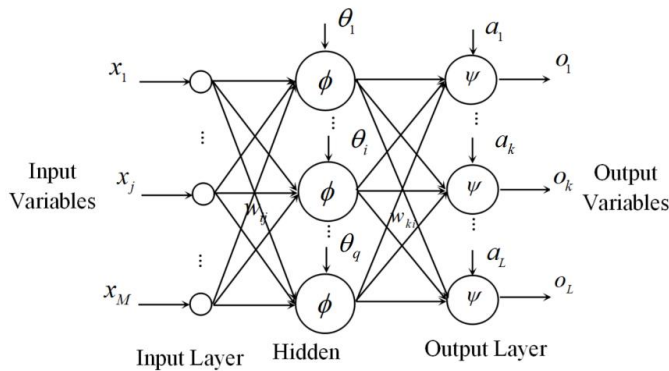


Figure 3. BPNN structure

After learning the sample set, the weights and thresholds used by the neural network are the internal representations obtained by the neural network through adaptive learning that best meet the data requirements. Then input the feature data of the sample set to be identified into the trained network, and the network can automatically reason and identify the output results of the samples [27]. If the neural network is regarded as a mapping from input data to output data, the mapping is highly nonlinear.

IV. RESULT ANALYSIS AND DISCUSSION

The physical and mental health is directly related to the rise and fall of the nation, the strength of the country and individual success. In order to prove the comprehensive effectiveness of the proposed optimization model of PMH quality based on fuzzy clustering, an experiment is needed. The quality of data plays a decisive role in the degree of network convergence in the training of neural networks. Good data can make the network achieve a better convergence balance in the shortest time. On

the contrary, if the data is improper, it is difficult to achieve the results required by the network no matter how the network parameters are modified. Build a modeling and simulation platform of PMH quality in Matlab environment. The survey results were divided into sample set and test set.

The PMH data clustering analysis model based on BPNN and the traditional PMH data clustering analysis model are used to carry out statistical experiments on the optimization of PMH quality. Comparing the statistical efficiency (%) of the two models, the comparison result is shown in Figure 4.

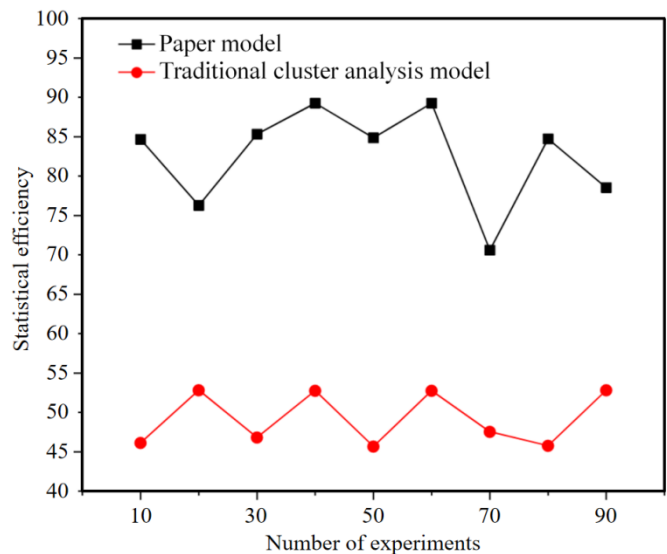


Figure 4. Comparison of statistical efficiency of different models

A person's psychological condition will change constantly with the increase of age and the different objective environment of participating in various activities. In this period, the psychological status is influenced by many factors such as the surrounding environment and the pressure, which often leads to some negative psychological problems. People's psychological state is dynamically changing, especially some hidden psychological problems can only happen under specific circumstances. Static psychological state data lags behind the changes of psychological state seriously. The final result is that the static stored psychological file data is different from the students' real psychological state, which affects the timeliness of psychological crisis identification. The living environment, family experience and so on interweave and influence psychological changes. The dynamic changes of these factors will lead to inferiority, anxiety and depression. Only by dynamically monitoring the changes of psychological trajectory and providing these data to the decision-makers of psychological education or other universities can the psychological crisis prevention work be carried out accurately and effectively.

The statistical efficiency of using this model to optimize PMH quality is higher than that of the traditional cluster analysis model. Using this model and the traditional cluster analysis model, the statistical experiments of university students' psychological hindrance to behavior were carried out. Compare

the statistical stability (%) of psychological hindrance to behavior with the two models. The comparison results are shown in Figure 5.

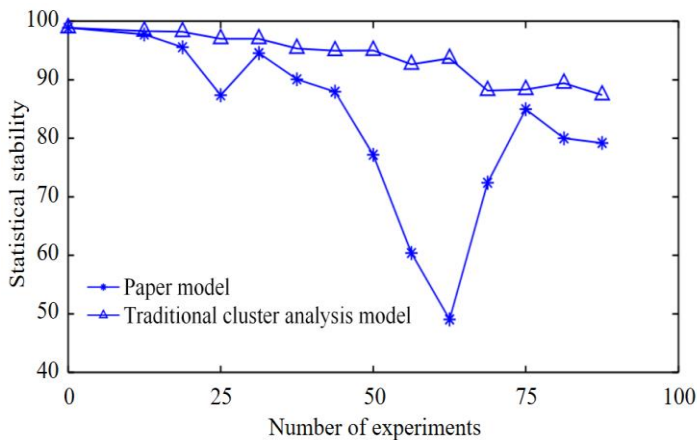


Figure 5. Comparison of statistical stability of different models

It can be seen from the figure that the statistical stability of psychological hindrance to behavior by using this model is better than that of the traditional model. Using this model and the traditional cluster analysis model, the statistical experiments of optimizing the PMH quality of university students are carried out. Compare the statistical time complexity (%) of the two models for optimizing the PMH quality, and the comparison results are shown in Figure 6.

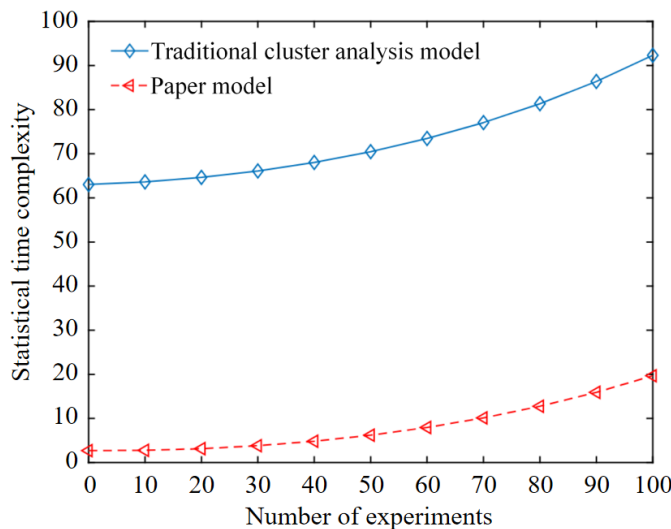


Figure 6. Comparison of statistical time complexity of different models

Psychological crisis intervention is a very complex work. Psychological counseling center only provides professional assistance to students who need treatment. In addition, it also needs the cooperation of school administrative departments, teachers, parents and student cadres. The administrative department of the school is the top decision maker of the work, and the rest are the early warning implementers. Every student's psychological problems are unrepeatable according to their different environment and social cognition. Therefore, it is necessary to formulate a personalized crisis intervention

program for each student, which requires a certain number of professional psychological intervention personnel. In the era of big data, the data sources of college students' psychological early warning include not only the psychological data obtained by routine methods such as daily psychological census, psychological investigation and outpatient inquiry, but also the behavior data related to study and life during school.

It can be seen that the statistical time complexity of using this model to optimize the PMH quality of university students is lower than that of the traditional cluster analysis model. In the process of network self-learning, the error is decreasing. Through the continuous training of neural network, the network can basically achieve the minimum error prediction. It can be seen that the prediction model has a better prediction effect. Through continuous training of neural network, the network can basically achieve the minimum error prediction. The influence of different integration scales on the accuracy of university students' psychological behavior operation is shown in Figure 7.

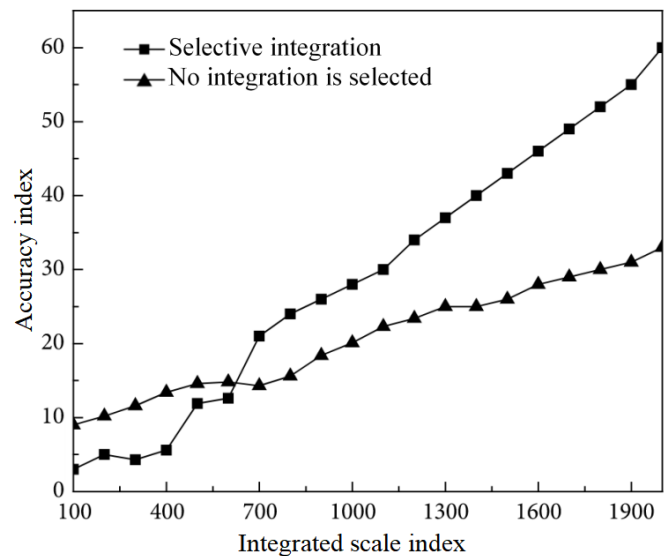


Figure 7. The impact of different integration scales on the accuracy of psychological behavior

Through the simulation test of several groups of data, it can be seen that the error between the predicted results of each group and the true values is very small, and the fitting degree of the input data basically meets the required requirements. The model can make a good prediction of the PMH status of university students. Artificial neural network is a mathematical model for information processing, which is connected by synapses similar to brain neuron structure. The ability of fuzzy clustering analysis and data prediction of optimized PMH data is obviously improved. The reason is that BPNN finds out the optimal weights and thresholds through fitness function, and then gives the weights and thresholds to neural network training instead of the initial weights and thresholds randomly generated by Matlab for each training, which greatly reduces the ability of self-adjustment of weights and thresholds. By using the principle of BPNN, a mathematical prediction network model is established on the basis of analyzing the psychological factors affecting university students. The calculation of the final result

of the prediction system is simple and flexible, which effectively improves the prediction efficiency and accuracy of the network model.

V. CONCLUSION

In the research of PMH, what we should do is to improve the measuring tools based on the existing measuring tools of PMH. In this paper, under the background of PMH stress, the clustering analysis based on BPNN is used to actively analyze the psychological state, and a prediction model of PMH state is established, which is simulated by Matlab and achieves good validity. Simulation results show that compared with traditional methods, the proposed model has great advantages in modeling efficiency, stability and time complexity, which provides a strong basis for improving PMH. BP algorithm has the ability of self-learning and self-adaptation. By comparing the errors between the actual output and the expected output of the network, the weights and thresholds of the network are adjusted repeatedly, and the errors are gradually reduced to achieve the desired accuracy. Therefore, using the trained network model to analyze and predict the PMH data of university students can obtain relatively reliable prediction results. Compared with the traditional prediction methods of PMH, the prediction method of BPNN avoids the complicated and uncertain routine modeling process.

In this paper, BP algorithm is used to optimize the fuzzy clustering analysis and prediction of PMH data, and good results are achieved. However, from the in-depth study and practical application, there are still many aspects to be further studied in the future.

REFERENCES

- [1] Wang Yuxing, Huang Jun, Pan Yingjie. Application of GA-BP Neural Network in the Prediction of Elderly Negative Emotions. *Small Microcomputer System*, 2020, 41(08):136-140.
- [2] Yu Baolu, Duan Xun, Wu Yun. Establishment and application of BP neural network data prediction model. *Computer and Digital Engineering*, 2016, 44(003):482-486+545.
- [3] Yi Hongrui. Research on Shenzhen Residents' Health Level Evaluation and Measurement and Control Model. *Science & Technology Information*, 2020, 018(012):180-182.
- [4] Zhang Qingshi. Human feature prediction based on BP neural network. *Shandong Chemical Industry*, 2019, 048(015):139-141.
- [5] Ruspini E H, Bezdek J C, Keller J M. Fuzzy Clustering: A Historical Perspective. *IEEE Computational Intelligence Magazine*, 2019, 14(1):45-55.
- [6] Hua Qiqin. Research on Optimization Modeling and Simulation of College Students' Mental Health Quality. *Computer Simulation*, 2016, 033(012):314-317.
- [7] Zhong Weiwei, Liu Liping, Wang Fangzheng. Research on College Students' Network Sentiment Analysis Based on Fuzzy Clustering. *Computer Knowledge and Technology*, 2019, 15(28):232-234+239.
- [8] Chen Qiuwu, Wei Huimei. The application of K-means clustering algorithm in analyzing the mental health of college students *Digital World*, 2020, 174(04): 150-151.
- [9] Cui Yuan, Zhang Junpeng. Cluster analysis of factors affecting the mental health of the elderly in Chengdu. *Chinese Electronic Journal of Geriatric Research*, 2017, 004(004): 38-42.
- [10] Shen Yang, Zeng Qingqi, Jiang Ying, et al. The mental health level and its influencing factors of a large state-owned enterprise in western China. *Occupation and Health*, 2016, 32(05):595-598.
- [11] Allahyari E, Moshtagh M. Predicting mental health of prisoners by artificial neural network. *BioMedicine*, 2021, 11(1):26-33.
- [12] Geraci J, Wilansky P, Luca VD, et al. Applying deep neural networks to unstructured text notes in electronic medical records for phenotyping youth depression. *Evid Based Ment Health*, 2017, 20(3):83-87 .
- [13] Shatte A, Hutchinson D M, Teague S J. Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, 2019, 49(9):1-23.
- [14] Shi Songhe, Feng Shaozhen, Zhang Zhimin, et al. Effect analysis of psychological intervention for middle-aged and elderly immigrants from the South-to-North Water Transfer Project. *Modern Preventive Medicine*, 2016, 43(6): 1051-1056.
- [15] Chang Qiang, Zhao Wei, Zhao Yangjie. Data classification prediction and realization based on neural network. *Software*, 2018, 39(12):207-209.
- [16] Deng Xiaowei, Zhang Chengbin. Research on Statistical Arbitrage Based on Mixed Neural Network Model Forecasting. *Statistics and Decision*, 2019, 035(001):174-177.
- [17] Gu Youzhong, Liu Jianli, Gao Weidong. Yarn strength prediction based on genetic simulated annealing neural network. *Shanghai Textile Science and Technology*, 2016, 391(01): 45-49.
- [18] Yang Huanhuan, Zhang Yanxin. Research on Mental Health Evaluation of College Students Based on WOA Improved Random Forest. *Microcomputer Applications*, 2019, 035(009): 37-40.
- [19] Wang Yueyi, Li Jingqiang, Wang Bei. Research on the Safety Psychological Evaluation Model of Civil Aviation Pilots. *Safety and Environmental Engineering*, 2017, 24(6):81-85.
- [20] Liu Yuanyuan, He Xingshi. BP neural network stock price prediction based on adaptive firefly algorithm. *Journal of Weinan Normal University*, 2019, 034(002): 87-96.
- [21] Qin Lina. Network security situation prediction based on neural network. *Journal of Anyang Normal University*, 2018, 112(02):72-74.
- [22] Karimu R Y, Azadi S. Diagnosing the ADHD Using a Mixture of Expert Fuzzy Models. *International Journal of Fuzzy Systems*, 2017, 20(4):1-15.
- [23] Arifoglu D, Bouchachia A. Detection of Abnormal Behaviour for Dementia Sufferers using Convolutional

- Neural Networks. *Artificial Intelligence in Medicine*, 2019, 94(3):88-95.
- [24] Nishida M, Kikuchi S, Fukuda K, et al. Jogging Therapy for Hikikomori Social Withdrawal and Increased Cerebral Hemodynamics: A Case Report. *Clinical Practice & Epidemiology in Mental Health Cp & Emh*, 2016, 12(1):38-42.
- [25] Lin Kun, Lei Yinjie. Research on palm vein recognition algorithm based on improved AlexNet convolutional neural network. *Modern Electronic Technology*, 2020, 558(07):60-64.
- [26] Xiao Tingyang, Zhang Di. Design and experiment of centimeter-level ultra-wideband ranging error correction model based on BP neural network. *Global Positioning System*, 2021, 46(4):1-7.
- [27] Wang Dan, Zou Rulian, Zhang Junying. Research on the Significance of Gender Differences in Mental Health Quality of Normal Students. *Management Observation*, 2019, 713(06):111-112.

**Creative Commons Attribution License 4.0
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US