A Comparative Analysis of Extreme Gradient Boosting, Decision Tree, Support Vector Machines, and Random Forest Algorithm in Data Analysis of College Students' Psychological Health

Qiong Li[1,2*], Jian Zhou[3]
1School of Mathematics and Physics, Hubei Polytechnic University, Huangshi 435000, China
[2]Hubei Key Laboratory of Intelligent Convey Technology and Device (Hubei Polytechnic University), Huangshi 435000, China
[3]Wuhan Raycus Fiber Laser Technologies Co, LTD, Wuhan 430074, China
E-mail of corresponding author: liqiong872979454@163.com
[*]Corresponding author

Keywords: decision tree, support vector machine, random forest, XGBOOST, machine learning, mental health, influence factor

Received: August 27, 2024

To solve the problem of identifying the mental health status of college students, this study investigated the psychological conditions of students in a certain department of a university in Hubei Province through a questionnaire survey using the SCL - 90 scale. It combined machine learning algorithms to analyze the applicability of the model and explore the differences between students with healthy and sub - healthy mental states. Data (including basic information) of 500 students were randomly collected. A self - compiled questionnaire was used in combination with on - site scoring by psychological teachers to classify the mental states of the 500 students into healthy and sub - healthy states. Questionnaire data were analyzed through decision tree, support vector machine, random forest, and XGBOOST algorithms to quickly identify the healthy and sub - healthy states and to mine the behavioral characteristics that have a certain correlation with the mental health status of students. The data information of 500 students was modeled respectively, and the classification effects of the models were evaluated through accuracy, precision, recall, F1 - score, and AUC. The results showed that among the four methods, the random forest had the best classification effect, with an R2 score of 0.8891, which was higher than the R2 score of 0.8393 for the decision tree, the R2 score of 0.8840 for the support vector machine, and the R2 score of 0.8618 for the XGBOOST algorithm. Considering the advantages of the random forest in terms of classification performance, modeling time, interpretability, feature selection, and simplicity, we recommend using the random forest model to assist in the diagnosis of mental health status classification. The experimental results on the SCL - 90 scale survey and the student basic information dataset show that the proposed model has high accuracy and can converge quickly, enabling more effective and accurate prediction of students' mental health status.

Povzetek: Primerjava učinkovitosti algoritmov strojnega učenja, kot so Extreme Gradient Boosting (XGBOOST), odločitveno drevo, podporni vektorski stroji (SVM) in naključni gozd, pri analizi psihološkega zdravja študentov. Uporablja podatke iz vprašalnika SCL-90 in ugotavlja, da naključni gozd dosega najboljše rezultate pri klasifikaciji mentalnega zdravja, s poudarkom na njegovi natančnosti, hitrosti in interpretaciji. Študija priporoča uporabo naključnega gozda za pomoč pri diagnozi mentalnega zdravja.

1 Introduction

The accuracy of psychological problem diagnosis and the effectiveness of intervention have always been major challenges in the field of mental health, largely due to the lack of scientific prediction tools [1-3]. With the increasing attention of the country and society to mental health issues, researchers are increasingly integrating modern scientific and technological methods (such as machine learning) into the research of mental health problems. As the core of artificial intelligence, machine learning has demonstrated unique advantages in many fields (such as finance and medicine), and the application of machine learning in the field of mental health is an inevitable trend.

Current research on mental health mainly focuses on two aspects. One is the text analysis and treatment of psychological problems, and the other is the analysis and prediction of influencing factors of mental health. For the first aspect of research, Ahmed et al. [4] proposed a deep adaptive clustering model based on an interpretable attention network based on the text descriptions of psychological patients. The experimental results showed that this method helps to label texts and improve the recognition rate of mental disorder symptoms. Li [5] designed a mental health education system based on artificial intelligence online technology, which can conveniently and quickly query daily psychological problems according to the text description information of psychological patients and provide intelligent online psychological counseling to help patients overcome psychological problems. For the second aspect of research, Chen and Jiang [6] analyzed the influencing factors of mental health based on cognitive computing, combining the symptom self - assessment scale (SCL_90), social rating scale, and health cognition questionnaire (HCQ_127) test data of students from different regions in China and the United States. The research showed that students in developed regions are more likely to have psychological problems. Zhou [7] used the symptom self - assessment scale (SCL_90) data of 1,264 freshmen at Guizhou University in China and classified the mental health status of freshmen using the decision tree C4.5 algorithm. The classification results showed that somatization and hostility are important factors affecting students' mental health. Liu and Xu [8] applied a BP (BackPropagation) neural network to predict students' psychological problems based on students' psychological, personal basic information, and socioeconomic data. The experimental results showed that this method has a high prediction accuracy and can effectively predict students' psychological problems.

However, the accuracy of parameter estimation of traditional machine - learning models for psychological diagnosis often depends on a large sample size, which is not suitable for the small - sample situations that sometimes occur in reality. Therefore, in recent years, some researchers have proposed machine - learning methods that can extract key information from small samples for psychological diagnosis [9-10], and these methods have good robustness for various types of data, enabling relatively accurate classification. However, these studies have not compared different machine - learning methods, and previous studies have rarely involved the four machine - learning methods of decision tree, support vector machine, random forest, and XGBOOST algorithm. These methods each have their own advantages in simple classification tasks. For example, XGBOOST uses the gradient - boosting method, which can continuously optimize the model. At the same time, it has a built - in regularization term to effectively prevent over - fitting and supports parallel computing, with a fast-training speed. The learning process of the decision - tree model automatically extracts important features, and in the case of less data, it has an advantage over other complex models for simple classification problems, and the interpretation of the results does not require a mathematical background [11]. When dealing with non - linear problems, SVM is suitable for many features; it can handle the interaction of nonlinear features with strong generalization ability; and it is suitable for small samples. The random - forest algorithm, due to its method of using multiple trees for voting, has a good resistance to noise and can adapt to various data types. Therefore, this study intends to use four machine - learning methods (decision tree, support vector machine, random forest, and XGBOOST algorithm) to compare the prediction of mental health status by different machine - learning methods, fit the psychological diagnosis data of college students under samples of different sizes, study the impact of small samples on the prediction results of different machine - learning methods, and thus obtain the classification accuracy of different machine - learning methods under small samples. This provides an effective theoretical method for the application of psychological diagnosis and evaluation under small samples.

2 Methods

The dataset used in this paper is derived from a mental health questionnaire for students at a certain university in Hubei Province. To ensure satisfactory experimental results, data pre - processing operations are required. Data pre - processing includes methods such as data cleaning, data integration, data transformation, and data reduction. To guarantee the accuracy of the experimental results, data with incomplete scale responses are deleted, ensuring that the scale assessment results of the data used in the experiment are all complete responses. Also, multiple assessment data are removed to ensure that each subject in the dataset corresponds to only one assessment result. Association rule algorithms are used to discover the relationships between scale factors, screen frequent item sets and association rules, and perform data discretization on scale factor scores and scale scores. The basis for this is the scale's scoring criteria. Taking the SCL - 90 scale as an example, if the factor standard score is greater than 2 points, the factor is considered positive with symptoms; otherwise, it is negative without symptoms. In addition, data discretization is also required to assess the prediction accuracy of the model. The data required in this paper include scale scores and factor scores. Taking the SCL - 90 scale as an example, the screened dataset includes scale scores, somatization (SOM) factor scores, obsessive - compulsive symptoms (OBS) factor scores, interpersonal sensitivity (INT) factor scores, depression (DEP) factor scores, anxiety (ANX) factor scores, hostility (HOS) factor scores, phobia (PHO) factor scores, paranoid ideation (PAR) factor scores, psychoticism (PSY) factor scores, and other (MIS) factor scores. In this research, the psychological assessment scale is simplified from the perspective of reducing the number of scale item groups.

According to the needs of the experimental design, the data required for this experiment are selected from the dataset to form a new dataset. The new dataset includes the total scale score, scores of each factor, the label of the total score, and the labels of each factor. In the experimental part of this paper, factors are represented by their English abbreviations.

This research applies four classical machine - learning regression algorithms to predict scale scores, including decision tree, support vector machine, random forest, and XGBOOST algorithm. By comparing the performance of scale - score prediction models based on these four classical algorithms, the optimal algorithm is selected.

In the experiments of this paper, to avoid the adverse effects of a small number of training samples on the model and the impact caused by the random division of the training set and the test set, the five - fold cross - validation method is adopted. This method reduces the impact of randomness on the model performance and ensures the stability and accuracy of the prediction results. The method divides the dataset into five equal parts. Each time, four of these parts are used as the training set, and the remaining one part is used as the validation set. Model training is carried out five times, and the model evaluation indicators of the five experiments are obtained. Finally, the average value of the five model evaluation indicators is returned as the model evaluation index. The entire sample is trained using the four algorithms, and appropriate parameters are selected through the cross - validation method to establish the optimal model. Finally, the model is verified and analyzed according to the mental health diagnosis results. The experimental results show that the mental health status assessment method based on the random forest can achieve an average accuracy of 96.67%, which can effectively diagnose the mental health status of college students.

2.1 Decision tree algorithm

The decision tree algorithm (DT) is a method for approximating discrete function values. It is a typical classification method. First, it processes the data, uses an inductive algorithm to generate readable rules and a decision tree, and then uses the decision to analyze new data. Essentially, the decision tree is a process of classifying data through a series of rules. The decision - tree algorithm constructs a decision tree to discover the classification rules contained in the data. How to construct a decision tree with high accuracy and small scale is the core content of the decision - tree algorithm. The construction of the decision tree can be carried out in two steps. The first step is the generation of the decision tree: the process of generating a decision tree from the training sample set. The second step is the pruning of the decision tree: The pruning of the decision tree is a process of testing, correcting, and modifying the decision tree generated in the previous stage. Mainly, the data in the new sample dataset (referred to as the test dataset) is used to verify the preliminary rules generated in the process of decision - tree generation, and the branches that affect the prediction accuracy are removed. Compared with Logistic Regression, the decision - tree model can make a non - linear segmentation of the dataset well. At the same time, the tree - shaped model is closer to the human way of thinking, generates visual classification rules, and the generated model is interpretable, which can solve the classification problem of complex decision boundaries. As a typical machine - learning method, trees reflect a mapping relationship between object attributes and object values, and are not affected by the empirical distribution assumptions of the dataset itself.

2.2 Support vector machine

The support vector machine (SVM) is a machine - learning method based on statistical learning theory, mainly used for classification and regression analysis. Its core idea is to find an optimal hyperplane to separate samples of different classes and maximize the margin between the two classes. The basic principle of the support vector machine is to map the data into a high - dimensional space and construct an optimal hyperplane to achieve efficient data classification. Its core idea is to maximize the classification margin, so that data points of different classes are as far away from the decision boundary as possible, thereby improving the generalization ability of the classifier. In the case of linear separability, SVM finds the optimal hyperplane by solving a quadratic programming problem. In the case of non - linear separability, SVM maps the data into a high - dimensional space by introducing a kernel function to make it linearly separable.

2.3 Random forest

The random forest (RF) adopts the idea of ensemble learning, using multiple decision trees as weak classifiers, and then combines them through a certain strategy to form a strong classifier with better results. It generates multiple decision trees by randomly selecting some variables and data, and then aggregates the results of multiple decision trees through the voting method of the decision trees in the forest. Compared with the single judgment of a single decision tree, the random forest has higher accuracy, and this algorithm has very good stability. As a type of probabilistic graphical model, it represents the relationships between observed variables, predictive variables, and labels in an intuitive and easy - to - understand tree - diagram model. Each node branch represents the process of a parent node splitting into child nodes, and each path from the root node to the leaf node is an accurate description of a decision rule. As the number of base learners in the random forest increases, the random forest often converges to a lower generalization error. At the same time, different from the decision tree in Bagging, which selects the optimal splitting attribute from all attribute sets, the random forest only selects the splitting attribute from a subset of the attribute set, so the training efficiency is higher.

2.4 XGBoost

Ensemble learning is an effective strategy that combines multiple machine - learning algorithms. A single machine - learning algorithm can solve only limited problems and has poor generalization and application ability. However, combining multiple machine - learning algorithms to complete a certain learning task often produces better results. Each learner can be regarded as a basic learning unit, and through their combination, a powerful whole is finally integrated, which can be used to solve more complex problems.

Ensemble learning has advantages such as improving model performance, reducing overfitting, reducing variance, providing higher prediction accuracy, and handling linear and non - linear data. The extreme gradient boosting algorithm (XGBoost) uses multi - threading to accelerate the construction of trees, uses tree models as basic classifiers to form a powerful classifier, and integrates multiple basic classifiers together. This has the advantages of high efficiency, accuracy, and good interpretability in classification tasks.

The characteristics and comparison of the four classical machine-learning algorithms of decision tree, support vector machine, random forest, and ensemble learning are shown in Table 1.

Table 1: The characteristics and comparison of the four classical machine-learning algorithms

Algorithm	Advantages and Disadvantages
Decision Tree	Advantages: (1) The calculation is simple, and the results can be interpreted and visualized. (2) It is not sensitive to missing values. (3) It can handle irrelevant features. Disadvantages: (1) There may be an overfitting problem. (2) It performs poorly when dealing with highly correlated features. (3) It is difficult to handle continuous data.
Support Vector Machine	Advantages: (1) Applicable to a variety of features. (2) Capable of handling the interactions of non - linear features and has relatively strong generalization ability. (3) Suitable for small - sample data. Disadvantages: (1) Not suitable for large - sample data. (2) Sensitive to missing data.
Random Forest	Advantages: (1) Strong generalization ability. (2) Applicable to various features. (3) Insensitive to missing values and capable of handling imbalanced data. Disadvantages: (1) Not very effective when solving regression problems. (2) Low performance when there is a large amount of noise.
XGBoost	Advantages: (1) It adopts the gradient boosting method, which can continuously optimize the model. (2) It has a built - in regularization term, which can effectively prevent overfitting. (3) It supports parallel computing, resulting in a fast training speed. Disadvantages: (1) It has poor interpretability. (2) If the parameters are not set appropriately, the model may become too complex and lead to overfitting. (3) It has a high memory footprint. (4) It does not support real - time prediction.

3 Experiments

3.1 Data

From January to December 2023, the mental health situation of students in a university in Hubei Province was investigated. The results of 500 valid questionnaires were studied. Among the 500 students, 238 were confirmed to have mental sub - health and 262 were confirmed to have good mental health by the psychological counseling center. Based on these data, follow - up modeling and analysis research were carried out.

3.1.1 Investigation on influencing factors of college students' mental health

The survey tools are as follows:

(1) Self made questionnaire. There are 14 factors in total, including gender (male and female), grade (freshman, sophomore, junior and senior), subject (liberal arts and Science), place of origin (rural and urban), only child (yes or no), parental rearing style (democratic, authoritarian, doting and neglecting), parents' education level (primary school and below, middle school and high school, University and above), love (yes or no), weekly exercise time (0-2 hours, 2-5 hours, and more than 5 hours), family economic health (poor, average, and good), community activities (never, rarely, generally, and often), whether to fail the course (no, yes), whether to experience recent stress events (no, yes), and whether the family is single parent (yes, no). If the number of item options is n, the score will be 1 to n, for example, 0-2 hours, 2-5 hours and more than 5 hours of exercise per week will be scored as 1, 2 and 3 respectively. Each record is allowed to have at most 2 missing values. Otherwise, as invalid data, the missing values are filled with mode.

(2) Communication ability score. The professional psychological counseling room scores the students' communication ability by communicating with each student for three to five minutes, and then includes the communication ability as one of the analysis data of the students' overall psychological situation.

3.1.2 Investigation on college students' mental health

The Symptom Checklist 90 (SCL - 90) test was used, with scores of 0 (none); 1 (very light); 2 (moderate); 3 (heavier); 4 (severe). The scoring method involves statistical analysis of the main indicators and factor data. The overall evaluation of the mental health status of college students in this school is based on the screening criteria: a total score greater than 160 points, a positive number of items greater than 43, or any factor score greater than 2 points.

3.2 Experimental scheme

In the experiment, 80% (400) of the mental health data of the 500 college students were randomly selected as training samples, and 20% (100) of the data were selected as test samples. Taking R2 as the index, the gridsearchcv function was used to search for parameters such as learning_rate, n_iterations, and Max_Depth (the maximum depth of the tree) to find the best model. The main function of the gridsearchcv function is to automatically adjust parameters, specify the parameter range, and find the optimal result and its corresponding parameters, which is suitable for small datasets. By repeatedly changing the sample data and setting the parameter range, the final parameters were determined.

The experimental evaluation indicators include accuracy (a), precision (p), recall (r), F1 - score (F1), and the area under the receiver operating characteristic (ROC) curve (AUC) to evaluate the performance of the machine - learning model.

4 Experimental results

4.1 Performance comparison and verification of multiple algorithms

In the research on the mental health prediction model for college students in this paper, prediction models were established using random forest, XGBoost, decision tree, and SVM algorithms respectively. The gridsearchcv function was used to optimize the parameters of the decision tree and the other three methods. In each experiment, 80% of the total samples were randomly selected for training and 20% for testing. The prediction performance was quantified and ranked using accuracy, precision, recall, F - score, and AUC. The experimental results are shown in Table 2. By comparing the performance of different algorithms, the algorithm most suitable for a specific psychological assessment can be determined, thereby improving the accuracy and reliability of the assessment.

Table 2: Experimental results and comparisons of the four models

Methods	Accuracy	Precision	Recall	F1	AUC
DT	0.9533	0.9740	0.9375	0.9554	0.9923
SVM	0.9605	0.9625	0.9625	0.9625	0.9528
RF	0.9733	0.9871	0.9625	0.9746	0.9973
XGBoost	0.9666	0.9870	0.9524	0.9681	0.9957

As can be seen from Table 2, the random forest model achieved the best results among the four evaluation indicators. Its accuracy in predicting the mental health status of students is 0.97, the precision is 0.98, and the AUC is 0.996. While the decision tree model performed the worst, with values of 0.96, 0.96, and 0.95 respectively. According to the feature importance of the random forest model, the features affecting mental health are ranked as shown in Figure 1.

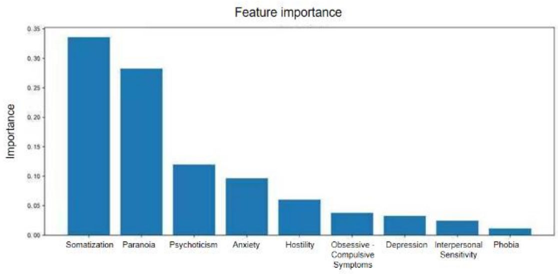

Figure 1: Feature importance ranking

The classification results showed that somatization, paranoia and psychoticism are important factors affecting students' mental health.

Using statistical methods, based on the proportion of the number of students with sub - mental - health in each group of susceptible factors, the important susceptible factors affecting college students' mental health are ranked as follows: parents' education mode; whether to experience recent stress events; communication ability score; grade; love; family economic status; whether the family is a single - parent family; only child; weekly exercise time; whether to fail a course; ,parents' education level; community activities; place of origin; gender; subject. This provides decision - making support for the formulation of college students' mental health training strategies.

The characteristics related to the family environment are: parents' education mode, family economic status, whether the family is a single - parent family, and whether the student is an only child. Their importance are ranked first, sixth, seventh, and eighth respectively, indicating that the family environment has a significant impact on the mental health of college students. Relevant educators can timely understand the changes in students' family environment by establishing an effective mechanism for regular family contact. Other important factors are: whether to experience recent stress events, communication ability, grade, love situation, weekly exercise time, and community activities. Therefore, encouraging students to participate in sports, cultivating a good learning atmosphere, and various community activities is conducive to the cultivation of college students' mental health. Grade is ranked fourth, indicating that as college students grow, their mental state constantly changes, demonstrating the dynamic nature of college students' mental health. Hence, regularly implementing mental health education is essential to ensure that this change is in the direction of improving mental health.

4.2 Model comparison in the case of small samples

The number of samples was reduced to 100, 150, and 200. The simulated data was still divided into training and test sets using 5 - fold cross - validation. The experimental results of Xgboost, decision tree, SVM (Support Vector Machines), and random forest on small samples were compared (see Table 3). The results show that the random forest model still achieved the best results in the four evaluation indicators of accuracy, precision, recall, and F1 - score. This indicates that the random forest model can more effectively predict the mental health status of students based on the SCL - 90 scale, helping schools to promptly identify students with psychological problems and carry out relevant interventions and counseling, enabling them to return to normal life and study, which has certain practical significance.

Table 3: Comparison of the effects of machine learning methods in the case of small samples

Algorithm	Sample size	Accuracy	Precision	Recall	F1	AUC
Decision Tree	100	0.9	0.8333	1	0.9091	0.9
	150	0.9556	0.9167	1	0.9565	0.9565
	200	0.95	0.9697	0.9412	0.9552	0.9514
Support Vector Machine	100	0.9667	0.9375	1	0.9677	1
	150	0.9778	0.9565	1	0.9778	0.998
	200	0.9833	1	0.9706	0.9851	0.9989
Random Forest	100	0.9333	0.8824	1	0.9375	0.9889
	150	0.9556	0.9167	1	0.9565	0.998
	200	0.9833	1	0.9706	0.9851	0.9989
XGBoost	100	0.9333	0.9333	0.9333	0.9333	0.9956
	150	0.9556	0.9167	1	0.9565	1
	200	0.9667	0.9706	0.9706	0.9706	0.9989

5 Discussion and suggestions

The research results show that:

The various indicators of Xgboost, decision tree, SVM (Support Vector Machines), and random forest are very similar and the scores are relatively high. This indicates that these models all fit the data well and can be considered as model options. Overall, the random forest model has a slightly higher score, showing the best scoring performance.

The classification accuracy of machine learning methods is not significantly affected by the sample size. The decision tree model performs poorly in classifying data, while the other models perform well.

In small - sample simulated data, the random forest algorithm and the support vector machine show good classification accuracy. The random forest algorithm exceeds the model with a sample size of 500 when the sample size is 200.

In actual small - sample psychological health diagnosis, it is recommended to use the support vector machine and the random forest algorithm, but the total number of attributes is required to be small. Through this research, the following conclusions can be drawn: Machine learning algorithms can provide rapid, consistent, and unbiased evaluations. Model algorithms can identify risk factors and predict outcomes, facilitating early intervention and prevention. They can also help customize treatment plans and optimize treatment effects.

From the comparative analysis of the performance of the four models in Table 2 and Table 3, it can be seen that the random forest model performs the best overall. The reasons are as follows:

(1). Reducing the impact of outliers: By constructing multiple decision trees and taking the average or majority - voting results, the random forest can effectively reduce the impact of individual outliers on the overall prediction results. In contrast, decision trees are prone to overfitting with relatively low prediction accuracy. The random forest greatly improves the model performance by integrating multiple decision trees.

(2). Reducing the risk of overfitting: Since the random forest uses a part of the samples and features to construct each decision tree, it reduces the possibility of overfitting, enabling the model to have better generalization ability. The random forest can handle thousands of input variables, determine the most important variables, and has strong processing ability for high - dimensional data. Therefore, it is considered a good dimensionality reduction method.

(3). Parallelization and distributed implementation: The highly parallel nature of the random forest makes it easy to implement in a distributed manner, which is very beneficial for large - scale data processing. When dealing with non - linear problems, SVM needs to select an appropriate kernel function, which has a high computational complexity and a slow training speed on large - scale data. The random forest performs better in handling high - dimensional features.

(4). Providing feature importance: The random forest can output the importance degree of features, which is very helpful for understanding and analyzing data features.

Therefore, the random forest shows certain advantages over decision trees and SVM in many aspects. However, in practical applications, the choice of which algorithm still needs to be determined according to the characteristics of specific problems and datasets.

This study also fits the psychological diagnostic data under small samples, and compares the prediction of results by different machine - learning methods, so as to explore the classification accuracy of different machine - learning methods under small samples. Although this study has examined the following aspects: (1) Explore the susceptible factors that affect the mental health level of college students with positive psychological symptoms; (2) Explore the effectiveness of machine - learning - based modeling in identifying and predicting college students with potential psychological problems; (3) Apply the effectiveness and accuracy of each machine - learning model in simulated small - sample data., limitations and directions of improvement are as follows: (1) The research objects are limited to a certain university in Hubei Province, so the sample representativeness is limited. (2) The predictive variables (susceptible factors) of the model in this study are measured by self - rating scales. Research shows that certain physiological susceptible factors also have a certain predictive effect on mental health. In the future, physiological indicators can be added to establish a prediction model to improve the accuracy of the model and provide more powerful auxiliary tools for the clinical work of college students' psychological health.

References

[1] Kolenik T, Gams M. Intelligent cognitive assistants for attitude and behavior change support in mental health: state-of-the-art technical review. Electronics 2021, 10: 1250. https://doi.org/10.3390/electronics10111250

[2] Kolenik T, Gams M. Persuasive technology for mental health: one step closer to (Mental Health Care) Equality? IEEE Technology and Society Magazine, 2021, 40(1): 80-86. https://doi.org/10.1109/MTS.2021.3056288.

[3] Kolenik T. Methods in digital mental health: smartphone-based assessment and intervention for stress, anxiety, and depression. Internet of Things Integrating Artificial Intelligence and IoT for Advanced Health Informatics, 2022: 105-128. https://doi.org/10.1007/978-3-030-91181-2_7.

[4] Ahmed U, Srivastava G, Yun U, et al. EANDC: An explainable attention network based deep adaptive clustering model for mental health treatment. Future Generation Computer Systems, 2022, 130: 106-113. https://doi.org/10.1016/j.future.2021.12.0

[5] Li LL. Online mental health education teaching mode and empirical research based on Artificial intelligence. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 2021, 40(2): 3467-3476. https://doi.org/10.3233/JIFS-189384

[6] Chen MM, Jiang S. Analysis and research on mental health of college students based on cognitive computing. Cognitive Systems Research, 2019, 56: 151-158. https://doi.org/10.1016/j.cogsys.2019.03.003

[7] Zhou LJ. Evaluation research on data processing of mental health of college students based on decision tree algorithm. Journal of Computational Methods in Sciences and Engineering, 2019, 19(4): 1101-1108. https://doi.org/10.3233/JCM-193866

[8] Liu HY, Xu JL. Research on the use of neural network for the prediction of college students' mental health. Scientific Programming, 2023. https://doi.org/10.1155/2023/9840394.

[9] Zhang Y, Naparstek S, Gordon J, et al. Machine learning-based identification of a psychotherapy-predictive electroencephalographic signature in PTSD. Nature Mental Health, 2023, 1: 284-294. https://doi.org/10.1038/s44220-023-00049-5.

[10] Koppe G, Meyer-Lindenberg A, Durstewitz D. Deep learning for small and big data in psychiatry. Neuropsychopharmacology, 2021, 46: 176-190. https://doi.org/10.1038/s41386-020-0767-z.

[11] Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics, 1991, 21(3): 660-674. https://doi.org/10.1109/21.97458.