Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Application of Intelligent Algorithms in Student Mental Health Management and Risk Early Warning Model Research

Min Qin[1], Yang Li[2], Jihua Cao[3,] ✉

[1] *School of Design and Creativity, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China*

[2] *Basic Teaching Department, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China*

[3] *Student Work Office, Guilin University of Electronic Technology, Guilin, Guangxi, 541000, China*

ABSTRACT

There is a close relationship between adolescent mental health and physical health, so it is of great practical significance to explore the specific influencing factors and early warning model of students' mental health. In this paper, the early warning model of students' mental health risk is constructed. Firstly, the association rules and Apriori algorithm are used to explore the relationship between the important influencing factors of students' mental health and common psychological problems, and then the CMA-ES-XGBoost prediction model is proposed to address the defects of the XGBoost prediction model that has high complexity and low prediction accuracy. It adopts the hyperparameters of CMA-ES optimization algorithm to find the optimal hyperparameter solution, and solves the fuzzy phenomenon existing in the early warning of mental health risks by fuzzy logic method, which reduces the error of prediction results. It is experimentally verified that the mental health prediction method based on CMA-ES-XGBoost performs well on the task of students with mental health risk, and the prediction accuracy is 89.66%, which is better than the comparison model. It can accurately detect the mood fluctuations of students with different types of personality when they are exposed to multiple extroverted stimuli, and accurately predict the emotional risk. It shows that the model in this paper realizes the function of predicting students' mental health status and achieves the expected goal of model design.

Keywords: Association rules; Apriori algorithm; CMA-ES-XGBoost prediction model; Fuzzy logic; Mental health management

✉ Corresponding author.
E-mail address: caojihua2024@126.com (J. Gao).

Received 12 January 2024; Revised 25 March 2024; Accepted 25 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-302

1. Introduction

Adolescents are at a critical stage of physical and mental development and are under pressure from various aspects such as academics, interpersonal relationships, and self-perception [1-2]. These pressures, if not properly handled, may trigger anxiety and depression in students and even hinder their healthy growth. Traditional psychological management methods are subjective and inefficient, which are contrary to the personalized and precise management required by modern education [3]. For this reason, digital technology is introduced into the field of mental health education by collecting and analyzing multi-dimensional data such as students' academic performance, homework completion, attendance, online behavior and social interactions, and then comprehensively describing the students' knowledge abilities, interests, behavioral habits and psychological characteristics [4-7]. While improving the accuracy and efficiency of mental health assessment, it also provides strong data support for education and teaching [8].

In order to pay comprehensive attention to students' mental health and detect potential psychological problems in a timely manner, more and more schools have begun to build a risk warning mechanism. By integrating students' historical data and current behavioral characteristics and using advanced algorithms and technologies, this mechanism aims to predict possible future psychological problems of students [9]. This early warning mechanism provides schools with a powerful tool to help them intervene before problems arise, thus effectively preventing them from occurring or worsening [10-12]. The risk prediction model is the core of the risk early warning mechanism. This model is capable of deep analysis and learning based on multidimensional data from students [13]. Through machine learning algorithms, the model can identify key features and patterns associated with students' psychological problems [14]. With the continuous progress of technology and the increasing abundance of data, the risk early warning mechanism will play a more important role in students' mental health education [15]. To this end, schools can further improve the function and accuracy of the early warning system, strengthen cooperation and communication with parents, communities and other parties to build a more comprehensive and efficient mental health education system.

The application of intelligent algorithms provides students with a comprehensive and objective psychological assessment, which helps teachers and parents to better understand students' psychological status and needs so as to formulate more effective educational strategies, and some scholars have studied the design scheme of related risk warning models. Feng, Y. et al. established an early warning model for college students' mental health based on long short-term memory (LSTM) network, and utilized the temporal characteristics of LSTM model to capture its change trend in time to achieve accurate prediction and intervention of mental health status [16]. Liu, Z. et al. introduced the application of distributed multi-agent system (MAS) algorithm in the construction of mental health early warning model for college students, which can identify multiple potential mental health problems by collecting relevant psychological data and analyzing them, and effectively improve the mental health and quality of life of college students [17]. Ye, L. proposed a deep learning algorithm model for college students' mental health early warning, which can utilize the expandability of deep learning algorithms when dealing with large-scale mental health data, and synthesize the learning results of each intelligent body to provide more comprehensive support for college students' mental health management [18]. Chen, Y. et al. constructed a multivariate decision tree model for college students' psychological crisis behaviors, combined with the goal equation calculation to obtain the optimal warning results, and realized the risk warning of college students' psychological crisis behaviors [19]. Xiong, W. developed a big data analysis system for college students' mental health education, and the altered system is based on massive student behavioral data on the Internet, which can predict the trend of students' psychological behaviors, so as to build a personalized emergency management mechanism for psychological crisis [20]. Wang, J. et al. designed a psychological crisis indicator system for college students based on constructivist psychology and other theories, and

developed a psychological crisis early warning model for college students based on this combined with a genetic BP neural network, which embodied high practical value [21]. Although the above machine learning prediction capabilities may be very powerful, and even their accuracy may be infinitely close to 100%, there are still drawbacks such as lagging, susceptibility to interference, and passivity in screening and risk prediction of mental health problems, and further research is urgently needed.

In this paper, an association rule mining algorithm is used to mine the important factors affecting students' mental health status and their degree of association. A risk warning model based on XGBoost algorithm for predicting mental health status is designed, and the covariance matrix adaptive evolution strategy is used to optimize the XGBoost algorithm. The parameter range setting and spatial structure under the search of the improved method are more scientific. In order to reduce the misjudgment of the mental health risk warning model, fuzzy logic is combined with the model in this paper. Finally, the prediction accuracy and optimal performance conditions of the model are verified through comparative experiments and cross-sectional optimal parameter search experiments. The validity of the model is verified in practical application, and in this way, the supplementary strategy for the application of the mental health risk warning model in this paper is proposed.

2. Selection and exploration of factors influencing mental health

2.1. Association rules

In order to make the subsequent descriptions of the algorithms related to association rule mining clearer and more accurate, some of the basic concepts and knowledge involved in association rule techniques are listed below.

Definition 1: Let there exist a set $IL=\{I_1,I_2,...,I_m\}$, each element of IL is called an item $(Item,I)$. The transaction database is denoted by $D=\{T_1,T_2,...,T_n\}$, each element of D is called a transaction $T(Transaction)$, also known as the set of items, T satisfies $T\subseteq I$, and each T has a non-repeating identifier, denoted as TID.

Definition 2: Let itemset X satisfy $X\subseteq I$ and X is composed of k items, then we can call X as K itemset.

Definition 3: Let item set X appear in the transaction database D times for M, D in the transaction T of the total number of N, then the value of MN is called the item set X of the degree of support, notation $Support(X)$, which X appears in D times is also called the degree of support for the number of counts, notation $Support\_count(X)$. In general, the user can set their own minimum threshold of support, notation min_supp, and $min\_supp\in[0,1]$.

Definition 4: Let the minimum support threshold be min_supp. If the support of itemset X is greater than or equal to min_supp, X is called a frequent itemset, and if X consists of k items, then X is a frequent K itemset.

Definition 5: If itemset $X\subseteq T$, itemset $Y\subseteq T$, and $X\cap Y\neq\varnothing$, a relationship $X\Rightarrow Y$ such as this can be called an association rule, which is used to indicate that a transaction T that contains X is likely to also contain Y.

Definition 6: With an association rule $X\Rightarrow Y$, if the itemsets containing both X and Y are denoted as $X\cap Y$, then the total number of times $X\cap Y$ occurs in D compared to the total amount of data in D is called the support of $X\Rightarrow Y$ and is denoted as $Support(XarrowY)$. The computational expression for this is:

$$Support(XarrowY)=P(XY)=\frac{N(X\cap Y)}{N(D)}=Support(X\cap Y) \tag{1}$$

Definition 7: Given an association rule $X\Rightarrow Y$, if the set of items containing both X and Y is denoted as $X\cap Y$, then the support of $X\cap Y$ over X is called the confidence level of $X\Rightarrow Y$,

denoted as $Confidence(X \Rightarrow Y)$, and min_conf is called the minimum confidence threshold, and $min\_conf \in [0,1]$. Its computational expression is:

$$Confidence(XarrowY) = P(Y \mid X) = \frac{N(X \cap Y)}{N(X)} = \frac{Support(X \cap Y)}{Support(X)} \tag{2}$$

Definition 8: With an association rule $X \Rightarrow Y$ and a minimum support threshold min_supp and a minimum confidence threshold min_conf, an association rule $X \Rightarrow Y$ is called a strong association rule if it satisfies the following two equations.

$$Support(X \Rightarrow Y) \geq min\_sup \tag{3}$$

$$Confidence(X \Rightarrow Y) \geq min\_conf \tag{4}$$

2.2. Apriori Algorithm

The Apriori [22] algorithm is simple and clear. If the algorithm exists frequent K itemsets, it will read the k transaction database, and frequent reading of the transaction database will increase the I/O load consuming a lot of time; at the same time, the algorithm connects through the frequent itemsets themselves, which will produce a large number of redundant candidate itemsets.

1) Algorithm idea. The main idea of Apriori algorithm to mine frequent itemsets is to search layer by layer and find frequent itemsets in a sequential loop. First, read all the transactions in the transaction database, calculate the number of times each item occurs, according to the user-specified value of min_supp, to find out all the frequent 1-item set L_1. Then, let every two items in the frequent 1-item set be connected to form a 2-item set C_2, and then reread all the data in D, according to the user-specified min_supp -value, delete the items that do not meet the value of min_supp, to find out all the frequent 2-item set L_2. next. Loop the above operation, so that the itemsets in L_{k-1} obtained in the previous time are connected two by two to form C_k, read all the data in D, according to the user-specified min_supp value, find out all the L_k. Repeat the above process until no new frequent itemsets are generated, and the algorithm ends.

2) Algorithm flow. The flow chart of Apriori algorithm is shown in Fig. 1:

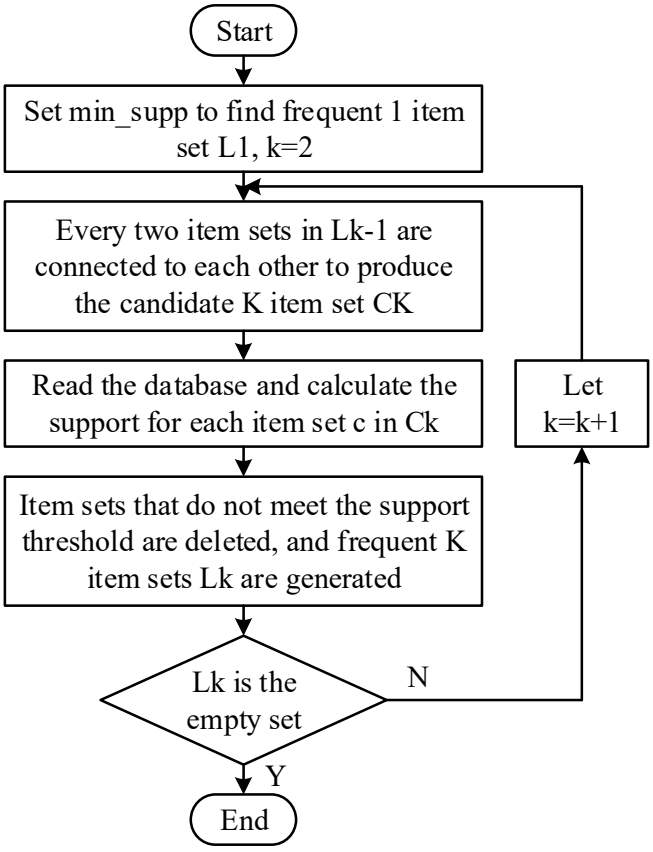

Fig. 1. Flowchart of the Apriori algorithm

2.3. XGBoost model construction

The basic idea of integrated learning is to combine a series of weak learning models into a strong learning model to improve the performance of machine learning and thus provide better prediction results than a single model. As an integrated learning method based on boosting method, XGBoost method integrates these weak classifiers by increasing the weights of samples with high error rate and weak classifiers with low classification error and selecting the classifier that minimizes the loss function as a new classifier in each iteration and integrating these weak classifiers by certain methods until the end of iteration condition is reached.

XGBoost [23] can be termed as an emerging branch of Boosting algorithm, which can be literally understood as Extreme Gradient Boosting. Extreme Gradient Boosting model is an improved Gradient Boosting Decision Tree (GBDT) proposed in Kaggle Machine Learning Competition with excellent performance and is a scalable tree augmentation system and hence it is very popular in data mining and machine learning competitions around the world.XGBoost model algorithm is more complex.The approximate workflow of XGBoost algorithm is as follows:

For a training set given N sample and M conditions, $|D|=n, x_i \in R^m, y_i \in R$. The XGBoost model is trained to generate a decision tree with K categorical regression decision tree (CART), which is mathematically expressed as follows:

$$\hat{y}_i = \varphi(x_i) = \sum_{i=1}^{K} f_i(x_i) f_i \in F \tag{5}$$

where $\hat{y}_i$ is the predicted value of sample x_i and F is the set of CARTs:

$$F = \{f(x) = \omega_{q(x)}\}(q : R^M arrow T, w \in R^T) \tag{6}$$

Where q is the structure of the tree, T is the number of leaf nodes, and w is the vector of weight coefficients of the leaf nodes. When a sample x_i is input to the XGBoost model, multiple CARTs in the model map the sample to the leaf nodes of each tree, and then the values on the leaf nodes of the CARTs are weighted and summed according to the leaf node weight vectors w to obtain the predicted values. When setting the parameters of the XGBoost algorithm, the CART can also be constrained, because it can be seen that when the depth of the tree is deeper, it means that the classification in its learning is more detailed, although it can ensure the accuracy of classification learning, but makes its generalization performance become lower. As a supervised learning method, its objective function is:

$$MinG = \sum_{i=1}^{n} Il(y_i, \hat{y}_i r) + \sum_{t=1}^{K} \Omega(f_i r) \tag{7}$$

where l is the loss function, which is used to measure the gap between the predicted score and the true score. Ω is the penalty function that incorporates a regular term, which serves to prevent the risk of model overfitting, and has the following expression:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{K} w_j^2 \tag{8}$$

where T is the number of leaf nodes and w_j is the j leaf node weights.

To construct the final XGBoost model to calculate the f_i of each tree, the t th tree is trained by the forward distribution algorithm with continuous iteration. Since the predicted values and residuals of the first $k-1$ trees have no effect on the optimization of the objective function, they can be directly removed, and the penalty function is expanded to obtain the final objective function as:

$$L^{(K)} = \sum_{J=1}^{T} gl\left[\left(\sum_i g_i\right) w_j + \frac{1}{2}\left(\sum_i h_i + \lambda\right) w_j\right] + \gamma T \tag{9}$$

The objective function can take a uniquely determined minimum value if and only if w_j takes the optimal value:

$$obj^* = -\frac{1}{2}\sum_{j=1}^{T} \frac{G_j}{H_j + \lambda} + \gamma T \tag{10}$$

The results of the decision tree are evaluated using this objective function, with smaller values representing a better tree structure. In general, setting the threshold value in advance during the calculation process can prevent the tree from growing too deep and is used to control the overfitting of the model. When the information gain cost exceeds the set threshold, the decision tree stops splitting, and the final objective function is the structural evaluation score of the decision tree. The objective function is solved optimally to obtain the optimal sub-model $f_K(x_i)$ at each step, and the overall XGBoost model can be obtained according to the principle of additive modeling.

Generally speaking, in a supervised machine learning algorithm, two parts need to be predefined: the parameters of the machine learning algorithm and the objective function of the machine learning algorithm. The parameters of the model are adjusted to make the model more suitable for processing data. Using the objective function as a supervisor allows the model to iterate to find the model that makes the objective function perform best and further process the data. As mentioned earlier, the penalty constant in support vector regression, the constant is used to ensure that the model can achieve a balance between prediction accuracy and generalization of prediction by setting a penalty value for classification error in the objective function. As for the XGBoost method, in order to make the model process data with high accuracy, the method of minimizing the penalty function is used to achieve it. At the same time, in order to improve the generalization performance of the model to

prevent overfitting, a regular variable is added, which can be interpreted as a penalty for the complexity of the model.Each iteration of XGBoost forms a classifier that minimizes the penalty function, and when the end condition of the iteration is reached, it forms a model that is more balanced in terms of the generalization performance and the accuracy of the data to be processed.

Overall, the extreme gradient boosting model is a highly scalable tree-structured augmentation model that can handle sparse and missing data well, and can greatly improve the speed of the algorithm and compress computational memory in large-scale data training.

2.4. Selection of Impact Factors

The Apriori algorithm was used to scan the data of five common psychological problems variables to obtain the set of items with a support degree greater than 0.8%, and sorted according to the support degree, and the sorting results are shown in Table 1, for example. It can be seen that school life maladjustment, learning stress and the combination of the two are the most frequent psychological problems of the students in this survey, 69.41% and 65.28% of the students have maladjustment and learning stress problems, respectively, and among them, 46.52% have both the sense of learning stress and maladjustment problems, which means that most of the students who have the sense of learning stress or maladjustment problems have both of them at the same time.

Table 1. Common psychological problem series

Serial number	Support	Item set
1	0.6941	Maladjustment
2	0.6528	Learning pressure
3	0.4652	Adapt to bad, study pressure
4	0.1235	Interpersonal sensitivity
5	0.1204	Anxiety
6	0.1194	Depression
7	0.1112	Adapt to bad, interpersonal tension and sensitivity
8	0.1009	Interpersonal stress is sensitive and learning pressure
9	0.0956	Depression, learning pressure
10	0.0954	Anxiety, learning pressure
11	0.0923	Adapt to bad, interpersonal tension and sensitivity, learning pressure
12	0.0815	Adapt to bad, anxiety
13	0.0803	Adapt to bad, depression

According to the rule that the output confidence is greater than 0.8 and the enhancement is greater than 1.125, the items are sorted according to the confidence level, and the sorting results are shown in Table 2. It can be seen that the causes of frequent item sets of school life maladjustment, learning stress and the combination of the two are interpersonal tension sensitivity, depression and anxiety. The rule with the highest confidence level is (interpersonal tension sensitivity, academic stress) → (maladjustment) with a confidence level of 0.9115, which indicates that 91.15% of the students who develop interpersonal tension sensitivity and academic stress experience maladjustment in school life. The support level of this rule is 0.0916, and the support level is located in the 4th place, which indicates that the probability of its occurrence is high. At the same time, this rule has the highest degree of enhancement of 1.3207, which indicates that when the combination of interpersonal sensitivity and academic stress occurs, it leads to an increase in the probability of the emergence of maladjustment in school life by 32.07%. The rule with the highest confidence in the combination of psychological

problems with the highest support is (interpersonal sensitivity) → maladjustment with a support of 0.1123, which has the highest probability of occurrence of this rule. From the above analysis, it can be seen that interpersonal tension sensitivity, depression and anxiety are 3 fundamental and important psychological problems that lead to 5 common psychological problems in students, and we need to focus on these 3 important psychological problems and explore the important influencing factors that lead to them.

Table 2. Common psychological problem strong association rule

Serial number	Causal set	Result item set	Support	Confidence	Degree of ascension
1	Interpersonal stress is sensitive and learning pressure	Maladjustment	0.0916	0.9115	1.3207
2	Interpersonal sensitivity	Maladjustment	0.1123	0.8982	1.3023
3	Depression, anxiety	Learning pressure	0.0503	0.8257	1.2649
4	Adapt to bad, interpersonal tension and sensitivity	Learning pressure	0.0915	0.8211	1.2579
5	Adapt to bad, depression	Learning pressure	0.0644	0.8103	1.2407
6	Interpersonal sensitivity	Learning pressure	0.1005	0.8097	1.2406
7	Depression	Learning pressure	0.0958	0.8075	1.2365
8	Adapt to bad, depression	Learning pressure	0.0647	0.8001	1.2263

Before the data mining began, first, the 23 influencing factor variables were correlated with the 3 important psychological problems variables to explore their linear correlations. Serial numbers 1-26 are teacher-student relationship, classmate relationship, parent-child relationship, parent-child communication frequency, self-confidence, physiological health, communication smoothness, physical exercise, study load, school physical environment, academic performance, living environment, family economic conditions, study atmosphere, sleep time, precautionary awareness, teacher quality, school management, foreign language proficiency, classroom discipline, mutual assessment game, introverted personality, depression, anxiety , interpersonal tension and sensitivity issues, relationship richness. The heat map of Pearson correlation coefficient between important psychological problem variables and important influencing factor variables is shown in Figure 2, which shows that there is a certain linear correlation between each influencing factor variable and important psychological problem variables, but the linear correlation is not too obvious.

From the correlation between the important psychological problem variables and the influencing factor variables, there is a relatively obvious negative correlation between the level of depression problem (yy) and parent-child relationship (qzgx), self-confidence (zxx), and physical health (sljk). There is a relatively significant negative correlation between the level of anxiety problems (jl) and self-confidence (zxx), physiological health (sijk), and a relatively significant positive correlation with introversion (nxxg). There was a relatively significant negative correlation between the level of interpersonal tension and sensitivity problems (rjgxjzmg) and peer relationships (txgx) and parent-child relationships (qzgx).

Therefore, in order to further explore the influencing relationships and influencing rules between the 23 influencing factor variables and the important psychological problem variables, further mining of the data is needed. The correlations within the 23 influencing factor variables can be observed, and

it can be seen that the relationship between classmates (txgx) and the relationship between teachers and students (ssgx), the frequency of parent-child communication (ffyhzgtpl) and the parent-child relationship (qzgx), the physical environment of the school (xxwzhj) and the burden of study (xxfd), and the family's economic conditions (jtjjtj) and the living environment (jzhj), There is a relatively strong positive correlation between school management (xxgl) and teacher quality (lssz). There is a relatively strong negative correlation between study atmosphere (xxfw) and study load (xxfd). Accordingly, the following conclusions can be drawn, there is a mutually reinforcing effect between classmate relationship and teacher-student relationship. Increasing the frequency of parent-child communication helps the parent-child relationship. Students in schools with better physical environments have relatively heavier study loads. Students from better-off families have a better living environment. Teachers in schools with quality management are more qualified. Cultivation of good learning atmosphere helps to reduce the burden of learning on the students.

There are many missing values in the data of the questionnaire data for each influence factor. Since linear models such as logistic regression and linear regression are sensitive to outliers and missing values, tree-based models have the advantage of being resistant to outliers and can deal with missing values. Based on this feature of this dataset, decision tree and random forest algorithms are selected for this paper.

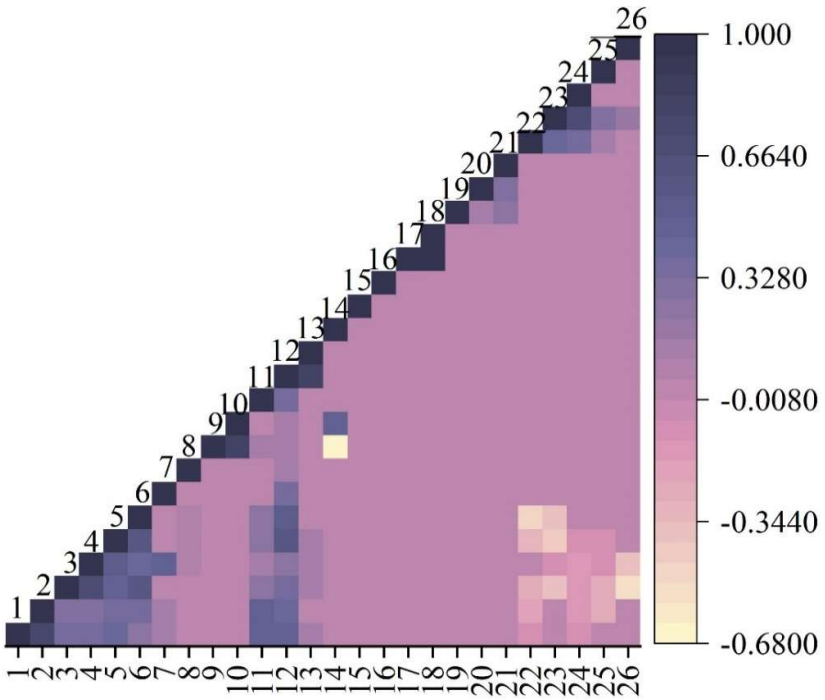

Fig. 2. The correlation coefficient of the variable and psychological problem

3. Early warning model of mental health risk based on improved XGBoost

The previously mentioned mental health risk warning model uses the XGBoost algorithm, which operates by combining classification and regression as a meta-learner, constructing subsequent trees based on the training residuals of previous trees, and gradually enhancing its predictive power through multiple training iterations. Progressively optimizing the objective function, the model strives to achieve the most accurate predictive value and build a powerful learner. In order to further improve the reliability of the model samples and the accuracy of the prediction results, this section proposes an optimization scheme for the mental health state model based on the CMA-ES-XGBoost algorithm.

3.1. CMA-ES hyperparameter optimization design

Hyperparameter optimization is the process of finding the hyperparameters that perform best for a machine learning model on a validation dataset. In machine learning, hyperparameters are parameters that need to be set manually and cannot be determined automatically through learning, and the optimal values need to be determined through trial and error. Different values of hyperparameters may have a significant impact on the performance and effectiveness of the model, so some specific methods are needed to search for the range of hyperparameter values, such as lattice search, stochastic search, and Bayesian optimization.

Traditional hyperparameter optimization methods mainly include grid search and stochastic search, and the two methods differ in the way they search for the optimal hyperparameters. Grid search finds the optimal hyperparameter combination by exhaustively enumerating all possible combinations within a given hyperparameter range. Random search, on the other hand, tests a certain number of parameter combinations by randomly sampling them within a specified parameter range. From the point of view of computational efficiency, grid search has a higher computational complexity and requires testing all possible parameter combinations. In contrast, the computational complexity of random search is lower and only a certain number of parameter combinations need to be tested. For the case of high parameter space dimensionality, the number of parameter combinations that need to be tested for grid search grows exponentially. In contrast, stochastic search is able to sample the parameter space more extensively, thus making it easier to find the globally optimal solution.

The core of the CMA-ES-XGBoost algorithm is to optimize the parameters of XGBoost with CMA-ES [24] to optimize the parameters of XGBoost to a better initial position, preventing it from falling into a local optimum and failing to train the machine learning model effectively. Its main features and advantages include, first, CMA-ES has a high global search capability and can search for a better combination of parameters in a shorter time. Second, CMA-ES performs well in dealing with high-dimensional, non-convex, and non-linear optimization problems, while there is a linear or non-linear relationship between the influencing factors of mental health assessment, and the hyperparameters of XGBoost are usually multi-dimensional as well, and CMA-ES can better adapt to this complexity. Third, CMA-ES can accelerate the search process through parallelization, the hyperparameter tuning of XGBoost usually requires a large number of iterations and evaluations, and parallelization can accelerate the process of model tuning. The specific implementation process of the CMA-ES-XGBoost algorithm is shown in Figure 3.

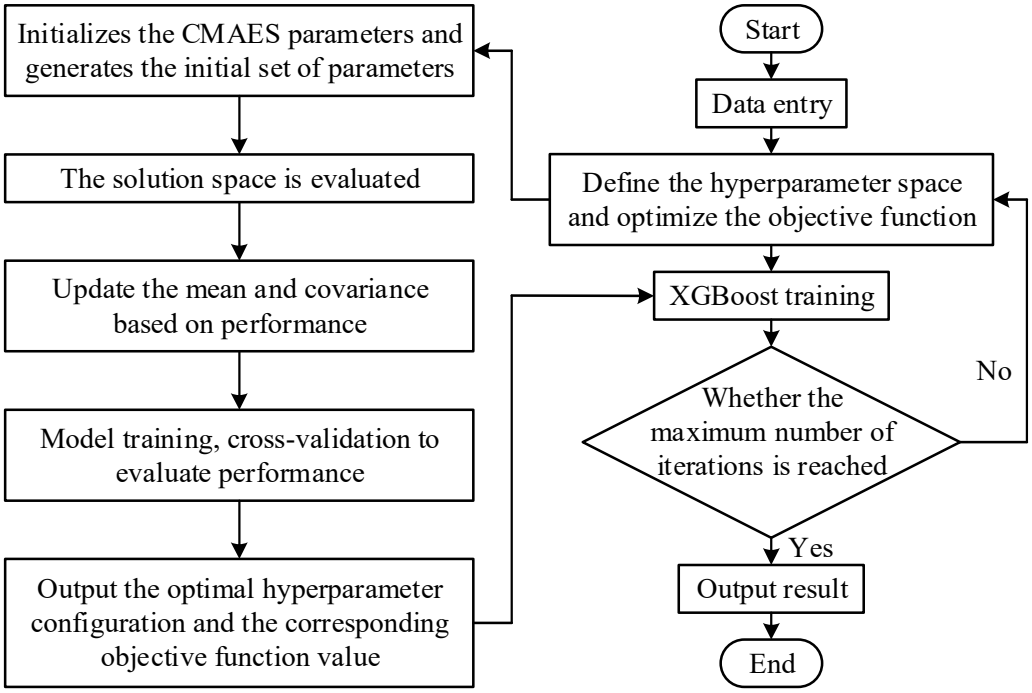

Fig. 3. Flowchart of the CMA-ES-XGBoost algorithm

3.2. Fuzzy_Logic probabilistic prediction

Fuzzy logic is a method and tool that can be used to mimic human thought patterns in order to more accurately represent, analyze and resolve incomplete and inaccurate information. In fuzzy set theory, the relationship between an element z in the domain Z and a certain set X on the domain Z is fuzzy and this affiliation can be represented by the following mapping:

$$Z \to [0,1] \tag{11}$$

$$\mu_{X'} : Z \to \mu_{X'}(z) \tag{12}$$

where X' is a fuzzy set, μ_x, is the affiliation function of fuzzy set X'; $X'(z)$ denotes the affiliation of element z in the domain Z to fuzzy set X. It can be seen that the degree of affiliation of element z to fuzzy set X' can vary between 0 and 1, instead of the non-zero-is-one case in classical set theory. By using the affiliation function, it is possible to quantitatively describe the phenomenon of some things whose boundaries are fuzzy and uncertain.

The principle of mental health risk early warning model based on CMA-ES-XGBoost algorithm is binary prediction, on the one hand, the result of binary model itself will have judgment error. On the other hand, under normal circumstances, there is a certain degree of ambiguity in the psychological assessment, and it is difficult to make a specific judgment on the “tiredness” and “relaxation” of the students, and the assessment scale is based on the score system, so there is a deviation between the psychological assessment data and the actual situation. Different groups may feel a certain symptom differently at different times, for example, during the Xin Guan epidemic, students are more sensitive to anxiety symptoms, depression, etc. The results of the assessment data will affect the accuracy of the trained model, which will ultimately lead to the actual prediction of the results will also have some deviation. The fuzzy inference process of the prediction model is shown in Figure 4.

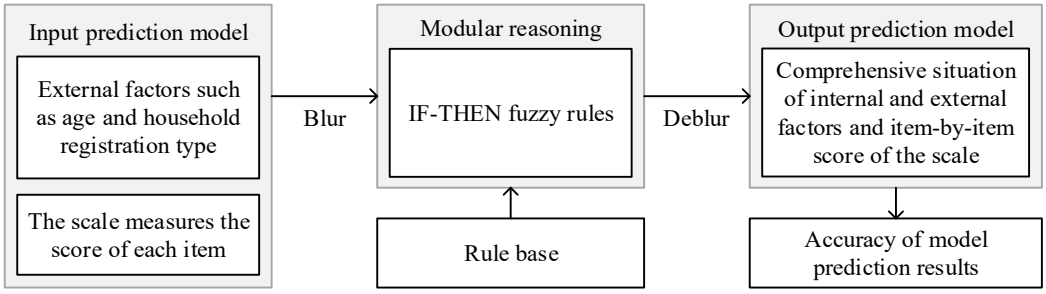

Fig. 4. Fuzzy reasoning process of mental health state prediction model

In order to solve the above problem of the prediction result deviation of the psychological assessment data, this paper improves the mental health risk early warning model of anxiety state by adding fuzzy logic, fuzzy processing the information of the mental health assessment data obtained by the system, establishing a fuzzy rule base and carrying out defuzzification of the model, and ultimately improving the accuracy of the model's prediction results.

The fuzzy inference process of the mental health risk early warning model in this chapter includes the following steps:

Step 1: Select the input and output variables that can reflect the mental health risk early warning model.

Step 2: Determine the input and output affiliation functions. In general, the triangular affiliation function is mainly selected to describe the linguistic variables. In this experiment, the triangular affiliation function is chosen for the study.

Step 3: Fuzzy rules are a set of logical relationship expressions for causal reasoning, the Symptom Self-Rating Scale SCL-90 includes 10 factors, 90 items, and each assessment item is scored as 1-5 points, using the If-Then principle to construct the corresponding fuzzy rule base for a certain measurement of psychological dimensions, and its fuzzy variables include 10 items of the topic, each variable includes Score situation: 1, 2, 3, 4, 5 five terms, based on 9 fuzzy variables, close to 50 fuzzy terms, a total of 450 fuzzy rules are defined. Mamdani fuzzy inference method is adopted in this paper.

Step 4: Fuzzy reasoning and defuzzification after fuzzification of input variables. The process of converting fuzzy quantities after fuzzy reasoning into precise clear quantities is defuzzification. In this experiment, the center of gravity method is used to defuzzify the output of the mental health risk warning model, and after defuzzification, the output of the mental health risk warning results in the form of probability is obtained.

4. Mental health model predictions

This chapter for this experiment using a well-known university students in the campus of the full amount of data shall prevail, involving a total of 36598 students, after data cleaning, data governance and other data pre-processing stage, to obtain valid student data 32858 people, including for the marking of psychological assessment of the abnormal students there are 456 people. Through the train_test_split() function in accordance with the ratio 7:3 each time to divide the corresponding test set for testing.

Within the field of machine learning, evaluation metrics play a crucial role when assessing the performance of a model. These metrics give this chapter the ability to differentiate between the performance of different models and identify the most appropriate model for a given problem. In this chapter, particular attention is paid to metrics that correctly identify individuals without and with psychological problems.

The error matrix is a widely used assessment tool for classification tasks. It usually takes the form of an n*n matrix, with the true categories represented through the rows of the matrix and the predicted

categories through the columns. Such matrices are often used to graphically demonstrate algorithm performance. Accordingly, the following four concepts are also involved in the error matrix:

True Positive (TP): a sample of correctly identified positive categories. True Negative (TN): correctly identified negative class samples. False Positive (FP): negative class samples incorrectly identified as positive. False Negative (FN): positive class samples incorrectly determined as negative class.

Generally, precision, recall, and F1-Measure are key metrics commonly used to measure the quality of a model in the performance evaluation of classification models, and are especially useful in unbalanced category problems. These metrics provide a comprehensive perspective for evaluating model performance in different aspects.

Among them, each assessment metric is defined as follows:

Precision:

DEFINITION: Precision is the proportion of samples predicted by the model to be positive cases that are actually positive cases. The relevant formula for its calculation is as follows:

$$\Pr ecision = \frac{TP}{TP + FP} \tag{13}$$

Recall:

Definition: recall is the proportion of samples that are actually positive cases that are correctly predicted by the model to be positive cases. The relevant formula is as follows:

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

F1-Measur:

DEFINITION: The F1 value is the reconciled average of precision and recall, combining the accuracy and completeness of the classifier. Its related formula is as follows:

$$F1 = \frac{2 * Accuracy * \mathrm{Re}call}{Accuracy + \mathrm{Re}call} \tag{15}$$

Taken together, the precision rate is concerned with the accuracy of the model's prediction as a positive example, the recall rate is concerned with the model's ability to recognize a positive example, and the F1 value finds a balance between the two. In different application scenarios, it may be necessary to adjust the weights of precision rate and recall rate according to the importance of the problem. In mental health prediction, the correct rate of detection of mental health anomalies is of more concern, i.e., among the individuals judged by the model to have mental health problems, what percentage of them actually have the problem. Thus, this chapter proposes the expression for calculating the accuracy rate as:

$$Accuracy = \frac{TN}{TN + FN} \tag{16}$$

The experiments conducted in this chapter aim to evaluate the performance of the proposed heterogeneous multiple machine learning algorithms integration framework in the task of student mental health warning. Specifically, the research looks at utilizing multiple classical classification algorithms for learning and improving the prediction accuracy by integrating the learning approach in order to identify students' mental health conditions more effectively. Several classical and more adaptable algorithms including SVM, C4.5 decision tree, KNN, plain Bayes, CNN, and DeepPsy were used in this experiment to ensure the comprehensiveness and reliability of the experiment. The performance of each algorithm in the experiment is shown in Table 3. max-depth is the maximum depth of the tree, for after using the algorithm of this paper, the accuracy confusion matrix is obtained as shown in Table 4 in the above experiment by making max-depth=10.

The effective data of the experiment is 32858 people, the data set will be in accordance with the ratio of training set and test set 7:3 to complete the experiment, the test set contains effective data for 9857 people, of which the number of people labeled as abnormal mental health status is 151 people, according to the results of the experiment labeled with TP=9691,FN=15,FP=21,TN=130.According to the formula to calculate the correctness rate of the mental health detection is 99.78%, and the correct rate of psychological abnormality detection is 89.66%. The algorithm in this paper has achieved a significant improvement in accuracy, with an average increase of more than 20%. Particularly noteworthy is that the algorithm performs well in detecting students with psychological problems, with a correct rate of 89.66%.

Table 3. The algorithm is shown in the experiment

Algorithm	Accuracy/%
SVM	58.25
C4.5	63.21
KNN	67.98
Plain beth	60.56
CNN	71.88
DeepPsy	78.54
This article improves the method	89.66

Table 4. Accuracy confusion matrix

	FP	TP
TN	21	9691
FN	130	15

In order to evaluate the adaptability and robustness of the algorithm, further horizontal comparisons with the cases of max-depth=9, max-depth=11, max-depth=13, and max-depth=15 can be obtained in more detail, and the results of the comparison experiments are shown in Fig. 5. Numbers 1-7 represent SVM, C4.5 Decision Tree, KNN, Plain Bayes, CNN, DeepPsy and the methods in this paper, respectively. This helps this chapter to gain a deeper understanding of the impact of max-depth on the performance of the algorithm, thus providing a more scientific basis for choosing the appropriate max-depth. Such a comparative analysis helps this chapter to better understand the characteristics of the algorithms and provide more reliable guidance for practical applications.

It can be found in this chapter that the accuracy of the algorithm in this paper tends to decrease with the increase of max-depth. This may be due to the fact that the increase of max-depth causes the algorithm to become too flexible in dealing with mental health anomaly detection, which is affected by the algorithm that does not have a high accuracy rate, thus affecting its performance. The highest accuracy rate of 89.66% was found at max-depth=10, and the accuracy rate still remained relatively high at max-depth=9 & max-depth=11, but the algorithm's performance showed a significant decrease at max-depth=13 & max-depth=15. The experimental results suggest that choosing an appropriate max-depth is critical to the performance of this paper's algorithm in the mental health anomaly detection task. Further research and tuning may need to consider different combinations of max-depth and other parameters to find the optimal model configuration.

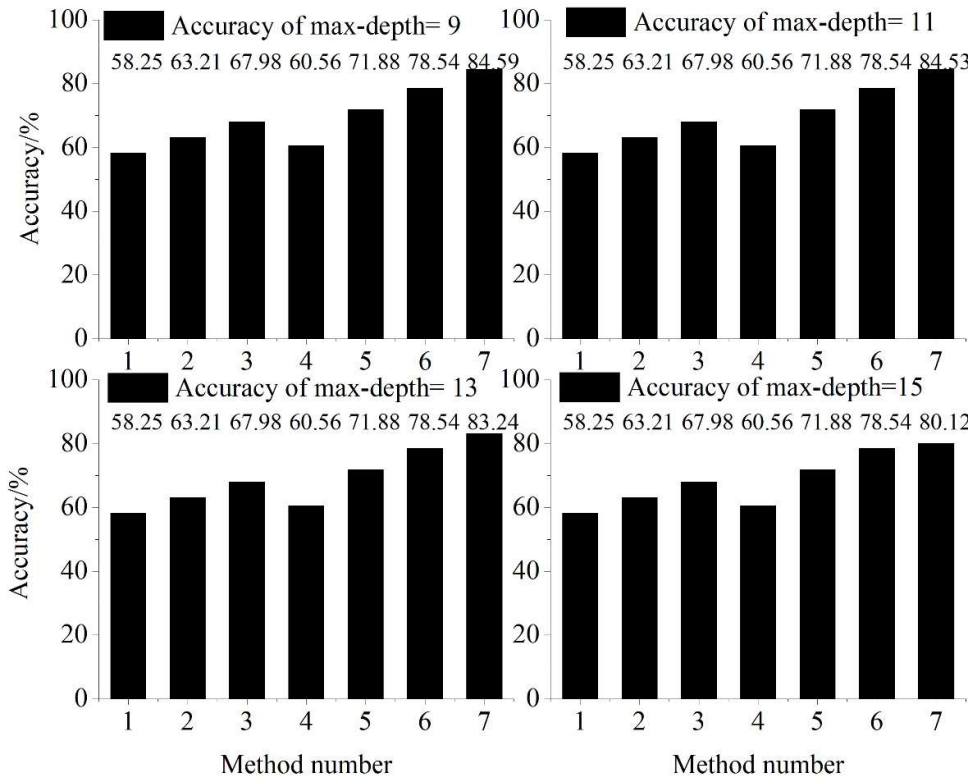

Fig. 5. The max-depth depth value accuracy comparison

5. Effectiveness of the application of early warning of students' mental health risks

Simulation is a computer simulation method performed before entering the real experiment. The operating system of the simulation machine used in this study was Windows 10 and the CPU model was 5-11400h. The simulation program was written in Python and the software package was Numpy. in order to carry out the simulation experiment, the study randomly set the personality space of the open personality versus the personality space of the neurotic personality. At the same time, the positive stimulus component and the negative stimulus component were randomly set, which in turn simulated the emotional changes of the open personality and the neurotic personality when they were subjected to multiple positive and negative stimuli, respectively. In this study, the intensity of the positive emotion is closer to 1, which means that it is stronger, and closer to 0, which means that it is weaker. Negative mood tends to be stronger as it approaches -1 and weaker as it approaches 0. Positive emotion intensity values above the positive emotion threshold indicate that the emotion is activated, and negative emotion intensity values below the negative emotion threshold indicate that the emotion is activated.

In real life, students are often faced with multiple stressful events, so this study conducted simulation experiments by randomly selecting two students with open personality and neurotic personality who passed the personality analysis questionnaire in the senior year of a university to simulate the process of emotional changes in different situations stimulated by the open personality and neurotic personality, respectively.

Figure 6 shows that the open personality was subjected to a positive stimulus at the moment t = 3, and t is the time to monitor the personality's mood change. This is followed by the process of mood change at moment t = 7 when the negative stimulus is again suffered. After suffering a positive stimulus at moment t = 3, the intensity of positive emotions first gradually increased and the intensity of negative emotions first gradually decreased. As time passes, the intensity of positive emotions

decreases, the intensity of negative "anger" and "disgust" emotions tends to decrease, and the intensity of "sadness" and "fear" emotions increases. Fear" emotions increased. At t = 7, the intensity of positive emotions gradually weakened and the intensity of negative emotions gradually strengthened, and broke the threshold of negative emotion activation and manifested itself. As time passes, the intensity of negative emotions gradually decreases, the intensity of positive emotions gradually increases, and the intensity of "pleasure" is gradually greater than the threshold of positive emotion activation.

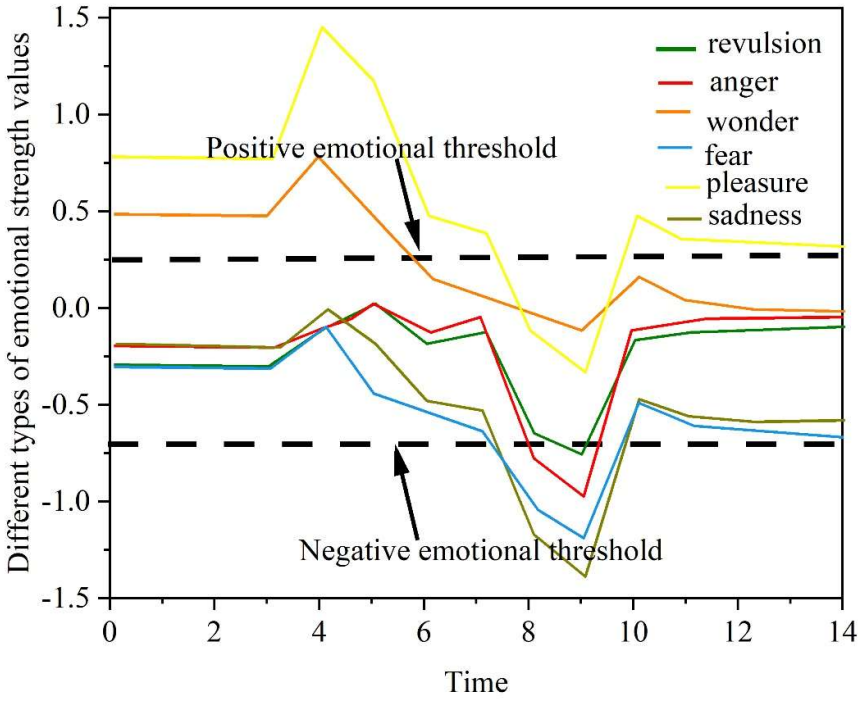

Fig. 6. The open personality is stimulated by multiple stress events

Figure 7 shows the process of emotional change in neurotic personality when subjected to positive stimuli at moment t = 3, followed by negative stimuli at moment t = 7. After being exposed to a positive stimulus at moment t = 3, the intensity of positive emotions gradually increased and the intensity of negative emotions gradually weakened. With the passage of time, the intensity of positive emotion gradually weakened to the point where it was no longer expressed, and the intensity of negative emotion tended to increase. After the negative stimulus at t = 7, the intensity of positive emotion gradually weakened, the intensity of negative emotion gradually increased, and broke through the threshold of negative emotion activation. Subsequently, the intensity of positive emotions increased, but remained below the activation threshold and was not activated again, while the intensity of negative emotions gradually weakened, but the "fear" emotion broke through the threshold again.

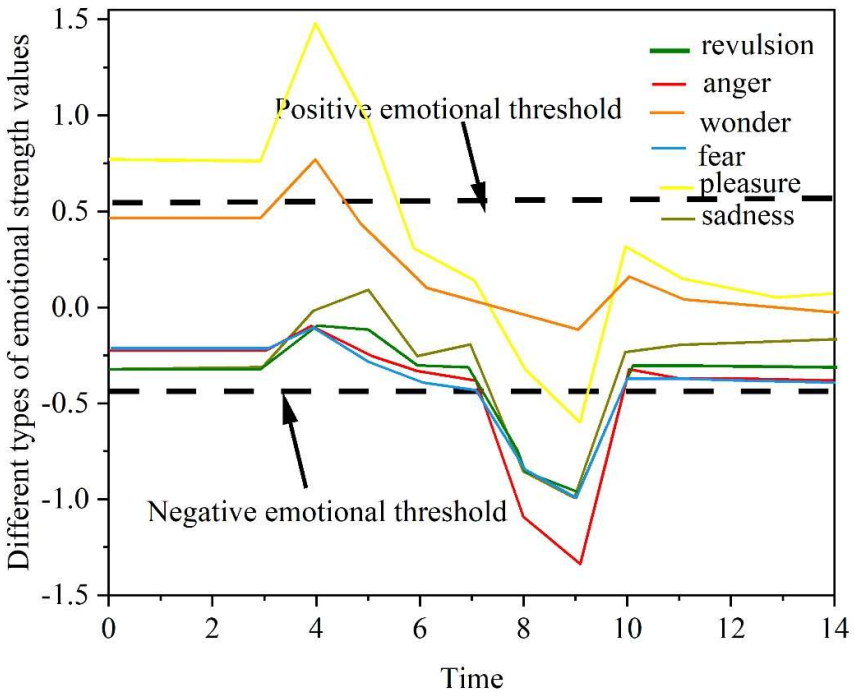

Fig. 7. Neurotic personality is stimulated by multiple stress events

6. Early warning strategies for students' mental health risks

Although the models and algorithms constructed in this study are based on certain psychological theories, psychological crisis involves psychology, medicine, sociology, philosophy and other disciplinary fields, and only by further integrating multidisciplinary fields can we better reveal the nature of psychological crisis. At the same time, the testing of models and algorithms requires long-term data tracking. The algorithm proposed in this study may face the following problems in real applications to be solved by subsequent research.

First, students may not often express their emotions through a certain social media, and when this situation is encountered, it is necessary to expand the collection of different social media to collect as much information about students as possible. Accurate psychological crisis warning is not realized by social media big data alone, but requires the integration of heterogeneous data from multiple sources.

At the same time, information stored in the school information management system, such as students' grades and family situations, is also a very important data source. How to integrate students' data in different social media and in different information systems, and synthesize multi-featured data for psychological crisis early warning is the focus of subsequent research that needs to be paid attention to.

Second, students' words describing stressful events in the text do not exist in the dictionary, which can lead to stressful events not being accurately recognized. To solve this kind of problem, we mainly rely on the function of recognizing unregistered words in the current lexical software. Take the mainstream “stuttering lexical” software as an example, the software identifies words that are not in the lexicon of the stressful event in the current utterance through the Hidden Markov Algorithm and automatically reminds the analysts to identify them manually, and then manually add new words to the lexicon if they are determined to be the new words describing the stressful event. If it is determined to be a new word describing a stressful event, it is manually added to the lexicon. With the continuous improvement of the lexicon, the accuracy of stressful event recognition will increase.

Third, the psychological crisis warning algorithm proposed in this study will inevitably involve students' personal accounts and other private information in practical applications, and how to protect students' private data is also a key concern for future research.

7. Conclusion

Taking students' mental health as the research background, this study systematically explores the algorithms based on students' mental health management and constructs a mental health risk early warning model based on CMA-ES-XGBoost algorithm.

The article focuses on multiple influencing factors such as peer relationship, teacher-student relationship, parent-child communication frequency, parent-child relationship, school physical environment, and study load with 3 common and important psychological problems of students, such as interpersonal tension sensitivity, depression, and anxiety, to provide more comprehensive information for the prediction of the mental health risk early warning model.

Subsequently, the performance and effectiveness of this paper's CMA-ES-XGBoost-based algorithm is demonstrated by comparing the accuracy of various classical algorithms, such as SVM, C4.5 Decision Tree, KNN, Plain Bayes, CNN, and DeepPsy. This paper's algorithm performs well in detecting students with mental emergence health risks, with 89.66% correct prediction rate. The optimal max-depth value of this paper's algorithm is determined to be 10 through cross-sectional experiments, which provides a strong support for the risk warning and intervention of students' mental health.

Through the model of this paper, we analyze the different types of emotional intensity changes of individual students with different types of personality when they are stimulated by multiple stressful events in the outside world, so as to judge their emotional performance. Based on the time series of students' emotional changes, the open personality was found to have weakened the intensity of positive emotions and strengthened the intensity of negative emotions after being exposed to negative stimuli, but there was no warning of negative emotions. The neurotic personality, on the other hand, after being exposed to negative stimuli, the intensity of positive emotions was weakened, the intensity of negative emotions was enhanced, and there was a negative emotion warning. It shows that the model in this paper realizes the effective early warning of students' psychological crisis.

References

[1] Sheldon, E., Simmonds-Buckley, M., Bone, C., Mascarenhas, T., Chan, N., Wincott, M., ... & Barkham, M. (2021). Prevalence and risk factors for mental health problems in university undergraduate students: A systematic review with meta-analysis. Journal of affective disorders, 287, 282-292.

[2] Han, K., Ji, L., Chen, C., Hou, B., Ren, D., Yuan, F., ... & He, G. (2022). College students' screening early warning factors in identification of suicide risk. Frontiers in genetics, 13, 977007.

[3] Zheng, H., Xu, Z., Dai, X., & Yang, Y. (2024). Research on Early Warning Indicators of College Students' Psychological Anxiety. International Journal of Trend in Scientific Research and Development, 8(6), 232-242.

[4] Wang, Y., Wen, J., Zhou, W., Tao, B., Wu, Q., Fu, C., & Li, H. (2023). An early warning method for abnormal behavior of college students based on multimodal fusion and improved decision tree. Journal of Intelligent & Fuzzy Systems, (Preprint), 1-23.

[5] Li, B. (2022). Design of early warning system for mental health problems based on data mining and database. Revista Brasileira de Medicina do Esporte, 29, e2022_0153.

[6] Yang, B. (2023). Model innovation of students' mental health education from the perspective of big data. Expert Systems, 40(4), e12948.

[7] Liu, H., & Jiao, N. (2020). Research on Students' Campus Behavior Analysis and Warning System Based on Big Data. In Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery: Volume 2 (pp. 392-400). Springer International Publishing.

[8] Sullivan, R. (2017). Early Warning Signs. A Solution-Finding Report. Center on Innovations in Learning, Temple University.

[9] Hong, Y. Z., Rani, M. N. A., Radzuan, N. F. M., Yen, L. H., & Nagalingam, S. A. R. A. S. V. A. T. H. I. (2024). An early warning system for students at risk using supervised machine learning. Journal of Engineering Science and Technology, 19(1), 131-139.

[10] Quan, M. (2024). Construction of Early Warning and Intervention Mechanisms for College Students' Mental Health Issues. Academic Journal of Humanities & Social Sciences, 7(5), 224-229.

[11] Lu, H., & Mingjian, W. (2023, August). Early Warning Mechanism of College Students' Psychological Crisis Based on Clustering Extraction Algorithm. In EAI International Conference, BigIoT-EDU (pp. 297-304). Cham: Springer Nature Switzerland.

[12] Zhang, Z. (2024). Early warning model of adolescent mental health based on big data and machine learning. Soft Computing, 28(1), 811-828.

[13] Li, X. (2022). Early Warning Method of College Students Mental Subhealth Based on Internet of Things. In IoT and Big Data Technologies for Health Care (pp. 49-60). Cham: Springer Nature Switzerland.

[14] Chen, X., Xu, L., & Pan, Z. (2022, February). Design of depression assessment and early warning system for college students based on machine learning and daily health data. In Thirteenth International Conference on Graphics and Image Processing (ICGIP 2021) (Vol. 12083, pp. 55-59). SPIE.

[15] Du, H. (2023). Role of the Psychological Test and Evaluation System Based on the Internet of Things in the Early Warning of Psychological Dangers of College Students. International Journal of Information and Communication Technology Education (IJICTE), 19(1), 1-19.

[16] Feng, Y., Li, J., Liu, T., Wei, Y., & Li, N. (2024). Construction of a mental health risk model for college students with long and short-term memory networks and early warning indicators. Journal of Intelligent Systems, 33(1), 20230318.

[17] Liu, Z., Zhao, X., Liu, C., Li, Y., & Yin, W. (2024, May). Early Warning Algorithm of College Students' Mental Health Based on Distributed Multi-Agent System. In 2024 International Conference on Telecommunications and Power Electronics (TELEPE) (pp. 310-315). IEEE.

[18] Ye, L. (2024, September). College student mental health early warning algorithm based on distributed multi-agent system. In AIP Conference Proceedings (Vol. 3131, No. 1). AIP Publishing.

[19] Chen, Y., & Ke, J. (2024). Multivariate Decision Tree-Oriented Early Warning Method for College Students' Psychological Crisis Behavior. International Journal of High Speed Electronics and Systems, 2440120.

[20] Xiong, W. (2020). Identification and early warning of college students' psychological crisis based on big data. In Data processing techniques and applications for cyber-physical systems (DPTA 2019) (pp. 23-28). Springer Singapore.

[21] Wang, J., Zhang, Z., Luo, H., Liu, Y., Chen, W., & Wei, G. (2019). Research on early warning model of college students' psychological crisis based on genetic BP neural network. American journal of applied psychology, 8(6), 112-120.

[22] Tiantian Li,Fang Liu,Xiaobin Chen & Chao Ma. (2024). Publisher Correction: Web log mining techniques to optimize Apriori association rule algorithm in sports data information management. Scientific Reports(1),30568-30568.

[23] Geetika Dhand,Meena Rao,Parul Chaudhary & Kavita Sheoran. (2025). A secure routing and malicious

node detection in mobile Ad hoc network using trust value evaluation with improved XGBoost mechanism. Journal of Network and Computer Applications104093-104093.

[24] Seyfullah Berk & Ertan Alptekin. (2024). Optimization of NOX emissions of a CRDI DIESEL engine using CMA-ES method. International Journal of Engine Research(12),2184-2203.