J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 6747--6768

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

Combinatorial Press

Optimization of Identification and Intervention Path of Psychological Problems of Secondary School Students in Cultural Education Based on Dijkstra's Algorithm

Lili Liu[1,✉]

[1] *School of Marxism, North University of China, Taiyuan, Shanxi, 030051, China*

ABSTRACT
The recent frequent occurrence of students' psychological crisis events has drawn widespread attention to mental health education in colleges and universities. Based on students' behavioral data, we use big data and data mining technology to model and analyze students' daily behaviors, complete the construction of students' social intimacy features based on Dijkstra's algorithm, use the C4.5 decision tree improvement algorithm based on variable-precision rough set to realize the identification of students' psychological problems, and analyze the intervention paths of students' psychological problems and the evaluation of the results of the intervention. The proposed method can recognize students' psychological problems more accurately, and the recognition accuracy of different levels of psychological problems reaches more than 72%, which is significantly higher than other classification methods. Learning anxiety, loneliness tendency and terror tendency of students in the intervention group were significantly reduced after the psychological intervention ($P < 0.05$), and the overall factor scores decreased by 9.85%, and the level of mental health was answered to be improved, which reflected the effectiveness of the proposed mental health intervention. The experiment proves that the model in this paper can effectively identify students with psychological abnormalities, and the proposed intervention path for students' psychological problems has a positive impact on the development of students' mental health.

Keywords: Dijkstra's algorithm; C4.5 decision tree; variable precision rough set; student psychological problem identification; psychological intervention

1. Introduction

✉ Corresponding author.
E-mail address: 13593146365@163.com (L. Liu).
Received 05 March 2024; Revised 25 May 2024; Accepted 05 December 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-369

With the rapid development of athletics and education, the quantity and quality of student training are steadily increasing, and the state and society are paying more and more attention to higher education. However, excessive tension, anxiety, and depression caused by academic pressure, social pressure, and campus violence, loneliness and confusion caused by family environment, interpersonal discomfort, and illness, and psychological disorders such as compulsion caused by Internet addiction have become common psychological problems among students nowadays [1-4]. Frequent dropouts and suicides in recent years have proved that many students have more or less psychological problems, and these psychological problems have aroused widespread concern in society. Globally, more than 50% of college students currently have mental health problems. According to statistics, among them, 17-69% of college students suffer from different levels of depression, 15-61% of college students suffer from anxiety disorders, more than 10% of college students suffer from different levels of pessimism, and more than 10% of college students harbor a certain degree of hatred towards society and even towards their families [5-8]. In addition, some scholars have pointed out that for students aged 18-28 years old, their mental health level is mostly low [9]. Mental health problems are seriously harmful. For individuals, mental health problems can cause students to have some negative effects, which can make them less adaptable in society, reduce their physiological health, and may even produce serious threats to their lives [10-11]. For families, mental health problems can cause psychological burdens on family members, as well as increase the economic burden of the family. For society, mental health problems can disrupt public order, deplete social resources, and exacerbate society's disease burden, with the negative impact of mental health problems accounting for 37% of all disease-related harms [12-13]. However, due to the lack of mental health knowledge and awareness, discriminatory attitudes toward mental patients and the limited economic conditions of students and many other reasons, a large number of students with mental health problems do not have the motivation to seek professional help, delaying the optimal time for treatment, the psychological problems are becoming more and more serious, and the risk of individual harm and social harm is exacerbated [14-17]. At all stages of life, people may have psychological problems for a variety of complex reasons, and early identification of these psychological problems can have a huge role in protecting personal safety and other aspects.

Usually the identification of psychological problems relies heavily on psychological questionnaires, but in questionnaires of varying quality, people tend to reflect a better version of themselves, which is usually not a good representation of a person's true inner thoughts, a strong sense of subjectivity, and the intensity of quantification of certain issues varies from person to person, and these limitations lead to the questionnaire data do not portray the subject well, especially serious when the amount of data is small [18-21]. In addition, each student's mental health status changes over time, usually lasting half a month or a month, and if one wants to know what the mental health status is after that, one needs to be tested again, which is labor-intensive and costly. Educators and therapists are beginning to realize that it is difficult to accurately identify students' psychological problems using this noisy questionnaire data. The various social networks in students' daily life can reflect students' social status most realistically, and many psychological problems are closely related to the social status, and students' daily performance and body physiological data can reflect students' psychological status to a certain extent, so in recent years the use of social information and other multi-source data to analyze people's psychological status is also more and more, and has achieved better results [22-25].

Thanks to the rapid development of computer science and mathematics and other subject areas, deep learning, data mining, big data and other technologies are also making rapid progress and becoming more and more deeply integrated into people's daily lives [26]. More and more algorithms have emerged in the field of computer science to solve certain specific problems, which are superior to traditional methods in many aspects. For psychological problem identification, He et al [27] reported to the effectiveness of the critical value method and the Patient Health Questionnaire-9 diagnostic algorithm in the diagnosis of major depression, with the former having better diagnostic results than the latter. Llano et al [28] compared the performances of support vector machine, logistic regression, random forest, Bayesian network and k-nearest neighbor algorithms in identifying psychological disorders, and logistic regression performed the best (with an accuracy rate of about 70%) and used it to construct an automatic diagnosis method. Vasha et al [29] analyzed the data text of the comment section in social media through data mining and machine learning, which can clearly identify the distinction between depressive manifestations and normal manifestations, with the best accuracy in support vector machines. Aljarah et al [30] mentioned that the use of machine learning and real-time data analysis can improve the efficiency of depression recognition and the correctness, and personalized interventions can also be proposed.Ding et al [31] used microblogging behavioral activity data as the basis for identifying depression in college students by capturing data features and classifying them using deep neural networks and deeply integrated support vector machine algorithms, respectively. Almars [32] used the attention mechanism and bi-directional long and short-term memory networks to weight the vocabulary of the user-posted text in a social network computation to explore the text hidden features so as to identify the depression. Tian [33] fused ECG, sequential backward selection algorithm, and convolutional neural network to collect information, downsize indicator features, and construct recognition model for college students' psychological problems, respectively, so as to improve the accuracy and feasibility of the recognition of psychological problems. For psychological interventions, Li and Shi [34] introduced to the effective mitigating effect of sports intervention methods on psychological crisis, but individualized design is needed, and it is difficult to be effective with set sports interventions. Oliveira et al [35] confirmed the effectiveness of cell phone apps to assist other psychological interventions in the treatment of psychological problems, and this type of mobile interference methods can be a campus psychological service center Räsänen et al [36] successfully alleviated student stress and depressive symptoms through an instructor-guided Acceptance and Commitment Therapy intervention by varying the Positive Thoughts Non-Response subscale. To summarize the above studies, it is evident that the identification of psychological problems is mostly on the consideration of depression, with less identification of other emotions, while the identification of the single data considered, resulting in the identification of the accuracy needs to be improved. In the psychological intervention, a few intervention methods have been proposed, but the specific intervention path has not been studied. Dijkstra's algorithm (Dijkstra) is the shortest path algorithm from one vertex to the rest of the vertices, which solves the shortest path problem in the entitled graph [37]. It has the potential to be applied in the exploration of psychological problem identification and intervention paths.

In this paper, the relevant features for students' psychological problem identification are selected from the three levels of individual, family and school, by modeling the relationship distance of students' social network, using Dijkstra's algorithm to obtain the length of the distance to reflect the students' social intimacy, and analyzing the quantitative processing methods of other students' features. Aiming at the problems of complex construction and low classification accuracy of C4.5 decision tree, it is improved by combining with variable precision rough set, and the improved C4.5 decision tree algorithm is used as a method of recognizing and classifying students' psychological problems. Taking the data related to students' mental health of CQ University as the dataset, the comparison experiments of different classification algorithms are conducted, and the model's effect of recognizing psychological problems is checked under different mental health levels and different students' majors. On this basis, the intervention path of students' psychological problems is explored. Finally, the intervention group and the control group were selected to carry out psychological intervention experiments, comparing the changes of the eight mental health test factors before and after the intervention experiments, and evaluating the effect of students' psychological interventions.

2. Dijkstra-based identification of students' psychological problems

This chapter builds a comprehensive model for identifying students' psychological problems with the help of data mining methods, which aims to assess students' psychological status more comprehensively and accurately, and provides a theoretical basis and practical support for scientific research and practical intervention of students' mental health problems.

2.1. Feature selection

Hierarchical analysis was used to construct an index system for the identification of psychological problems among college students, which covered three core indicators at the feature level and was subdivided into nine specific assessment indicators at the base level, as follows:

Personal factor level: physical condition, interpersonal communication (social intimacy), personality orientation.

At the family level: family economic situation, upbringing, and nature of household registration.

At the school level: academic performance, diligence in school, participation in group activities.

2.2. Social intimacy

The algorithm is based on students' behavioral record data at school, and extracts all students' behavioral records for N consecutive months for behavioral trajectory data mining. The number of co-occurrences of any two students is counted and transformed into the student relationship index; the more common actions/co-occurrences, the more similar the similarity/relationship between students. According to the students' relationship index, a personal social graph is constructed for each student, and the social network of the whole school is formed by the combination of social graphs, with student objects as nodes in the network and relationship index as distances in the network.

2.2.1. Parameter Variable Descriptions and Definitions:

1) Co-occurrence. Co-occurrence, which specifically refers to the simultaneous existence of behavioral records of students S_A and S_n at a location L_i within a given time window $[T-t, T+t]$, has an association rule $S_A \rightarrow S_B$, which denotes that there exists one co-occurrence of students S_A and S_B at location L_i have one co-occurrence.

2) Time Window. The time window is used to determine whether a student's behavioral record can be classified as a “co-occurrence” in the range of $|T_A - T_B| \leq 2 * t_{L,h}$. Where T_A is the time of a behavioral record of student A, T_B is the time of a behavioral record of student B, $2 * t_{l,h}$ is the size of the interval of the judgmental condition, and the parameter $t_{L,h}$ is the threshold value of the location L_i in the hourly segment of $h \in [1, 24]$.

3) Relationship index. Set the closeness of any two students as $R_A(B)$, i.e., for student A, the closeness of student B to him/her is $R_A(B)$. The formula is:

$$R_A(B) = \sum_{i \in L} \left[\frac{O_A^i(B)}{C_A(i)} \times \frac{|S|}{S_A(i)} \right] \tag{1}$$

Where: L is the location where all student behaviors were recorded, $C_A(i)$ denotes the total number of times student A was recorded at location L_i in a given statistical time period, $O_A^i(B)$ denotes the number of times student A and student B were co-occurring at location L_i in a given statistical time period, and $|S|$ denotes the total number of students in the school, and $S_A(i)$ denotes the total number of students who co-occur with student A at location L_i.

4) Similarity Characterization. For a given any two students i and student j, the similarity between the two is calculated as:

$$\tau_{ij} = \sum_{l=1}^{L} \frac{\tau_{i,j}^{l}}{\max\limits_{j} \tau_{i,j}^{l}} \tag{2}$$

Where: $\tau_{i,j}^{l}$ denotes the number of co-occurrences of two students at location l, and $\max\limits_{j} \tau_{i,j}^{l}$ denotes the maximum number of times that student i co-occurs with other students at location l.

2.2.2. Calculating time window parameters. For any location L_i, there exists the parameter set $T_i = [t_1, t_2, \cdots, t_{24}]$. To determine the appropriate $t_{L,h}$, set the parameters LB_i, UB_i, and p_i to denote the lower, upper, and growth steps of $t_{L,h}$, respectively. For an h-hourly segment at location L_i, the alternative $t_{l,h}$ has the candidate set $TS = \left[LB_i, LB_i + p_i, LB_i + 2p_i, \cdots, UB_i\right]$. In seconds, then for each choice TS_i, the hourly segment can be divided into $\frac{3600}{TS_i}$ intervals, then for location l at h hours of time period d, there is:

$$W_{l,d,h,TS_i} = \left[n_{1,l,h}, n_{2,l,h}, \cdots, n_{\frac{3600}{TS_i},l,h} \right] \tag{3}$$

where n is the number of people whose behavior was recorded in the corresponding hourly period of the time period, i.e:

$$n_{i,l,h} = \sum_{k=1}^{d} o_{i,l,h,k} \tag{4}$$

where $o_{i,l,h,k}$ denotes the number of co-occurrences in the ith interval of the h-hourly segment at location i on the kth day of the time period.

The probabilities of the Poisson distribution are computed for the set W based on each interval of each location, hourly segment, to obtain the probabilities of $Possion\left(X = 2\right)$ and $Possion(X \geq 2)$.

2.2.3. Constructing a student co-occurrence matrix. For student m, the matrix is obtained by calculating the number of co-occurrences with other students:

$$S_m = \begin{bmatrix} O_{L_1,1m} & \cdots & O_{L_1,sm} \\ \vdots & \ddots & \vdots \\ O_{L_n,1m} & \cdots & O_{L_n,sm} \end{bmatrix} \tag{5}$$

where $O_{L_n,sm}$ denotes the number of co-occurrences of student m and student s at location L_n.

2.2.4. Calculating the Student Relationship Index Matrix. For student m, the matrix can be obtained by the proximity formula:

$$R_m = \left[R_m(S_1) \cdots R_m(S_x)\right] \tag{6}$$

$$R = \begin{bmatrix} R_1(S_1) & \cdots & R_1(S_{sx}) \\ \vdots & \ddots & \vdots \\ R_x(S_1) & \cdots & R_x(S_x) \end{bmatrix} \tag{7}$$

2.2.5. Dijkstra-based shortest distance. Based on the actual reality consideration, the relationship index matrix of all students in the school is likely to be a sparse matrix, in order to better explore the relationship between any two students, the social relationship of the students is constructed as a network, with the students as the nodes of the network, and the inverse of the relationship index is the distance between the nodes, and through the shortest path Dijkstra's algorithm, to obtain the social path and the shortest relationship distance between any two students.

Considering the intimacy formula $R_A(B)$ as a positive index, in order to support the shortest path algorithm, its inverse is used as the relationship distance to obtain the matrix:

$$D = \begin{bmatrix} \frac{1}{R_1(S_1)} & \cdots & \frac{1}{R_1(S_x)} \\ \vdots & \ddots & \vdots \\ \frac{1}{R_x(S_1)} & \cdots & \frac{1}{R_x(S_x)} \end{bmatrix} \tag{8}$$

After computed by the shortest path algorithm, the distance matrix D is updated to get the distance matrix D', where v_{mx} denotes the shortest relational distance between student m and student s and the path matrix P, and P_{mx} is the social path (nodes passed in the social network) from student m to student s. According to the closeness formula $R_A(B)$, it is known that the relationship distance is directed, so $P_{mx} \neq P_{xm}$:

$$D' = \begin{bmatrix} v_{11} & \cdots & v_{1s} \\ \vdots & \ddots & \vdots \\ v_{s1} & \cdots & v_{ss} \end{bmatrix} \tag{9}$$

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1s} \\ \vdots & \ddots & \vdots \\ P_{s1} & \cdots & P_{ss} \end{bmatrix} \tag{10}$$

$$P_{ms} = [S_m \cdots S_s] \tag{11}$$

Thus, for any student in the social network, there exists $N_s = [D'_s, P_s]$ denoting the student's social relationship profile.

2.2.6. Generating student personal social graphs. Access control gates, cafeteria consumption card and other equipment flow data can be used to reflect the social behavior of students, using the graph $G(V,E)$ to represent the social network, where V represents the set of nodes, E represents the set of connecting edges, the node represents the student, the connecting edges represent the social relationship between the students. The existence of a connecting edge between any two nodes needs to satisfy three conditions at the same time: ① being captured by the same device (swiping behavior or being recognized by the same face recognition camera), ② being captured by the same device at intervals less than 2 minutes, and ③ the number of times that conditions ① and ② are satisfied at the same time during the observation period is greater than a threshold value T.

The final social graph of a student is obtained by filtering the relationship distance D'_s in the student's social set N_s by setting a distance threshold H and obtaining objects with $D'_s < H$.

2.3. Other personal factors

2.3.1. Physical condition. Specific data describing the physical condition of students were selected from school hospital visits, annual physical education test scores, and Exercise APP Trail Luck Running punch card data. The physical condition of students was compared by calculating the position of each normal distribution, i.e., the standardized score, of a student in each of the three types of data as follows:

For the frequency of medical consultation p_A there is a total population of students $P = [p_1, p_2, \cdots, p_{n-1}, p_n]$, calculate the mean and variance or standard deviation of this distribution respectively, and then use the probability density function to calculate its Z-Score. Its probability density is as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma_0^2} \right\} \tag{12}$$

For frequency of medical visits, the mean μ and standard deviation S were calculated:

$$\bar{X}_{medical} = \frac{\sum_{i=1}^{n} p_i}{n} \tag{13}$$

$\bar{X}_{medical}$ denotes the sample size of the number of medical visits, and n is the number of such samples, as in the following equation:

$$S_{medical} = \sqrt{\frac{\sum (p_i - \bar{X}_{medical})^2}{n-1}} \tag{14}$$

where $S_{medical}$ denotes the standard deviation of this sample, which is calculated as follows:

$$Z_{medical} = \frac{X_{medical} - \bar{X}_{medical}}{S_{medical}} \tag{15}$$

Here, $Z_{medical}$ denotes the standardized score and $X_{medical}$ is the frequency of the student's medical visits p_A.

The same steps can be applied to sports test scores and sports app clocking data.

Finally, the average of the Z-scores for the three factors is calculated as follows:

$$Average\ Z-score = \frac{Z_{medical} + Z_{test} + Z_{app}}{3} \tag{16}$$

2.3.2. Personality orientation. After a complex psychometric evaluation of the questionnaire, data on five dimensions of personality orientation were obtained, corresponding to extroversion, introversion, emotional stability, responsibility, and openness. A clustering algorithm was used to categorize the students into different clusters based on the similarity of the personality dimensions. This helps to identify groups of students sharing similar personality traits.

2.4. Family factors

2.4.1. Family economic situation. Adopting the tier of students' identification of their family income to make data support for mental health prediction and judgment.

1) Collect family income data. Obtain income data from a large number of families to ensure the accuracy and comprehensiveness of the data. After data preprocessing, put the total annual household income into the set X and arrange the data set X in ascending order.

2) Calculate the median. Sort the total annual household income and identify the household income in the middle position.

If the size of X is odd (n is odd), the median M is the sorted middle value, i.e.:

$$M = X\left[\frac{n+1}{2}\right] \tag{17}$$

If the size of X is even (n is even), the median M is the average of the two values in the middle of the sorted order, ie:

$$M = \frac{X\left[\frac{n}{2}\right] + X\left[\frac{n}{2}+1\right]}{2} \tag{18}$$

3) Setting relative standards. A range of 76%-200% relative to the median is set for defining the middle-income group.

4) Classification of household income groups. Based on the relative criteria set, household income is classified into three income groups: high, middle and low.

2.4.2. History of growth and nature of households. Four levels of data, namely, family structure, whether they are left-behind children, arts and sciences, and the level of the city to which their high school belongs, were selected to construct a multidimensional data table portraying students' growth history.

The nature of hukou may have a potential impact on students' mental health by influencing the socioeconomic environment, cultural environment, and educational opportunities in which they live. The nature of hukou is categorized into rural hukou and non-rural hukou.

2.5. School factors

2.5.1. Academic performance. Students' mental health was predicted by analyzing the percentage of failed courses, individual course grades, course credits, and semester GPAs. In order to combine the four characteristics of failing a course, course score, course credits and semester credits GPA per semester, a composite scoring model was designed to highlight the ability to predict students' mental health. The following is the composite scoring model:

$$Score = 25 * \left[(1 - fail) + s * 0.01 + \frac{total\_credit}{\max\_credit} + \frac{GPA}{FULL} \right] \tag{19}$$

where 25 is a weighting factor to adjust the contribution of each characteristic to the composite score. fail denotes the percentage of failed courses, $fail = \text{Total number of failed subjects/total number of subjects}$, s denotes the average course score, $total\_credit$ denotes the total number of credits for the student, max_credit denotes the maximum total number of credits for the cohort of students in the same grade level, GPA denotes the student's GPA, and FULL denotes the number of full grades earned at the school.

2.5.2. Attendance at school. Design a comprehensive diligence assessment model based on learning behavioral characteristics to quantitatively analyze students' performance in academics and life using the principle of normal distribution. The model indicators include the frequency score of surfing the Internet, the class attendance score, the breakfast behavior score and the number of library entries score.

Attendance is calculated:

i. Frequency score of Internet access:

$$\begin{aligned} F_{\text{Frequency of access to the Internet}} &= ZScore(\text{Frequency of Internet behavior}) \\ &\times MaxScore \end{aligned} \tag{20}$$

ii. Class attendance ratings:

$$\begin{aligned} F_{\text{Attendance}} &= ZScor(\text{Number of times students attended classes}) \\ &\times MaxScore \end{aligned} \tag{21}$$

iii. Breakfast behavior scores:

$$\begin{aligned} F_{\text{Breakfast behavior}} &= ZScore(\text{Number of occurrences of breakfast behavior}) \\ &\times MaxScore \end{aligned} \tag{22}$$

iv. Library access scores:

$$F_{\text{Library accesses}} = ZScore(\text{Library accesses}) \times MaxScore \tag{23}$$

v. Comprehensive Diligence Score:

$$Q_{\text{Industriousness}} = \omega_1 * F_{\text{Frequency of Internet access}} + \omega_2 * F_{\text{Attendance}} + \omega_3 * F_{\text{Breakfast behavior}} + \omega_4 * F_{\text{Number of librarian entries}} \tag{24}$$

Where: $ZScore(\cdot)$ is the function for calculating the Z-score. $F_{\text{Learning frequency}}$, $F_{\text{Attendance}}$, $F_{\text{Breakfast behavior}}$, F_{Library} are the students' learning frequency, class attendance, breakfast behavior, and library entry scores, respectively. maxScore is the maximum value of the score and is usually set to 100. $\omega_1, \omega_2, \omega_3, \omega_4$ are weighting coefficients to adjust the contribution of each feature to the composite diligence score. The model was fitted by setting $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 0.25$.

2.5.3. Participation in group activities. An assessment model was constructed to quantify students' activeness in group activities as an important predictor of mental health. In order to more closely capture the associations between collective activity characteristics, covariance matrix and multivariate normal distribution were introduced to more fully consider the potential associations between characteristic items.

Collective activity score $= P$ (collective activity frequency, club organization participation rate, volunteer service hours, sports activity participation frequency), by introducing multivariate normal distribution, the model can better capture the associations between features. To wit:

$$f(x;\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{k}{2}} \cdot |\Sigma|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}(x-\mu)^T \cdot \Sigma^{-1} \cdot (x-\mu)\right) \tag{25}$$

where x is the eigenvector, μ is the mean vector, and Σ is the covariance matrix $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_k^2 \end{bmatrix}$. This yields the characteristic term probability modeling:

$$P(x) = f(x;\mu,\Sigma) \tag{26}$$

x is a vector of eigenterms, which includes group activity frequency, club organization participation, volunteer hours and sports activity participation frequency.

2.6. Methods of Identifying Psychological Problems

The research objective of this paper is a binary classification problem, and the C4.5 decision tree algorithm is chosen as a method for recognizing students' psychological problems. Aiming at the problems that the decision tree generated by this method does not have high classification accuracy, has more branches, is larger in size, and the classification accuracy decreases rapidly in the presence of noise, an improved algorithm for C4.5 decision tree based on Variable Precision Rough Set (VPRSC4.5) is proposed.

2.6.1. Variable Precision Rough Sets. Variable Precision Rough Set (VPRS), which introduces a threshold on the basis of the rough set model to allow a certain degree of classification error to exist, thus effectively solving the problem of data inconsistency and the classification of noisy data, and improving the classification accuracy and generalization ability of the decision tree.

$S=(U,A,V,f)$ is an information system where U is a non-empty finite set, i.e., an argument domain, A is a non-empty finite set of attributes, V is a set of value domains of attributes, and $f:u\times A\to V$ is an informational function that assigns an informational value to each attribute of each object, i.e., $\forall a\in A$, $x\in U$, $f(x,a)\in V_a$.

$\beta-$Majority inclusion relations. Given an information system $S=(U,A,V,f)$, for each subset $X\subseteq U$ and an equivalence relation R, $U/R=\{Y_1,Y_2,\cdots,Y_n\}$, the $\beta-$ majority inclusion relation is defined as follows:

The set of approximations on $\beta-$ of X is:

$$\bar{R}_\beta(X)=U\{Y_i\mid Y_i\in U/R:P(Y_i,X)<1-\beta\} \tag{27}$$

The set of approximations under $\beta-$ of X is defined as:

$$\underline{R}_\beta(X)=U\{Y_i\mid Y_i\in U/R:P(Y_i,X)\le\beta\} \tag{28}$$

where: $P(Y_i,X)=\frac{|Y_i\cap X|}{|Y_i|}$ is the rate of correct categorization of the set Y_i into X, and $\underline{R}_\beta(X)$, also known as the $\beta-$ positive domain, is denoted $POS_\beta(X,Y)$, is the maximal set of objects that are judged to belong to X with a correct rate no greater than β based on prior knowledge.

For a given threshold β, an argument domain U, an equivalence class X of conditional attributes, and an equivalence class Y of decision attributes, the approximate classification quality is defined as:

$$\gamma_\beta(X,Y)=\frac{|POS_\beta(X,Y)|}{|U|} \tag{29}$$

It can be seen that when $\beta=1$, VPRS is the classical rough set model. Under the same conditions, as β decreases, the positive domain of $\beta-$ gradually increases, and the approximate classification quality increases, thus providing some suppression of noise.

2.6.2. Information gain rate under VPRS. For a random variable X defined on a given domain U, let its probability distribution be $\{p_1,p_2,\cdots,p_n\}$ and $\sum_{i=1}^{n}p_i=1$, then the information entropy of X is $H(X)=-\sum_{i=1}^{n}P_i\log_2(p_i)$. In information theory, entropy can be used to measure the uncertainty of an information system.

In the domain U, let the divisions derived by P and Q on U be X and Y, $X=\{X_1,X_2,\cdots,X_n\}$, $Y=\{Y_1,Y_2,\cdots,Y_m\}$, then the probability distribution of P and Q in U is a subset composition is:

$$[X:P]=\begin{bmatrix} X_1X_2\cdots X_n \\ p(X_1)p(X_2)\cdots p(X_n) \end{bmatrix} \tag{30}$$

$$[Y:Q]=\begin{bmatrix} Y_1Y_2\cdots Y_n \\ p(Y_1)p(Y_2)\cdots p(Y_n) \end{bmatrix} \tag{31}$$

where: $p(X_i)=\frac{|X_i|}{|U|}(i=1,2,\cdots,n)$, $p(Y_j)=\frac{|X_j|}{|U|}(j=1,2,\cdots,m)$, with the symbol $|M|$ being the base of the set M, i.e., the number of elements contained in the set. The $\beta-$entropy $H^{\beta}(p)$ of the knowledge p is:

$$H^{\beta}(p)=-\sum_{i=1}^{n}\beta p\left(\frac{|X_i|}{|U|}\right)\log_2\left(p\left(\frac{|X_i|}{|U|}\right)\right) \tag{32}$$

Then the knowledge $Q\left(U\mid IND(Q)=\{Y_1,Y_2,\cdots,Y_m\}\right)$ is relative to the knowledge $P\left(U\mid IND(P)=\{X_1,X_2,\cdots,X_m\}\right)$ can be defined as:

$$H^{\beta}(Q\mid P)=-\sum_{i=1}^{n}P(X_i)\sum_{j=1}^{m}\beta p\left(Y_j\mid X_i\right)\log_2\left(p\left(Y_j\mid X_i\right)\right) \tag{33}$$

where: $p\left(Y_j\mid X_i\right)=\left|Y_j\cap X_i\right|/\left|X_i\right|(i=1,2,\cdots,n;j=1,2,\cdots,m)$.

Therefore, the information gain rate $\beta-ratio$ under variable precision rough sets is:

$$Ratio^{\beta}(C,U)=\left[\inf(X)-\sum_{i=1}^{n}\beta\frac{|X_i|}{|U|}\times\inf(X)\right]\Big/Split(C,U) \tag{34}$$

where: $Split(C,U)=-\sum_{i=1}^{n}\beta\frac{|X_i|}{|U|}\times\log_2\left(\frac{|X_i|}{|U|}\right)$, and β is the categorization threshold, generally specified by human experience, here $0.5\le\beta<1$.

2.6.3. VPRSC4.5 algorithm. The process of constructing a decision tree is mainly the selection of nodes. In the process of constructing the decision tree improvement algorithm using the variable precision rough set method, the measure of the approximate classification quality of each attribute $\gamma_{\beta}(X,Y)$ is computed according to Eq. (29), and the attribute with the largest classification quality measure is selected as the root node. By analogy, the middle node of the tree is the attribute that has the largest approximate classification quality measure for the samples contained in the subtree rooted at that node, until the samples contained in the branches cannot be further divided, generating leaf nodes. When two or more attributes have the same approximate classification quality, the $\beta-$information gain rate $(\beta-ratio)$ for this attribute with variable precision is calculated, and the attribute with the largest information gain rate is selected as the root node.

The algorithmic process is as follows:

a) Generate a node root based on the training sample set T.

b) If C is empty, label node root with the decision attribute value that accounts for the majority of the training instance set T and return. (c) If U takes exactly the same value on the decision attribute, the decision attribute value is used to label node root and return.

c) For each equivalence class $X_i = U / C_i (i = 1, 2, \cdots, n)$ of the conditional attribute C_i, compute its approximate categorization quality $\gamma_i (i = 1, 2, \cdots, n)$ according to equation (29).

d) Take the maximum value of γ_i, $\max \gamma_i (i = 1, 2, \cdots, n)$.

if $\max \gamma_i$ only then

Assign its corresponding attribute C_i to node root

else

e) Calculate the information gain rate $\beta-ratio(C_i, T)$ for all conditional attributes that have reached the maximum value according to equation (34), and assign the value of the attribute C_i corresponding to the maximum information gain rate to node root.

f) Add a new branch $root = B_j$ under root for each possible value of root, such that $T - B_j$ denotes the subset of T that satisfies the attribute c taking the value B_j.

if $T - B_j$ is empty then

return the value of the most decision-making attribute in T.

else

Add a subtree $VPRSC4.5(T - B_j, C - c, D, f)$ under this new branch.

g) If U is completely categorized by the decision tree branch, the algorithm ends.

2.7. *Experiments and analysis of results*

2.7.1. Experimental data. This paper investigates 300 undergraduate students at CQ University, of which 100 students were assessed as students with mental health problems by experts at the university's Psychological Center, and their mental health levels were classified as mild, moderate, and severe. In addition, 200 undergraduate students were randomly selected from the whole university. In this section of the experiment, the students were categorized into two groups; mild, moderate and severe were positive samples, and students without mental health problems were negative samples.

When designing machine learning experiments, the division of the dataset is particularly important, and the reasonable division of the training set and test set helps to train and validate the model. In this experiment, the data set sample is 300 students, of which 100 students are positive samples and 200 students are negative samples, that is to say, the ratio of positive samples and negative samples is 1:2. In order to make the ratio of the samples in the training set and the test set are 1:2, 50 and 100 students are randomly selected from the positive samples and the negative samples of the data set, respectively, to serve as the training set, and 25% of which are used as the validation set and ensure that the ratio of positive and negative samples are both 1:2.Also, the ratio of positive and negative samples in the test set is 1:2.

2.7.2. Comparative Experimental Analysis. The extracted features are fed into the classification algorithm for model training and testing, comparing four common machine learning classification algorithms, Random Forest (RT), Gradient Boosting Tree (GBDT), Plain Bayes (NB), Neural Network

(NN) and this paper's C4.5 Decision Tree based on Variable Precision Rough Sets (VPRSC4.5), and the results of different algorithms for recognizing the students with mental health problems are shown in Figure 1. First, it can be clearly observed that in all models, the test accuracy is higher than the recall rate, and the recall rate can only reach a maximum of 0.71, indicating that fewer samples can be correctly identified. Secondly, VPRSC4.5 has the best overall performance, especially the recall rate is improved more, therefore, the VPRSC4.5 algorithm is superior in this paper to be chosen for students' psychological problem recognition. In order to verify the generalization ability of the algorithm, the test dataset is inputted into the VPRSC4.5 algorithm, and the following experimental results are obtained: the Precision is 0.75, the Recall is 0.73, and the F1-Measure is 0.74. According to the experimental results of the test set, the VPRSC4.5 algorithm of this paper is able to identify 73% of the students with mental health problems.

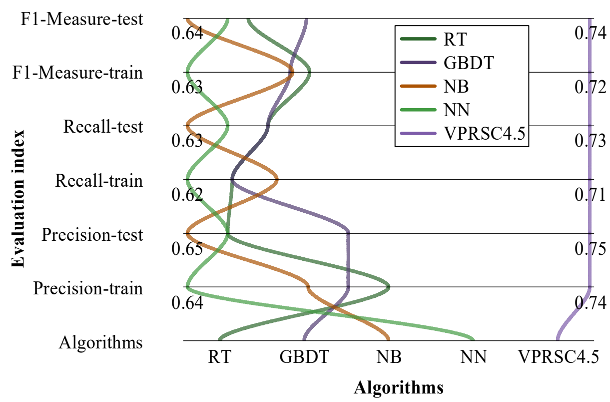

Fig. 1. Results of five classification algorithms to identify students with psychological

2.7.3. Results of identification of different psychological problems. Using the above five methods, the samples of three types of mental health problems in the test set, mild, moderate and severe, were identified, and the results of the identification of different psychological problems are shown in Figure 2. In this paper, the recognition accuracy of VPRSC4.5 algorithm for the three types of mental health problems, mild, moderate and severe, is 0.73, 0.72 and 0.75, the recall is 0.75, 0.74 and 0.77, and the F1 value is 0.74, 0.73 and 0.76, which are higher than the other four classification algorithms. Among them, the VPRSC4.5 algorithm was the most effective in recognizing students' severe mental health problems.

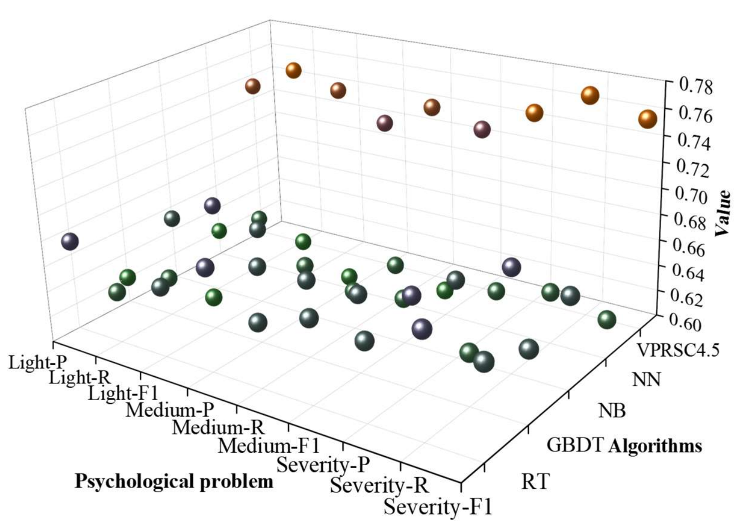

Fig. 2. Results of different psychological problem identification

2.7.4. Recognition performance in different specializations. The experiment analyzes the accuracy of different methods in recognizing the mental health of students in different majors, and the comparison results of the recognition accuracy in different majors are shown in Figure 3. The VPRSC4.5 method of this paper has higher accuracy than the other four methods, and the accuracy of mental health recognition of students in different majors is always higher than 74%, and the recognition accuracy of different majors has little fluctuation, and the stability of recognition is high. The recognition accuracy of the four comparative methods has more ups and downs, and the value of the recognition accuracy is generally low, and the recognition accuracy of different majors is lower than 70%. It can be seen that under different majors, the VPRSC4.5 method of this paper has high accuracy and stability of mental health identification, which can greatly improve the mental health identification effect of students in different majors.

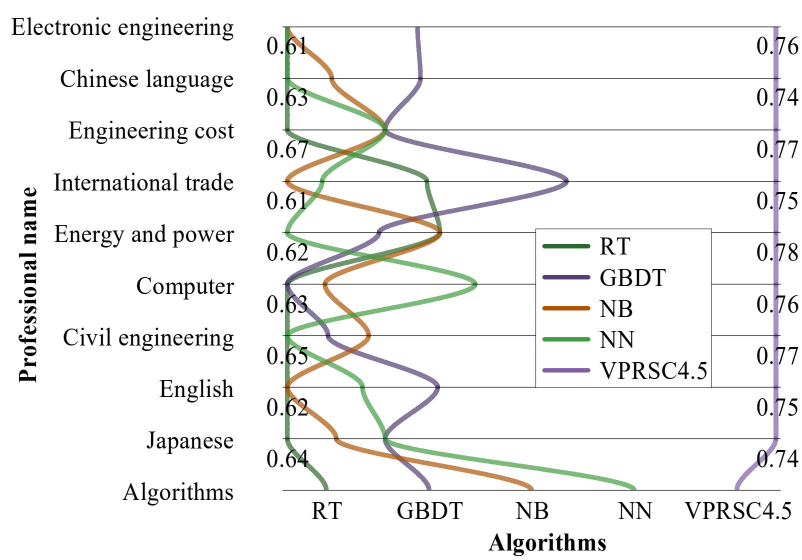

Fig. 3. Comparison of the identification accuracy of different majors

2.7.5. Characteristic importance analysis. Table 1 lists the 20 traits with high Pearson's correlations. 7, 6, and 7 of the 20 traits belonged to personal, family, and school factors, respectively, with a more balanced distribution. Two of the first three correlated traits were related to personality orientation, and three of the first seven traits were related to growth history, indicating that students' personality traits and growth history have the greatest impact on their mental health. In addition, students' academic performance and school diligence also have some influence on their mental health.

Table 1. The top 20 most relevant features

Number	Feature category	Feature description	Correlation coefficient
1	Personality orientation	High internal orientation	0.256
2	Growth process	Stay-at-home children	0.255
3	Personality orientation	Low openness	0.251
4	Family economy	Low income	0.249

5	Physical condition	High medical frequency	0.248
6	Growth process	Disabled family	0.233
7	Growth process	Expanding family	0.232
8	Academic performance	Poor academic performance	0.230
9	Diligence degree	Low attendance rate	0.221
10	Household registration	Farm home	0.218
11	Diligence degree	High Internet frequency	0.211
12	Participate in collective activities	Lower frequency of participation	0.209
13	Academic performance	High inferior lattice	0.205
14	Social intimacy	Weekends in the dormitory	0.203
15	Social intimacy	Lunch average	0.193
16	Growth process	The city level of high school	0.191
17	Academic performance	Semester credit point	0.188
18	Diligence degree	Breakfast habit	0.185
19	Physical condition	Low physical test scores	0.172
20	Social intimacy	Get up early on Monday	0.168

3. Intervention pathways for students with psychological problems

This chapter discusses the intervention path of students' psychological problems. Colleges and universities need to build students' psychological crisis early warning and intervention mechanism, combined with students' psychological problem identification model to establish a rapid response network of psychological crisis, to strengthen the various types of psychological crisis intervention and healing identification tracking, in-depth development of psychological crisis prevention, and to strengthen the organic combination of psychological crisis early warning and intervention.

3.1. Improving the organization of psychological interventions

The leading organization of the psychological crisis early warning and intervention system needs to regulate the internal and external relations of the school, and its composition should include leaders related to psychological crisis early warning and intervention, and assign tasks to subordinate organizations. The executive body performs its duties according to the work plan formulated by the leading body, including the daily maintenance of the psychological crisis intervention system, the monitoring of high-risk groups, and the tracking and monitoring of the results of the implementation of the intervention. The school must ensure the normal operation of the psychological crisis intervention center, and the psychological crisis management office is responsible for setting up work objectives, deploying the personnel needed for the system, and allocating resources according to the summary of all parties. The on-campus implementation team is the executive body of the psychological crisis intervention system on campus, the social support management team is the executive body of the psychological crisis intervention system off-campus, and the technical team is responsible for the construction and maintenance of the psychological crisis intervention system. The identification of students' psychological problems is the foundation of the psychological crisis intervention mechanism. Colleges and universities can screen out those who have signs of psychological crisis according to the identification model of students' psychological problems, classify them according to the degree of the outbreak of psychological crisis, and formulate a psychological crisis intervention program for the high-level group.

3.2. Safeguards for psychological interventions

The psychological crisis early warning and intervention mechanism for students in colleges and universities consists of subsystems such as crisis prevention and intervention and social support, and the crisis prevention system is established mainly to enhance the psychological tolerance of students, so that they can deal with crises partly caused by personal reasons by themselves. The purpose of crisis early warning is to issue early warning information to nip the crisis in the bud, and the social support system runs through the whole process of crisis intervention. Colleges and universities to build students' psychological crisis early warning and intervention home-school linkage mechanism, need to strengthen communication, education and guidance, clear division of labor between the family and the school, tracking attention to promote the linkage of long-term effectiveness. Psychological crisis early warning and intervention is a special psychological service, which is characterized by the timeliness and preventive nature of the help, which helps students to rationally treat and deal with the psychological crisis, and prevents accidents that are difficult to save. Most of the psychological crises of students in colleges and universities come from academic, family and employment pressure, etc. Colleges and universities need to subjectively train students' stress perception tendency to reduce the probability of the occurrence of students' psychological crises, and objectively reduce the intensity of the crisis and the damaging nature of the crisis events, and pay attention to the guiding attention to the poor students and the students who have difficulties in employment in the prevention.

4. Evaluation of the effectiveness of interventions for students' psychological problems

4.1. Objects and Methods

1) Subjects of the study. Four schools in a city were selected, two of which served as intervention schools and the other two as control schools, and the subjects of the study were all freshmen and sophomore students.

2) Instruments. The Mental Health Diagnostic Test (MHT) test was used for mental health diagnostic testing. The test has 100 items and contains 8 content scales, namely, learning anxiety, anxiety about people, loneliness tendency, self-blame tendency, allergy tendency, physical symptoms, phobic tendency, and impulsivity tendency. The standardized score of each content scale is less than 5 for normal mental health, between 5 and 8 for easy mental health problems, and more than 8 for serious mental health problems that require special guidance programs. If the standard score of the whole scale is greater than 65, it is a serious mental health problem.

3) Test Method. The MHT psychological test questionnaire was uniformly distributed, and the students in each school were tested by the main examiner according to the guidelines, and were asked to answer the questions on the questionnaire one by one.

4) Psychological intervention. Based on the mental health identification of students in the intervention schools using the psychological problem identification model, based on the intervention pathway for students' psychological problems, targeted mental health interventions were provided to students in the intervention group schools, and no interventions were taken in the control group schools.

5) Final Survey. At the end of the implementation of mental health interventions, students were tested again to compare the differences before and after the interventions.

6) Statistical analysis. EpiData 3.1 was used for data entry, and SPSS 18.0 was used for statistical analysis of the data, with independent samples t-test for quantitative data and chi-square analysis for qualitative data.

4.2. Assessment of the effectiveness of the intervention

The standardized score results for the full scale showed that the proportion of students with standardized scores ⩾65 on the entire test in the intervention group was 1.47% and 0.98% before and after the intervention, while the proportion of students with standardized scores ⩾65 on the entire test in the control group was 1.25% and 1.43% before and after the intervention, respectively. The proportion of students with more severe mental health problems decreased by 0.49% overall in the intervention group, while it increased by 18% in the control group.

Comparison of the factors of the mental health tests of the students in the control and intervention groups before and after the intervention is shown in Tables 2 and 3. After the intervention, the scores of the learning anxiety, loneliness tendency and terror tendency factors in the intervention group decreased compared to the pre-intervention period, and the difference was statistically significant ($P < 0.05$). Compared with the baseline survey, the scores of the learning anxiety factor and the loneliness tendency factor in the control group increased, and the difference was statistically significant ($P < 0.05$), while the other factors did not change much. Overall, the mean score of each factor for students in the control group increased by 0.22 points, up 5.21%. The mean score of each factor of the students in the intervention group decreased by 9.85%, indicating that the mental health of the students was improved after the mental health intervention.

Table 2. The comparison of the control group' psychological health test factors

Content	Base line		Final period		T value	P value
	Mean	SD	Mean	SD		
Learning anxiety	6.93	2.15	7.58	2.257	-3.162	<0.001
Interpersonal anxiety	3.92	2.26	3.96	2.15	0.156	0.152
Loneliness tendency	2.75	2.35	2.79	2.58	3.329	0.005
Self-blame tendency	5.62	2.26	5.89	2.36	0.475	0.228
Allergy tendency	5.23	2.40	5.27	2.79	-0.425	0.335
Physical symptom	4.19	2.16	4.54	2.27	0.756	0.195
Horror tendency	3.57	2.56	3.85	2.15	2.262	0.067
Impulsive tendency	2.13	2.25	2.25	2.26	-1.197	0.223

Table 3. The comparison of the intervention group' psychological health test factors

Content	Base line		Final period		T value	P value
	Mean	SD	Mean	SD		
Learning anxiety	7.28	3.24	6.39	3.12	3.278	0.002
Interpersonal anxiety	3.87	2.42	3.35	2.23	0.596	0.257
Loneliness tendency	2.35	1.77	1.92	1.54	3.156	0.006
Self-blame tendency	4.41	2.43	4.14	2.14	-0.178	0.416
Allergy tendency	4.68	2.35	4.31	2.22	-0.662	0.641
Physical symptom	3.94	2.59	3.85	2.28	-0.651	0.645
Horror tendency	3.43	2.62	3.12	2.43	2.402	0.033
Impulsive tendency	2.01	2.15	1.74	2.38	-3.618	<0.001

5. Conclusion

In this paper, we select multifaceted student behavioral features, combine Dijkstra's algorithm to construct students' social intimacy, and analyze the way of extracting other features, and use the C4.5 decision tree improvement algorithm based on variable-precision rough set as a classifier to construct a model of students' psychological problem identification. Following that, the intervention path of students' psychological problems is proposed, and experiments are set up to evaluate the effect of psychological intervention.

1) In this paper, the results of the indicators for recognizing students' psychological problems using the VPRSC4.5 algorithm are higher than those of the comparison methods, and 73% of students with mental health problems can be identified. In the identification of psychological problems of students with different mental health levels and different majors, the identification accuracy, recall and F1 value of the VPRSC4.5 algorithm are all above 0.72, while the indicator results of the comparison methods are less than 70%, which further verifies the superiority of the VPRSC4.5 algorithm. In addition, students' personality traits and growth history are important factors affecting their psychological health.

2) After psychological intervention, students in the intervention group showed significant decreases in learning anxiety, loneliness tendency and terror tendency ($P < 0.05$), and the average scores of each factor decreased by 9.85%, while the average scores of the control group improved by 5.21%, indicating that the psychological intervention plays a positive role in the psychological health of students.

3) Colleges and universities should carry out psychological crisis prevention education, set and evaluate psychological crisis warning and intervention indicators, make full use of students' psychological problem identification model, build a three-dimensional network of psychological crisis warning and intervention information, improve the operation mechanism of students' psychological crisis intervention, activate the all-around psychological support system of the school, the family and the society, improve the professional organization, clarify the functions and responsibilities, and optimize the students' psychological crisis intervention. Optimize the operation process of students' psychological crisis intervention.

References

[1] Nguyen-Thi, H. P., Truong, A. N., Le, V. T., Nguyen, X. T. K., & Tran-Chi, V. L. (2024). The Relationship between Academic Performance, Peer Pressure, and Educational Stress as It Relates to High School Students' Openness to Seeking Professional Psychological Help. Journal of Curriculum and Teaching, 13(1), 83-101.

[2] Zhao, N., Yang, S., Zhang, Q., Wang, J., Xie, W., Tan, Y., & Zhou, T. (2024). School bullying results in poor psychological conditions: evidence from a survey of 95,545 subjects. Frontiers in psychology, 15, 1279872.

[3] Aqeel, M., Abbas, J., Shuja, K. H., Rehna, T., Ziapour, A., Yousaf, I., & Karamat, T. (2022). The influence of illness perception, anxiety and depression disorders on students mental health during COVID-19 outbreak in Pakistan: a web-based cross-sectional survey. International Journal of Human Rights in Healthcare, 15(1), 17-30.

[4] Lebni, J. Y., Toghroli, R., Abbas, J., NeJhaddadgar, N., Salahshoor, M. R., Mansourian, M., ... & Ziapour, A. (2020). A study of internet addiction and its effects on mental health: A study based on Iranian University Students. Journal of Education and Health Promotion, 9(1), 205.

[5] Gao, L., Xie, Y., Jia, C., & Wang, W. (2020). Prevalence of depression among Chinese university students: a systematic review and meta-analysis. Scientific reports, 10(1), 15897.

[6] Ahmed, G., Negash, A., Kerebih, H., Alemu, D., & Tesfaye, Y. (2020). Prevalence and associated factors of depression among Jimma University students. A cross-sectional study. International journal of mental health systems, 14, 1-10.

[7] Jaafari, Z., Farhadi, A., Lari, F. A., Mousavi, F. S., Moltafet, H., Dashti, E., & Marzban, M. (2021). Prevalence of depression in Iranian college students: a systematic review and meta-analysis. Iran J Psychiatry Behav Sci, 15(1), e101524.

[8] Islam, S., Akter, R., Sikder, T., & Griffiths, M. D. (2020). Prevalence and factors associated with depression and anxiety among first-year university students in Bangladesh: a cross-sectional study. International Journal of Mental Health and Addiction, 1-14.

[9] Montagni, I., Tzourio, C., Cousin, T., Sagara, J. A., Bada-Alonzi, J., & Horgan, A. (2020). Mental health-related digital use by university students: a systematic review. Telemedicine and e-Health, 26(2), 131-146.

[10] Tang, F., Byrne, M., & Qin, P. (2018). Psychological distress and risk for suicidal behavior among university students in contemporary China. Journal of affective disorders, 228, 101-108.

[11] Li, Y., & Guo, K. (2023). Research on the relationship between physical activity, sleep quality,

psychological resilience, and social adaptation among Chinese college students: A cross-sectional study. Frontiers in psychology, 14, 1104897.

[12] Kohls, E., Baldofski, S., Moeller, R., Klemm, S. L., & Rummel-Kluge, C. (2021). Mental health, social and emotional well-being, and perceived burdens of university students during COVID-19 pandemic lockdown in Germany. Frontiers in psychiatry, 12, 643957.

[13] Li, T., Zhang, X., Chen, M., Wang, R., He, L., Xue, B., & Zhao, D. (2020). Psychological distress and its associated risk factors among university students. Revista da Associacao Medica Brasileira, 66, 414-418.

[14] Broglia, E., Millings, A., & Barkham, M. (2021). Student mental health profiles and barriers to help seeking: When and why students seek help for a mental health concern. Counselling and psychotherapy research, 21(4), 816-826.

[15] Priestley, M., Broglia, E., Hughes, G., & Spanner, L. (2022). Student perspectives on improving mental health support services at university. Counselling and Psychotherapy Research, 22(1).

[16] Posselt, J. (2021). Discrimination, competitiveness, and support in US graduate student mental health. Studies in Graduate and Postdoctoral Education, 12(1), 89-112.

[17] Thompson, M. N., Goldberg, S. B., & Nielsen, S. L. (2018). Patient financial distress and treatment outcomes in naturalistic psychotherapy. Journal of counseling psychology, 65(4), 523.

[18] Behera, S., Paluri, S. S. S. L., & Mishra, A. (2021). Mental health status of students pursuing professional training: A questionnaire-based study. Journal of education and health promotion, 10(1), 399.

[19] Gebremariam, H. T., Tassew, M. D., & Woldemaryam, F. S. (2024). Insights into psychosocial problems and associated factors among higher education students in Ethiopia: a cross-sectional study. BMC Public Health, 24(1), 2784.

[20] Bernaras Iturrioz, E., Insúa Cerretani, P., & Bully Garay, P. (2018). Prevalence and severity of psychological problems in university students. British Journal of Guidance & Counselling, 46(4), 418-428.

[21] Kiekens, G., Hasking, P., Bruffaerts, R., Alonso, J., Auerbach, R. P., Bantjes, J., ... & Kessler, R. C. (2023). Non-suicidal self-injury among first-year college students and its association with mental disorders: results from the World Mental Health International College Student (WMH-ICS) initiative. Psychological medicine, 53(3), 875-886.

[22] Sarafraz, M. R., Chavoshi, H., & Alinaghi, M. (2019). Social networks and students: educational performance, psychological well-being, and mental health. International Journal of School Health, 6(3), 1-6.

[23] Setko, A. G., Zhdanova, O. M., & Lukyanov, P. V. (2022). Features of physiological reactions to the learning load on students with different mental abilities. Hygiene and Sanitation, 101(2), 211-217.

[24] Bukhari, S. R., & Afzal, F. (2017). Perceived social support predicts psychological problems among university students. The International Journal of Indian Psychology, 4(2), 18-27.

[25] H. Birk, R., & Samuel, G. (2020). Can digital data diagnose mental health problems? A sociological exploration of 'digital phenotyping'. Sociology of Health & Illness, 42(8), 1873-1887.

[26] Webb, M., Davis, N., Bell, T., Katz, Y. J., Reynolds, N., Chambers, D. P., & Sysło, M. M. (2017). Computer science in K-12 school curricula of the 2lst century: Why, what and when?. Education and Information Technologies, 22, 445-468.

[27] He, C., Levis, B., Riehm, K. E., Saadat, N., Levis, A. W., Azar, M., ... & Benedetti, A. (2020). The accuracy of

the Patient Health Questionnaire-9 algorithm for screening to detect major depression: an individual participant data meta-analysis. Psychotherapy and psychosomatics, 89(1), 25-37.

[28] Llano, J., Ramirez, V., & Morillo, P. (2020). Benchmarking of Classification Algorithms for Psychological Diagnosis. In Smart Technologies, Systems and Applications: First International Conference, SmartTech-IC 2019, Quito, Ecuador, December 2-4, 2019, Proceedings 1 (pp. 188-201). Springer International Publishing.

[29] Vasha, Z. N., Sharma, B., Esha, I. J., Al Nahian, J., & Polin, J. A. (2023). Depression detection in social media comments data using machine learning algorithms. Bulletin of Electrical Engineering and Informatics, 12(2), 987-996.

[30] Aljarah, B., Alomari, G., & Aljarah, A. (2024). Synthesizing AI for Mental Wellness and Computational Precision: A Dual Frontier in Depression Detection and Algorithmic Optimization. AlgoVista: Journal of AI and Computer Science, 1(3), 592672.

[31] Ding, Y., Chen, X., Fu, Q., & Zhong, S. (2020). A depression recognition method for college students using deep integrated support vector algorithm. IEEE access, 8, 75616-75629.

[32] Almars, A. M. (2022). Attention-Based Bi-LSTM Model for Arabic Depression Classification. Computers, Materials & Continua, 71(2).

[33] Tian, Y. (2022). Identification and modeling of college students' psychological stress indicators for deep learning. Scientific Programming, 2022(1), 6048088.

[34] Li, J., & Shi, D. (2022, November). Sports Behavior Intervention of College Students' Psychological Crisis: Internal Logic and Path Selection. In 2022 International Conference on Sport Science, Education and Social Development (SSESD 2022) (pp. 166-172). Atlantis Press.

[35] Oliveira, C., Pereira, A., Vagos, P., Nóbrega, C., Gonçalves, J., & Afonso, B. (2021). Effectiveness of mobile app-based psychological interventions for college students: a systematic review of the literature. Frontiers in psychology, 12, 647606.

[36] Räsänen, P., Muotka, J., & Lappalainen, R. (2020). Examining mediators of change in wellbeing, stress, and depression in a blended, Internet-based, ACT intervention for university students. Internet interventions, 22, 100343.

[37] Bing, H., & Lai, L. (2022). Improvement and application of Dijkstra algorithms. Acad. J. Comput. Inf. Sci, 5(5), 97-102.