Zhou H and Huang D. (2024) PERSONALIZED ATHLETE MENTAL HEALTH CARE RECOMMENDATIONS AND EXAMINATION STRATEGIES.Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (95) pp. 75-92.
DOI: https://doi.org/10.15366/rimcafd2024.95.005

ORIGINAL

PERSONALIZED ATHLETE MENTAL HEALTH CARE RECOMMENDATIONS AND EXAMINATION STRATEGIES

Haoxing Zhou [1], Dewen Huang [1*]

[1] The Public Course Teaching Department, Guangdong Polytechnic, Foshan, 528041, China
E-mail: 13580454445@163.com

Recibido 04 de Junio de 2023 **Received** June 04, 2023
Aceptado 05 de Febrero de 2024 **Accepted** February 05, 2024

ABSTRACT

In recent years, the athlete mental health problems have become more and more prominent. Compared with high school, college athlete students are in a more complex environment and face more diversified difficulties, such as the constant emergence of problems in academics, interpersonal interactions, emotional life, employment, etc., which make college athletes overwhelmed, and the accumulation of pressures from many sides leads to a variety of psychological problems. Therefore, it is important to pay attention to the individual athlete mental health problems, and to detect and guide the development of athletes' mental health in time, which is crucial for the management of athlete students in colleges and universities. The purpose of this paper is to use data mining principles and methods to explore the factors leading to psychological problems of college athletes, so as to carry out psychological intervention purposefully. In this paper, we use data mining technology to set up a questionnaire with 90 questions reflecting mental health, and take SCL-90 Symptom Self-assessment Scale as the measurement standard to collect data, establish a data set, analyse the data using improved Apriori algorithm, and establish a decision model of obsessive-compulsive symptoms using decision tree ID3 algorithm, which provides a theoretical basis for the evaluation of the athletes mental health status of college athletes and decision-making of health checkups. It provides a theoretical basis for the evaluation of the mental health status of athletes and the decision-making of health examination.

KEYWORDS: Athlete Health Examination; Personalized Mental Health; Athlete Mental Health Care

1. INTRODUCTION

The development of today's world is changing day by day, and a strong talent is a strong country. Various kinds of excellent talents cultivated by universities and colleges are the inexhaustible power to promote the development of the society, and having healthy psychology and sound personality is the prerequisite and key to the comprehensive development of college athlete students and is also an essential requirement to adapt to the society (Kornadt, Hagemeyer, Neyer, & Kandler, 2018). At present, with the development of society and higher education, college athletes, as the main training object of higher education, their ideology, emotional cognition (M. Chen & Hao, 2019; Pruessner, Barnow, Holt, Joormann, & Schulze, 2020) and psychological state is also in a strong plasticity state, which is very easy to change. During their four-year college education, students experience a crucial period of independence from parental supervision. At this stage, they aspire to achieve personal excellence and distinction through their own endeavours. However, this journey is inevitably accompanied by a multitude of challenges, including academic examinations, social relationships, romantic pursuits, employment prospects, the pressure to pursue graduate studies, and career advancement. Currently, it is important for students to possess a strong psychological disposition in order to effectively confront and resolve many forms of stress and adversity. Currently, a significant number of college athlete exhibit suboptimal psychological well-being and lack resilience.

This is evident through the prevalence of various psychological issues among students, such as low self-confidence and a pronounced inferiority complex. Many students perceive themselves as incapable of accomplishing tasks and lack the courage to engage in proactive practice. Consequently, they tend to withdraw when faced with challenges. Additionally, a subset of students demonstrates shyness and timidity, inhibiting their ability to initiate social interactions with peers and seek guidance from teachers.

Consequently, these students find themselves in a state of stagnation amidst difficult circumstances. Certain students exhibit shyness and hesitancy, which inhibits their willingness to actively engage with peers and instructors, as well as seek guidance. Consequently, they find themselves in a state of isolation, lacking positive interpersonal connections. Additionally, some students possess a disposition that hinders their ability to establish harmonious relationships with others, leading to withdrawal and an inability to assimilate into collective environments. Furthermore, certain individuals experience psychological tendencies characterized by obsessive-compulsive anxiety, resulting in a perpetual state of apprehension. The presence of psychological challenges and unfavourable living circumstances, whether inside educational institutions or society at large, hinders the personal growth and self-actualization of college athletes, hence impeding their ability to attain a fulfilling and contented existence. In the face of so many possibilities of psychological problems, how to provide individualized mental health care advice to college athlete and how to build a decision support system for health checkups in order to guide college athlete towards a healthy and positive life has become an urgent issue (Bai, Kusi-Sarpong, Badri Ahmadi, & Sarkis, 2019; De Groot, Delmaar, & Lupker, 2000; Hoch & Schkade, 1996).

With the increase of pressure from employment, examination and other aspects, there are frequent suicide incidents of students jumping from buildings, which is heartbreaking, and at the same time, it also makes the whole society start to pay attention to the mental health of college athlete and start to think about how to let the students have a healthy mentality.

College athlete are the talents cultivated by the state and families for the society, they should make their own contribution to the country after graduation and pursue their own happy life, but due to all kinds of psychological problems, tragedy occurs, so it is very important for colleges and universities and parents to find out how to find out the mental health problems of college athlete in advance, and provide appropriate counselling (Pedrelli, Nyer, Yeung, Zulauf, & Wilens, 2015; Rickwood, Deane, & Wilson, 2007). In the face of the growing threat of psychological problems among contemporary college athlete, psychological counselling for college athlete is also getting more and more attention from the state. The major universities and colleges across the country generally set up specialized psychological counselling departments to manage their own school athletes' mental health counselling, research, guidance work, each student has his or her own psychological profile, psychological growth trajectory can be traced (Connell & Frye, 2006; Henrichs et al., 2010).

With the development of modern science and technology, all walks of life cannot be separated from the application of computer networks and information systems, general colleges and universities have a special decision support system for mental health checkups, psychological assessment of new students entering the school each year, including the basic situation of the athletes' mapping statistics as well as the most representative SCL-90 scale for assessment, for each student to produce an analysis of the report in the system can be exported to all the students with psychological problems tendency report. The system can derive reports on all students with a tendency to psychological problems and focus on counselling interventions (Cantor et al., 2021; Kadafi, Alfaiz, Ramli, Asri, & Finayanti, 2021; Lisanti, Golfenshtein, Min, & Medoff-Cooper, 2023). Currently, prominent institutions own their own decision support systems for mental health evaluations. These systems generate and gather a considerable quantity of data each year after the completion of psychological assessments. Consequently, the amassed data continues to grow with each passing year. The accumulation and consolidation of psychological evaluation data over an extended period of time has resulted in a distinctive data resource.

However, if this data is not used for focused analysis and application, it remains inert and devoid of value, while also occupying hardware and software resources. Hence, it is essential for us to get the crucial information concealed inside this data by using certain technological methods, and thereafter evaluate and use this data at an elevated level, thereby enabling it to manifest its inherent worth (Ribeiro Junior, Werneck, Oliveira, & Ibáñez, 2023).

Current psychological assessment systems typically possess capabilities such as data manipulation, including addition, deletion, modification, counting, and querying. However, these systems lack the ability to uncover valuable knowledge and underlying patterns concealed within the data. Furthermore, they are unable to predict the occurrence of diverse severe psychological issues among college athlete based on existing data. The use of data mining technologies has the potential to somewhat mitigate this issue. The psychological counselling system employed by colleges and universities has amassed a substantial volume of psychological assessment data from students over an extended period of time.

By leveraging appropriate data mining technologies and algorithms, there exists a significant potential to gain profound insights into the underlying factors contributing to students' issues. In order to achieve this objective, it is important to examine the extensive collection of psychological evaluation data pertaining to a substantial cohort of pupils.

Therefore, this paper utilizes data mining technology, sets up a questionnaire with 90 questions reflecting mental health as samples, uses the SCL-90 Symptom Self-Assessment Scale as the measurement standard, collects data, establishes a dataset, analyses the data by using improved Apriori algorithm, and establishes a decision model of obsessive-compulsive symptoms by using the Decision Tree ID3 Algorithm, which provides theoretical basis for the evaluation of the mental health status of college athlete and decision-making of health checkups. The main contributions are as follows:

(1) In this paper, the improved Apriori algorithm was used for data analysis, and the decision tree ID3 algorithm was used to establish a decision model of obsessive-compulsive symptoms, which provided a theoretical basis for the evaluation of the mental health status of college athlete and the decision-making of health examination.

(2) In this paper, we set up a questionnaire based on the SCL-90 Symptom Self-Assessment Scale to obtain data related to the mental health of college athlete, including sleep quality, dietary habits, exercise, emotional status, academic pressure, living conditions and other aspects of the data, screen the valid data, improve the quality of the data, and analyze and process them by using data analysis methods.

2. Related Works

2.1 Data Mining

The world has a large amount of data generated every day, covering various industries and aspects, the huge, complicated and redundant data make it difficult for us to directly and accurately find the data we need, and data mining can precisely solve the troubles in this regard. Data mining (Grossi, Tavano Blessi, Sacco, & Buscema, 2012; King & Resick, 2014; Krist et al., 2020), also known as data mining, as the name suggests, it is from the massive amount of data to dig out the hidden, with close correlation, the process of valuable data, data is like a big gold mine, and play the miner's we constantly dig its value.

After obtaining the data, through the integration and analysis of related data, inductive reasoning, you can convert these data into usable information, discover its hidden laws and apply it to various industries and fields, whether it is market analysis, operations management or scientific exploration, can give full play to its role and value, the main process as shown in Figure 1.

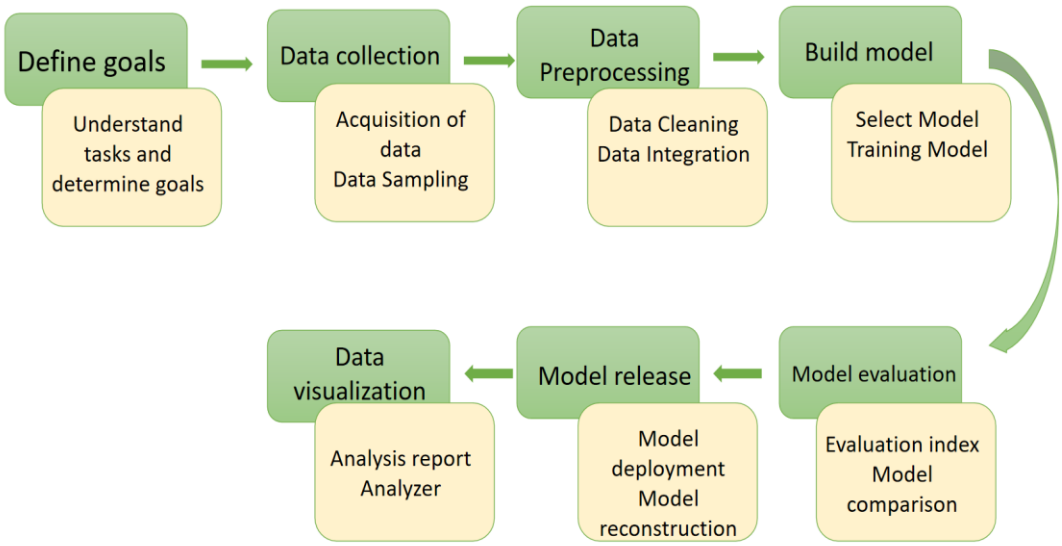

Figure 1: Schematic diagram of the data mining process

There are several common data mining tasks: (1) Classification and regression: classification is the use of already known attribute features to train and construct classification models to achieve the classification of unknown attribute data. Its training dataset is labeled, and its categories are determined before classification, such as decision tree models. Regression, on the other hand, is the use of previous attribute data for prediction, the first assumption that the function fits the target data, and then use the size of the error to determine the function that best fits the target, the smaller the error, the better, such as logistic regression.

(2) Cluster analysis: similar to the things to the class, people to the group, clustering is to the similarity of the data as a criterion for the division of the data, the same cluster of data is generally similar to a high degree of similarity, the different clusters have a large difference between the different clusters. Clustered data does not have labels, we do not know the specific classification of the data in advance and can only be divided according to the similarity or distance of the data attributes. Common algorithms (J. Chen, Han, Zhang, You, & Zheng, 2023; J. Chen, Li, et al., 2023; Li & Cao, 2023) are divided into algorithms based on split, hierarchy, density, network and model, such as K-means based on split, neural network based on model, etc.

(3) Association rules: it is to mine the hidden and not obvious relationship between data, and its main task is to reveal the degree of correlation between data and use its inner connection to realize greater value. For example, Apriori algorithm.

(4) Temporal Patterns: As the name suggests, temporal patterns are to look for patterns and trends in accordance with time or sequence and modeling, combining correlation patterns and time series patterns, exploring the correlation between data in the time dimension, emphasizing the importance of the temporal sequence of data, such as cycles, seasons, etc., and predicting the unknown with known data.

(5) Deviation detection: it is to find abnormalities and exceptions, analyze the data that differ greatly from the normal standard situation, these abnormal data should not be directly discarded, but should be further analyzed and studied to find out the hidden problems, and very often the abnormal data is more valuable for research than the normal data, such as fraud detection, which can remind us of the prevention of fraud, and take corresponding measures in time to avoid producing greater Losses.

2.2 Apriori algorithm

The Apriori algorithm is an algorithm used to find frequent itemsets for Boolean association rules. The core of the Apriori algorithm can be summarized as a layer-by-layer search iteration, exploring $(k+1)$ itemsets based on k itemsets, and so on. At the beginning, the database is scanned as a whole, the number of occurrences of each item is calculated, and the items that meet the minimum support are gathered together, so that the frequent 1 item set is obtained and marked with L_1. Using the same method, L_3 is found from L_2, and L_k is found based on the previous frequent itemset L_{k-1}. This loop is iteratively executed until no frequent L_1 itemsets can be found. Therefore, the system overhead of this process is relatively large, and finding each L_k requires a complete scan of the database. The Apriori algorithm has many advantages in the generation process of frequent itemsets: (1) It has clear algorithm execution steps and can generate smaller candidate sets according to the properties of the Apriori algorithm. (2) This algorithm can achieve the goal of reducing the size of the candidate item set based on a priori principles or reducing the number of comparisons. (3) All frequent item sets can be generated without duplication or omission.

3. Methodology

Most decision support systems for mental health checkups lack a data mining module. Consequently, these systems are limited to collecting and organizing data, as well as conducting basic statistical analysis. As a result, they are unable to analyze the potential relationships between the attributes of college athlete and their various psychological symptoms.

This absence of comprehensive analysis hinders college counsellors and psychological counsellors from effectively implementing preventive interventions. Therefore, the application of data mining technology to college athlete is currently lacking a reliable foundation. The user's text is already academic. Hence, the use of data mining technology in the analysis of psychological evaluation data pertaining to college athlete is deemed necessary and unavoidable. This research applies the association rule mining algorithm to the decision support system designed for the assessment of college athlete' mental health.

3.1 Association rule-based health data mining

Association analysis can be described as a method of discovering some latent valuable or regular connection in a target data set through the execution of some algorithm. Such discovered connections are generally represented by association rules which are generated on the basis of frequent item sets. Association rule mining has been widely used in many fields.

3.1.1 Association Rule

The association rule is generally expressed in the form of an expression in which X implies Y, where the intersection of X and Y is the empty set, i.e., X and Y are two sets with no common set. The relevance of an association rule can be measured in terms of its support and confidence. Support is a measure of how often the rule appears in a given dataset, and confidence is a measure of how many times Y has a chance of appearing in a transaction containing X. The following two equations intuitively represent the relevance of an association rule. The following two formulas visually represent the two metrics of support (S) and confidence (C):

$$S(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \qquad (1)$$

$$C(X \to Y) = \frac{\sigma(X \cup Y)}{\sigma_X} \qquad (2)$$

The support of a rule is the probability of simultaneous occurrence of the set of items within the rule, and the confidence is a measure of how trustworthy the rule is. The usefulness of a mined rule and its level of certainty are usually expressed in terms of support and confidence.

Pre-set support thresholds and confidence thresholds are usually determined by professionals and mining experts, and if the results are greater than or equal to the minimum support and confidence, then the association rule is said to be useful. Some statistical aspects of association analysis can also be performed to discover certain characteristic ways of association between related items.

3.1.2 Itemset

Suppose $I = \{i_1, i_2, \ldots, i_d\}$ is the set of all items in a transaction, and $T = \{t_1, t_2, \ldots, t_N\}$ is the set of all transactions, then the subset of I is the set of items contained in each transaction t_i. According to the concept in association analysis, a set containing any number of items is called an itemset. An itemset contains k items, and it is called a k -itemset. For example, (Kocjan, Kavčič, & Avsec, 2021) contains two items, then it is a 2-item set. The empty set contains no items. The number of items listed in a transaction can be called the transaction width. If transaction t_j includes itemset X, then it means that itemset X is a subset of transaction t_j.

3.1.3 Support count

Support count is important for item sets; it refers to the number of transactions that contain a specific itemset. The mathematical formula is used to express the support count $\sigma(X)$ of the itemset X as follows:

$$\sigma(X) = |\{t_i | X \subseteq t_i, t_i \in T\}| \qquad (3)$$

3.2 Decision Tree Algorithm

3.2.1 Split Rules

In the process of constructing the decision tree model, the split nodes need to be determined. The determination of the split nodes requires certain split rules, which are also called attribute partition metrics. Currently, there are two commonly used attribute classification measures, namely information gain and information gain rate. Suppose D is a training sample set containing class labels, and attributes with different class labels contain n different attribute values, then it is defined as n different classes $C_i (i = 1,2,\cdots,n)$, $C_{i,D}$ is a set of samples belonging to class C_i in D, and $|D|$ is The number of samples in

D, $|C_{i,D}|$ is the number of samples in $C_{i,D}$. D is all the sample sets stored in a node, then the expected values of the various categories of sample sets can be calculated by the following formula:

$$\text{Info}(D) = -\sum_{i=1}^{n} \frac{|C_{i,D}|}{|D|} log_2 \left(\frac{|C_{i,D}|}{|D|}\right) \qquad (4)$$

Where $\text{Info}(D)$ is called entropy. "Entropy" represents the degree of confusion. The larger the value, the more mixed it is. The smaller the value, the higher the purity. If the samples in D are divided according to attribute B, and assume that attribute B has m different discrete values $\{b_1, b_2, \cdots, b_m\}$. Then D will be divided into m subsets $\{D_1, D_2, \cdots, D_m\}$ by attribute B. The m branches generated from node N correspond to these m subsets. B_j is defined as the value of the sample in $D_j(h = 1,2,3, \cdots, m)$ on attribute B. The expectation of sample classification partitioning of D according to attribute B can be calculated using the following equation:

$$\text{Info}_B(D) = -\sum_{j=1}^{m} \frac{|D_j|}{|D|} \times \text{Info}(D_j) \qquad (5)$$

The reduction in entropy due to knowing the value of B is calculated as follows:

$$\text{Gain}(B) = \text{Info}(D) - \text{Info}_B(D) \qquad (6)$$

The process of classification is to make the constructed tree more and more orderly and regular, and the degree of chaos is continuously reduced. The more orderly an organized system is constructed, the less entropy it has. Therefore, the best solution is the splitting solution that maximizes the entropy reduction. In addition, the equation for calculating the information gain rate is as follows:

$$\text{GainRatio}(A) = \frac{\text{Gain}(A)}{\text{SplitInfo}(A)} \qquad (7)$$

3.2.2 Tree Pruning

In the process of constructing the above decision tree is to use the training set data, the decision tree constructed is more complex and luxuriant, in order to improve the classification efficiency and readability of the decision tree, the tree needs to be pruned. In this paper, the PEP leaf pruning algorithm is used to post prune the decision tree.PEP is based on the error rate before and after pruning to determine the pruning of the subtree. Before pruning, the error rate of one of the subtrees is estimated:

$$e = \frac{\sum E_i + 0.5 * L}{\sum N_i} \qquad (8)$$

where E_i is the number of errors at that node, N_i is the number of samples at that node and L is the number of leaf nodes. After pruning, its probabilistic misclassification rate e_t is (E +0.5)/N, so the mean value of the number of leaf node misclassifications E_t is:

$$E_t = N * e_t = N * \frac{E+0.5}{N} = E + 0.5 \qquad (9)$$

Before pruning, the misclassification mean and standard deviation are calculated as:

$$e_T = \frac{\sum E_i + 0.5 * L}{N} \qquad (10)$$

$$E_T = N * e_T = N * \frac{\sum E_i + 0.5 * L}{N} = \sum E + 0.5L \qquad (11)$$

$$\delta = \sqrt{N * e_T * (1 - e_T)} = \sqrt{N * E_T * (1 - E_T)} \qquad (12)$$

If $E_t < E_{T_i} + \delta$, then prune that subtree.

3.3 Improved Apriori Algorithm

Apriori algorithm adopts a step-by-step search method, which needs to enumerate all the frequent itemsets, and from the above learning process, we can learn that the algorithm's shortcomings are manifested in the following aspects. First, the database needs to be scanned repeatedly during operation, which will cause the algorithm to require excessive computer I/O load and long operation time and other shortcomings.

Secondly, when the transaction database is small, the frequent itemset pattern obtained from this kind of data is usually short, and ordinary association rule mining algorithms can get good performance. However, when the algorithm is applied to large transactional databases, such as the medical field, business field, etc., the performance of the algorithm decreases dramatically due to the large number of field patterns. This is due to the following three reasons If the length of the longest frequency set of the transactional database is n, then the algorithm needs to scan the database n times.

Scanning a large database multiple times greatly increases the I/O load on the algorithm. Generating a frequency set pattern means that a large number of additional frequency sets need to be generated and their support calculated, generating a large number of candidate sets and calculating their support consumes a lot of time. D is regarding any real-world significance, considering any item generated by a connection as a candidate item set requires doing useless searching and counting.

Therefore, in the frequent itemset-based discovery algorithm, the following three aspects should be considered to improve the efficiency of the algorithm Reducing the number of times of scanning the database D Reducing the number of candidate items in the candidate itemset C_k Reducing the generation of redundant rules. After comprehensively considering the advantages and disadvantages of Apriori algorithm, the improvement methods of frequent itemset discovery algorithm in related references, and the time efficiency and storage space requirement of the algorithm, this paper proposes the following improvement scheme, i.e., the matrix input method is adopted in the storage process of transaction database D.

In the optimization process of the algorithm, firstly, we should improve the efficiency of the algorithm from the following three aspects In the optimization process of the algorithm, the database is first encoded, and the matrix storage method is used to read all the data into the memory and store the itemsets at one time, so that the database can be scanned in the cache.

While scanning the database, record the support transactions of each item, and determine the support transaction set of the candidate k item set by the support transaction set of each item in the candidate k item set, so as to avoid repeated scanning of the database, effectively avoiding the execution of scanning the database for many times.

4. Experiment and Results

4.1 Questionnaire

The questionnaire is to provide investigators with valuable data by understanding targeted and purposeful questions, and to analyze and study the data to obtain the required basis. Therefore, the design of the questionnaire is particularly important. A good questionnaire can not only clearly convey the questions to the respondents, but also make the respondents enjoy it and gain the support of the respondents. Therefore, the setting of questionnaires must follow certain principles and pay attention to skills. The design principles of the questionnaire are as follows:

(1) Clear purpose. The main purpose of the questionnaire is to obtain the data required by the investigator. The questions set must be indispensable and irrelevant questions should not be asked. The theme is clear, the goals are clear, and the key points are highlighted; (2) Pay attention to logic. When setting questions, do not arrange them randomly. There must be a logical sequence,

easy first and then difficult. Set them in a regular order. Questions that are difficult to answer should be placed at the end; (3) Easy to understand. The questionnaire survey is aimed at the general population.

The questions should be set clearly and concisely. Do not beat around the bush or use professional vocabulary. If the questions are too advanced and difficult to understand, it will reduce the interest of the respondents and the respondents will easily refuse or give up. Therefore, use simple and easy-to-understand expressions; (4) The length is reasonable. The questionnaire should not be too long. If the content is too long, the respondents will lose their patience and may give up answering. They may also answer randomly to save time, which will reduce the credibility of the survey results. The design of the psychological questionnaire for college athletes is shown in Table 1.

Table 1: Design of psychological questionnaire for college athlete

ID	ITEMS	LEVEL
1	Do you often have headaches?	1~5
2	Are you often nervous?	1~5
3	Do you often have unnecessary thoughts?	1~5
4	Often feel dizzy	1~5
5	Decreased interest in the opposite sex	1~5
6	Seeking perfection from others' censure	1~5
7	Feeling that others can control one's thoughts	1~5
8	Often blaming others for trouble	1~5
9	Often forget things	1~5
10	Worry about one's appearance	1~5
11	Easily excited	1~5
12	Often feel chest tightness	1~5
13	Fear of open places	1~5
14	Decreased energy	1~5
15	Want to end one's life	1~5

4.2 SCL-90

The SCL-90 Symptom Self-Assessment Scale (SSAS) is intended for adults over the age of 16 years and is used to detect whether a person has a mental health problem and the severity of the problem, and it can be self-administered as well as tested on others, but it is not suitable for patients with mania and schizophrenia due to the lack of special items such as fluttering. The general timeframe for assessment is within a week of feeling, and too long a period of time will result in less credible data. The specific questions corresponding to each factor are shown in Figure 2.

F1:Somatization	1,4,12,27,40,42,48,49,52,53,56,58
F2:Obsessive-compulsive symptoms	3,9,10,28,35,45,46,51,55,65
F3:Interpersonal sensitivity	6,21,34,36,37,41,61,69,73
F4:Depression	5,14,15,20,22,26,29,30,31,32,54,71,79
F5:Anxiety	2,17,23,33,39,57,72,78,80,86
F6:Hostility	11,24,63,67,74,81
F7:Fear	13,25,47,50,70,75,82
F8:Paranoid	8,18,43,68,76,83
F9:Mental illness	7,16,35,62,77,84,85,87,88,90

Figure 2: Distribution of questions corresponding to each factor

4.3 Experimental results and analysis

From the rating scale of "SCL-90 Symptom Self-Rating Scale", it can be seen that when the existence of factor mean score is greater than 2 indicates that in this factor athletes may have greater distress, greater probability of psychological problems, therefore, the different influencing factors in this paper's mental health dataset are classified and counted to obtain the number of each factor with a mean score of greater than 2, and the data are visualized as shown in Figure 3.

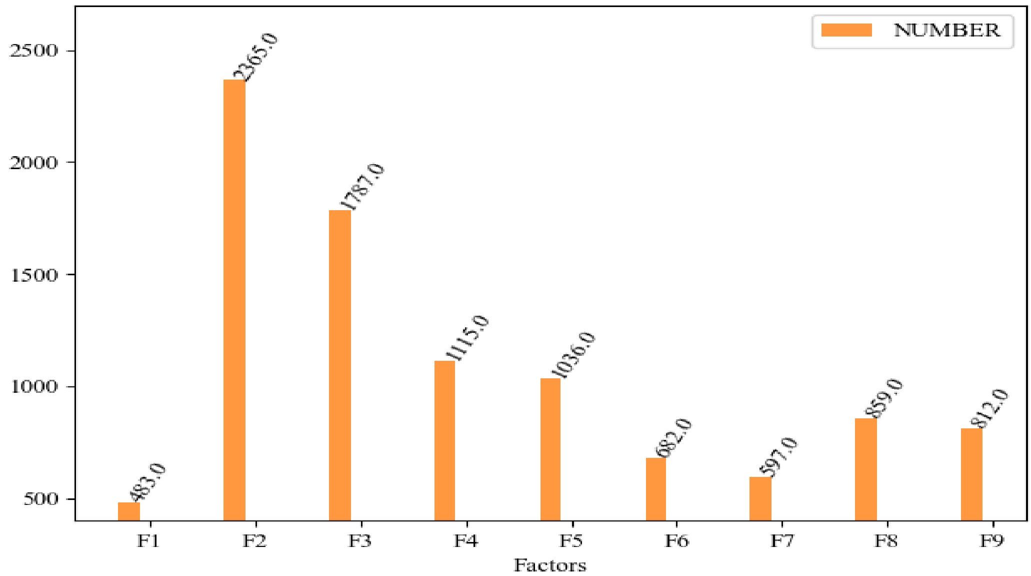

Figure 3: Distribution of factors for the presence of psychological problems.

As can be seen from Figure 3, among the factors affecting mental health, obsessive-compulsive symptoms are in the first place, with 2,365 people, accounting for 19.87% of the total number of samples, and it is the one that has the greatest impact on the mental health of college athlete, which can be seen from the fact that college athlete may have some obsessive-compulsive symptoms in a greater or lesser degree, and may often repeat some unnecessary things, unable to control themselves, and needing to focus their attention on it.

Secondly, interpersonal relationship sensitivity, with 177 people having troubles in this area, accounting for 15.01% of the total number of samples, nowadays the society is complicated and diverse, the importance of interpersonal relationship is more and more prominent, the athletes with better interpersonal skills can adapt to the surrounding environment faster and better, however, for some of the athletes who are not good at socializing, this is a big challenge, and the difficulty of changing the character will make some of the athletes more and more afraid of interpersonal relationship. Difficulty in changing personality will make some college athlete more and more afraid of interpersonal relationships, thus causing psychological pressure. The third one is depression, with 1,115 athletes, accounting for 9.35% of the total samples. Lack of self-confidence, low mood and irritability, poor sleep, etc. occur from time to time, which need to enhance the self-confidence of college athlete and improve the quality of sleep. These three factors are crucial to the mental health of college athlete and should be emphasized when focusing on the mental health of college athlete. The remaining factors are, in order, paranoia, psychoticism, other, hostility, terror, somatization, and anxiety.

Compared with the three items of obsessive-compulsive symptoms, interpersonal sensitivity, and depression, although the influence of the remaining factors is small, it does not mean that these factors are insignificant, and they are still an important part of the problem that cannot be ignored.

Although mental health problems may exist if a single factor score exceeds 2, however, this is an analysis from a localized perspective, and does not mean that mental problems do not exist if the mean score of a single factor does not exceed 2, so we also need to analyze it from a holistic perspective. The total scores obtained from the psychological surveys in the mental health dataset of this paper are categorized, and the higher the total score, the worse the level of mental health, and the more in need of teachers' and counsellors' focused attention. When the total score does not exceed 160 points, it means that the possibility of the college student suffering from mental health problems is small; when the total score exceeds 160 points but is not higher than 200 points. It means that the student may have a mild psychological problem, but not serious; when the total score exceeds 200 points but is not higher than 300 points, it means that the student may suffer from moderate mental health problems. When the total score is more than 200 but not higher than 300, the student may have moderate psychological problems; and when the total score is more than 300, the student is very likely to have serious mental illness, and the data are visualized as shown in Figure 4.

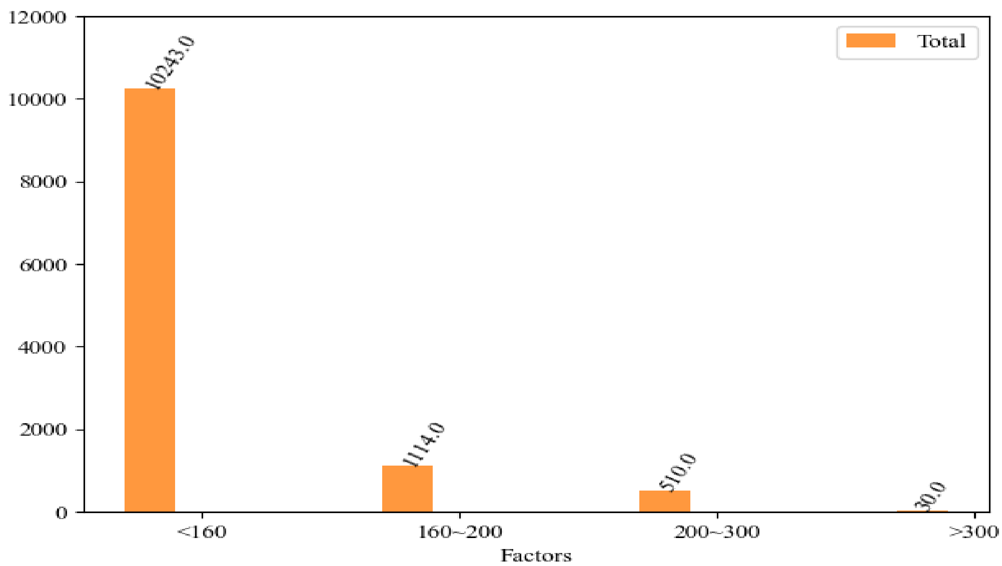

Figure 4: Distribution of total scores

Among the 897 college athlete' mental health samples in this paper, there are 243 athletes with a total score of not more than 160. From the point of view of the total score, these college athletes are mentally healthier, and the possibility of psychological problems is small, so they can be appropriately relieved from paying attention to the psychology of these college athlete; there are 1114 athletes with a total score of more than 160, but not higher than 200, accounting for 9.36% of the total sample, and these athletes have mild psychological problems.

There are 1114 athletes with a total score over 160 but not higher than 200, accounting for 9.36% of the total sample, these athletes have mild psychological problems, counsellors need to communicate with these athletes from time to time to understand their inner thoughts, find out where the problems lie, and regularly understand the state of the athletes, and timely counselling is needed in case of aggravation of the mental health problems; 510 athletes with a total score over 200 but not higher than 300, and 510 athletes with a total score over 200 but not higher than 300, are in a relatively low probability of having psychological problems. There are 510 athletes with a total score of more than 200 points but not higher than 300 points, accounting for 4.29% of the total number of samples, and these athletes suffer from moderate psychological problems. Compared with the athletes with mild psychological problems, they need the help of teachers and counsellors more than ever to help them channel their psychological problems in time and regulate their psychological state, and to tell the parents of the athletes that they should always know the psychological condition of the athletes, and that the parents and the school should work together to help the athletes, to alleviate the psychological pressure of the athletes and provide them with psychological counselling regularly, so as to prevent psychological problems from aggravating.

Psychological counselling, to prevent the aggravation of psychological problems; the total score of more than 300 points of 30 people, accounting for 0.25% of the total number of samples, these people are suffering from very serious psychological problems, psychological and emotional are in an unstable state of extremes, may have the idea of lightening the life of a person, the need for professional psychological counselling teachers to provide psychological assistance or to go to the psychiatric clinic to carry out a more systematic and professional treatment for athletes who may be mentally ill, counsellors should help the athletes with psychological counselling. For athletes who may have mental illnesses, counsellors should contact their parents in a timely manner and let the athletes go home for medical treatment to avoid accidents. Through the above analysis, the effectiveness of the health examination decision support system constructed in this article is fully proved.

5. Conclusion

In this paper, we utilize data mining principles and methods to excavate the factors leading to psychological problems of college athlete, so as to make psychological interventions purposefully. In this paper, we use data mining technology to set up a questionnaire with 90 questions reflecting mental health as samples, and the SCL-90 Symptom Self-assessment Scale as the measurement standard, collect data, establish a data set, analyze the data by using the improved Apriori algorithm, and establish a decision model of obsessive-compulsive symptom by using the Decision Tree ID3 algorithm, so as to provide a theoretical basis for the evaluation of college athlete' mental health status and decision-making of health checkups. This model provides a theoretical basis for the evaluation of the mental health status of college athlete and the decision-making of health examination.

References

Bai, C., Kusi-Sarpong, S., Badri Ahmadi, H., & Sarkis, J. (2019). Social sustainable supplier evaluation and selection: a group decision-support approach. *International Journal of Production Research, 57*(22), 7046-7067.

Cantor, A. G., Jungbauer, R. M., McDonagh, M., Blazina, I., Marshall, N. E., Weeks, C., . . . Chou, R. (2021). Counseling and behavioral interventions for healthy weight and weight gain in pregnancy: evidence report and systematic review for the US Preventive Services Task Force. *Jama, 325*(20), 2094-2109.

Chen, J., Han, P., Zhang, Y., You, T., & Zheng, P. (2023). Scheduling energy consumption-constrained workflows in heterogeneous multi-processor embedded systems. *Journal of Systems Architecture, 142*, 102938.

Chen, J., Li, T., Zhang, Y., You, T., Lu, Y., Tiwari, P., & Kumar, N. (2023). Global-and-local attention-based reinforcement learning for cooperative behaviour control of multiple UAVs. *IEEE Transactions on Vehicular Technology*.

Chen, M., & Hao, Y. (2019). Label-less learning for emotion cognition. *IEEE transactions on neural networks and learning systems, 31*(7), 2430-2440.

Connell, A. M., & Frye, A. A. (2006). Growth mixture modelling in developmental psychology: Overview and demonstration of heterogeneity in developmental trajectories of adolescent antisocial behaviour. *Infant and Child Development: An International Journal of Research and Practice, 15*(6), 609-621.

De Groot, A. M., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology: Section A, 53*(2), 397-428.

Grossi, E., Tavano Blessi, G., Sacco, P. L., & Buscema, M. (2012). The interaction between culture, health and psychological well-being: Data mining from the Italian culture and well-being project. *Journal of Happiness Studies, 13*, 129-148.

Henrichs, J., Schenk, J., Roza, S., Van den Berg, M., Schmidt, H., Steegers, E., . . . Tiemeier, H. (2010). Maternal psychological distress and fetal growth trajectories: the Generation R Study. *Psychological medicine, 40*(4), 633-643.

Hoch, S. J., & Schkade, D. A. (1996). A psychological approach to decision support systems. *Management Science, 42*(1), 51-64.

Kadafi, A., Alfaiz, A., Ramli, M., Asri, D. N., & Finayanti, J. (2021). The impact of islamic counseling intervention towards students' mindfulness and anxiety during the covid-19 pandemic. *Islamic Guidance and Counseling Journal, 4*(1), 55-66.

King, M. W., & Resick, P. A. (2014). Data mining in psychological treatment research: a primer on classification and regression trees. *Journal of consulting and clinical psychology, 82*(5), 895.

Kocjan, G. Z., Kavčič, T., & Avsec, A. (2021). Resilience matters: Explaining the association between personality and psychological functioning during the COVID-19 pandemic. *International Journal of Clinical and Health Psychology, 21*(1), 100198.

Kornadt, A. E., Hagemeyer, B., Neyer, F. J., & Kandler, C. (2018). Sound body, sound mind? The interrelation between health change and personality change in old age. *European Journal of Personality, 32*(1), 30-45.

Krist, A. H., Davidson, K. W., Mangione, C. M., Barry, M. J., Cabana, M., Caughey, A. B., . . . Kubik, M. (2020). Behavioral counseling interventions to promote a healthy diet and physical activity for cardiovascular disease prevention in adults with cardiovascular risk factors: US Preventive Services Task Force recommendation statement. *Jama, 324*(20), 2069-2075.

Li, Y., & Cao, J. (2023). Adaptive binary particle swarm optimization for WSN node optimal deployment algorithm. *IECE Transactions on Internet of Things, 1*(1), 1-8.

Lisanti, A. J., Golfenshtein, N., Min, J., & Medoff-Cooper, B. (2023). Early growth trajectory is associated with psychological stress in parents of infants with congenital heart disease, but moderated by quality of partner relationship. *Journal of Pediatric Nursing, 69*, 93-100.

Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College athlete: mental health problems and treatment considerations. *Academic psychiatry, 39*, 503-511.

Pruessner, L., Barnow, S., Holt, D. V., Joormann, J., & Schulze, K. (2020). A cognitive control framework for understanding emotion regulation flexibility. *Emotion, 20*(1), 21.

Ribeiro Junior, D., Werneck, F., Oliveira, H., & Ibáñez, S. (2023). RELATIVE AGE EFFECT ON BRAZILIAN BASKETBALL OVER TIME. *Revista multidisciplinar de las Ciencias del Deporte, 23*(89).

Rickwood, D. J., Deane, F. P., & Wilson, C. J. (2007). When and how do young people seek professional help for mental health problems? *Medical Journal of Australia, 187*(S7), S35-S39.