Application of ID3 Algorithm in College Students' Mental Health Education

Jie Hou(✉)

College of General Aviation, Jingchu University of Technology, Jingmen 448000, China
houjie8445@163.com

Abstract. This paper studies the application of ID3 algorithm in college students' mental health education. ID3 algorithm is a classical decision tree data mining algorithm for data classification and analysis, which is widely used. College Students' mental health and school mental health education have once again become a hot spot in the society. In order to make the mental health education in Colleges and universities embark on the road of healthy development, the first and basic thing is to have a clear understanding of the existing model of mental health education in Colleges and universities, and establish a new model of mental health education suitable for yourself on this basis. In this way, it is conducive to the practical and effective development of all aspects of mental health education. The ID3 algorithm is applied to the evaluation system of College Students' mental health education. By establishing the decision tree model of the attribute of the teaching quality of mental health education, the mental health education is classified according to the teaching quality, and the incentive effect is formed through affirmation or negation, so as to promote the improvement of the teaching quality of mental health education and the development of education and teaching.

Keywords: ID3 algorithm · Mental health education · Data processing

1 Introduction

Colleges and universities undertake many functions such as scientific research, talent education and training, cultural communication and exchange. It is an important foundation and guarantee for social development. The occupation of mental health in Colleges and universities has the characteristics of comprehensiveness, creativity and competitiveness. Therefore, how to make full use of the database Data mining and other technologies provide more effective support for the evaluation and management of mental health teaching quality, which has important theoretical research and application value.

It is necessary for college students to take the course of mental health. First of all, as college students, we are just coming of age. Although the appearance of college students looks very mature, in fact, their hearts have not reached the appearance of maturity. As the saying goes: you just look mature. Then, the university is equivalent to a small society. College students who have just entered the society may not adapt to the university life that is inconsistent with the learning and life mode of junior and senior high school

M. A. Jan and F. Khan (Eds.): BigIoT-EDU 2022, LNICST 465, pp. 286–297, 2023.
https://doi.org/10.1007/978-3-031-23950-2_31

[1]. The university is not a place that only tells you about your academic achievements, but a place that allows you to develop comprehensively. In high school, as long as you study hard and get good grades, you can get the likes of teachers and classmates and get scholarships, but not in college. Therefore, some students can't adapt to the rhythm and mode of college life after entering the University for several months. Secondly, in many key universities, every year, more or less, students end their precious lives because they are too stressed to find someone to talk to (they are likely to suffer from depression). In fact, this is the psychological pressure is too great, leading to their psychology began to become unhealthy. And now the psychological pressure of college students comes from all aspects, so now many universities hold psychological drama competitions every year in order to make students strengthen their awareness of mental health.

In fact, mental health and physical health are equally important to people. Physical health, mental health is not healthy, sooner or later will be stuffy out of the disease, but mental health is beneficial to physical health (mainly good mentality).

Strengthening college students' mental health education is not only an important way to improve college students' psychological quality, promote their healthy growth and cultivate all-round socialist builders and successors, but also one of the important tasks to strengthen and improve college students' Ideological and political education. As a college student, we should understand the meaning and standards of mental health, study the characteristics and influencing factors of College Students' mental development, clarify the theories and principles of College Students' mental health, and grasp the materials and methods of College Students' mental health education [2]. People with mental health should have relatively correct cognition, good emotions, strong will and healthy personality psychology.

When dealing with the new environment, there will be all kinds of emotions, such as irritability, depression, depression, etc. when dealing with problems that cannot be solved by yourself, don't bear alone, let alone dig into the horns, be enterprising, take the initiative to adjust your psychology and understand how to adapt to the environment; We should correctly deal with setbacks and hardships in life; We should learn to change the way of learning, make learning simple and pleasant, and then determine a realistic goal; Know what you know and know what you don't know. Stop without knowing, and act with knowing. We should learn to understand everything in reality from the heart, and be at ease once, so as to make ourselves have strong adaptability.

As early as the late 1970s and early 1980s, machine learning researcher J. Ross Quinlan developed the decision tree algorithm (iterative dichotomy). It can be used as a classification algorithm in the form of rule-free induction or prediction tree. Among many classification methods in the field of data mining, decision tree classification method has been welcomed for its advantages of high speed, high precision, intuitive and easy to understand and simple generation mode.

This paper applies the decision tree algorithm to the evaluation system of mental health teaching quality in Colleges and universities, constructs the decision tree of related attributes of mental health teaching quality, classifies mental health according to the level of teaching quality, and forms an incentive effect through affirmation or negation, so as to promote the improvement of mental health teaching quality and the development of education and teaching.

2 Related Work

2.1 ID3 Algorithm Principle

Decision tree is a machine learning model, which is very intuitive and easy to understand, and is closely combined with data structure. Our learning threshold is also very low, which is much simpler than those models that are prone to a pile of formulas.

In fact, we often use decision trees in our lives, but we don't find them ourselves. The essence of decision tree is a bunch of if else combinations. For example, we go to a small stall to buy watermelon. How do the fruit vendors do it? Pick up the watermelon, roll it around, take a look, and then stretch out your hand to pat it. You will know whether the watermelon is sweet or not. We remove the factors related to these actions and extract the core essence [3]. Basically, there are three:

1. The color on the surface of watermelon is often sweet if it is bright The sound of watermelon beating is crisp and often sweet. 3 Whether there are melons and vines in watermelon, those with vines are often sweet.

These three are obviously not equal, because the sound of beating is the most important, perhaps followed by the surface color, and finally melons and vines. So when we choose, we must first listen to the sound, then look at the melons and vines, and finally look at the color. We abstract the logic and organize it into a tree structure, so it becomes a decision tree.

The word "information entropy" is very puzzling. Its original English text is information entropy, which is actually equally difficult to understand. Because entropy itself is a concept in physics and thermodynamics, which is used to measure the non-uniformity of object dispersion. In other words, the larger the line, the greater the degree of dispersion of the object, which can be simply understood as the more scattered. For example, if we overturn a box of sorted table tennis balls in the room, the table tennis balls inside will obviously be scattered to all parts of the room [4]. This scattered process can be understood as the process of direct increase.

Information line has the same meaning, which is used to measure the disorder of a piece of information. The larger the number, the more disorganized the information. Otherwise, the more conditioned the information is. Information is derived from the famous information science masterpiece "information theory", whose author is the famous Shannon. But this word is not original by Shannon. It is said that it was taken by Feng Neumann, the father of computer. The meaning of his name is also very simple, because everyone doesn't understand what this word means.

We have explained this concept in detail when we introduced cross line before. Let's briefly review it. For an event x, assuming that its probability of occurrence is p (x), the amount of information of the event itself is:

$$I(X) = -\log_2 P(X) \tag{1}$$

Then the meaning of information entropy is actually the expectation of the amount of information, that is, multiplying the amount of information by its probability:

$$H(X) = -P(X)\log_2 P(X) \tag{2}$$

Similarly, suppose we have a data set in which there are k samples, and the proportion of each type of sample is p(k), then if we regard this proportion as probability, we can write the information of the whole set:

$$H(D) = -\sum_{i=1}^{K} P(K)\log_2(P(K)) \tag{3}$$

According to the definition of information entropy, we can know that if the data becomes pure, the information entropy should be reduced. The more you reduce, the better the effect of segmentation. So we found a way to measure the segmentation effect, which is information gain. According to the definition of information gain, we can easily figure out the whole process of decision tree establishment. Every time we select the segmentation feature, we will traverse all the features. Each value of the feature corresponds to a subtree. We find the feature with the largest gain after segmentation by calculating the information gain [5]. After the upper structure is created, continue to build the tree in the form of recursion until the data set after segmentation becomes pure or all features are used.

This algorithm is called ID3 algorithm, which is also the most basic construction algorithm of decision tree. Here is a small detail. According to the definition of ID3 algorithm, each segmentation selects the feature rather than the value of the feature. And each value of the feature selected as the segmentation feature will establish a subtree, that is, each feature will appear only once in the decision tree. Because after one use, all values of this feature will be used up. But expand the dict to get the tree structure in Fig. 1 below:

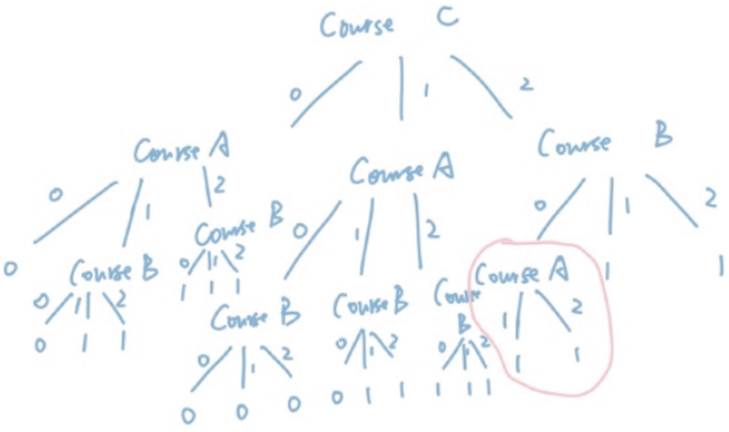

Fig. 1. Dict expanded decision tree structure

In the decision tree algorithm, the classification weight of the root node is the highest and decreases successively downward; Selecting the feature with the strongest classification ability as the root node can greatly improve the classification efficiency. The classification ability of each feature is quantified by information gain. The greater the

feature information gain, the stronger the classification ability, that is, calculate the information gain of each feature point in the data set. The feature point with the largest information gain is used as the root node of the decision tree and recurs downward in turn.

2.2 Students' Mental Health

Mental health is an important guarantee for college students to master cultural and scientific knowledge. With a good attitude, they can not only achieve good learning results, but also benefit their lifelong development. College students should achieve mental health, strive for physical and mental health, be able to read relevant psychological counseling books, carry out psychological counseling if necessary, carry out regular physical exercise, have a good time, enhance their psychological quality, and meet the challenge with good psychological quality [6]. Figure 2 below shows the characteristics of College Students' mental health data.

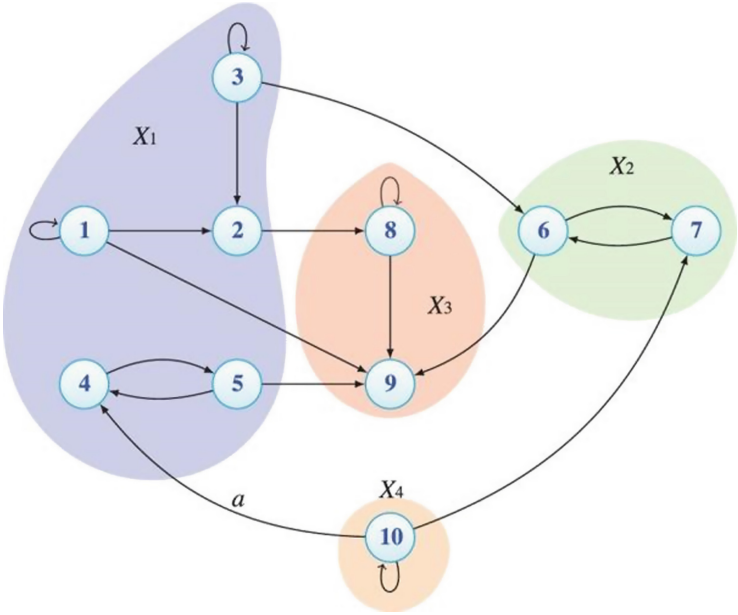

Fig. 2. Characteristics of mental health data of College Students

Good interpersonal relationship enables people to obtain a sense of security and belonging, give people spiritual pleasure and satisfaction, and promote physical and mental health; Bad interpersonal relationships make people feel depressed and nervous, bear loneliness and loneliness, and hurt their physical and mental health. Therefore, we college students should cultivate good communication skills. We should adhere to harmonious interpersonal relations, be willing to communicate with others, and be able to live in harmony with others with an attitude of understanding, tolerance, friendship,

trust and respect. Through interpersonal communication, we should recognize the social responsibilities of college students and cultivate the habit of abiding by discipline and social ethics. Enhance the ability of psychological adaptation, be able to work together and cooperate with others, adhere to a coordinated relationship with groups, and ensure the healthy psychological development of college students.

Love has become a common phenomenon among modern college students. College students must establish a correct outlook on love. In love, we should follow: take love as the foundation, take noble interests as the driving force for the development of love, respect each other's choices, freedom, rights and personality in love, abide by their duties, be loyal and single-minded, and treat each other sincerely. We should take understanding as the premise, dedication as the material and being loved as the result. College students should avoid the negative impact of love, learn to use the skills of love, learn to clarify emotional values, make love decisions, pay attention to love communication, and perform the duty of love.

Psychology is the subjective response of organisms to the objective material world. Psychological phenomena include psychological process and personality. Human psychological activities will have a process from occurrence to development and then to disappearance. Psychology is also the function of our human brain and the subjective reflection of our human brain on the objective material world. People's psychology includes psychological process and personality psychological characteristics [7]. The process of psychological development is an active process, not what we think of as a passive process. So what is mental health and what are the standards of our college students' mental health? In a broad sense, mental health is an efficient, satisfactory and sustainable mental state. In a narrow sense, mental health refers to the process of people's basic psychological activities, which has content integrity and coordination, that is, from recognition to emotion, then to will, behavior and personality integrity and coordination, so that people can adapt to the society and keep pace with the society. A healthy person should have normal intelligence, be able to control emotions, be emotionally stable, have a strong will, be normal, adapt well, have good self-awareness and interpersonal relationships, and have a healthy personality. Some people think that psychological problems are physiological diseases when they have not been exposed to the discipline of psychology. In fact, psychological problems are not different from physiological diseases. Psychological problems are a series of problems caused by people's internal mental factors, precisely a series of problems caused by the control system of the central nervous system of the brain. It will indirectly change people's character, world outlook and emotion, so, Psychological problems also have a great impact. We should pay attention to psychological problems. In case of psychological problems, please consult a psychologist more often, so as to make our new youth grow healthily.

3 Student Mental Health Data Processing

The influence of parental rearing patterns on children's psychological and behavioral development and health can not be ignored. A survey of 165 freshmen to junior students in the branch of Medical College of Zhengzhou Railway Vocational College shows that the parental rearing style has a direct impact on the mental health of children and adolescents.

Intimacy, care and consideration make children form personality characteristics such as enthusiasm, self-confidence, self-esteem, independence, perseverance and positive emotion, which promotes the development of mental health. Exclusion and negation make children form the psychology of humiliation, inferiority, self sin, helplessness, caution, indecision, fear of being rejected, don't know what love is, and don't know how to be loved. Parents' excessive intervention in the way of education makes children form personality characteristics such as lack of self-confidence, excessive self-control and excessive dependence. The harsh educational methods make parents feel inferior, helpless and insecure. Severe punishment is especially easy to lose children's self-esteem and produce anxiety and obsessive-compulsive symptoms.

Using the clinical symptom Checklist-90 (SCL-90) and self-made family information questionnaire, 5090 freshmen of grade 2020 in a university in Chengdu were investigated. The results showed that family structure had a significant impact on the overall level of mental health and the three factor scores of interpersonal sensitivity, depression and psychosis. In terms of student communication, students from sojourning families are more sensitive than students from other families [8]. There are more students from single parent families and their overall mental health is poor. College students from single families have serious psychological barriers, such as depression, loneliness, inferiority complex and rebellious psychology. Due to the imperfect family structure, children often lose the possibility of communicating with their parents, and can only get partial care compared with children in ordinary families; Their hearts will be covered, prone to sensitivity and inferiority.

A harmonious family atmosphere can make children happy physically and mentally, reduce psychological pressure, and is conducive to the healthy development of body and mind. In disharmonious families, quarrel, doting, indifference and domestic violence do harm to people's physical and mental health in varying degrees, especially the two family atmospheres of doting and violence.

The survey found that the mental health level of students with financial difficulties was significantly lower than that of students with financial difficulties. The lower the income, the lower the mental health level of college students; Most of the students from poor families come from rural areas. They are under great economic and psychological pressure. Most poor college students are unable to change the current situation for a time due to financial constraints, limited vision and lack of communication ability. They are prone to sensitivity, inferiority complex, depression, anxiety, interpersonal tension and other characteristics. At the same time, college students living in affluent families and poor students will also have more mental health problems.

Over fitting means that the model performs very well in the training data set, but the data outside the training data set can not fit the data well. It may be that the number of samples is too small, and the attributes of the samples cannot replace all the overall data [9]. Therefore, even if the samples can be well distinguished, the data outside the samples may not be well distinguished.

Decision tree algorithm is easy to understand and interpretable. It is a very common and excellent machine learning algorithm, which can be classified or regressed. Now many of the best integration models are also based on decision trees. Therefore, the

series of decision tree algorithms can not be bypassed by machine learning. It requires very systematic and profound learning and understanding.

In information theory, the greater the information gain of an attribute, the stronger the ability of the attribute to reduce the entropy of the sample. In other words, determining this attribute will make the system more stable and orderly (the smaller the entropy, the more stable the system is), and the higher the purity of the partition.

No matter how many features a data set has, only one feature can be selected each time the data set is divided. Then which feature is selected as the reference attribute for division for the first time to classify the data faster?

The answer must be the feature with the best classification ability, but the question comes, how to judge which feature has the best classification ability? A more perceptual concept can be introduced, that is, purity. The purer after splitting, the better. There are three common methods to measure purity. Different measurement methods may lead to different splitting [10].

ID3 algorithm: iterative dichotomizer 3, the third generation of iterative binary tree, is the first proposed decision tree algorithm, which uses information gain as the splitting criterion.

4 Application of ID3 Algorithm in College Students' Mental Health Education

Modern college students are active, eager for knowledge, pursuing progress, enterprising, ambitious, and able to integrate their personal ideals and aspirations with national rejuvenation The great cause of national prosperity has become the mainstream of College Students' psychological development. However, with the advent of knowledge economy and information society, great changes have taken place in people's ideas and lifestyles. Science and technology are changing with each passing day. With the intensification of social competition, college students are facing great psychological pressure in the period of "psychological weaning". College students are increasing day by day because of the extreme psychological problems. According to the relevant survey of the Ministry of education, 20.33% to 25.63% of the students in school have varying degrees of psychological problems, psychological problems and psychological problems.

There are two parts of the data from the school's educational administration system and the performance management system (one part is the data from the school's examination and the other part is the performance management system); The other part is the questionnaire of students, including the mastery of College Students' mental health course before learning, learning interest, learning attitude, learning habits, theoretical learning effect, computer learning effect and computer time after class. A total of 321 samples were collected. Data preprocessing can ensure the data set quality of data mining. The method of data preprocessing includes data cleaning, data reduction and data transformation. In the student survey data, some data items are not filled in, resulting in the invalidity of the questionnaire. Among the scores of College Students' mental health courses derived from the educational administration management system, there are two kinds of special data. One is the students who have been retaken. Their poor completion of various teaching tasks usually leads to the poor scores of College Students' mental

health courses; The other is the absence or delay of a small number of students due to illness, taking various examinations and other special reasons. The final score of College Students' mental health course is 0. This kind of data will be deleted. There are many attributes in the collected data, but not every attribute is closely related to the task of data mining. Data reduction can obtain a smaller attribute set than the original data, and the new attribute set does not affect the results of data mining. First, convert the usual score and total evaluation score of College Students' mental health course scores into characters. The conversion rules are as follows: 80–100 are represented by a, 60–79 by B and 0–59 by C. Secondly, delete the student number and name attributes irrelevant to data mining.

Calculate the information gain [11]. Taking "College Students' interest in mental health courses" as an example, the calculation of information gain is as follows: the first step is to calculate the subset information entropy divided by attribute values.

Attribute "College Students' interest in mental health courses" = "a", of the 31 samples, 28 samples with a total rating of "yes" and 3 samples with a total rating of "no".

IS11,S12) = I(28,3) == 0.4586.

Attribute "College Students' interest in mental health courses" = "B", I (S21, S22) = 0.2602.

Attribute "College Students' interest in mental health courses" = "C", 1s31, S32) = 0.4395.

Attribute "College Students' interest in mental health courses" = "d", 1s41, S42) = 0.

Calculation information expectation:

$$\max \sum_{I} \left[U^{I}(X^{I}) - C^{I}(X^{I}) \right] \tag{4}$$

E (College Students' interest in mental health courses) = S11, S12 + I (S21, S22) + I (S31, S32) + I (S41, S42) = 0.3315.

Calculate information gain:

$$O_v = \sum_{u \in N[v]} w_{u,v} x_u \tag{5}$$

Gain (College Students' interest in mental health courses = I (S1, S2) – e college students' interest in mental health courses) = 0.3982–0.3315 = 0.0667 step 5: determine the root node of the decision tree. According to the principle of ID3 algorithm and the above calculation results, the test attribute is "usual score". Each attribute value corresponds to a branch, and other attributes are also divided and constructed in this way [12].

The sixth step is to further divide the branch nodes according to the above method. In order to prevent the training data from over fitting and delete the leaf nodes with low probability, the post pruning is used to prune the generated decision tree to form the final decision tree, as shown in Fig. 3.

Students' identification with the school is an important indicator to evaluate a school's management level and education level, and improving college students' identification

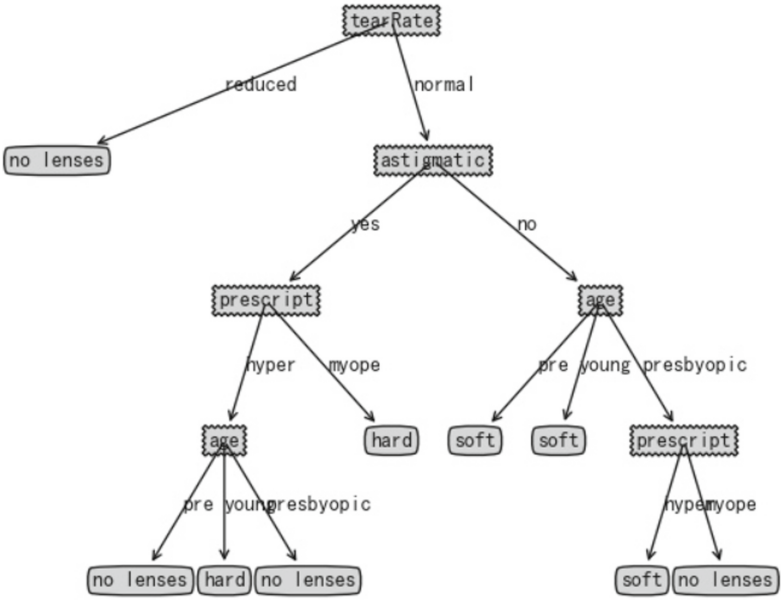

Fig. 3. Final decision tree of mental health education

with the school through mental health education is also one of the ways. First, the school should improve the relationship with students, strengthen interaction and pay attention to students' learning, life and psychological state [13]. For example, a long-term activity of "President and Secretary listen to me" carried out by a university in Guizhou allows students to fully express their needs for the school and implement the reform, which solves the problems that may arise in students' life and study, At the same time, it strengthens students' participation and sense of belonging. Second, colleges and universities should strengthen the linkage mechanism between Counselors and student work offices, pay attention to the psychological state of each student, and once problems are found, they should jointly solve them in time according to the specific situation of students. For example, when counselors find that students are frustrated in love and start to get tired of learning, they need to link with the student office to seek the advice of psychological counselors and give corresponding psychological counseling, Provide targeted help to students. Third, some students from rural areas are prone to inferiority complex in the face of many students with superior urban conditions due to family conditions [14]. However, we should not pay too much attention, otherwise it will be counterproductive, which requires the school to carry out some inspirational lectures or publicity, establish examples of excellent character and learning, recognize the aura on ourselves through the comparative effect, and establish the life idea of "struggling life is valuable", so as to gradually eliminate the psychological state of inferiority and improve the sense of recognition and belonging to the school.

5 Conclusion

Different mental health education models are put forward by experts through practice and summary in different periods and under different educational conditions. They provide guidance and models for the development of teachers' mental health education. However, in the current situation, we still need to explore more practical and effective methods of mental health education for college students in China. However, we still need to explore more practical and effective methods of mental health education for college students. This paper introduces data mining technology into the field of mental health education in Colleges and universities, classifies psychological data by using the classification technology of data mining, and realizes classification mining by using ID3 algorithm. The experimental results achieved the expected goal, excavated a series of psychological classification rules, provided a scientific reference for the mental health education in our school, and made a certain contribution to the improvement of the level of mental health education.

References

1. Hu, N.: Constructing a new mode of mental health education for college students under the network environment. **2018**, 5528–5539 (2021)
2. Xie, W.: Big data analysis on the management content of college students' mental health education. In: 2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) (2021)
3. Pottinger, A.M., Passard, N., Gordon Stair, A.: Using faith-based organisations to promote mental health education to underserved populations in Jamaica. Health Educ. J. **80**(5), 461–471 (2021)
4. Liu, Q., Liao, X.: Research on university mental health education based on computer big data statistical analysis. In: 2021 2nd International Conference on Big Data and Informatization Education (ICBDIE) (2021)
5. Javed Mehedi Shamrat, F.M., Ranjan, R., Hasib, K.M., Yadav, A., Siddique, A.H.: Performance evaluation among ID3, C4.5, and CART decision tree algorithm. In: Ranganathan, G., Bestak, R., Palanisamy, R., Rocha, Á. (eds.) Pervasive Computing and Social Networking. Lecture Notes in Networks and Systems, vol. 317. Springer, Singapore (2022)
6. Arlis, S., Defit, S.: Machine learning algorithms for predicting the spread of Covid19 in Indonesia. TEM J. **10**(2), 970–974 (2021)
7. Li, P., Zhou, Y.: An analysis of english listening and speaking teaching mode based on ID3 algorithm. In: Xu, Z., Parizi, R.M., Loyola-González, O., Zhang, X. (eds.) CSIA 2021. AISC, vol. 1343, pp. 773–777. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-69999-4_107
8. Kuang, Y.: Innovative development of combination of digital media technology and virtual reality technology. In: Zheng, X., Parizi, R.M., Loyola-González, O., Zhang, X. (eds.) Cyber Security Intelligence and Analytics: 2021 International Conference on Cyber Security Intelligence and Analytics (CSIA2021), vol. 2, pp. 210–216. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-69999-4_28
9. Gu, Z., He, C.: Application of fuzzy decision tree algorithm based on mobile computing in sports fitness member management. Wirel. Commun. Mob. Comput. **2021**(6), 1–10 (2021)
10. Singh, G., Sharma, S.: Impact of digital transformation of education system on the mental health of students. In: Innovative Teaching Practices for Education Par Excellence in Online Era (2021)

11. Scholarworks, S., Shulski, A.: How to reduce mental health crises by improving training and education for law enforcement first responders (2021)
12. Deshpande, D., Mhatre, C.K.: A study of impact of online education on mental health and academic performance of children of project affected people studying at undergraduate level in Navi Mumbai. Revista Gestão Inovação e Tecnologias **11**(4), 3866–3875 (2021)
13. Stegenga, S.M., Sinclair, J., Knowles, C., Storie, S.O., Seeley, J.R.: Lived experiences of mental health in higher education: a comparative analysis of determinants to supports and services. Am. J. Orthopsychiatry **91**(6), 738–750 (2021). https://doi.org/10.1037/ort0000575
14. Bynum, W.E., Sukhera, J.: Perfectionism, power, and process: what we must address to dismantle mental health stigma in medical education. Acad. Med. **96**(5), 621–623 (2021)