East Asian Math. J.
Vol. 38 (2022), No. 1, pp. 065–084 YNMS-logo
http://dx.doi.org/10.7858/eamj.2022.006

RESEARCH ON THE DEVELOPMENT OF COLLEGE STUDENT EDUCATION BASED ON MACHINE LEARNING

—TAKE THE PHYSICAL EDUCATION OF YANBIAN UNIVERSITY AS AN EXAMPLE

Yu Quan, Wei-Jie Guo, Lin He, and Zhe-Zhi Jin*

Abstract. This paper is based on Yanbian University's physical test data, and uses statistical analysis methods to study the relationship between college students' physical test scores to promote college physical education. Firstly, using gender as categorical variables, we conduct a general analysis of students in different majors and different grades, and obtain the advantages and disadvantages of male and female college students; then we use Decision Trees and Random Forest algorithms to conduct modeling analysis to provide valuable suggestions for relevant departments of the university. the aiming of this research analyzing about the undergraduates physical test is that giving universities the targeted suggestions to improve the college graduate rate and promote the overall development of higher education, lay the foundation for achieving universal health.

1. Introduction

With the rapid development of China's economy, people's living standards are gradually improving, so the people's demands on the quality of life are also increasing. The continuous development of health care and education has led to a gradual enhancement in people's health awareness. College education, which can provide a steady stream of spiritual power and vitality for Chinese society, is an indispensable source of continuous development and progress in China' future. Therefore, all aspects of college students' education play roles profound and far-reaching.

In the pursuit of universal health today, health education for college students is also the focus of national attention. Due to the limitations of environment and conditions, physical education was neglected by the public in the past. Nowadays, the most intuitive phenomenon is that undergraduates are poor in health

Received June 2, 2021; Revised September 27, 2021; Accepted January 11, 2022.
2010 *Mathematics Subject Classification.* 11A11.
Key words and phrases. all-round higher education, college students physical test, DT, RF.
*Corresponding Author.

(pISSN 1226-6973, eISSN 2287-2833)

and lack of health awareness. In order to speed up the overall development of higher education, the Ministry of Education has promulgated a series of reform policies for physical education teaching in colleges and universities, such as the major decision that undergraduates need to be qualified for physical education in order to graduate, and physical education must be regarded as an essential part of higher education.

There are currently relevant researches on physical education for college students, of which [1, 2, 3] are analyses of the current states of physical health test scores and their influencing factors, but in those papers, there is no scientific systematic analysis of the college students physical test itself.[4] has selected 10 factors affecting physical fitness performance, and used Factor Analysis to judge the main factors of the main factors relatively.[5], through gender differences in detailed indicators of physical test, puts forward targeted advice for the corresponding physical education.[6] based on fuzzy theory and physical measurement indicators, quantify and evaluate students' physical health status to establish a reasonable classification model. And Decision Tree, Random Forest and other machine learning methods are applied suitably to the physical health test analysis of college students for [7, 8].

2. Selection of indicators and processing of data

2.1. Selection of indicators

The result of college students' physical education is reflected in the annual physical health test of college students. The physical health test of college students is an important part of students' comprehensive development of core literacy, because of its mandatory and universal nature, its data is also very useful for measuring the health level of college students. As the physical health test of college students' indicators is drawn up by the *National Health and Physical Fitness Standard*, this paper is to classify the indicators, and the results of the selection of the indicators are as follows:

Table 1. classification of indicators

physical measurement indicator classification		test metrics in detail
morphological indicator		body mass index (BMI)
functional indicator		vital capacity
physical fitness indicators	explosive quality	50-meter dash/standing long jump
	flexibility quality	Sit and Reach
	strength quality	pull-ups(male)/one-minute sit-ups (female).
	endurance quality	1000-meter race(male)/800-meter race(female).

2.2. Data pre-processing

The data, highly accurate, was exported by the Physical Fitness Testing And Health Center of Yanbian University in Jilin Province in China. So there's no need to take into account duplicate values and outlier points. Because some physical testing programs for males and females are different, the results of physical fitness tests are divided into two parts based on gender. And the data types of the selected target variables are different, to be detail, except gender is qualitative data, the remaining indicators data are quantitative.

(1) Missing value processing

1000-meter race is for males, 800-meter race is for females respectively in terms of endurance quality. In terms of strength quality, males test pull-ups, males test one-minute sit-ups. So the boys' 800-meter race, pull-ups results, girls' 1000-meter run, sit-up scores are empty, each data by gender will correspond to two empty data set values, the empty values be filled with 0.

(2)Non-feed variable processing

In this paper, the physical analysis of college students is divided into grades and majors, and the names, classes and other unimportant fields in the physical fitness test data need to be deleted.

3. general feature analysis

3.1. Test subjects and research methods

(1) Test object

The study was conducted on all students who took part in the physical health test of college students from the first to fourth undergraduate levels at present of Yanbian University, including 5476 males and 9404 females, the total number is 14880.

(2) Research methods

a. Literature Review Method

The literature review method is to grasp the latest research contents and achievements of the physical health test of college students through the way of literature search websites, such as CNKI and VIP, and to lay the foundation for the research of this paper by combining the theory and research significance of the actual background analysis.

b. Statistical Analysis Method

Statistical analysis is mainly based on the physical evaluation criteria of college students to obtain the health and physical fitness test data from a statistical perspective to carry out a comprehensive analysis, and the use of statistical computing software such as SPSS, python and others to proceed original data and study through modeling and other ideas for further processing and analysis.

c. Machine Learning Methods

In order to increase the graduation rate of college students, improve the undergraduate education rate and promote the all-round development of universities education, this paper uses Decision Tree (DT) and Random Forest (RF), flexible machine learning methods, to find out the relatively important factors that affect college students physical test scores.

3.2. the general situation of undergraduates physical fitness test

The overall physical test of college students explains by the table 2 below:

TABLE 2. score level of test set about male and female

gender	good	pass	Failed
male	63(1.15%)	2296(41.93%)	3117(56.92%)
female	2009(2.13%)	4010(42.64%)	5194(55.23%)

According to table 2, we can conclude that boys and girls in the overall performance of the failure rate is relatively close, women scored slightly better than men, but neither of them get the outstanding performance score. In a word, the overall performance is not satisfactory.

3.3. The comparison of upper grades and lower grades according to the gender

In order to development higher education, this paper regards the upper grades(junior and senior) as a reference and provides suggestions for the graduation rate of the lower grades(freshmen and sophomore). Therefore, we make a comparison in the performance of the upper and lower grades in the college students physical health test firstly.

(1)Senior grades vs. Junior grades in Physical Test Scores in males

TABLE 3. normality test (standing long jump as an example)

	Kormogolov-Sminova			Shapiro Wilk		
	statistics	degree of freedom	Significance	statistics	degree of freedom	Significance
standing long jump of junior grades	.060	2653	.000	.985	2653	.000
standing long jump of senior grades	.049	2653	.000	.984	2653	.000

a. Riley's significance correction

First of all, doing the normality test of the target variables, taking the standing long jump as an example (as is shown in the table 3 above), concludes that the variable standing long jump does not satisfy the normal distribution. It is

same for males that 50-meter dash, pull-ups, 1000-meter race, Sit and Reach are all disobedient to the normal distribution. So in order to make a comparison of lower grades (freshmen and sophomore) and upper grades (junior and senior) about male in the college students health test, we use Mann-Whitney U test, which is one of the non-parameter tests and does not require samples to satisfy normality and homogeneity of variance, to test indicators for contrast, and the test results are as follows:

TABLE 4. hypothesis test summary(standing long jump as an example)

Null Hypothesis	test	significance	result
In the category of grade, the distributions of long standing jump are the same	Independent sample Mann-Whitney U test	.000	reject the null hypothesis

The non-parametric test results for each target variables are as follows:

TABLE 5. hypothesis test summary (male)

category	project	test	significance	result
lower grades / upper grades	50 meters	independent sample Mann-Whitney U test	.000	reject the null hypothesis
lower grades / upper grades	standing long jump		.000	reject the null hypothesis
lower grades / upper grades	Sit and Reach		.000	reject the null hypothesis
lower grades / upper grades	1000-meter race		.000	reject the null hypothesis
lower grades / upper grades	pull-ups		.000	reject the null hypothesis

Every variable in table 5 rejects the original assumption that there is no difference between the physical test scores of upper graders and those of lower graders.

According to the conclusion above, we reuse the data to contrast the grade level of the lower and upper grades by the chi-square test and the results are shown in the table 6 below:

Refer to *the college students sports health tutorial*, we know that track and field sports include four aspects: sprinting, long-distance running, jumping, throwing. In the physical health test of college students, the proportion of track and field sports is relatively large, while in the physical education curriculum of college students, track and field is only a part of the sports.

TABLE 6. the detail score level and the chi-square test in each project(male)

project	grade	Excellent(%)	Good(%)	Pass(%)	Failed (%)	χ^2	p
50-meter dash	lower grades	253(8.96%)	178(6.31%)	2018(71.48%)	374(13.25%)	587.3	0.000
	upper grades	49(1.85%)	77(2.90%)	1468(55.33%)	1059(39.92%)		
standing long jump	lower grades	8(0.28%)	29(1.03%)	911(32.27%)	1875(66.42%)	16.36	0.001
	upper grades	12(0.45%)	32(1.21%)	984(37.09%)	1625(61.25%)		
Sit and Reach	lower grades	259(9.17%)	422(14.95%)	1639(58.06%)	503(17.82%)	41.19	0.000
	upper grades	294(11.08%)	301(11.35%)	1696(63.93%)	362(13.64%)		
1000-meter race	lower grades	23(0.815%)	62(2.20%)	1310(46.405%)	1428(50.58%)	89.22	0.000
	upper grades	27(1.018%)	28(1.055%)	938(35.356%)	1660(62.571%)		
pull-ups	lower grades	7(0.248%)	19(0.673%)	220(7.793%)	2577(91.286%)	12.93	0.005
	upper grades	7(0.26%)	16(0.60%)	143(5.4%)	2487(93.74%)		
score	lower grades	0(0%)	50(1.77%)	1266(44.85%)	1507(53.38%)	44.16	0.000
	upper grades	0(0%)	13(0.49%)	1030(38.82%)	1610(60.69%)		

Combined with table 5 and table 6, the physical health test results of undergraduate boys are analyzed, and we know that the failure rates of the pull-up ares the highest whether in the lower grades or upper ones, and they are both higher than 90%. The project pull-ups need the upper limb strength, which indicates that the deficiency of upper limb strength training for undergraduate males is eye-catching, which also shows that there are in-comprehensive exercises and imbalance phenomena in the physical education curriculum of college students.

Compared with the others, males have the highest excellent rate in Sit and Reach and a relatively lower failure rate in it. And undergraduate males perform best in this project on the whole. Sit and Reach is a measure of the flexibility of the joints and body flexibility of physical exercise projects, combined to that college students are in their youth, we can conclude that Sit and Reach is of the least difficult as a physical test project.

In track and field, the outstanding rate of male students in the junior and senior grades of the 50-meter dash is lower than that of male students in the lower grades, and the failure rate is about three times that of lower grade students, and the overall performance of male students in the lower grades is better than that of male students in the senior grades. At the same time, it

also shows that seniors neglect the exercise of the extreme intensity work item after entering universities. 50-meters dash is closely related to explosiveness, flexibility, and speed, so for college senior graders, it is necessary to strengthen the extreme work program exercise. As one of the track and field projects, the excellent rate of the long jump is also very low, indicating that the male lower limb exercise still needs to be strengthened.

(2)Senior grades vs. Junior grades in Physical Test Scores in females:

Contrast of different grades in females by the chi-square test and the results are shown in the table 7:

TABLE 7. the detail score level and the chi-square test in each project(female)

project	grade	Excellent (%)	Good(%)	Pass(%)	Failed(%)	χ^2	p
50-meter dash	lower grades	25(0.5%)	117(2.5%)	3089(65.6%)	1479(31.4%)	245.5	0.000
	upper grades	11(0.2%)	50(1.1%)	2443(52.0%)	2190(46.7%)		
standing long jump	lower grades	31(0.7%)	88(1.9%)	1594(33.8%)	2997(63.6%)	56.83	0.000
	upper grades	29(0.6%)	150(3.2%)	1858(39.6%)	2657(56.6%)		
Sit and Reach	lower grades	872(18.5%)	735(15.6%)	2669(56.7%)	434(9.2%)	43.69	0.000
	upper grades	996(21.2%)	853(18.2%)	2543(54.2%)	302(6.4%)		
1000-meter race	lower grades	100(2.1%)	255(5.4%)	3098(65.8%)	1257(26.7%)	283.6	0.000
	upper grades	94(2.0%)	121(2.6%)	2483(52.9%)	1996(42.5%)		
one-minute sit-up	lower grades	44(0.9%)	126(2.7%)	3162(67.1%)	1378(29.3%)	152.4	0.000
	upper grades	20(0.4%)	66(1.4%)	2694(57.4%)	1914(40.8%)		
score	lower grades	0(0%)	122(2.6%)	2826(60.0%)	1762(37.4%)	108.9	0.000
	upper grades	0(0%)	78(1.7%)	2368(50.4%)	2248(47.9%)		

After the analysis of undergraduate girls' physical health test scores, we can see that the failure rate of one-minute sit-up for lower grade females is significantly lower than that of upper ones, but the excellent rate of sit-ups whether girls in junior or senior grades are very low similarly, and the scores are mostly concentrated in the interval 60-70. Sit-up needs waist and abdomen force, indicating that undergraduate female students for the waist belly training can be further strengthened.

Similar to boys, for Sit and Reach, girls have the highest pass rate and the highest rate of excellence compared with other projects. This also shows that girls are very good at flexibility.

In track and field, the pass rates of junior and senior about the projects: 50-meter dash and 800-meter race, in females, are inferior to the ones of lower grades. It is shown that junior and senior neglect the exercises of the extreme intensity work items and the strength endurance projects after university admission. Therefore, like males in junior and senior, female students in the same grades should also strengthen the extreme work project exercises and pay more attention to the usual endurance exercise at the same time. As one of the track and field events, the excellent rate of the girls in the standing long jump is also very low, the failure rate is high on the contrary. Overall, the failure rate for girls in the long jump is highest, about 60%, indicating that female students need to further increase their lower limbs and jump training.

4. theories of DT

A decision tree algorithm is a tree structure similar to a flowchart, whose rule is *if... then...* thought. Common Decision Tree algorithms are ID3 algorithms, C4.5 algorithms, and CART algorithms.

In this paper, a better CART decision tree algorithm is selected, which is also known as the classification regression tree algorithm.

4.1. Gini Index

The field selection indicator used by the CART decision tree algorithm is the Gini index, and the Gini Importance method is the default method for calculating relative importance used by the random forest model in scikit-learn python package. The formula for the Gini index can be expressed as

$$Gini\left(p_1, p_2, \ldots, p_K\right) = \sum_{k=1}^{K} p_k\left(1 - p_k\right) = \sum_{k=1}^{K}\left(p_k - p_k^2\right) = 1 - \sum_{k=1}^{K} p_k^2.$$

For an event, it has K possible different values, and p_k represents the probability that the k^{th} possible values will occur.

In practical applications, the empirical probability is used instead of the probability, so the empirical Gini coefficient can be expressed as

$$Gini(D) = 1 - \sum_{k=1}^{K}\left(\frac{|C_k|}{|D|}\right)^2.$$

Where the quotient value represents how often the k^{th} possible value occurs, $|D|$ represents all the sample points in the event, and $|C_k|$ represents the number of times the k^{th} possible value of the event occurs.

For Yanbian University junior and senior students physical fitness test of the overall score grade, there are two possible values, divided into pass and fail, so $K = 2$. Thus, the data set contains a sample of 7,633 students who are juniors or seniors, of which 4,264 passed and 3,009 failed, so its Empirical Gini index for the event of student physical performance passing is:

$$Gini(sum\_score\_grade) = 1 - \left(\frac{4264}{7633}\right)^2 - \left(\frac{3009}{7633}\right)^2 = 0.464.$$

4.2. Conditional Gini Index

When selecting a variable as a root node or intermediate node, you need to calculate the conditional Gini index, which is the expectation of the conditional Gini index under each value of a variable. Condition The dilytic principle of the Gini index is used, but discrete variables are calculated differently from continuous variables.

4.2.1. *Discrete Variable.* As a qualitative variable of physical fitness test score, Province, which is divided into Jilin Province and non-Jilin Province two values, can be directly divided into two categories, but for three or more different values of variables, it can not be divided into two categories, then can be packaged and processed, resulting in a number of different combinations. And then we need calculate each combination of the conditional Gini index, from which the minimum condition Gini index corresponding to the combination as the second class of the variable division.

The formula for the conditional Gini index can be expressed as:

$$\begin{aligned} Gini_A(D) &= \sum_{i,k} P(A_i)\, Gini(D_k \mid A_i) \\ &= \sum_{i=1}^{2} P(A_i) \left(1 - \sum_{k=1}^{K} (p_{ik})^2\right) \\ &= \sum_{i=1}^{2} P\left(\frac{|D_i|}{|D|}\right) \left(1 - \sum_{k=1}^{K} \left(\frac{|D_{ik}|}{|D_i|}\right)^2\right), \end{aligned}$$

Among them, $P(A_i)$ represents the probability of variable A in group i under binary division, the corresponding empirical probability is $P\left(\frac{|D_i|}{|D|}\right)$, and $Gini(D_k \mid A_i)$ represents the conditional Gini index of the k^{th} value in variable D, wherein $\frac{|D_{ik}|}{|D_i|}$ represents the frequency of the k^{th} value taken in the variable D within the group A_i.

4.2.2. *Continuous numeric variables.* The principle of Dichotomy is applied because the corresponding values of each variable are continuous, but the classification method of discrete variables is different from the former.

Suppose a numeric variable contains n observations, Sort it firstly, and then calculate the mean between the two adjacent values $\bar{x}_l = \frac{x_i + x_{i+1}}{2}$ so $n-1$ averages are got. This means, as a classification judgment value, the data set is split into two parts, and then calculating the size of the Gini index drop value can get $n-1$ Gini index drop values, from which the largest Gini index

drop value is selected, its corresponding mean is the second-class division of the variable.

During the growth process, the Decision Tree contains a decreasing amount of information from the root node to the last leaf node, and the amount of information from the root node decreases to 0 of each leaf node.

The decrease speed of Gini index of the dependent variable can measure the degree of influence of the independent variables on the dependent variable. To be explained, faster the decline, stronger the influence of independent variable on the dependent variable.

The speed of the descent can be expressed in the following formula:

$$\Delta Gini(D) = Gini(D) - Gini_A(D),$$

As can be seen from the above formula, for event A which is known, if the impact of the event A is greater on the event, the conditional Gini index is the smaller correspondingly, so the difference will be more obvious, which illustrates that the event's Gini index has dropped more.

Whether the dependent variable is discrete or continuous, in the selection process of the variable of the root node or intermediate node, the dependent variable is selected under the influence of the independent variables, which causes the Gini index of the dependent variable to drop the most.

After the Decision Tree growth is completed, the proportion of sample data in the final leaf node determines the prediction result. According to the principle of minority obeying majority, the leaf node is divided into the category which type of leaf node has the largest sample size. In this article, if the student physical test scores and pass the number of students accounted for a relatively large, it will be judged as passing, otherwise it is a failure.

4.3. Decision Tree pruning

Decision Tree in the modeling process is easy to result in over-fitting problems. In other words, the model in the training set has a higher prediction accuracy, but the effect on the test set is not very ideal. The training error is very small in the specific operation result, but the test error is much larger than the training error, which is a very large error value.

To solve overfitting, the decision tree is usually pruned.

There are usually two ways to make a tree's pruning, one is pre-pruning and the other is post-pruning. In this paper, pre-pruning method is used to solve the problem of fitting. Pre-pruning is to pruning the tree during its growth, such as limiting the maximum depth of the tree, limiting the minimum sample size contained in the middle or leaf nodes of the tree, and limiting the number of leaf nodes generated by the decision tree. The pre-pruning of the Decision Tree is already reflected in the corresponding python codes.

5. Theories of RF

5.1. The idea of a random forest

Random forests are integrated algorithms and made up of several fully grown CART decision trees. The data that makes up the multiple decision trees is generated at random, using Bootstrap sampling.

The basic idea of the algorithm is to use the multiple decision tree voting mechanism to solve the classification problem. Use the judgment of multiple decision trees as a vote to determine the type of sample to which the sample belongs, based on the principle of minority obedience to majority.

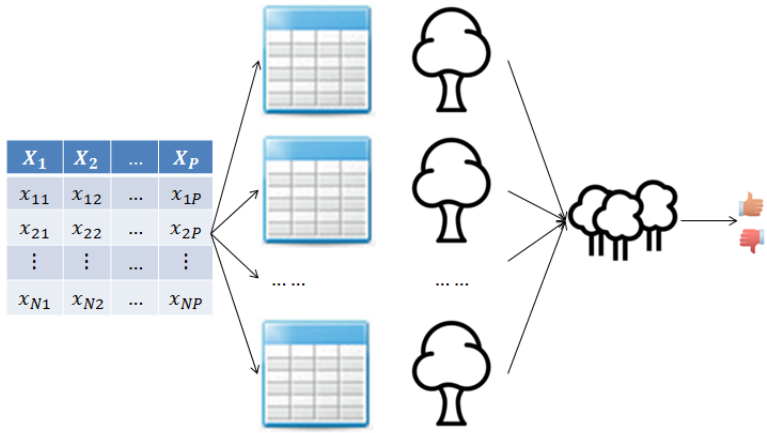

FIGURE 1. Algorithm of RF

Method steps:

1.Using Bootstrap sampling, k data sets are generated from the original data set, each containing N observations and P arguments.

2.For each data set, construct a CART decision tree, and in the process of building a sub-tree, the arguments used as the selection of node fields are randomly selected p arguments, not all arguments.

3.Let each decision tree grow as fully as possible, so that each node in the tree is as pure as possible, i. e. each sub-tree in a random forest does not need to be pruned.

4.For the random forest of k CART trees, the highest-vote category is used for the final judgment of the classification question using the voting method.

5.2. Randomness of random forests

Randomness of random forest is reflected in two aspects: One is that the training samples for each tree are random, and the other is that the classification

fields for each node in the number are also randomly selected. The introduction of two randomness can solve the problem that random forest is prone to over-fitting.

Decision tree and random forest algorithm, select 'BMI', 'VC BMI' (vital capacity body mass index), '50M' (50-meter dash), 'SLJ' (standing long jump), 'SAR' (Sit and Reach), 'SUS' (one-minute sit-ups), 'PU' (pull-ups), '800M' (800-meter race), '1000M' (1000-meter race), 'Male', 'Jilin, 'AGE' 12 variables as modeling variables.

The positive correlation between vital capacity and BMI was studied, and in order to reduce error and avoid multiple co-linearity, the vital capacity body mass index 'VCBMI' can be used.

In order to check the presence or absence of multicollinearity, the correlation heat map of the test values of each index of men and women in upper grades is done as follows:

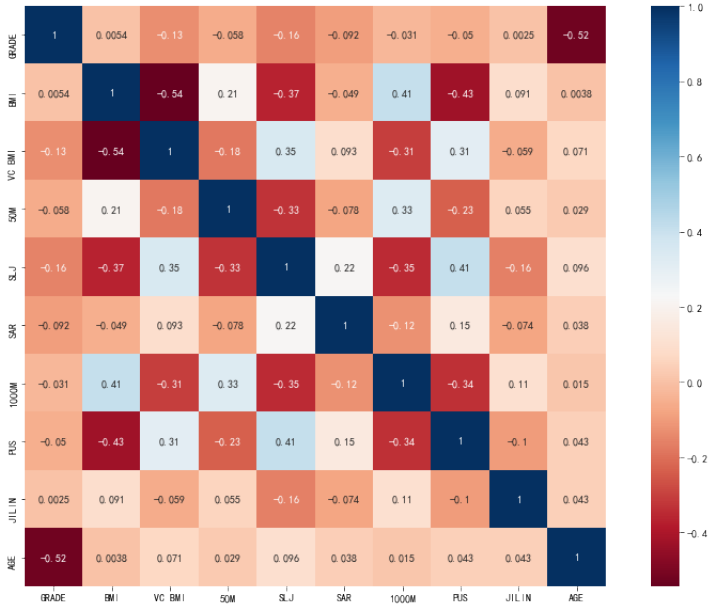

FIGURE 2. correlation heat map of senior grades(males)

Correlation heat maps of each test indicator of male and female students in the senior grades can be seen that the correlation coefficient between the variables of each indicator is less than 0.5, and there is no significant correlation variable, that is, there is no multicollinearity. The rationality of the selection and change of variables is reasonable.

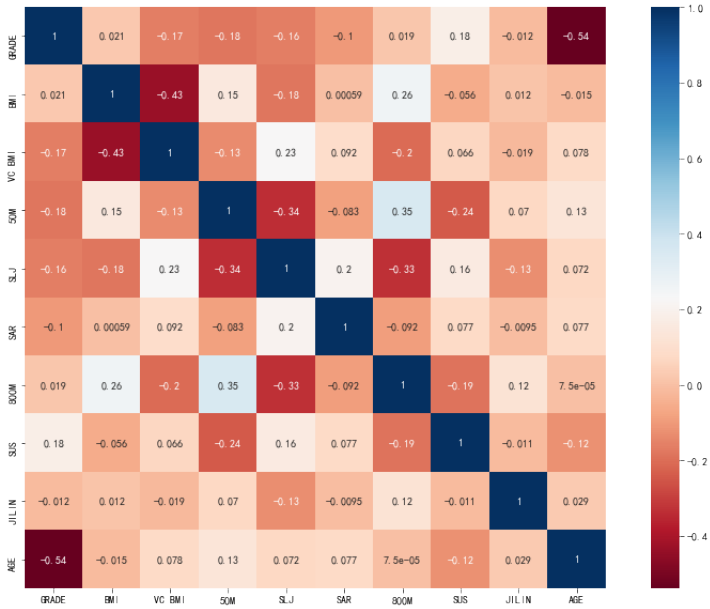

FIGURE 3. correlation heat map of senior grades(females)

5.3. Decision Tree Results Accuracy

(1)Accuracy

Firstly, we divide the data set into train data set and test data set, and their ratios are 0.75 and 0.25 respectively for model building.

And then the prediction accuracy of the decision tree model on the test set is 86.26 percent, the precision is 85.67 percent, the recall is 93.91 percent, the f1 score is 93.91 percent, so the prediction accuracy of the model is still very high. The confusion matrix of the decision tree model is shown in the figure 4,

And the number of students who passed the physical test correctly predicted is 85.67 percent in the actual number of passing students, and the proportion of students who failed the physical test correctly predicted is 86.73 percent in the actual number of students who failed, both of which exceed 80 percent and the ratio is very close,indicating that the model's matching effect on passing and failing characteristics is similar.

(2)ROC curve

The AUC value is 0.93, that means the area below the curve is as high as 0.93, which exceeds the common evaluation standard of 0.8. It can be considered that the fitted decision tree model is very reasonable and can better characterize the data.

(3)Part of the decision tree

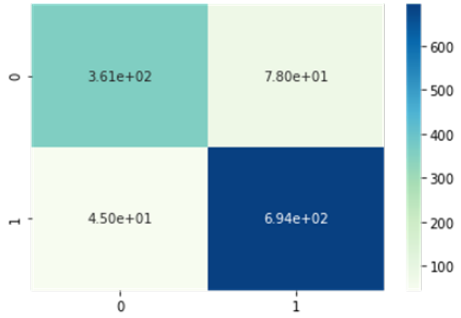

FIGURE 4. confusion matrix of the decision tree

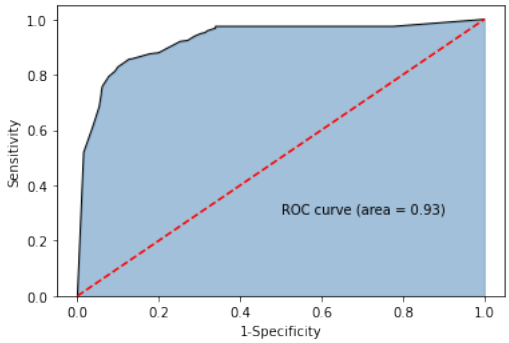

FIGURE 5. ROC curve of DF

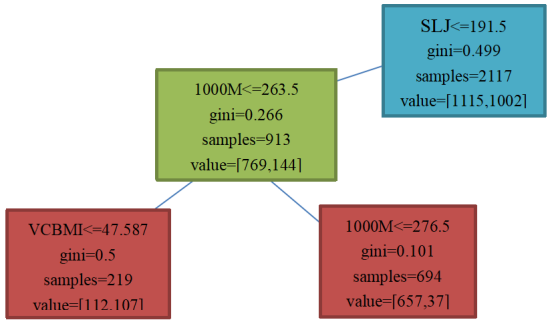

FIGURE 6. Part of DF

The variable selected by the male root node is standing long jump, and takes 191.5 as the split point, the corresponding left branch node is that standing

long jump is less than 191.5 cm, and the right branch node is that standing long jump is greater than 191.5 cm,

In the far left branch, for example, if the boy runs 1000 m for less than 263.5 seconds and had a VCBMI of less than 47.587, he will be a boy who passed the physical test.

To better reduce the likelihood of over-fitting problems, the integrated random forest algorithm is once again used for fitness test data. Random forest algorithm can not only speed up the operation speed, but also improve the accuracy of prediction, and better improve the prediction effect of the model.

5.4. Random Forest results

(1)Accuracy

The prediction accuracy of random forest model on test set obtained by python's random forest output is 92.07 percent, the precision is 91.08 percent, the recall is 94.99 percent, the f1 score is 94.42 percent, so the prediction accuracy of the model is very high, and better than that of decision tree, which is one of the advantages of random forest, that is, the prediction accuracy is higher. The confusion matrix of the random forest is shown in the figure 7,

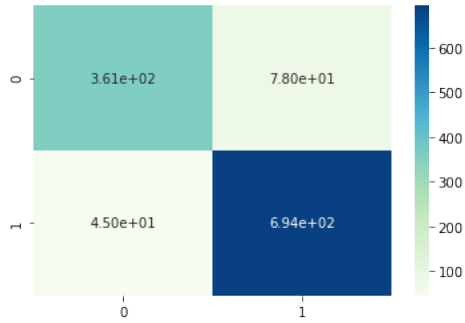

FIGURE 7. confusion matrix of the random forest

And the correct prediction of the number of students who pass the test in the actual number of passing students accounted for 91.08 percent, the correct prediction of the number of students who failed the test in the actual number of students without passing 92.86 percent, both ratios are more than 90 percent, indicating that the model of the physical fitness test data prediction effect is very good.

(2) ROC curve

According to figure 8, AUC value is 0.98, that is, the area under the curve as high as 0.98, not only far beyond the commonly used evaluation standard 0.8, but also close to 100%, it can be considered that the fitted random forest model is very reasonable, can exactly depict the characteristics of the data.

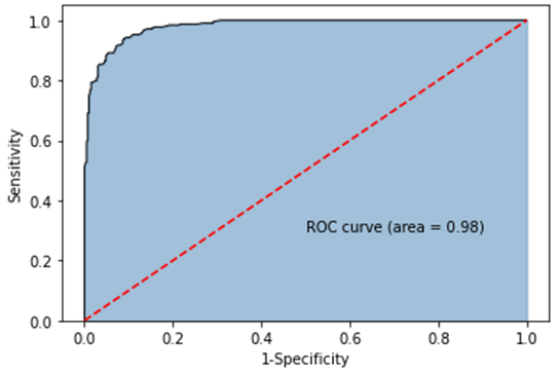

FIGURE 8. ROC curve of RF

5.5. Variable importance

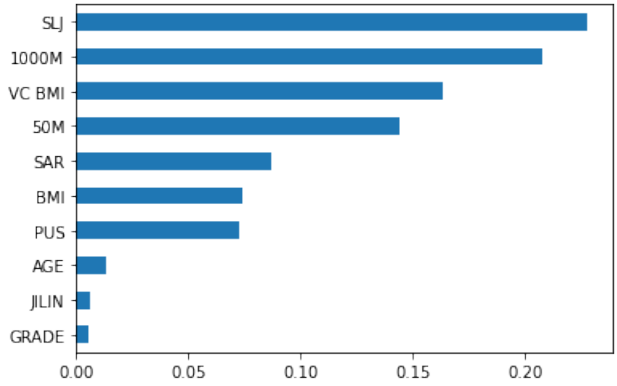

FIGURE 9. relative importance of various indicators(males)

Through the random forest model, it can be clearly seen that the standing long jump, 1000-meter race and vital capacity body mass index, 50-meter dash are the main test indicators that affect the physical health test scores of males in upper grades in college, and because the age gap between undergraduate students in adjacent grades is small, that grade and age account for relatively low degree of relative importance is more reasonable. And whether Jilin Province on the impact of physical test scores is also low, indicating that the physical test performance difference of the undergraduate senior boys from Jilin Province and the rest of the senior undergraduate boys is not big.

The figure 9 can be used to divide the influencing factors about males in senior grades of college students physical test into three gradients:

TABLE 8. three gradients of influencing factors(males)

relatively important variables gradient	corresponding variables
first gradient (>0.2)	standing long jump, 1000m race
second gradient (>0.15)	vital capacity body mass index, 50m dash
the third gradient (>0.05)	Sit and Reach, pull-ups, BMI

From the table above we can see that, in order to improve the pass rate of college students and the comprehensive quality of college students all-round development and the undergraduate graduation rate, regarding the upper grades (junior and senior) as a reference, the school should focus on increasing the college students standing long jump, 1000-meter race related lower limb exercise, endurance and other projects training for male college students.

The females' decision tree is similar to the random forest algorithm analysis, there's no need to narrate elaborately, so the relative importance of the 10 variables that affect females' physical performance are given as follows. In this way, the key factors affecting girls' physical performance are also judged.

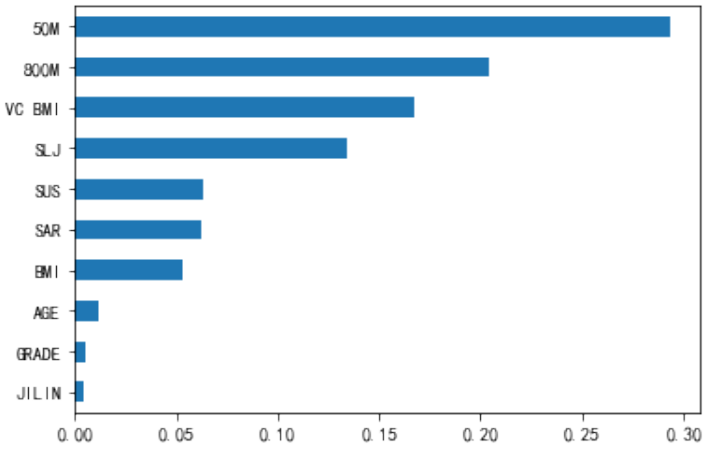

FIGURE 10. relative importance of various indicators(males)

Through the random forest model, it can be clearly seen that 50-meter dash, 800-meter race and vital capacity body mass index are the main test indicators that affect the physical health test scores of college students in the senior grade. Among them, 50-meter dash on the physical test results far greater than the remaining two, which shows the extreme importance of 50-meter dash for senior girls physical testing. Age, grade and province also had a smaller impact on physical performance. The indicator gradient is shown in the table below:

TABLE 9. three gradients of influencing factors(females)

relatively important variables gradient	corresponding variables
first gradient (>0.2)	50m dash, 800m race
second gradient (>0.15)	vital capacity body mass index, standing long jump
the third gradient (>0.05)	one-minute sit-up, Sit and Reach, BMI

It can be seen that females should increase the 50-meter dash, 800-meter race project training, to improve explosiveness and endurance.

Through the analysis of the physical examination of male and female students in the senior grades, it is obtained that both boys and girls need to improve their physical test scores by setting the four evaluation criteria of long jump, 1000-meter race, 50-meter dash and vital capacity. What is different is that the problem that being lack of explosiveness is more prominent for females, while that of males is more lacking in physical strength and endurance.

6. Conclusions and Suggestions

(1)Conclusions

I. The training of upper limb strength in physical education needs to be strengthened;

II. Current physical education in higher grades neglects the exercise of extreme intensity work items;

III. Girls have insufficient lower limb explosive power and physical coordination;

IV. Among the tested items, the most important factors affecting college male students physical test scores are standing long jump and 1000-meter running;

V. Among the tested items, the most important factors affecting college female students physical test scores are 50-meter dash and 800-meter running;

VI. Concluded from Random Forest, pull-ups is unimportant in those indicators that affecting physical test scores. According to the analysis results of general analysis, the possible reason is that boys perform badly in pull-ups and the scores they get in pull-ups are relatively concentrated.

(2)Suggestions

I. It is the performance of college physical education. For the weak projects of senior students, the junior students can be given the corresponding professional guidance and movement essentials, we should stimulate the students to participate in sports and improve the training. In this way, they can get rid of the "shackles" of the phone, and improve their health.

This teaching method is also applicable for teaching in other disciplines. According to the students' wishes and homework situation, combined with the previous teaching difficulties, it is the special focus of the subject of teaching

for junior students, and constantly practice and update. For subjects that students are good at, timely training exercises and scoring are needed, which can stimulate students' enthusiasm for learning. And usually scores is the embodiment of students' learning attitude and ability, and that training scores as part of the usual score is a very fair choice.

II. Many colleges and universities physical education compulsory courses are only offer in the lower grades, but overall the physical fitness of the upper grades compared to the lower grades has decreased, which shows that students' awareness of extracurricular exercise is not strong. Based on this, we recommend to set up some physical education elective courses in the upper grades, to provide targeted advice and guidance to college students.

In the curriculum aiming at junior and senior, we can arrange the basic course elective courses and increase the frequency of review of the basic knowledge points, which can not only help the advanced curriculum study, but also strengthen the students' self-planning and management consciousness.

III. For the arrangement of university curriculum, we can try to develop comprehensive quality evaluation based on full research, break the boundaries of the subject to evaluate students' text expression, logical reasoning, image thinking, innovation and so on, so as to provide reference for students' course selection and curriculum design, and formulate more scientific and diversified training programs in combination with students' interests.

References

[1] C.Zhang, *A Study on the Present Situation of College Students' Body Testing* , Colleges and Universities Journal of Jilin normal University (Natural Science Edition) **4** 154–156.

[2] S. Li, *Analysis on the Factors Affecting College Students' Body Test Achievement*, Contemporary Sports Science and Technology **10**(29):8–10.

[3] D. Chen. *Analysis on Influencing Factors and Countermeasures of College Students' physical Test* , Contemporary Sports Science and Technology **10**(18):62–64.

[4] G. Wang, *A Study on the Influencing Factors of College Students' Body Test Results Based on Factor Analysis*, Journal of Tong-Hua Normal University (2016), **37**(2):128–130,134.

[5] H. Wei, J. Xiao and Q. Shao, *Sex Difference Analysis of College Students' Health in Local Colleges—Based on the Data of 2019 Student Physical Test in a certain college*, Journal of Hubei University of Science and Technology **40**(05):85–89.

[6] J. Wang, *A Study on Classification and Comprehensive Evaluation of College Students' Physical Health Based on Fuzzy Theory*, Journal of Shang-Qiu normal University **36**(12):84–87.

[7] S. Zhao, *The Application of Decision Tree Technology—takes Wuhan Light Industry University as an example*, Contemporary Sports Technology (2019), **9**(25):28–29.

[8] M. Feng, H. Feng, Y. Zhang and L. Wang, *Random forest model analysis of physical health :568 students from Tongji University*, China Organizational Engineering Research (2019), **23**(23):3722–3728.

Yu Quan
Department of Economics and Management of Yanbian University, Yanji, China

Wei-Jie Guo
College of Science, Yanbian University, Yanji, China

Lin He
College of Science, Yanbian University, Yanji, China

Zhe-Zhi Jin
Department of Economics and Management of Yanbian University, Yanji, China
E-mail address: jinzhezhi@sina.com