WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n1.21
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Confirmatory Study on Prediction of Maximal Oxygen Uptake by Ensemble Learning Algorithm

Yao Tong [1], Danlan Wu [2, *], Jiaxu Zhang [1], Yefeng Yang [2], Briyanna Christina Etienne [3]

[1] School of Physical Education and Health, Zhaoqing University, Guangdong Province, China

[2] School of Computer Science and Software, School of Big Date, Zhaoqing University, Guangdong Province, China

[3] Graduate School of Business, Chung-Ang University, Korea

ABSTRACT

Objective: The prediction model of maximal oxygen uptake (VO_2max) was established by different ensemble learning methods. Compared with multiple linear regression, a common prediction method of VO_2max, the effectiveness of the ensemble learning model is verified, and screen out the optimal VO_2max prediction model. Methods: A total of 84 male college students, were recruited, informed of the purpose and procedure of the experiment, and passed the PAR-Q questionnaire screening. Cardiopulmonary Exercise Test (CPET), lung function, body composition test, VO_2max, minute ventilation volume (VE), tidal volume (Vt), weight, and other related indicators. After the correlation analysis of the experimental results, a VO_2max multiple linear regression model and an ensemble learning prediction model are computed. Results: The indicators that finally enter the multiple linear regression equation include: weight, Vt, VE. The regression equation is: VO_2max (ml/kg/min) = 52.455+0.005×VT (ml/min) + 0.111×VE (L/min)-0.441×weight (kg), and the regression coefficient test results is extremely significant (P<0.01). The VO_2max prediction model is computed by using the ensemble learning algorithm and among the five prediction models, it is found that Weighted average (R^2=0.7715) has the best effect, and higher than multiple regression model, there is no significant difference between the actual value and the predicted value VO_2max (P>0.05). Conclusion: In the case of a small sample size, the weighted average model predicts VO_2max the most accurately, compared with multiple regression models, the ensemble learning model can include more basic indicators, the evaluation of VO_2max in the body is more comprehensive, and the prediction results are more accurate.

KEYWORDS

Ensemble learning; Multiple linear regression; Maximum oxygen uptake; Predictive models

1. INTRODUCTION

Oxygen uptake, also known as oxygen consumption, refers to the body's ability to absorb oxygen from the outside world and it is closely related to the respiratory system. People with high oxygen uptake display better cardiac reserve and lung ventilation ability. Clinically, maximal oxygen uptake (VO_2max) was mainly used to determine whether the cardiopulmonary function was sound, reflecting the normal intake and utilization of oxygen in the cardiopulmonary function. In addition, people with high oxygen uptake exhibited better cardiac reserve and lung ventilation ability. Therefore, VO_2max is universally regarded as an independent predictor and gold standard for evaluating cardiopulmonary fitness (Powell et al, 1994). The measurement methods of VO_2max can be divided into direct measurement and indirect measurement. The direct measurement of VO_2max requires expensive instruments and high professional and technical requirements, and the exercise intensity in the test process has certain safety risks, easy to be affected by the subjective feeling of the tester. Indirect measurements include a 12-minute run, 6-minute step-up and step-down experiment, 20-meter run back and forth, etc (Black et al, 2016). but each method has its own specificity when it is used alone to predict VO_2max.

Another method is to construct the VO_2max prediction equation, which has been widely used in recent years. Due to the degree of obesity has a great influence on the accuracy and variability of the traditional prediction equation (the FRIEND equation, Wasserman equation, and European equation), which increases the variability of the predicted value of VO_2max (Myers et al, 2017) In order to ensure a more accurate prediction equation, other significant factors affecting VO_2max should be included to build a multiple linear regression model (Dong YaNan et al, 2017).

The latest research found that ensemble learning in machine learning algorithms can be used to build prediction models with the influence of multiple complex factors, and the prediction efficiency is better than that of traditional statistical methods, but different prediction models have differences in the reliability, accuracy and operability of the model prediction of each impact factor (Krittanawong et al, 2020; Baashar et al, 2022). Ensemble learning was originally proposed by HANSEN, aiming at a certain problem by building machine learning models multiple times to obtain the base model and integrating multiple base models to obtain the ensemble learning model. This method can significantly reduce the generalization error of the classifier (HANSEN et al, 1990). Common methods include boosting, weighted average, bagging, stacking, etc.

Based on the influencing factors of VO_2max, it is directly related to lung ventilation function and closely related to lipid metabolism, body composition, and other indicators reflecting the comprehensive function of the body (Warren et al, 2010). This study aims to push show VO_2max prediction models using different ensemble learning methods according to body composition and lung ventilation-related indicators and compare the differences between multiple linear regression methods and ensemble learning methods in predicting VO_2max, to screen out the optimal VO_2max prediction model, to accurately predict individual VO_2max level, and assess cardiovascular health status. This study also aims to provide low-risk, simple, and effective prediction methods, and further expand the application of artificial intelligence technology in the field of cardiovascular systems.

2. OBJECTS AND METHODS

2.1. Subject

Approved by the Ethics Association, 84 male college students in their junior year of university were randomly selected in this study. Their basic information was as follows: age (20.88±1.057) years old, weight (79.614±19.8006)kg, BMI (24.965±5.1961)kg/m2.

Inclusion criteria: good health, no recent illness, no disease or discomfort in respiratory system, cardiovascular system, or endocrine system; no other diseases, good health level, and PAR-Q questionnaire screening. Before the experiment, the physical test results of the subjects were collected and sorted, and the subjects signed the informed consent form. The staff informed the subjects of the purpose and procedure of the experiment, and fully informed the subjects of the problems that may occur during the experiment and the precautions for exercise.

2.2. Measurement Methods of Relevant Indicators

2.2.1. Exercise cardiopulmonary test

The subjects had no vigorous exercise 48 hours before the test and walked to the laboratory on an empty stomach in the morning. ERGOLINE sit-down power bicycles (Ergoline 100, Ergoline Gmbh, Lindenstraße 5, Deutschland and AEI MAX-II Metabolic Carts (AEI Technologies Inc.) were used to measure body gas metabolism during progressive load exercise.

First, according to the instructions, preheat the instrument for 30-60 minutes, and explain the test method, principle, and precautions to the subject. Secondly, the AEI MAX-II Respiratory gas metabolism analyzer is calibrated, including gas composition calibration and flow calibration, and air composition determination is performed to ensure the accuracy of O_2 and CO_2 content in the standard gas bottle. In the final test process, the subjects were asked to warm up for 15 minutes before starting the formal test. The basic information of the subjects was inputted and the height of the power bicycle was adjusted. The proper height was that after the leg was pushed straight, the angle between the thigh and the vertical plane was 15°. The starting power of the power bicycle was 25W and the movement frequency was 60r/min, which remained unchanged. The entire formal test took about 10 minutes and the tightness of the breathing mask was always paid attention to during the test to ensure the accuracy of the test results.

The VO_2max evaluation criteria are: the oxygen uptake peak occurs with the increase of exercise intensity, the respiratory quotient is above 1.15, and when the subjects feel exhausted subjectively. Stop testing when the above three points are reached.

2.2.2. Measurement of lung function

The index of lung function was tested by FGC-A+ portable lung function instrument. Subjects' Vital capacity (VC), Forced vital capacity (FVC), percentage of forced vital capacity in the first second (FEV1%), and Maximal voluntary capacity were measured ventilation (MVV), Minute ventilation volume (VE), Tidal volume (Vt), and vital capacity body mass index were calculated.

2.2.3. Body composition

Body composition was measured by TEZEWAT6200 human body composition analyzer on the day before the start of the experiment and the day after the end of the experiment.

Detection procedure: The subjects took off their clothes, shoes, and all other metal objects that may affect the results, wiped their hands and feet with electrolytic wet wipes, and stood on the foot electrode of the body composition analyzer, holding the electrode rod in hand, and keeping the arms in an extended state. Main outcome measures: Body weight, Body Mass Index (BMI), Waist to Hip Ratio (WHR), Body Fat Rate (BF%), and basal metabolic rate.

3. CORRELATION ANALYSIS AND PREDICTION MODEL

3.1. Correlation Analysis Between Indicator Factors and Maximal Oxygen Uptake

Through the correlation analysis of various index factors affecting the maximal oxygen uptake in this study, it can be seen from Table (1) that: VT, VE, and FVC were significantly correlated with maximal oxygen uptake ($P<0.01$), VC and BMI were significantly correlated ($P<0.05$).

Table 1. Pearson correlation analysis between index factor and maximal oxygen uptake

Independent variable	Dependent variable	Mean value ± Standard deviation	Person correlation	Sig.
VO_2max (ml/kg/min)	Age	20.88±1.06	.060	.586
	VT (ml/min)	2290.74±426.68	.527**	.000
	VE (L/min)	102.26±22.60	.529**	.000
	FVC (ml/min)	3464.39±930.51	.393**	.000
	VC (ml/min)	5363.81±967.66	.262*	.016
	Weight (kg)	79.61±19.80	.213	.051
	BMI (kg/m^2)	24.97±5.20	.264*	.015

Note: * At 0.05 level (two-tailed), the correlation was significant; ** At level 0.01 (two-tailed), the correlation was significant.

3.2. Multiple Linear Regression Prediction Model

In this study, the relative value of VO_2max was used as the dependent variable, and body weight, BMI, Vt, VE, FVC, and VC were used as independent variables for correlation analysis. The multiple linear regression model was constructed according to the correlation strength of each indicator, and the Forward method (F=0.05 entered, F=0.1 excluded) was adopted to enter the equation. The indicators that could enter the model eventually included: weight, Vt, and VE.

As shown in Table (2), the results of multiple linear regression analysis show that both the regression equation and the regression coefficient test results are extremely significant (P< 0.01), and there is a significant linear relationship between the explanatory variable and all the explained variables. After the adjustment of the fitting degree, the coefficient is 0.669, the fitting result is good, the Durbin-Watson coefficient is 2.32, and the multi collinearity is weak.

Therefore, the VO_2max(ml/kg/min) regression equation is finally obtained as follows: VO_2max (ml/kg/min) =52.455+0.005×VT (ml/min) +0.111×VE (L/min)-0.441×weight (kg)

Table 2. Multiple linear regression analysis of relative VO_2 Max (ml/kg/min)

P value			F inspect (F value/P value)	Degree of fitting Adjust R^2	Equation constant	Durbin-Watson
Weight(kg)	VT(ml/min)	VE(L/min)				
0.000	0.001	0.001	46.602/0.000	0.675	52.445	2.163

3.3. Ensemble Learning Prediction Model

3.3.1. Ensemble Learning

Ensemble learning in machine learning is to improve the performance of machine learning model by combining multiple base learners. It is a technique of combining multiple classifiers or regressors to achieve better performance than a single classifier or regressor. The advantage of ensemble learning is that it can reduce overfitting and improve prediction accuracy by combining multiple base classifiers or regressors. However, ensemble learning also has some disadvantages, such as the need for more computational resources, time, and data, and higher requirements for the quality of base classifiers or regressors. Common ensemble learning methods include bagging, boosting, stacking, and weighted averaging.

3.3.2. Ensemble learning model predic outcomes

Observing the various evaluation indicators in the table 3 and Figure 1-5 model test results, it was found that the weighted average method yielded the best results. In this paper, a simple analysis is conducted on this. When the differences between the base models are significant, stacking and blending methods are more suitable. This is because the weighted average method may not fully utilize the characteristics of each base model, while stacking and blending methods can train a secondary model by using the prediction results of multiple base models as features, to fully utilize the relationship between the base models. Bagging, stacking, and blending methods require a relatively large amount of data, while the amount of data used in this experiment is small, so the final results are relatively poor compared to boosting and averaging.

Evaluation index:

$$\text{RMSE: } \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\hat{y}_{\mathrm{i}}-y_{\mathrm{i}})^2}$$

$$\text{MAE: } \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|$$

$$R^2\text{: } 1-\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{\sum_{i=1}^{n}(y_i - \bar{y})}$$

Table 3. Results analysis of ensemble learning prediction models

	Advantage	RMSE	MAE	R^2	Shortcoming
Bagging	The variance is reduced, the efficiency of parallel implementation is higher than that of other integrated algorithms, and the sensitivity to outliers is reduced.	360.9785	289.7521	0.4264	When the amount of data is small, the improvement effect is not obvious and the efficiency is poor.
Boosting	The generalization error is low, easy to implement, the classification accuracy is high, and the adjustable parameters are few.	235.4094	176.9555	0.7561	Sensitive to outliers
Weighted Averaging	Reduce overfitting and make boundaries smoother.	227.8407	184.2711	0.7715	The smaller the differences between the models, the worse the fusion results.
Stacking	Use multiple cross-validations to make the model more robust.	272.3037	223.4549	0.6736	The requirements for data set size are higher.
Blending	Easy to use	293.1409	235.7127	0.6217	Using part of the data as a set aside verification, the efficiency of data utilization is relatively low. Easy to overfit

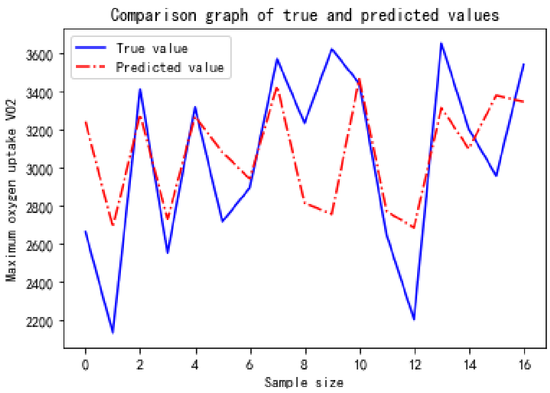

Figure 1. Bagging in Ensemble Learning

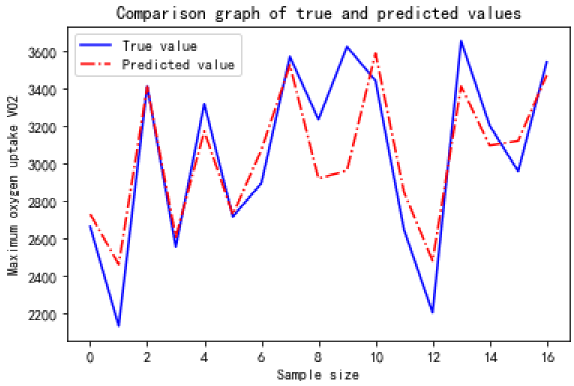

Figure 2. Boosting in Ensemble Learning

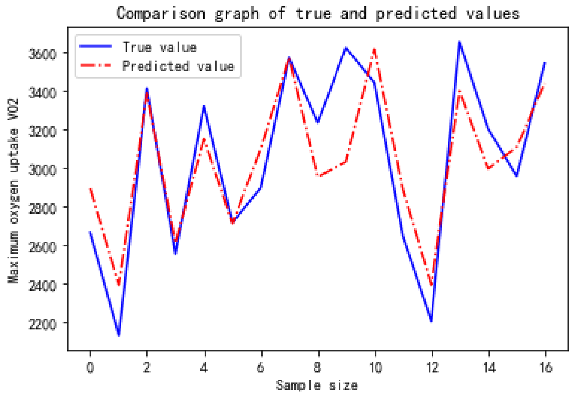

Figure 3. Weighted Averaging in Ensemble Learning

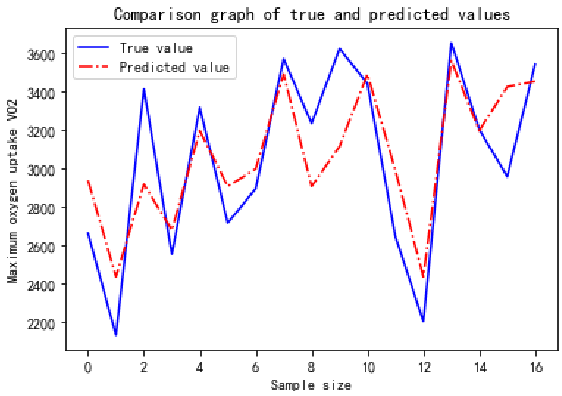

Figure 4. Stacking in Ensemble Learning

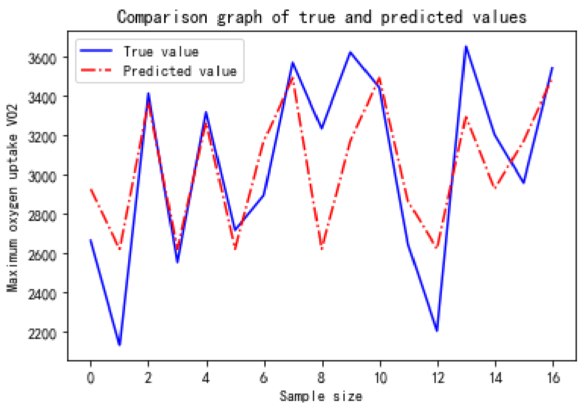

Figure 5. Blending in Ensemble Learning

4. MODEL TEST

4.1. Multiple Linear Regression Model Test

In order to test the accuracy of the multiple regression model, the VO_2max-related data of 17 subjects were selected and brought into the equation to calculate the predicted value. In order to facilitate the comparison with the ensemble learning model later, the relative value of VO_2max was converted into absolute value and the difference between the measured value and the predicted value was analyzed by paired sample T-test using SPSS. It can be seen from Table (4) that there is no significant difference between the predicted value calculated by introducing the equation and the measured value ($P > 0.05$), indicating that the prediction equation of the relative value of male college students VO_2max constructed in this study is valid (see Figure 6).

Table 4. Overview of prediction accuracy of multiple linear regression model

Prediction object	True value	Predicted value	P value
1	2663	2441	0.910
2	2134	3081	
3	3411	3781	
4	2554	2951	
5	3317	3049	
6	2716	2846	
7	2894	2522	
8	3568	2852	
9	3233	3200	
10	3621	3454	
11	3440	3280	
12	2646	2887	
13	2205	2883	
14	3651	3424	
15	3202	2896	
16	2957	3041	
17	3540	3342	

Note: A P value greater than 0.05 indicates that there is no significant difference between the measured value and the predicted value

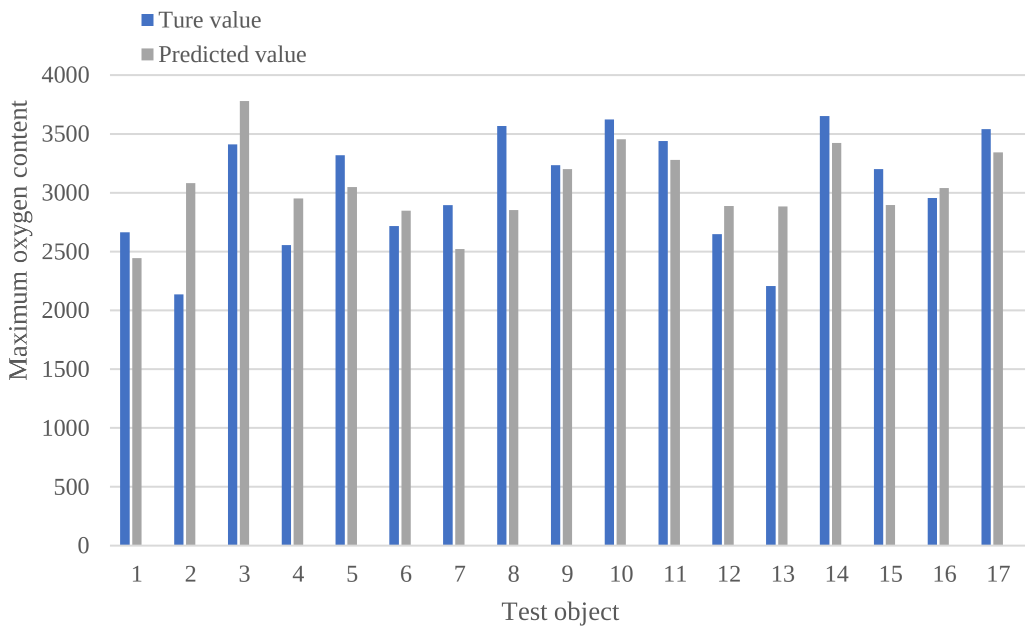

Figure 6. Comparative analysis of measured and predicted VO2max values of multiple linear regression model

4.2. Ensemble Learning Training Model Analysis

In the ensemble learning model of this study, the programming environment of Python 3.8.13 was used to train each ensemble learning model: bagging, boosting, weighted average, stacking, and blending. In this model training, 84 datasets were divided into 67 training sets and 17 test sets. In the end, the weighted average model was found to have the best prediction effect. Then, SPSS was used to conduct paired sample T-test for the measured value and predicted value to analyze the prediction accuracy. As can be seen from Table (5), there is no significant difference between the actual value and the predicted value VO_2max ($P > 0.05$), suggesting that the weighted average model is accurate in predicting VO_2max (see Figure 7).

Table 5. Weighted average fusion prediction accuracy overview

Prediction object	True value	Predicted value	P value
1	2663	2893	0.957
2	2134	2393	
3	3411	3385	
4	2554	2610	
5	3317	3147	
6	2716	2709	
7	2894	3094	
8	3568	3569	
9	3233	2953	
10	3621	3030	
11	3440	3613	
12	2646	2887	
13	2205	2393	
14	3651	3397	
15	3202	2995	
16	2957	3105	
17	3540	3437	

Note: A P value greater than 0.05 indicates that there is no significant difference between the measured value and the predicted value

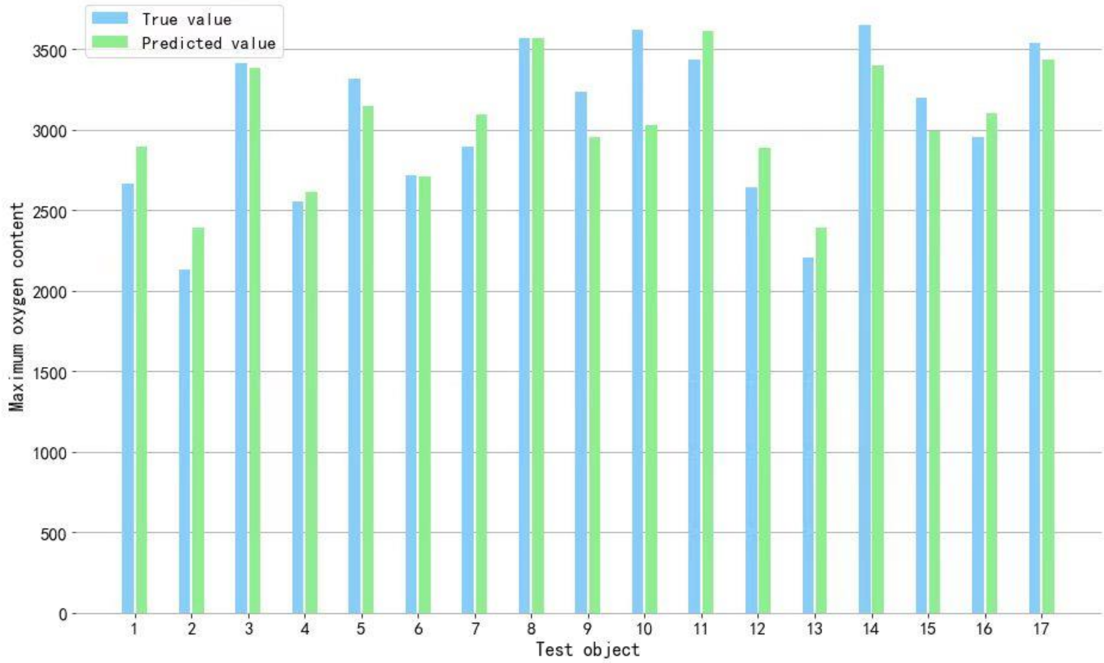

Figure 7. Weighted average model VO_2max measured value compared with predicted value

4.3. Comparative Analysis of Models

R^2 is one of the of the commonly used performance indicators in machine learning, used to evaluate the degree of model fitting. It represents the degree to which the model fits the data, also known as the coefficient of determination. On a scale of 0 to 1, it measures the difference between the observed and predicted values as a percentage of the total difference.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{\sum_{i=1}^{n}(y_i - \bar{y})}$$

The closer R^2 is to 1, the better the model fits the data; the closer R^2 is to 0, the worse the model fits the data.

The comparison results of different models show that (Table 6), boosting and weighted average model values are 0.7561 and 0.7715, respectively, indicating that the model fits well. Stacking and multiple linear regression model are in second place, and the values are 0.6736 and 0.675, respectively. The values of the bagging, blending, and blending models are all less than 0.65, and the fitting effect is worse than that of the above four models.

Table 6. Comparative analysis of fitting degree of each model

	Bagging	Boosting	Weighted Average	Stacking	Blending	Multiple linear regression
R^2	0.4264	0.7561	0.7715	0.6736	0.6217	0.675

5. DISCUSSION

5.1. VO_2max Prediction Model

The cardiorespiratory function is the core element of monitoring various components of physical health and is highly correlated with all-cause mortality and cardiovascular disease mortality in all populations. Those with high cardiorespiratory endurance levels show a lower incidence of mortality and have a higher correlation with metabolic syndrome which is mainly represented by obesity (Delgado-Floody et al, 2019). VO_2max is a golden index for evaluating cardiopulmonary function, but there are many inconveniences and safety risks in the measurement process. The purpose of constructing a prediction model is to simplify the measurement process of VO_2max on the premise of improving the prediction accuracy. Since there are many influencing factors and related indicators of VO_2max, the regression equations of VO_2max in different populations are different. Many studies have demonstrated the correlation between VO_2max and various indicators, including body composition (BMI, body fat percentage, lean body mass), heart function (heart rate index, blood pressure, cardiac output), lung ventilation function (tidal volume, vital capacity, ventilate per minute), etc. Through comprehensive analysis of these indicators, it can indirectly reflect the VO_2max level of the body (Wu Dong-zhe et al, 2023; Liang et al, 2022; Agostoni et al, 2017).

Current studies have shown that the addition of morphological indicators (BF%, BMI) and physiological indicators (HR, VE, Vt) to construct multiple linear regression equations can improve the accuracy of VO_2max prediction (Fudge et al, 2007). Gender, body weight, BMI, maximum heart rate, and load were selected to construct a regression model for predicting adolescent VO_2max. The multiple correlation coefficient was 0.884 and the good fit was 0.781, indicating that the model could effectively predict adolescent VO_2max and the maximum load was closely related to VO_2max. However, there is a multicollinearity problem between maximum heart rate and other indicators (Liang et al, 2022). It is suggested that heart rate is affected by age, gender, health level, individual exercise basis, and other factors, which limits the accuracy of estimating VO_2max with heart rate (Achten et al, 2003). In the study of constructing the VO_2max regression equation model based on heart rate index (HRI), there is no significant difference between the predicted data and the measured value at 40%VO_2max, but with the increase of VO_2max intensity, there is a phenomenon of high predicted value when 60% and 80% VO_2max exercise (Kang et al, 2020). The prediction result of this model takes into account variables such as age, gender, and weight, and is obviously restricted by exercise intensity. However, it is certain that for people who only need low-intensity exercise, measuring HRI value can more easily predict the oxygen uptake level of the body.

The effective capacity of the lungs determines the total oxygen volume of the body for pulmonary ventilation, pulmonary ventilation, and gas exchange, which directly affects the oxygen uptake level and blood oxygen-carrying capacity. Adding the indicators (VE and Vt) to the model to evaluate pulmonary ventilation function can improve the accuracy of the VO_2max prediction equation (Neder et al, 2006). In this study, a multivariate linear regression model of male college students' VO_2max was constructed based on Vt, VE, and body weight, and the regression equation and regression coefficient test results were extremely significant. There is a significant linear relationship between VO_2max and Vt, VE, and body weight, and the coefficient is 0.675. The model has a high fitting result, and the regression equation of maximal oxygen uptake (ml/kg/min) is calculated as follows: 52.455+0.005×VT (ml/min) + 0.111×VE (L/min)-0.441×weight (kg). The results of the model test show that there is no significant difference between the predicted value and the measured value, which further proves the effectiveness of adding the related index of pulmonary ventilation function into the regression equation. Consistent with the studies of Policarpo Barbosa et al (Policarpo Barbosa et al, 2020), in addition to the basic indicators, heart rate, and load indexes corresponding to minute ventilation (E) at the second ventilatory threshold were added. The VO_2max multiple linear regression model was constructed to evaluate the exercise-based and sedentary population of different genders (subjects aged 18-36 years), and the goodness of fit > 0.8. the mean Standard error of the estimate was 4.9ml/kg-1min-1, indicating high prediction accuracy. The difference is that the regression model in this study only covers male college students in predicting VO_2max. More targeted and accurate prediction equations need to be established in predicting VO_2max for different genders and different age groups.

5.2. Evaluation of Application Effectiveness of Ensemble Learning Model

The nonlinear and cross-learning modes used by machine learning make up for the shortcomings of traditional analysis methods in this respect and gradually become an important research method in the field of medical treatment and health (Schwalbe et al, 2020). Currently, it is widely used in the construction of cardiovascular disease risk prediction models (Wang Suhuai et al, 2021). Ensemble learning is a kind of machine learning, through the integration of existing machine learning models, to obtain higher performance analysis models. Motwani (Motwani et al, 2017) predicted the all-cause mortality of suspected CHD patients by Boosting the ensemble learning algorithm, and Boosting method enhanced the prediction effect of the ensemble model, with the model predicting all-cause mortality (ACM) with a Brier score of 0.08 and an AUC of 0.79. In predicting 2-year all-cause mortality, it was better than logistic regression analysis (AUC of 0.68), and in predicting 5-year all-cause mortality, Better than the Framingham risk score (FRS) and coronary computed tomographic angiography (CCTA) assessments alone. Frade (Frade et al, 2023) included 11 indicators that could explain the ability in four aspects: hemodynamics, anthropometry, physical activity, and lung function. The Support vector Regression (SVR) method in the Support vector machine (SVM) algorithm was used to construct the VO_2max prediction model for people with different aerobic capacity levels, and the average error was very low (SVR =0.038L/min). The model is reliable, and the reliability of the prediction model can be verified by a low MAE value and high R-value. In addition, in a study on the correlation between physical health parameters and workability performance of firefighters, the predictive variables included body weight, VO_2max, BF%, upper limb muscle strength, and lower limb muscle strength. The linear correlation was analyzed by the partial least-squares regression (PLSR) algorithm. Many-to-multiple linear regression modeling can be carried out, and the accuracy of the model for the prediction of the training set and the test set can reach 98.76% and 98.75% (Xu et al, 2020).

After the observation and the analysis of various evaluation indicators in Figure 1-7, we found that the weighted average method performed best in terms of effect. To further explore this phenomenon, this paper makes a simple analysis.

When there are significant differences between models, the methods of stacking and blending are more applicable. This is because the weighted average method may not make full use of the characteristics of each model, while the stacking method and blending method can make full use of the relationships between models by taking the predictions of multiple models as features and training a secondary model. It should be noted that the methods of bagging, stacking, and blending have higher requirements on the amount of data, while the amount of data used in this experiment is less. Therefore, compared with boosting and averaging methods, these methods performed poorly in this experiment. To sum up, the weighted average method has the best performance in the comprehensive evaluation, and its value is 0.7715, but when there are significant differences between models, the stacking method and blending method are more suitable choices. However, these methods have high requirements for data volume and need enough data support to give full play to their advantages.

In statistical methods, multiple linear regression algorithms can handle the linear fit degree in model prediction, so it has been used as the main algorithm for predicting VO_2max in many studies (Liu et al, 2014; Aandstad et al, 2021). Therefore, this paper verifies the multiple linear regression model and ensemble learning model respectively, and according to the comparison of model fit degree (see Table 6), Weighted average (R^2 =0.7715) and boosting (R^2 =0.7561) and the model have better prediction effect than the multiple linear regression model (R^2 =0.675). The fitting degree of the model is good and the Stacking model (R^2 =0.6736) is inferior to that of the multiple linear regression model, while the values of the bagging, blending, and blending models are all less than 0.65, indicating a poor fitting degree. In the selection of indicators, the five models of ensemble learning selected six indicators (body weight, BMI, Vt, VE, FVC, VC) that were significantly correlated with VO_2max for training, while there were only three types of multiple linear regression models (body weight, Vt, VE). For explanatory variables, ensemble learning models could include more basic indicators. The evaluation of VO_2max is more comprehensive and the prediction result is more accurate. From the perspective of sample detection data, the predicted value of the multiple linear regression model is low, 7 are higher than the measured value, and 10 are lower than the measured value. In addition, the predicted value obtained by the weighted average model is not significantly higher or lower. In addition, because ensemble learning has certain requirements for the amount of data, the small sample size in this study may be the main reason for the poor prediction effect of bagging and blending models.

6. DEFICIENCY AND PROSPECT

(1) Due to the large difference between male and female VO_2max, this study only models male VO_2max and more research on female VO_2max prediction models can be done in the future.

(2) Obesity is one of the important factors affecting VO_2max. However, due to the limited sample size, there is no separate comparison for the differences in the degree of obesity. In the future, the sample size will be further expanded to refine the prediction model. In addition, the sample size is too small, which may be the main reason that the results of other ensemble learning prediction models are not within the ideal range.

ACKNOWLEDGEMENTS

Funding: Zhaoqing University research fund project (No. QN202340)

REFERENCES

[1] Powell KE, Blair SN. The public health burdens of sedentary living habits: theoretical but realistic estimates. Med Sci Sports Exerc. 1994; 26(7):851-856.

[2] Black NE, Vehrs PR, Fellingham GW, George JD, Hager R. Prediction of VO2max in Children and Adolescents Using Exercise Testing and Physical Activity Questionnaire Data. Res Q Exerc Sport. 2016; 87(1):89-100. doi:10.1080/02701367.2015.1124969

[3] Myers J, Kaminsky LA, Lima R, Christle JW, Ashley E, Arena R. A Reference Equation for Normal Standards for VO2Max: Analysis from the Fitness Registry and the Importance of Exercise National Database (FRIEND Registry). Prog Cardiovasc Dis. 2017; 60(1):21-29. doi:10.1016/j.pcad.2017.03.002

[4] Dong YaNan, Qin Fei, Qu Chaoyi, et al. Research status and prospect of maximum oxygen uptake assessment and application. Chinese Journal of Sports Medicine. 2017; 36(8) : 731-739.

[5] Krittanawong C, Virk HUH, Bangalore S, et al. Machine learning prediction in cardiovascular diseases: a meta-analysis. Sci Rep. 2020; 10(1):16057. Published 2020 Sep 29. doi:10.1038/s41598-020-72685-1

[6] Baashar Y, Alkawsi G, Alhussian H, et al. Effectiveness of Artificial Intelligence Models for Cardiovascular Disease Prediction: Network Meta-Analysis. Comput Intell Neurosci. 2022; 2022:5849995. Published 2022 Feb 24. doi:10.1155/2022/5849995

[7] HANSEN L, SALAMON P. Neural network ensembles. IEEE Transaction on Pattern Recognition and Machine Intelligence. 1990; 12(10): 993-1001.

[8] Warren TY, Barry V, Hooker SP, Sui X, Church TS, Blair SN. Sedentary behaviors increase risk of cardiovascular disease mortality in men. Med Sci Sports Exerc. 2010; 42(5):879-885. doi:10.1249/MSS.0b013e3181c3aa7e

[9] Delgado-Floody P, Latorre-Román P, Jerez-Mayorga D, Caamaño-Navarrete F, García-Pinillos F. Feasibility of incorporating high-intensity interval training into physical education programs to improve body composition and

cardiorespiratory capacity of overweight and obese children: A systematic review. J Exerc Sci Fit. 2019; 17(2):35-40. doi:10.1016/j.jesf.2018.11.003

[10] Wu Dong-zhe, GAO Xiao-Lin, LI Chuang-tao, WANG Hao. The prediction model of VO2 Max of backpropagation neural network was constructed according to the optimum point of heart and lung. Chinese Tissue Engineering Research. 2023; 27 (8): 1224-1231.

[11] Liang Y, Yu Q. Study on Maximal Oxygen Uptake of Respiration and Heart Rate in Exercise Training Based on Regression Equation. J Healthc Eng. 2022; 2022:5961197. Published 2022 Mar 10. doi:10.1155/2022/5961197

[12] Agostoni P, Vignati C, Gentile P, et al. Reference Values for Peak Exercise Cardiac Output in Healthy Individuals. Chest. 2017; 151(6):1329-1337. doi:10.1016/j.chest.2017.01.009

[13] Fudge BW, Wilson J, Easton C, et al. Estimation of oxygen uptake during fast running using accelerometry and heart rate. Med Sci Sports Exerc. 2007; 39(1):192-198. doi:10.1249/01.mss.0000235884.71487.21

[14] Achten J, Jeukendrup AE. Heart rate monitoring: applications and limitations. Sports Med. 2003; 33(7):517-538. doi:10.2165/00007256-200333070-00004

[15] Kang J, Ratamess NA, Faigenbaum AD, et al. Use of Heart Rate Index to Predict Oxygen Uptake - A Validation Study. Int J Exerc Sci. 2020; 13(7):1705-1717. Published 2020 Dec 1.

[16] Neder JA, Stein R. A simplified strategy for the estimation of the exercise ventilatory thresholds. Med Sci Sports Exerc. 2006; 38(5):1007-1013. doi:10.1249/01.mss.0000218141.90442.6c

[17] Policarpo Barbosa F, Silva PE, Guimarães AC, Pernambuco CS, Dantas EH. Prediction of maximum oxygen uptake through incremental exercise testing using ventilometry: a cross-sectional study. Braz J Phys Ther. 2020; 24(4):365-372. doi:10.1016/j.bjpt.2019.07.002

[18] Schwalbe N, Wahl B. Artificial intelligence and the future of global health. Lancet. 2020; 395(10236):1579-1586. doi:10.1016/S0140-6736(20)30226-9

[19] Wang Suhuai, Li Jingjie. Advances in the clinical application of machine learning in cardiovascular diseases. Advances in Cardiology. 2021;42 (02): 144-147. The DOI: 10.16806 / j.carol carroll nki. Issn 1004-3934.2021.02.013.

[20] Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J. 2017; 38(7):500-507. doi:10.1093/eurheartj/ehw188

[21] Frade MCM, Beltrame T, Gois MO, et al. Toward characterizing cardiovascular fitness using machine learning based on unobtrusive data. PLoS One. 2023; 18(3): e0282398. Published 2023 Mar 2. doi:10.1371/journal.pone.0282398

[22] Xu D, Song Y, Meng Y, István B, Gu Y. Relationship between Firefighter Physical Fitness and Special Ability Performance: Predictive Research Based on Machine Learning Algorithms. Int J Environ Res Public Health. 2020; 17(20):7689. Published 2020 Oct 21. doi:10.3390/ijerph17207689

[23] Liu X. A study on the correlation between 800 meters, 1000 meters and 20mSRT and maximal oxygen uptake. Shanghai: East China Normal University. 2014.

[24] Aandstad A. Estimation of maximal oxygen uptake from the 3,000 m run in adult men and women. J Sports Sci. 2021; 39(15):1746-1753. doi:10.1080/02640414.2021.1898106