Predicting academic achievement from the collaborative influences of executive function, physical fitness, and demographic factors among primary school students in China: ensemble learning methods

Zhiyuan Sun[1,2], Yunhao Yuan[3], Xuan Xiong[4], Shuqiao Meng[5], Yifan Shi[1,2] and Aiguo Chen[6*]

Abstract

Background Elevated levels of executive function and physical fitness play a pivotal role in shaping future quality of life. However, few studies have examined the collaborative influences of physical and mental health on academic achievement. This study aims to investigate the key factors that collaboratively influence primary school students' academic achievement from executive function, physical fitness, and demographic factors. Additionally, ensemble learning methods are employed to predict academic achievement, and their predictive performance is compared with individual learners.

Methods A cluster sampling method was utilized to select 353 primary school students from Huai'an, China, who underwent assessments for executive function, physical fitness, and academic achievement. The recursive feature elimination cross-validation method was employed to identify key factors that collaboratively influence academic achievement. Ensemble learning models, utilizing eXtreme Gradient Boosting and Random Forest algorithms, were constructed based on Bagging and Boosting methods. Individual learners were developed using Support Vector Machine, Decision Tree, Logistic Regression, and Linear Discriminant Analysis algorithms, followed by the establishment of a Stacking ensemble learning model.

Results Our findings revealed that sex, body mass index, muscle strength, cardiorespiratory function, inhibition, working memory, and shifting were key factors influencing the academic achievement of primary school students. Moreover, ensemble learning models demonstrated superior predictive performance compared to individual learners in predicting academic achievement among primary school students.

Conclusions Our results suggest that recognizing sex differences and emphasizing the simultaneous development of cognition and physical well-being can positively impact the academic development of primary school students. Ensemble learning methods warrant further attention, as they enable the establishment of an accurate academic early warning system for primary school students.

*Correspondence:
Aiguo Chen
agchen@yzu.edu.cn
Full list of author information is available at the end of the article

Sun *et al. BMC Public Health*   (2024) 24:274

Page 2 of 13

Background

Academic success is a crucial predictor of students' future opportunities and aspirations [1]. Exceptional academic achievement not only boosts the self-confidence of primary school students [2] but also fosters a genuine interest in learning, thereby significantly contributing to their overall academic development [3]. In China, academic achievement holds immense importance as it determines the continuation of students' education, serves as an important benchmark for evaluating educational performance, and influences further educational pursuit [4]. Consequently, academic achievement has been a focal point of research within the Chinese educational landscape.

Executive function, as the core of primary school students' cognitive, emotional and social functions, has a significant impact in shaping their future quality of life [5]. Its abnormal development has been linked to various public health issues, such as autism spectrum disorders [6]. Psychological research emphasizes the critical influence of executive function on the academic achievement of primary school students [7]. Notably, primary school students with superior executive function tend to exhibit higher academic achievement [8], with the inhibition, working memory, and shifting aspects of executive function demonstrating positive correlations with academic performance [9]. Cognitive training has been shown to effectively enhance executive function and subsequently improve academic achievement [10]. Besides, physical fitness is an essential indicator of primary school students' health status, and its decline poses risks to cardiovascular health and overall future well-being [11, 12]. Existing research underscores a close association between physical fitness and academic achievement [13]. Factors such as body mass index, muscle strength, and cardiorespiratory function have proven to be reliable predictors of academic success [14–16]. Engaging in physical exercise has been identified as an effective strategy for improving academic achievement by enhancing physical fitness [17]. Additionally, demographic factors, such as sex, have been identified as contributors to variations in academic achievement [18], with girls generally outperforming boys, especially in Chinese language proficiency [19].

While executive function, physical fitness, and demographic factors have been extensively studied in relation to students' learning, it is essential to acknowledge that learning outcomes cannot be attributed solely to any individual factor. Academic achievement is a complex outcome that emerges from the interaction and integration of multiple factors operating as a cohesive unit or system [20]. In this respect, Gouveia et al. identified an interaction between physical fitness and demographic factors, indicating their joint predictive capacity for changes in the academic achievement of primary school students [21]. Another study also demonstrated that executive function acts as a fully mediating factor in the relationship between physical fitness and academic achievement [22]. However, existing research has predominantly approached the significance of executive function, physical fitness, and demographic factors from isolated perspectives, with limited exploration of their collaborative influences, emphasizing the need for further investigation in this area. In particular, it is not clearly elucidated which factors among executive function, physical fitness, and demographic variables play pivotal roles in collaboratively influencing the academic achievement of primary school students. Thus, this study will comprehensively examine multiple factors, including executive function, physical fitness, and demographic information, to identify the key factors that shape the academic achievement of primary school students.

Machine learning, a pivotal research approach in artificial intelligence, is designed to gather knowledge and patterns from intricate data, facilitating the prediction of future behaviors and trends [23]. While machine learning may pose challenges in establishing causal inferences compared to traditional statistical methods, its prowess in predicting complex data far exceeds statistical methods [24]. It focuses on achieving actual predictions and obtaining better prediction performance [25]. Over the past few years, machine learning has demonstrated remarkable success in the fields of psychology and sports science, gaining widespread recognition from researchers [26, 27]. Moreover, machine learning methods have been effectively employed in predicting academic achievement, enabling the early identification of students facing academic challenges [28]. Noteworthy applications include the use of machine learning algorithms like Naive Bayes and Decision Tree (DT) for predicting graduation grades [29], as well as the application of Support Vector Machine (SVM) and DT algorithms to distinguish students across various academic achievement groups [30]. These findings underscore the efficacy of machine

Sun *et al. BMC Public Health*   (2024) 24:274

Page 3 of 13

learning methods in predicting academic achievement and providing timely warnings for students at risk. However, relying on individual learners based on single algorithm often struggles to capture the comprehensive relationships among variables for optimal predictive performance. To address these limitations, scholars have proposed a practical solution: ensemble learning methods [31].

Ensemble learning, a model fusion approach that combines multiple learners, commonly employs methods such as Bagging, Boosting, and Stacking [32]. Ensemble learning methods have demonstrated superior generalization ability compared to individual learners [33]. In scenarios with limited sample sizes, selecting an appropriate learner can be challenging, and ensemble methods mitigate learning risks through the collective voting of each learner [34]. Guerrero-Higueras et al. utilized interactive learning platform data to predict academic achievement, finding that Bagging and Boosting methods outperformed individual learners like Naive Bayes and Linear Discriminant Analysis (LDA) [35]. Ban et al. similarly confirmed the effectiveness of machine learning in predicting academic achievement, with the Stacking method enhancing the predictive performance of individual learners, such as DT and Logistic Regression (LR) [36]. However, further validation is necessary to ascertain whether ensemble learning methods outperform individual learners in predicting the academic achievement of primary school students using executive function, physical fitness, and demographic factors. Our research has the potential to translate findings into practical applications, offering early academic warnings for primary school students.

As previously, the academic achievement of primary school students is intricately linked to their executive function, physical fitness, and demographic factors. However, it is unclear which factors among them play key roles in collaboratively influencing the academic achievement of primary school students. Additionally, while machine learning methods have gained widespread use in predicting academic achievement, there is a need for further validation to ascertain whether ensemble learning methods outperform individual learners when predicting the academic achievement of primary school students based on executive function, physical fitness, and demographic factors. Hence, this study sought to investigate the key factors that collaboratively influence academic achievement, drawing from executive function, physical fitness, and demographic factors. Subsequently, ensemble learning methods will be employed to predict academic achievement, and their predictive performance will be compared with that of individual learners. The primary assumptions of this study are enumerated as follows:

1) Sex, body mass index, muscle strength, cardiorespiratory function, inhibitory control, working memory, and shifting factors are key factors that collaboratively influence the academic achievement of primary school students.
2) Ensemble learning methods perform better than individual learners in predicting the academic achievement of primary school students using executive function, physical fitness, and demographic factors.

The outcomes of this study will contribute additional evidence to clarify the complex relationship among executive function, physical fitness, demographic factors, and academic achievement. Moreover, this study will provide effective methods for identifying academically disadvantaged primary school students, thereby contributing valuable insights to educational practices.

Methods

Participants

Participants were sourced from the Physical Fitness Monitoring Project in Huai'an City, China, comprising ordinary primary school students in grades 4–6 receiving compulsory education. Utilizing the cluster sampling method, 24 test groups were selected from five schools, yielding a total of 360 primary school students. The distribution across grades included 135 students in the fourth grade, 165 students in the fifth grade, and 60 students in the sixth grade. Each group met the following criteria: (i) willingness to disclose academic achievement, (ii) participation in executive function measurement, and (iii) completion of the physical fitness test. Seven students were excluded from the analysis due to missing data, as they did not fulfill the required measurement and testing procedures. Consequently, data from 353 participants were included in subsequent processing and analysis.

Measurements

Basic demographic factors (including sex and grade) were collected. Executive function, a multifaceted process encompassing distinct sub-functions such as inhibition, working memory, and shifting [37], was assessed using three computer-based neuropsychological assessments [38]. The stimulus presentation and response data collection were performed with the E-prime software platform. The flanker task served as an assessment of the inhibition aspect of executive function. This task comprised two trial types: congruent (e.g., LLLLL

Sun *et al. BMC Public Health*   (2024) 24:274

Page 4 of 13

or FFFFF) and incongruent (e.g., LLFLL or FFLFF). Participants were tasked with swiftly and accurately responding to the middle letter. Evaluation indicators for the flanker task included mean reaction time and mean accuracy in both congruent and incongruent trials. To evaluate the working memory aspect of executive function, the 1-back task was employed. In this task, participants carefully observed a letter (one of B, D, L, Y, O) presented on the screen and promptly judged whether the current letter matched a previously presented one. Evaluation indicators for the 1-back task comprised mean reaction time and mean accuracy. The shifting aspect of executive function was assessed using the more-odd shifting task. This task included two trial types: homogeneous (e.g., big/small or odd/even judgment) and heterogeneous (e.g., big/small-odd/even judgment shifting). Participants were required to quickly and accurately respond to different conditions. Evaluation indicators for the more-odd shifting task included mean reaction time and mean accuracy in both homogeneous and heterogeneous trials.

We primarily consulted the Chinese National Student Physical Fitness Standard (CNSPFS) [39, 40] and selected body mass index, muscle strength, cardiopulmonary function, speed, aerobic endurance, and flexibility to evaluate the physical fitness of primary school students. Body mass index was calculated as the ratio of weight (kg) to height (m^2). Muscle strength was assessed across three dimensions: upper limb, trunk, and lower limb. The standing long-jump test (cm) gauged lower limb strength, with participants placing their feet together behind the starting line and leaping forward as far as possible. For upper limb strength, the push-up test (times) required participants to maintain a straight body line, arms in a chest position, and hands slightly wider than shoulders, completing as many push-ups as possible. Trunk strength was evaluated through the sit-up test (times), where participants in a supine position with knees bent and feet flat completed as many sit-ups as possible within 1 min. Cardiopulmonary function was measured using the vital capacity test (ml), involving participants taking deep breaths and exhaling toward the blowing mouth until they were unable to exhale. Speed was assessed through the 50-m sprint test (s), where participants ran 50 m in a straight line on a flat playground runway at their maximum speed. Aerobic endurance was determined via the 50-m × 8 shuttle run test (s). This involved participants running back and forth between two parallel lines drawn 50 m apart, completing the course as quickly as possible four times, crossing each line with their feet. Flexibility was evaluated using the sit and reach test (cm). Participants sat on the ground, and the evaluator, ensuring their legs were straight, observed the extent of their forward reach as they slowly extended as much as possible.

Academic achievement assessments and grouping

In China, primary school education centers around core subjects such as Chinese, mathematics, and foreign languages. Thus, we gathered the final exam scores for these three subjects and aggregated them into a total score to represent the academic achievement of primary school students. To ensure comparability, we transformed the total scores into standard scores (with an average of 0 and a variance of 1) based on school and grade parameters [41, 42]. Accordingly, primary school students whose academic achievement standard score was below the average were identified as non-high-score (NHS) students and the remaining as high-score (HS) students (56.66%, $n = 200$) [42]. In addition, to avoid possible learning risks due to different sample sizes, we randomly sampled 150 students from each group [43]. Among the 300 participants, there were 151 boys and 149 girls. The majority were in fifth grade ($n = 136$, 45.33%), followed by fourth grade ($n = 111$, 37%), and sixth grade ($n = 53$, 17.67%).

Recursive feature elimination cross-validation

We applied the recursive feature elimination cross-validation (RFECV) method [44] to identify the key factors that collaboratively influence the academic achievement of primary school students [20]. This method involves ranking factors based on the coefficients or important properties of different models. In each iteration, it recursively eliminates a small number of redundant or less significant factors, ultimately retaining the optimal set of factors. Additionally, to mitigate potential collinearity impacts on predictive performance, we initially excluded highly correlated factors. Correlation coefficients between each pair of factors were computed using Pearson and Spearman correlation analyses. In instances where the absolute value of the correlation coefficient exceeded 0.7, only one factor was retained.

Single machine learning algorithms

A published systematic review highlighted the significant contributions of SVM and DT algorithms as individual models in predicting academic achievement using machine learning methods, and the LR algorithm was also reported to have good predictive performance [45]. In addition, the LDA algorithm has been proven to be applicable for academic achievement prediction [35]. Hence, we opted for four single algorithms—SVM, DT, LR, and LDA—to establish individual learners for the prediction of academic achievement.

Sun *et al. BMC Public Health* (2024) 24:274

Page 5 of 13

The SVM algorithm is employed for classifying academic achievement by identifying the optimal hyperplane with the largest classification margin. Notably, it incorporates a diverse range of powerful kernel functions, enabling the processing of datasets using highly complex structured methods [46]. On the other hand, the DT algorithm excels at summarizing decision rules from datasets with features and labels, presenting these rules in the form of a tree graph. Recognized for its ease of understanding, suitability for various types of data, and wide applications across fields, the DT algorithm is a versatile tool [47]. LR algorithms are designed to predict the probability of future outcomes based on existing data performance. With a regularization term integrated into the model, LR can effectively reduce model complexity and counteract overfitting concerns [48]. Lastly, the LDA algorithm plays a pivotal role in projecting data into a low-dimensional space. It aims to maximize the differences between categories while minimizing differences within categories, thus effectively achieving the objectives of classification and discrimination [49].

Ensemble learning methods

Ensemble learning methods, including Bagging, Boosting, and Stacking, have made remarkable contributions to predicting academic achievement [45]. The Bagging method [50] enhances predictive performance by reducing classification error variance. A notable member of the Bagging category is the Random Forest (RF) algorithm [51], which has demonstrated robust predictive capabilities in academic achievement [52]. Bagging employs the Bootstrap method to randomly select M samples from the primary school students' dataset, with replacement, ensuring eligibility for subsequent sampling. This process iterates to train N weak classifiers, and their outputs are combined, each receiving equal weighting.

Boosting is another ensemble method that elevates weak classifiers to robust classifiers. The eXtreme Gradient Boosting (XGB) is a powerful member of the Boosting family [53] and has been applied successfully in predicting academic achievement [54]. The boosting process starts by training a weak classifier using the dataset. It then adjusts the sample distribution based on the weak classifier's performance, assigning greater weight to previously misclassified samples. The next weak classifier is then trained according to the adjusted sample distribution [55]. Finally, these weak classifiers are combined, with those performing better assigned higher weights.

The Stacking method [56] constructs several base classifiers using the dataset. It generates a meta-dataset and passes it to the next layer, where the meta-classifier of the final layer produces the ultimate prediction outcome. The outputs of the base classifiers within the meta-dataset are considered new features, while the raw dataset labels are retained as sample labels. For the stacking ensemble learning method in this study, four selected single algorithms—SVM, DT, LR, and LDA—were used to establish a Stacked ensemble learning model for predicting academic achievement (Fig. 1).

Evaluation and validation

The dataset was randomly divided into a training set (80%, $n = 240$) and a test set (20%, $n = 60$) to facilitate machine learning model training and validation. To enhance predictive performance reliability, we employed the repeated five-fold cross-validation method. This method involves randomly dividing the training set into five folds, with four folds used for training in each iteration and the remaining fold reserved for validation. The average accuracy following cross-validation was reported as an assessment of predictive performance [57]. Accuracy is defined as the proportion of correctly classified samples and is specifically expressed as:

$$Accuracy = (TP + TN)/(TP + FP + FN + TN)$$

where TP (true positive) refers to the number of samples whose actual value is positive, and the model predicts them as positive. TN (true negative) is the number of samples whose actual value is negative, and the model predicts them as negative. FP (false positive) is the number of samples whose actual value is negative, and the model predicts them as positive. FN (false negative) is the number of samples whose actual value is positive, and the model predicts them as negative.

Besides, the trained models were applied to the test samples, and multiple indicators such as accuracy, precision, and recall were used to re-evaluate predictive performance [45]. Precision and recall are defined as follows:

$$Precision = TP/(TP + FP)$$

$$Recall = TP/(TP + FN)$$

The accuracy, precision, and recall indicators ranged from 0 to 1, with values closer to 1 indicating better predictive performance. Additionally, the permutation test method was applied to measure the probability that the final predicted results occurred by chance.

Sun *et al. BMC Public Health*   (2024) 24:274

Page 6 of 13
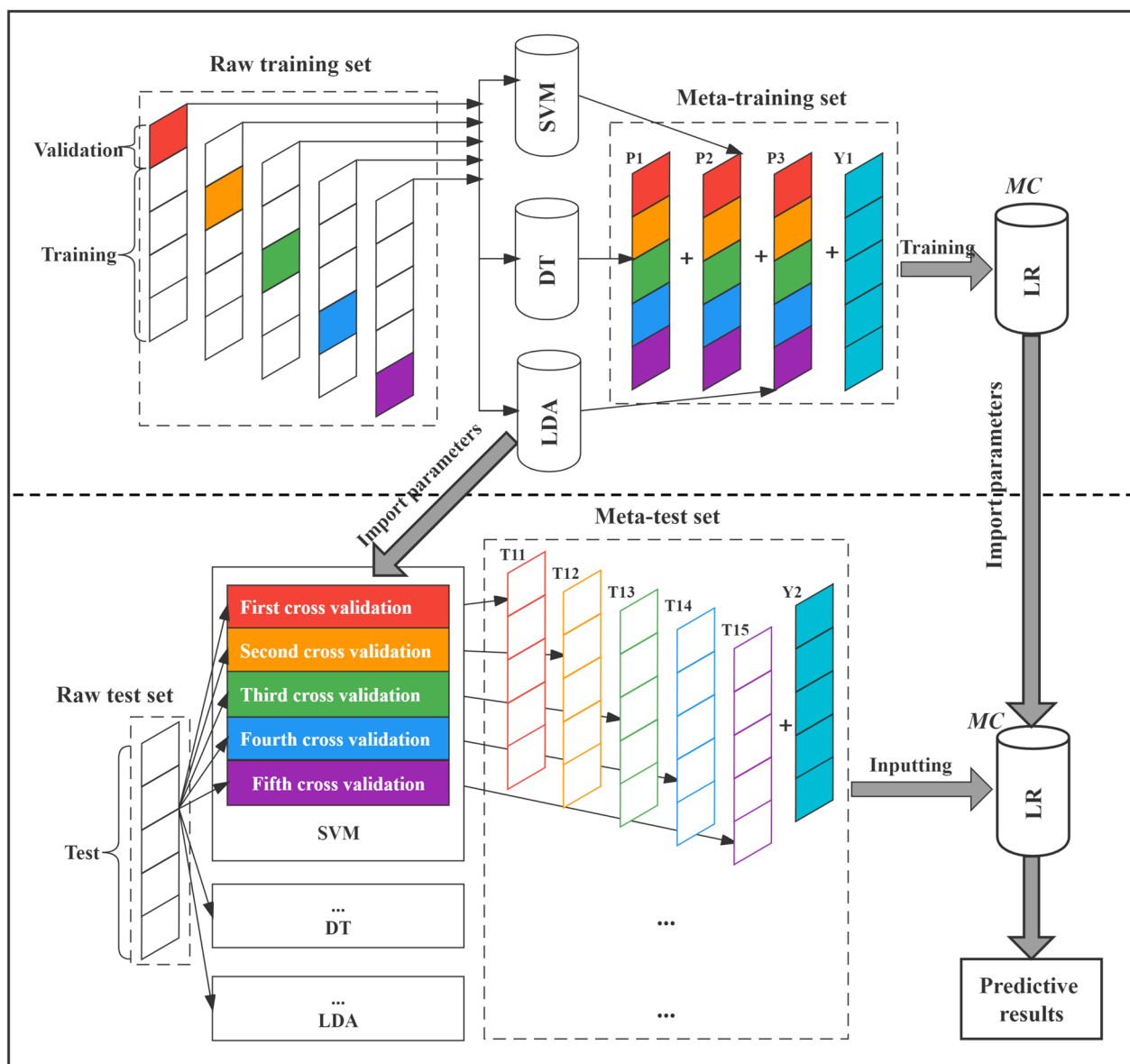
Fig. 1 The principle of Stacking ensemble learning model Note. SVM is Support Vector Machine; DT is Decision Tree; LDA is Linear Discriminant Analysis; LR is Logistic Regression; MC is Meta-Classifier
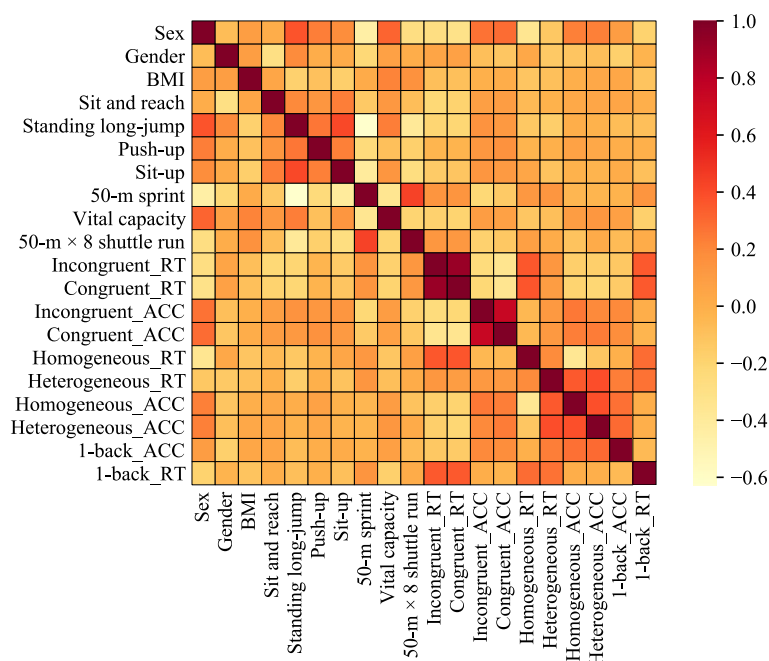
Results

Key factors

To mitigate potential collinearity effects, we excluded highly correlated factors based on the results of Pearson and Spearman correlation analyses. As illustrated in Fig. 2, the average reaction time and accuracy in congruent trials exhibited strong correlations with those in incongruent trials, with correlation coefficients exceeding 0.7. Consequently, we excluded the mean reaction time and accuracy in incongruent trials from further analysis. In total, 18 factors were retained for subsequent analysis, encompassing sex, grade, body mass index, sit and reach, standing long-jump, push-up, sit-up, 50-m sprint, vital capacity, 50-m × 8 shuttle run, and the mean reaction time and accuracy in congruent trials, homogeneous trials, heterogeneous trials, and 1-back task.

We utilized the RFECV method to investigate the key factors that collaboratively influence primary school students' academic achievement. In Fig. 3(a), the XGB algorithm identified a crucial factor set consisting of eight factors. Sex, push-up, vital capacity, mean

Fig. 2 The results of correlation analysis for factors Note. Incongruent_RT and Incongruent_ACC are the mean reaction time and accuracy in incongruent trials, respectively. Congruent_RT and Congruent_ACC are the mean reaction time and accuracy in congruent trials. Homogeneous_ RT and Homogeneous_ACC are the mean reaction time and accuracy in homogeneous trials. Heterogeneous_RT and Heterogeneous_ACC are the mean reaction time and accuracy in heterogeneous trials. 1-back_RT and 1-back_ACC are the mean reaction time and accuracy in 1-back task

reaction time in congruent and 1-back task, and mean accuracy in homogeneous and heterogeneous trials were identified as influential factors. Notably, sex was ranked as the most significant factor.

In Fig. 3(b), the RF algorithm selected a significant factor set comprising five factors. Vital capacity, mean reaction time in congruent trials, heterogeneous trials, and 1-back task, along with mean accuracy in heterogeneous trials, were identified as substantial contributors to academic achievement. Mean reaction time in heterogeneous trials was considered the most crucial.

Figure 3(c) illustrates that the SVM algorithm chose a factor set of 17 factors, excluding only the mean reaction time in heterogeneous trials. Sex was identified as the most critical factor.

Figure 3(d) displays that the DT algorithm opted for a factor set of 14 factors. Grade, 50-m sprint, 50-m × 8 shuttle run, and mean accuracy in 1-back task were not considered key factors. Mean reaction time in heterogeneous trials was regarded as the most important.

In Fig. 3(e), the LR algorithm determined that the optimal factor set should include six factors. Sex, vital capacity, mean reaction time in congruent and homogeneous trials, and mean accuracy in homogeneous trials and 1-back task played pivotal roles. Mean accuracy in homogeneous trials was identified as the most important.

As shown in Fig. 3 (f), the LDA algorithm determined the optimal factor set, which included 12 factors. This algorithm excluded sit and reach, push-up, sit-up, 50-m sprint, 50-m × 8 shuttle run, and mean reaction time in 1-back task. It considered mean accuracy in homogeneous trials as the most important factor.

The results indicated that sex, body mass index, muscle strength, cardiorespiratory function, inhibition, working memory, and shifting were considered key factors by no less than half of the machine learning algorithms. Additionally, the shifting aspect of executive function was identified as the most important factor by most algorithms.

Predictive performance

We constructed two ensemble learning models, XGB and RF, using the training set. Additionally, SVM, DT, LR, and LDA models were established, and a Stacking ensemble learning model was created based on their prediction results. We utilized the cross-validation method to obtain ten predictive accuracy for each model and compared their average accuracy. Table 1 displays the accuracy of seven machine learning models in predicting primary school students' academic achievement using their optimal feature subset.

Sun *et al. BMC Public Health*   (2024) 24:274

Page 8 of 13
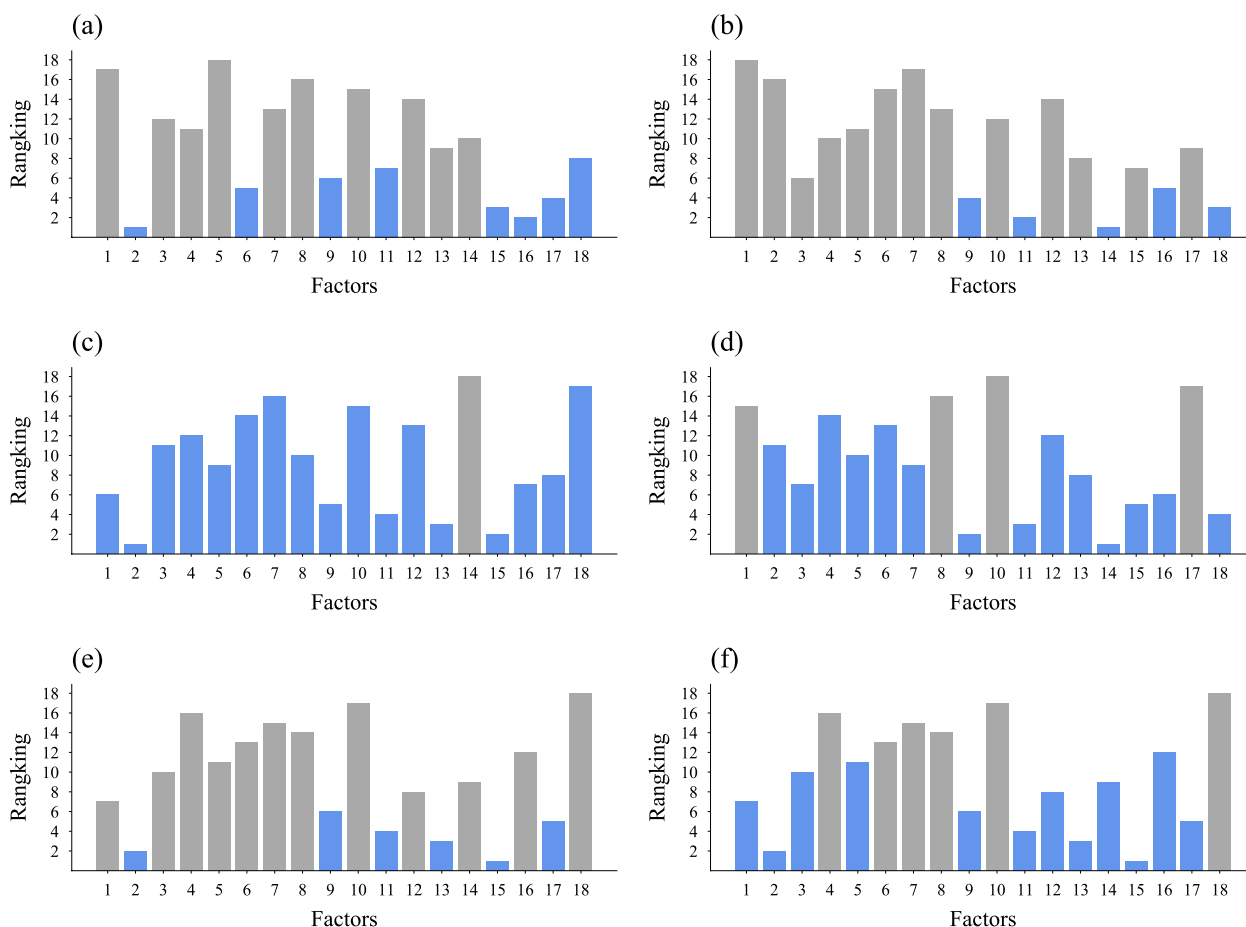
Fig. 3 The best factor subset and feature importance ranking of each algorithm Note. **a** represents the eXtreme Gradient Boosting algorithm; **b** is the Random Forest algorithm; **c** is the Support Vector Machine algorithm; **d** is the Decision Tree algorithm; **e** is the Logistic Regression algorithm; **f** is the Linear Discriminant Analysis algorithm. Blue color indicates the factors selected as key, while gray indicates exclusion. Due to space constraints, all factor names couldn't be labeled in the diagram. The factors in each subfigure are in the same order: sex, grade, body mass index, sit and reach, standing long-jump, push-up, sit-up, 50-m sprint, vital capacity, 50-m × 8 shuttle run, mean reaction time in congruent trials, mean accuracy in congruent trials, mean reaction time in homogeneous trials, mean reaction time in heterogeneous trials, mean accuracy in homogeneous trials, mean accuracy in heterogeneous trials, mean reaction time in 1-back task, and mean accuracy in 1-back task

As shown in Table 1, the average accuracy of XGB, RF, SVM, DT, LR, LDA, and Stacking models was 70.21% (permutation test, iterations = 1000, $p < 0.001$), 70.42% ($p < 0.001$), 61.88% ($p < 0.001$), 61.46% ($p < 0.001$), 61.46% ($p < 0.001$), 63.33% ($p < 0.001$), and 68.33% ($p < 0.001$), respectively. The highest accuracy, occurring three times, was 79.17%, with two instances in the RF model and one in the XGB model. The lowest accuracy, recorded once, was 52.08% in the LR model. The minimum accuracy of XGB, RF, SVM, DT, LR, LDA, and Stacking models exceeded the baseline accuracy (50%), with the highest accuracy exceeding it by 29.17%, 29.17%, 20.83%, 18.75%, 18.75%, 22.92%, and 25%, respectively. Notably, the lowest, highest, and average accuracy of the three ensemble learning models exceeded those of the four

individual learners, suggesting that ensemble learning methods outperform traditional machine learning methods. The results indicated that machine learning models could predict academic achievement of primary school students using executive function, physical fitness, and demographic factors, with the accuracy of ensemble learning models surpassing that of individual learners.

We inputted test samples into the machine learning models and used accuracy, precision, and recall indicators to re-evaluate their prediction performance, directly confirming their generalization capability. In Fig. 4, the XGB model achieved accuracy, precision, and recall of 68.33%, 70.37%, and 63.33% ($ps = 0.004$), respectively. The RF model achieved 66.67%, 64.71%, and 73.33% ($ps = 0.003$), while the Stacking model

Table 1 The accuracy of machine learning models in the training set (%)

Subsets	XGB	RF	SVM	DT	LR	LDA	Stacking
5-fold CV (1st)							
First	72.92	75	66.67	62.5	66.67	62.5	64.58
Second	64.58	62.5	66.67	56.25	68.75	66.67	75
Third	62.5	60.42	60.42	64.58	62.5	60.42	68.75
Fourth	72.92	72.92	54.17	58.33	68.75	54.17	62.5
Fifth	79.17	79.17	62.5	64.58	52.08	72.92	70.83
5-fold CV (2nd)							
First	68.75	62.5	58.33	60.42	54.17	64.58	66.67
Second	75	75	58.33	68.75	60.42	64.58	68.75
Third	70.83	79.17	62.5	58.33	58.33	62.5	72.92
Fourth	70.83	66.67	58.33	64.58	60.42	68.75	64.58
Fifth	64.58	70.83	70.83	56.25	62.5	56.25	68.75
Repeated fivefold CV							
Average	70.21	70.42	61.88	61.46	61.46	63.33	68.33

CV is cross-validation. XGB is eXtreme Gradient Boosting model; RF is Random Forest model; SVM is Support Vector Machine model; DT is Decision Tree model; LR is Logistic Regression model; LDA is Linear Discriminant Analysis model; Stacking is Stacking ensemble learning model
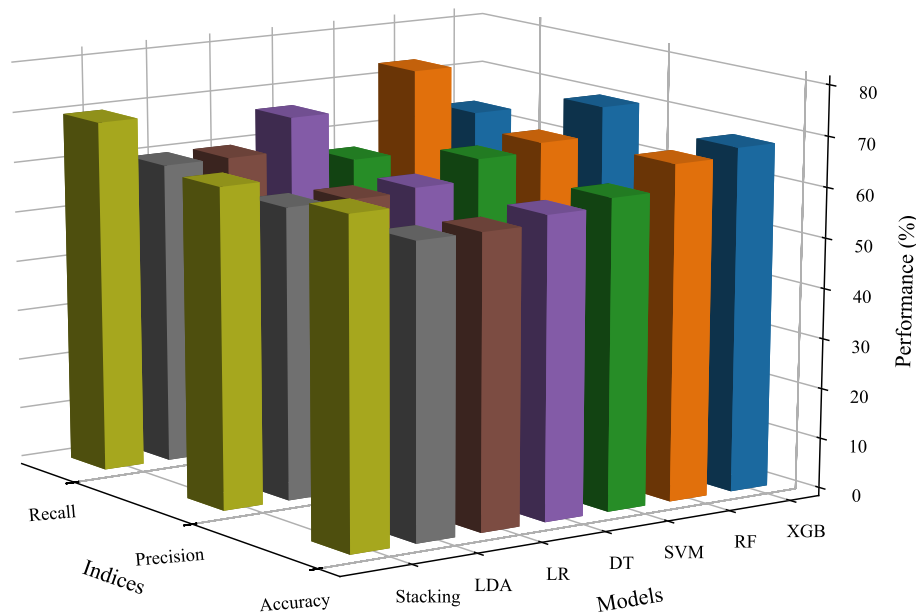
Fig. 4 The performance of machine learning models in test samples Note. XGB is eXtreme Gradient Boosting model; RF is Random Forest model; SVM is Support Vector Machine model; DT is Decision Tree model; LR is Logistic Regression model; LDA is Linear Discriminant Analysis model; Stacking is Stacking ensemble learning model

achieved 65%, 63.64%, and 70% ($ps = 0.023$). In addition, the evaluation indices scores of the three ensemble learning models in the test samples were generally higher than those of the four individual learners. The results once again validated that the ensemble learning model yielded stronger performance than individual learners in predicting the academic achievement of primary school students.

Discussion

We sought to determine which factors from executive function, physical fitness, and demographic factors play key roles in collaboratively influencing the academic achievement of primary school students. Our findings indicated that sex, body mass index, muscle strength, cardiorespiratory function, inhibition, working memory, and shifting were key factors that

Sun *et al. BMC Public Health*   (2024) 24:274

Page 10 of 13

collaboratively influence the academic achievement of primary school students. To begin with, we identified sex as a key demographic factor associated with academic achievement. Previous studies have demonstrated sex-based differences in academic performance [18]. This discrepancy may be attributed to girls having higher failure anxiety than boys, potentially leading to reduced academic success [58].

Besides, physical fitness-related factors, including body mass index, muscle strength, and cardiorespiratory function, emerged as significant contributors to academic achievement in primary school students. Firstly, Castelli et al. observed an inverse relationship between the overall academic achievement of third and fifth-grade students and their body mass index [59]. Research findings imply that elevated body mass index may contribute to issues of overweight/obesity, heightening the risk of unequal treatment and limiting access to learning resources for primary school students [60]. Secondly, Xu et al. found positive correlations between the upper limb strength, lower limb strength, and trunk strength of fifth-grade students with their academic achievement [42]. Scientific evidence supports the notion that strength training can enhance brain plasticity and various functions, including learning and memory [61]. Thirdly, Liang et al. uncovered a connection between poor vital capacity and subpar academic achievement [16]. A growing body of research suggests that aerobic exercise improves cardiorespiratory fitness, inducing changes in brain structure and function [62]. These alterations positively affect executive function [63], which is necessary for academic improvement.

Finally, executive function-related factors, namely inhibition, working memory, and shifting, were recognized as pivotal elements influencing academic achievement. Borella et al. employed Stroop task to assess the inhibition aspects of executive function and determined that inhibition could serve as a predictor for the academic achievement of primary school students [64]. A meta-analysis uncovered that primary school students with lower academic achievement exhibited poorer performance in working memory task compared to their higher-achieving peers [65]. Magalhães et al. further demonstrated that the shifting aspects of executive function could be predictive of academic achievement [66]. Another study found that persistent errors in the Wisconsin Card Sorting test inversely predicted mathematical ability [67]. The predictive influence of executive function on the academic achievement of primary school students should not come as a surprise. As an advanced cognitive process, executive function represents a critical cognitive ability essential for robust academic performance. Moreover, executive control is intricately linked to the functioning of the prefrontal cortex [68], a region closely associated with learning and memory processes.

Additionally, we conducted an exploration to assess whether ensemble learning methods could outperform individual learners in predicting the academic achievement of primary school students, considering executive function, physical fitness, and demographic factors. Our findings revealed that the accuracy of the XGB, RF, and Stacking models surpassed that of individual learners, a result consistently verified through test samples. Ding et al. used factors such as mental health status and coping styles to predict academic achievement [69]. Their results demonstrated that machine learning methods could accurately predict academic success, with Bagging exhibiting superior predictive performance compared to Naive Bayes and K-Nearest Neighbors algorithms. Another investigation collected demographic and personality factors to predict academic achievement [70]. The outcomes suggested that machine learning models could accurately forecast students' academic success, with the Stacking model demonstrating superior performance over individual learners. Ensemble learning operates on the principle of combining multiple weak learners to create a strong learner. Bagging employs the Bootstrap method to generate each base learner from a random subset of the dataset. Boosting establishes multiple base learners by adjusting sample weights and assigning higher weights to those with superior performance. Stacking selects heterogeneous individual learners as base learners, allowing them to observe data from various data spaces and structures. Ensemble learning methods employ diverse strategies to mitigate the predictive errors of base learners, ultimately enhancing predictive performance and generalization capabilities.

Our findings yield profound insights for the field of education, poised to instigate positive changes in practice. In practical applications, this study establishes a cornerstone for personalized education. Schools can forge customized learning plans by consistently assessing key factors, thereby better catering to the individual needs of each primary school student. Additionally, this research strengthens the potential for early intervention measures. Educators can promptly identify academic challenges faced by primary school students and offer targeted support, fostering an environment where students can more effectively realize their potential [28]. In terms of education policy, our study prompts contemplation on resource allocation and policy formulation. Governments and school administrators, armed with an understanding of key factors, can allocate resources more precisely to ensure schools provide effective support and education. Furthermore, these findings may provoke considerations among policymakers, leading

to adjustments that better integrate data science into educational practices and propel the education system toward a more intelligent and personalized direction [71]. In summary, this research makes substantial contributions to the field of education, laying a robust foundation for advancing the academic development and personalized learning of primary school students. It offers education policymakers a fresh perspective, inspiring potential adjustments in policies to better align with the needs of primary school students and optimize the entire education system.

It is crucial to acknowledge the limitations inherent in our study. Firstly, while past research has effectively utilized smaller datasets to construct machine learning models for accurate academic achievement prediction [72], and ensemble learning methods are deemed suitable for situations with smaller samples [34], it remains imperative to conduct further research on more extensive datasets. Moreover, our study exclusively focused on executive function, physical fitness, and demographic factors as predictive variables. To enhance the predictive performance of our models, it is worthwhile to explore additional avenues. One plausible approach involves the inclusion of variables from diverse categories, such as cortical thickness [73] and hippocampal volume [74]. This strategy aims to encompass a broader spectrum of factors associated with academic achievement, leveraging the robust capabilities of machine learning methods for intricate data analysis and, consequently, improving predictive performance.

Conclusion
Our findings indicated that sex, body mass index, muscle strength, cardiorespiratory function, inhibition, working memory, and shifting were key factors that collaboratively influence the academic achievement of primary school students. Additionally, ensemble learning models demonstrated superior performance compared to individual learners in predicting academic achievement of primary school students. These findings underscore the importance of recognizing sex differences and highlighting the interconnected development of cognition and the body, which can positively impact the academic development of primary school students. Moreover, our results advocate for greater attention to ensemble learning methods, emphasizing their utility in establishing an accurate academic early warning system for primary school students.

Abbreviations
NHS   Non-High-Score
HS   High-Score
DT   Decision Tree

SVM   Support Vector Machine
LDA   Linear Discriminant Analysis
LR   Logistic Regression
RF   Random Forest
XGB   eXtreme Gradient Boosting
RFECV   Recursive Feature Elimination Cross-Validation

Authors' contributions
Conceptualization, A. C.; methodology, software, investigation, visualization, formal analysis, and data curation, Z. S., X. X., S. M., and Y. S.; writing—original draft preparation, Z. S.; writing—review and editing, Y. Y. and A. C.. All authors read and approved the final manuscript.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This work has been approved by the Huaian Education Bureau in China and has obtained informed consent from all participants and/or their guardians.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]College of Physical Education, Yangzhou University, Yangzhou 225127, China. [2]Institute of Sports, Exercise and Brain, Yangzhou University, Yangzhou 225127, China. [3]School of Information Engineering, Yangzhou University, Yangzhou 225127, China. [4]Department of Physical Education, Nanjing University, Nanjing 210033, China. [5]Department of Physical Education, Xidian University, Xian 710126, China. [6]Nanjing Sport Institute, Nanjing 210014, China.

References
1. Beaujean AA, Firmin MW, Attai S, Johnson CB, Firmin RL, Mena KE. Using personality and cognitive ability to predict academic achievement in a young adult sample. Pers Individ Differ. 2011;51:709–14.
2. Abdullah G, Isnanto, Vidiyanti NPY. In *Student's Self-Confidence and Their Learning Achievement on Elementary Schools*, 5th International Conference on Education and Technology (ICET 2019), Faculty of Education UM, Atlantis Press: Faculty of Education UM, 2019.
3. Killer O, Baumer J. Does Interest Matter? The Relationship between Academic Interest and Achievement in Mathematics. J Res Math Educ. 2001;32:448–70.
4. Liang X, He J, Zhou J, Liu P. The Influence of Cognitive Ability on Academic Performance of Junior Middle School Students: A Mediated Moderation Model. Psychol Dev Educ. 2020;36:449–61.
5. Chen A, Yin H, Yan J, Yang Y. Effects of Acute Aerobic Exercise of Different Intensity on Executive Function. Acta Psychol Sin. 2011;43:1055–62.
6. Li X, Liu J, Yang W, Cao B, He X, Li Z, et al. Executive Function, Theory of Mind, and Symptom in Children with High Functioning Autism. Chin Ment Health J. 2012;26:584–9.

Sun *et al. BMC Public Health*   (2024) 24:274

Page 12 of 13

7. Stad FE, Van Heijningen CJM, Wiedl KH, Resing WCM. Predicting school achievement: Differential effects of dynamic testing measures and cognitive flexibility for math performance. Learn Individ Differ. 2018;67:117–25.
8. Kieffer MJ, Vukovic RK, Berry D. Roles of Attention Shifting and Inhibitory Control in Fourth-Grade Reading Comprehension. Read Res Q. 2013;48:333–48.
9. Lubin A, Regrin E, Boulc'h L, Pacton S, Lanoë C. Executive Functions Differentially Contribute to Fourth Graders' Mathematics, Reading, and Spelling Skills. J Cogn Educ Psychol. 2016;15:444–63.
10. Ren S, Cai D. Effects of Executive Function Training on Mathematics Skills for Math Learning Difficulty Students in a Primary School. Chin J Spec Educ. 2019;63–71.
11. Bel-Serrat S, Heinen MM, Mehegan J, O'Brien S, Eldin N, Murrin CM, et al. Predictors of weight status in school-aged children: a prospective cohort study. Eur J Clin Nutr. 2019;73:1299–306.
12. Huang C-P, Chen W-L. Relevance of Physical Fitness and Cardiovascular Disease Risk. Circ J. 2021;85:623–30.
13. Górska P, Krzysztoszek J, Korcz A, Bronikowski M. Does fitness enhance learning/academic performance? Biomedical Human Kinetics. 2018;10:163–8.
14. Lv B, Lv L, Bai C, Luo L. Body mass index and academic achievement in Chinese elementary students: The mediating role of peer acceptance. Child Youth Serv Rev. 2020;108:104593.
15. Van Dusen DP, Kelder SH, Kohl HW, Ranjit N, Perry CL. Associations of Physical Fitness and Academic Performance Among Schoolchildren. J Sch Health. 2011;81:733–40.
16. Liang Z, Zhang Y. Empirical Research of Physical Shape, Cardiopulmonary Function and Academic Scores based on the Comparison between National Guideline of Student Physical Fitness 2014 and 2007. Sports Sci. 2016;37:89–97.
17. Cheng G, Chen J, Yao Z. The Influence of Physical Health on Academic Performance. Education Research Monthly. 2021;74–83.
18. Eveland-Sayers BM, Farley RS, Fuller DK, Morgan DW, Caputo JL. Physical Fitness and Academic Achievement in Elementary School Children. J Phys Act Health. 2009;6:99–104.
19. Lei S, Lei S, Liu Z. Analysis of the Age and Gender Difference of School performance about primary school pupils. Theory Pract Educ. 2004;24:45–6.
20. Hu J, Dong X, Peng Y. Discovery of the key contextual factors relevant to the reading performance of elementary school students from 61 countries/regions: insight from a machine learning-based approach. Read Writ. 2021;35:93–127.
21. Gouveia ÉR, Gouveia BR, Marques A, Lopes H, Rodrigues A, Peralta M, et al. Physical Fitness Predicts Subsequent Improvement in Academic Achievement: Differential Patterns Depending on Pupils' Age. Sustainability. 2020;12:8874.
22. Zhong Y, Xiong Q, Tang H. Physical Fitness, Executive Functioning, and Academic Achievement in Primary School Children. Chin J Health Psychol. 2016;24:1096–100.
23. Sun Z, Yuan Y, Dong X, Liu Z, Cai K, Cheng W, et al. Supervised Machine Learning: A New Method to Predict the Outcomes following Exercise Intervention in Children with Autism Spectrum Disorder. Int J Clin Health Psychol. 2023;23:100409.
24. Sun Z, Herold F, Cai K, Yu Q, Dong X, Liu Z, et al. Prediction of Outcomes in Mini-Basketball Training Program for Preschool Children with Autism Using Machine Learning Models. Int J Ment Health Promot. 2022;24:143–58.
25. Kagiyama N, Shrestha S, Farjo PD, Sengupta PP. Artificial Intelligence: Practical Primer for Clinical Research in Cardiovascular Disease. J Am Heart Assoc. 2019;8:e012788.
26. Su Y, Liu M, Zhao N, Liu X, Zhu T. Identifying psychological indexes based on social media data: A machine learning method. Adv Psychol Sci. 2022;29:571–85.
27. Rossi A, Pappalardo L, Cintia P, Iaia FM, Fernàndez J, Medina D. Effective injury forecasting in soccer with GPS training data and machine learning. PLoS ONE. 2018;13:e0201264.
28. Quilez-Robres A, González-Andrade A, Ortega Z, Santiago-Ramajo S. Intelligence quotient, short-term memory and study habits as academic achievement predictors of elementary school: A follow-up study. Stud Educ Eval. 2021;70:101020.
29. Asif R, Hina S, Haque SI. Predicting Student Academic Performance using Data Mining Methods. Int J Comput Sci Netw Sec. 2017;17:187–91.
30. Xu X, Wang J, Peng H, Wu R. Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. Comput Hum Behav. 2019;98:166–73.
31. Nordhausen K. Ensemble Methods: Foundations and Algorithms by Zhi-Hua Zhou. Int Stat Rev. 2013;81:470.
32. Wen L, Hughes M. Coastal Wetland Mapping Using Ensemble Learning Algorithms: A Comparative Study of Bagging. Boosting Stacking Techn Remote Sensing. 2020;12:1683.
33. Polikar R. Ensemble based systems in decision making. IEEE Circuits Syst Mag. 2006;6:21–45.
34. Dietterich TG. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. Mach Learn. 2000;40:139–57.
35. Guerrero-Higueras ÁM, Fernández Llamas C, Sánchez González L, Gutierrez Fernández A, Esteban Costales G, Conde González MÁ. Academic Success Assessment through Version Control Systems. Appl Sci. 2020;10:1492.
36. Ban W, Jiang Q, Zhao W. Research on Precise Prediction of Online Learning Performance Based on Multi-algorithm Fusion Strategy. Modern Distance Education. 2022;37–45.
37. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. Cogn Psychol. 2000;41:49–100.
38. Chen A, Yan J, Yin H, Pan C, Chang Y. Effects of acute aerobic exercise on multiple aspects of executive function in preadolescent children. Psychol Sport Exerc. 2014;15:627–36.
39. Zhu Z, Yang Y, Kong Z, Zhang Y, Zhuang J. Prevalence of physical fitness in Chinese school-aged children: Findings from the 2016 Physical Activity and Fitness in China-The Youth Study. J Sport Health Sci. 2017;6:395–403.
40. Liao Y, Liu H, Xu X, Zhu j, Zhou R, Liu J, et al. Influence of Teenager Physical Fitness on Academic Achievement in Adolescents: A Moderated Mediation Model Journal of Wuhan Institute of Physical Education. 2022;56:70-78.
41. Zhang L, Lu C, Chen B. The Pathway Regarding the Influence of Physical Activity on Middle School Students' Academic Performance. Youth Studies. 2021;70–82+93.
42. Xu K, Sun Z. Predicting Academic Performance Associated with Physical Fitness of Primary School Students Using Machine Learning Methods. Complement Ther Clin Pract. 2023;51:101736.
43. He H, Garcia EA. Learning from Imbalanced Data. IEEE Trans Knowl Data Eng. 2009;21:1263–84.
44. Sharma NV, Yadav NS. An optimal intrusion detection system using recursive feature elimination and ensemble of classifiers. Microprocess Microsyst. 2021;85:104293.
45. Balaji P, Alelyani S, Qahmash A, Mohana M. Contributions of Machine Learning Models towards Student Academic Performance Prediction: A Systematic Review. Appl Sci. 2021;11:10007.
46. Xu K, Sun Z, Qiao Z, Chen A. Diagnosing autism severity associated with physical fitness and gray matter volume in children with autism spectrum disorder: Explainable machine learning method. Complement Ther Clin Pract. 2024;54:101825.
47. Zhang W, Wang Y, Wang S. Predicting academic performance using tree-based machine learning models: A case study of bachelor students in an engineering department in China. Educ Inf Technol. 2022;27:13051–66.
48. Owusu-Boadu B, Nti IK, Nyarko-Boateng O, Aning J, Boafo V. Academic Performance Modelling with Machine Learning Based on Cognitive and Non-Cognitive Features. Appl Comput Syst. 2021;26:122–31.
49. Graf R, Zeldovich M, Friedrich S. Comparing linear discriminant analysis and supervised learning algorithms for binary classification—A method comparison study. Biom J. 2022;1–20.
50. Breiman L. Bagging Predictors. Mach Learn. 1996;24:123–40.
51. Breiman L. Random Forests. Mach Learn. 2001;45:5–32.
52. Zaffar M, Savita KS, Hashmani MA, Rizvi SSH. A Study of Feature Selection Algorithms for Predicting Students Academic Performance. Int J Adv Comput Sci Appl. 2018;9:541–9.
53. Wang Z, Zhang Q, Lan K, Yang Z, Gao X, Wu A, et al. Enhancing instantaneous oxygen uptake estimation by non-linear model using

cardio-pulmonary physiological and motion signals. Front Physiol. 2022;13:897412.

54. Lenin T, Chandrasekaran N. Learning from Imbalanced Educational Data Using Ensemble Machine Learning Algorithms. Webology. 2021;18:183–95.

55. Chan JC-W, Huang C, DeFries RS. Enhanced algorithm performance for land cover classification from remotely sensed data using bagging and boosting. IEEE Transactions on Geoscience and Remote Sensing. 2001;39:693–695.

56. Breiman L. Stacked Regressions. Mach Learn. 1996;24:49–64.

57. Lenhard F, Sauer S, Andersson E, Månsson KN, Mataix-Cols D, Rück C, et al. Prediction of Outcome in Internet-Delivered Cognitive Behaviour Therapy for Paediatric Obsessive-Compulsive Disorder: A Machine Learning Approach. Int J Methods Psychiatr Res. 2018;27:e1576.

58. Wang Y, Xu Z. The Influence of Academic Achievement on Failure Anxiety from the Perspective of Growth Mindset:An Analysis of Gender Differences. Educ Sci Res. 2022;39–46.

59. Castelli DM, Hillman CH, Buck SM, Erwin HE. Physical Fitness and Academic Achievement in Third- and Fifth-Grade Students. J Sport Exerc Psychol. 2007;29:239–52.

60. Lumeng JC, Forrest P, Appugliese DP, Kaciroti N, Corwyn RF, Bradley RH. Weight status as a predictor of being bullied in third through sixth grades. Pediatrics. 2010;125:e1301–7.

61. Cassilhas RC, Viana VAR, Grassmann V, Santos RT, Santos RF, Tufik S, et al. The Impact of Resistance Exercise on the Cognitive Function of the Elderly. Med Sci Sports Exerc. 2007;39:1401–7.

62. Chaddock L, Pontifex MB, Hillman CH, Kramer AF. A Review of the Relation of Aerobic Fitness and Physical Activity to Brain Structure and Function in Children. J Int Neuropsychol Soc. 2011;17:975–85.

63. Chaddock-Heyman L, Erickson KI, Kienzler C, King M, Pontifex MB, Raine LB, et al. The Role of Aerobic Fitness in Cortical Thickness and Mathematics Achievement in Preadolescent Children. PLoS ONE. 2015;10:e0134115.

64. Borella E, Carretti B, Pelegrina S. The Specific Role of Inhibition in Reading Comprehension in Good and Poor Comprehenders. J Learn Disabil. 2010;43:541–52.

65. Xie Y, Chen F, Yang J, Jin T. The Relationship between Working Memory and the Academic Performance of Primary and Middle School Students: A Meta-Analysis. Chin J Health Psychol. 2016;24:134–7.

66. Magalhães S, Carneiro L, Limpo T, Filipe M. Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. Child Neuropsychol. 2020;26:934–52.

67. Li M, Shen D, Bai X. A Research on Cognitive Flexibility of Students in Different Grades. Chin J Special Educ. 2007;80–86.

68. Funahashi S. Neuronal mechanisms of executive control by the prefrontal cortex. Neurosci Res. 2001;39:147–65.

69. Ding X, Nie J, Zhang B. Using Demographic Information, Psychological Assessment Data and Machine Learning to Predict Students' Academic Performance. J Psychol Sci. 2021;44:330–9.

70. Adejo OW, Connolly T. Predicting student academic performance using multi-model heterogeneous ensemble approach. J Appl Res High Educ. 2018;10:61–75.

71. Wakelam E, Jefferies A, Davey N, Sun Y. The potential for student performance prediction in small cohorts with minimal available attributes. Br J Edu Technol. 2019;51:347–70.

72. Lau ET, Sun L, Yang Q. Modelling, prediction and classification of student academic performance using artificial neural networks. SN Appl Sci. 2019;1:1–10.

73. Meruelo AD, Castro N, Nguyen-Louie T, Tapert SF. Substance use initiation and the prediction of subsequent academic achievement. Brain Imaging Behav. 2020;14:2679–91.

74. Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, et al. Exercise training increases size of hippocampus and improves memory. Proc Natl Acad Sci USA. 2011;108:3017–22.

Publisher's Note