



Original article

Supervised machine learning: A new method to predict the outcomes following exercise intervention in children with autism spectrum disorder[☆]Zhiyuan Sun^{a,b}, Yunhao Yuan^c, Xiaoxiao Dong^{a,b}, Zhimei Liu^{a,b}, Kelong Cai^{a,b}, Wei Cheng^{a,b}, Jingjing Wu^{a,b}, Zhiyuan Qiao^{a,b}, Aiguo Chen^{a,b,d,*}^a College of Physical Education, Yangzhou University, Yangzhou 225127, China^b Institute of Sports, Exercise and Brain, Yangzhou University, Yangzhou 225127, China^c School of Information Engineering, Yangzhou University, Yangzhou 225127, China^d Nanjing Institute of Physical Education, Nanjing 210014, China

ARTICLE INFO

Keywords:

Supervised machine learning

Prediction

Outcomes

Exercise intervention

Autism spectrum disorder

ABSTRACT

The individual differences among children with autism spectrum disorder (ASD) may make it challenging to achieve comparable benefits from a specific exercise intervention program. A new method for predicting the possible outcomes and maximizing the benefits of exercise intervention for children with ASD needs further exploration. Using the mini-basketball training program (MBTP) studies to improve the symptom performance of children with ASD as an example, we used the supervised machine learning method to predict the possible intervention outcomes based on the individual differences of children with ASD, investigated and validated the efficacy of this method. In a long-term study, we included 41 ASD children who received the MBTP. Before the intervention, we collected their clinical information, behavioral factors, and brain structural indicators as candidate factors. To perform the regression and classification tasks, the random forest algorithm from the supervised machine learning method was selected, and the cross validation method was used to determine the reliability of the prediction results. The regression task was used to predict the social communication impairment outcome following the MBTP in children with ASD, and explainable variance was used to evaluate the predictive performance. The classification task was used to distinguish the core symptom outcome groups of ASD children, and predictive performance was assessed based on accuracy. We discovered that random forest models could predict the outcome of social communication impairment (average explained variance was 30.58%) and core symptom (average accuracy was 66.12%) following the MBTP, confirming that the supervised machine learning method can predict exercise intervention outcomes for children with ASD. Our findings provide a novel and reliable method for identifying ASD children most likely to benefit from a specific exercise intervention program in advance and a solid foundation for establishing a personalized exercise intervention program recommendation system for ASD children.

Introduction

Autism spectrum disorder (ASD) is a life-long neurodevelopmental disability that manifests in early childhood (Lord et al., 2020, 2018). The main symptoms of ASD are social communication impairments (SCI) and restricted and repetitive behaviors (RRB). The prevalence of ASD among children in the United States has increased from 1 in 5000 in 1975 to 1 in

36 in 2023, according to data from the Centers for Disease Control and Prevention (Maenner et al., 2023). According to a nationwide multi-center population study, the prevalence of ASD in Chinese children was approximately 1 in 143 (Zhou et al., 2020). ASD has evolved from a rare disease to a primary public health concern impacting children's survival and development.

Physical exercise can promote the development of children's

[☆] This research was supported by grants from the National Natural Science Foundation of China (31771243) and the Fok Ying Tong Education Foundation (141113). Thanks to professors Chen and Yuan for their help during manuscript preparation. The contributions of other authors in data collection, manuscript revision, etc. are also indispensable. All authors have read and agreed to the published version of the manuscript, and have reported that there are no competing interests to declare.

* Corresponding author at: College of Physical Education, Yangzhou University, Yangzhou 225127, China.

E-mail address: agchen@yzu.edu.cn (A. Chen).

<https://doi.org/10.1016/j.ijchp.2023.100409>

Received 23 May 2023; Accepted 22 August 2023

Available online 7 September 2023

1697-2600/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

physical fitness (Clemente et al., 2022), cognitive function (Anzeneder et al., 2023), and brain plasticity (El-Sayes et al., 2019) and have significant advantages, including low cost, simple implementation, and limited side effects (Chen et al., 2021). In recent years, physical exercise has gradually attracted the attention of researchers and has been applied to the rehabilitation and treatment of children with ASD (Haghighi et al., 2022). A growing body of research suggests that exercise intervention can improve behavioral outcomes, including stereotypical behavior, social-emotional function, cognitive ability, and attention, in children with ASD (Bremer et al., 2016). A non-randomized crossover study indicates that ball-tapping exercise intervention can effectively improve stereotypical behaviors in children with ASD (Tse et al., 2018). Zhao et al. (2021) found that compared to the control group, children with ASD who participated in the therapeutic horseback riding program showed significant improvement in social interaction and communication skills. Toscano et al. (2022) reported that structured exercise programs could improve social interaction and stereotypical behavior in children with ASD, making it a potent adjunctive therapy.

Although there is sufficient evidence for the positive effect of exercise intervention on symptom performance in children with ASD, these interventions are neither equally effective nor effective for all participants. A study investigating the impact of yoga intervention programs for children with ASD found that not all participants could benefit equally from this program. Especially, 19% of participants did not receive satisfactory benefits (Rosenblatt et al., 2011). In recent years, precision medicine, aimed at determining individual uniqueness and predicting favorable or unfavorable effects, has become a major goal of healthcare. Understanding which individuals are most likely to have beneficial effects before intervention is crucial for maximizing the benefits of exercise intervention and achieving personalized treatment for children with ASD. However, previous studies on exercise intervention in children with ASD have often been limited by traditional statistical analysis methods, focusing only on whether a specific exercise program has significantly improved the ASD child population while ignoring the possible differences in outcomes caused by individual uniqueness within the group. Therefore, it is necessary to investigate a new method to predict possible outcomes for children with ASD before receiving exercise intervention.

Machine learning is an important approach to artificial intelligence in research. It aims to learn knowledge and laws from complex data to predict future behavior results and trends (Xu & Sun, 2023). Based on different learning methods, machine learning can be divided into supervised machine learning and unsupervised machine learning (Nematollahi et al., 2023). The supervised machine learning is mainly used for completing regression and classification tasks. Its input and output are specified, and it attempts to learn from the input the pattern of the expected output (Popović Krneta et al., 2023). The unsupervised machine learning is mainly used for completing clustering and dimension reduction tasks. Its output is not specified, and it aims to find connections between input data and discover potential patterns (Castiglioni et al., 2021). In recent years, supervised machine learning has been applied to predict mental disease outcomes, which helps predict which patients can benefit from specific intervention methods and which patients need to choose alternative treatment approaches (Sun et al., 2022). In the depression, Pearson et al. (2019) performed regression task using an ensemble of elastic net and random forest (RF) to predict which depression patients can benefit from internet intervention. The results found that the ensemble method predicted more variance than the benchmark model. In the obsessive-compulsive disorder (OCD), Lenhard et al. (2018) performed classification task to predict the outcome of internet-delivered cognitive behavior therapy in OCD adolescents. The results indicated that all machine learning algorithms performed well in the prediction of treatment response. In the attention deficit hyperactivity disorder (ADHD), Kim et al. (2015) also performed classification task to predict the methylphenidate response of ADHD. The results found that the prediction performance of support vector machines was

well and could assist in predicting treatment response in ADHD. However, few studies have applied the supervised machine learning method to predict the outcomes of an exercise intervention to improve the symptom performance of ASD children. The efficacy and practical value of this method have not yet attracted significant attention from researchers.

In recent years, the positive effects of the mini-basketball training program (MBTP) for children with ASD have been confirmed. This program is characterized by its simplicity, consistency, and enjoyment (Cai et al., 2020). Therefore, using the study of the MBTP to improve the symptom performance of ASD children as an example, we predicted the possible intervention outcomes from the individual uniqueness in ASD children using the supervised machine learning method. Currently, it is unknown which variables may be reliable predictors of the outcomes of exercise interventions (including the MBTP). Reviewing existing research findings, we found that the core symptoms of children with ASD are strongly correlated with symptomatic severity (Bodfish et al., 2000). Compared with normally developing children, ASD children are more susceptible to physical health problems (Toscano et al., 2019; Tyler et al., 2014). Cognitively, ASD children differ significantly from children with normal development, and it has been observed that deficits in executive function contribute, at least in part, to SCI and RRB (Li et al., 2012). In addition, the SCI of children with ASD is associated with the gray matter volume (GMV) of the left dorsolateral superior frontal gyrus (Cheng et al., 2023). The RRB is associated with the atypical development of GMV in the cortex and subcortical region (Eisenberg et al., 2015). The MBTP positively affects the executive function (Wang et al., 2020) and physical health (Cai et al., 2021) of ASD children. The improvement of RRB following MBTP in children with ASD correlates with an increase in the GMV (Dong et al., 2020). The evidence presented above suggests that clinical information, physical fitness, executive function, and differences in GVM in children with ASD in pre-intervention may influence the outcomes of MBTP and should be considered important predictive variables.

As previously stated, the individuality of children with ASD may make it difficult for them to achieve comparable benefits from specific exercise interventions. Therefore, a new method for predicting the possible outcomes of children with ASD must be explored to maximize the benefits of exercise intervention. In recent years, supervised machine learning has been applied to predict the outcomes of behavioral interventions for mental disorders; however, few studies have used it to predict the outcomes of exercise interventions for children with ASD. Therefore, we used the study of the MBTP improving the symptom performance of children with ASD as a reference and applied the supervised machine learning method to predict the possible intervention outcomes based on the uniqueness of children with ASD and investigate and validate the efficacy of this method. Our findings will help in identifying children with ASD who are most likely to benefit from a specific exercise intervention program and will serve as a foundation for the development of a personalized exercise intervention recommendation system for children with ASD.

Methods

Data

Longitudinal data were obtained from a long-term research project that explored the effects of exercise intervention on symptom performance in children with ASD (Cheng et al., 2023; Dong et al., 2020; Wang et al., 2020). This project has administered the MBTP to ASD preschoolers (ages 3–6) and school-aged children (7–12). In both studies, the same evaluation method and process, were used. To maximize the sample size included in the analysis, all ASD children from both studies who participated in the MBTP were combined for analysis. Finally, we included data from 41 children with ASD, including 37 boys and 4 girls, aged around 6 years (mean age: 6.2 years old, range: 3–12).

Table 1
Mini-basketball training program.

Phase	Goal	Content	Duration
Phase I	Standardize classroom routines; Increase children's interest in mini-basketball	Classroom routines (taking turns, waiting, obeying, etc.); Simple basketball training (roll and throw the ball, etc.)	2 weeks
Phase II	Improve children's mini-basketball skills; Improve their social communication skills	Basic basketball skill (dribbling, passing, etc.); Peer coordination training (passing and catching the ball, relay racing, etc.)	8 weeks
Phase III	Improve children's cooperative ability, social skills, and collectivization	Group game based on mini-basketball (basketball-dribbling relay, basketball-passing relays, basket-moving shooting, etc.)	2 weeks

Studies inclusion criteria were: (i) Han nationality, (ii) children diagnosed with ASD via DSM-5, (iii) all ASD children need the consent of their guardians to participate in the study, (iv) having the scanning conditions for magnetic resonance imaging, such as no implanted metal in the body. Exclusion criteria were: (i) history of head trauma, (ii) additional psychiatric and/or neurological diseases including epilepsy and phenylketonuria, (iii) visual and/or hearing impairments, (iv) involvement in a basketball training program in the past six months, (v) physical disability or other health conditions where physical exercise is not allowed (acute phase after fracture). Studies have obtained ethical approval from the Ethics and Human Protection Committee of the Affiliated Hospital of Yangzhou University and have been registered with the Chinese Clinical Trial Registry (ChiCTR1900024973). In addition, written informed consent from the parents/guardians of each ASD child was also obtained.

Mini-basketball training program (MBTP)

Two certified physical education teachers conducted the MBTP, and operational details can be found in published articles based on the data from these studies (Cai et al., 2020; Wang et al., 2020). The MBTP is divided into three phases, as shown in Table 1. A single exercise session consists of four parts, in the order of introduction, warm-up exercise, MBTP, and relaxation. The MBTP was conducted five times per week for 12 weeks. Each session lasted 40 min, and the exercise intensity level was moderate.

Candidate predictors

At baseline, we collected basic demographic information (including age and sex), clinical information, physical fitness, executive function, and brain structure indicators of children with ASD as candidate factors. Clinical information included not only symptomatic severity, but also SCI and RRB. They are important information that can be seen as the severity of core symptoms in children with ASD before intervention. Physical fitness included body mass index (BMI), speed-agility, muscle strength, flexibility, and balance. The executive function included working memory, regulation and inhibition. The brain structure data included the GMV of the entire brain.

Children with ASD had their symptomatic severity evaluated using the Childhood Autism Rating Scale (CARS) and clinical assessment reports (Schopler et al., 1980). The total score of this scale is 60 points, and the scoring criteria are: no ASD (total score < 30 points), mild to moderate ASD (30–37 points), and severe ASD (37–60 points, with at least 5 items scoring above 3 points). The Social Response Scale Second Edition (SRS-2) was used to evaluate SCI in children with ASD. The reliability of this scale has been confirmed (Bruni, 2014). The scale calculates the total score, and the severity of the SCI increases as the score increases.

The Repetitive Behavior Scale-Revised (RBS-R) was used to evaluate the RRB of children with ASD (Bodfish et al., 2000). The scale was revised by Bodfish et al. with excellent reliability and validity (Martínez-González & Piqueras, 2018). On the scale, the total score is recorded, and the severity of the RRB increases as the score increases.

China's manual of physical fitness evaluated the physical fitness of ASD children (Cai et al., 2021). The BMI was defined as weight (kg)/height (m²). The 2 × 10-m shuttle run test was used to evaluate speed-agility. Two parallel lines were drawn on the floor, 10-m apart. ASD children should run back and forth as quickly as possible, crossing each line with their feet each time. The shorter the time, the better their speed and agility. This trial was conducted twice, and the best of the two attempts was recorded (s). The standing long jump test was used to evaluate muscle strength. ASD Children place their feet together behind the jump rope and spring as far forward as possible. Repeat the examination three times and record your highest score (cm). The sit-and-reach test is used to evaluate flexibility. ASD children sit on the floor, with the evaluator keeping their legs straight and slowly extending forward as much as possible, recording the best result of two attempts (cm). The balance beam test was used to evaluate the balance. ASD children stand on the starting point facing the balance beam (length 3 m, width 10 cm, height 30 cm) with their arms outstretched to the side. After hearing the 'start' command, move forward. The shorter the time it takes to cross the finish line, the better the balance, as measured by the stopwatch used by the evaluator. Repeat the test twice, keeping the best score (s). If the test cannot be completed or the time exceeds 30 s, record it as 30 s.

The Childhood Executive Functioning Inventory (CHEXI) was used to assess the executive function of children with ASD (Thorell & Nyberg, 2008). This scale has good reliability and validity, and can be used as an evaluation tool for executive function in Chinese children (Wei et al., 2018). There are 24 items on the scale, which was initially divided into four subscales: working memory, planning, regulation and inhibition. Wei et al. (2018) found that the correlation between working memory and planning ability in the four-factor model was 0.97, suggesting that the two factors can be combined into working memory, and the three-factor structure (working memory, regulation, and inhibition) fitted the data well. All questions are scored in reverse, so the higher the dimensional score, the greater the deficits in the respective executive function components.

The images were acquired using a 3.0T GE Healthcare whole-body high-speed imaging system with echo planar imaging (GE Discovery MR750w 3.0T). The scanning parameters of the T1-MPRAGE structural image were as follows: repetition time = 7.20 ms, echo time = 3.06 ms, thickness = 1.00 mm, flip angle = 12°, acquisition matrix size = 256 × 256, and field of view = 256 × 256 mm. The SPM-based CAT12 toolbox was used to preprocess the T1-MPRAGE structural image. The preprocessing mainly included: (i) image quality inspection; (ii) image registration to the standard space of the Montreal Neurological Institute (MNI) according to the DARTEL method; (iii) segmentation of the whole-brain into gray matter, white matter, and cerebrospinal fluid. Then, the DPABI software package was used to further process the preprocessed images. In the ROI signal extractor of DPABI software, a widely used Anatomical Automatic Labeling (AAL) template was selected as Mask to extract the GMV of 116 brain regions (Liu et al., 2020).

Outcomes

In previous studies that applied supervised machine learning methods to predict the outcomes following interventions in mental diseases, both regression models and classification models have been used in cases. In addition, the original form of the outcome variable in this study is the score of the scale, which is not only a continuous variable, but also can be transformed into a categorical variable. Therefore, both the two main tasks of supervised machine learning (regression and classification) will be discussed.

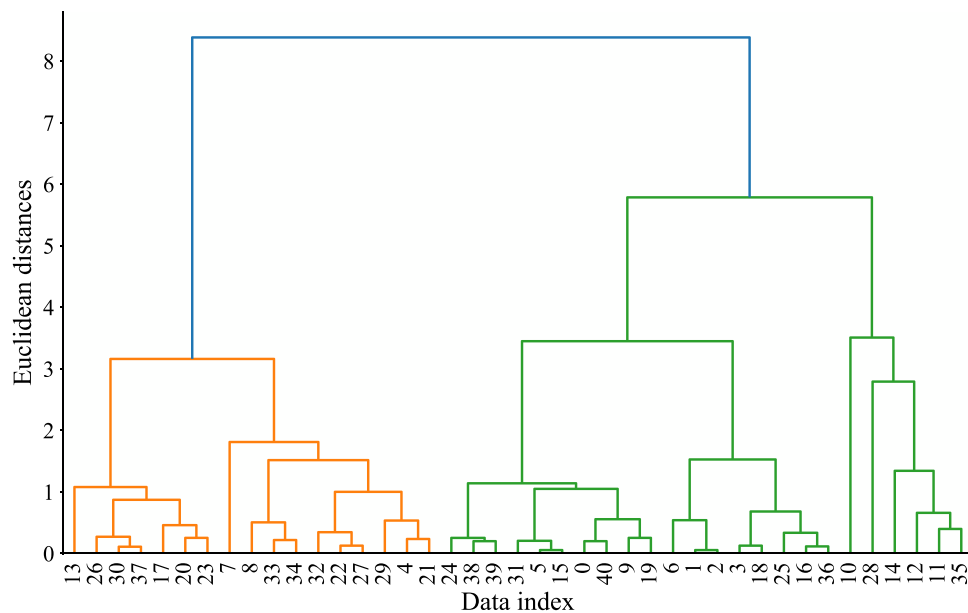


Fig. 1. The results of hierarchical agglomerative clustering.

The dataset's label (target variable) must be a continuous variable for the regression task. We define the label of the regression task as the difference between the core symptom performance of ASD children before and after exercise intervention. The SCI outcome of ASD children was defined as the difference in SRS-2 scale scores following the MBTP, and the RRB outcome was defined as the difference in RBS-R scale scores. To avoid similar redundant expressions, we only performed a regression task on the SCI outcome in children with ASD.

The classification task requires categorical variable labels in the dataset. There is currently no clear grading standard for the outcome of core symptoms in children with ASD. To investigate the possible varying degrees of benefits, hierarchical agglomerative clustering (HAC) analysis based on Ward linkage method was performed on the SCI and RRB outcomes following the MBTP in children with ASD (Bakkelund, 2021; Kruse et al., 2017). Based on HAC results (Fig. 1), we clustered the dataset into two clusters, with 24 samples clustered into cluster 1 (58.5%) and 17 samples clustered into cluster 2 (41.5%). The results of the Mann Whitney *U* test indicated that the SCI outcome ($z = -4.886$, $p < 0.001$) and RRB outcome ($z = -4.1$, $p < 0.001$) of cluster 2 were better than those of cluster 1. Therefore, we defined Cluster 1 as a fair outcome group and Cluster 2 as a better outcome group. Finally, the labels for the classification task were defined as two categories: fair outcome and better outcome.

Supervised machine learning method

Feature engineering

Firstly, we performed Z-score standardization on the data to eliminate the problem of dimensional differences between variables and then conducted the statistical analysis. In this study, $p < 0.05$ was considered statistically significant, and $p < 0.08$ was considered marginal significance. For the regression task, Pearson correlation analysis was used to determine the relationship between candidate factors and the SCI outcome in children with ASD. For the classification task, Spearman correlation analysis was performed to quantify the correlation between candidate factors and core symptom outcomes in children with ASD. We initially selected candidate factors with statistical significance (including marginal significance) as predictors and included them in the corresponding supervised machine learning task. Finally, if the feature dimensions initially included in the regression or classification task are large (for example, the sample size is far less than 10 times the number of

features), we will use principal component analysis (PCA) to reduce dimensionality (Greenacre et al., 2022) and extract principal components with an eigenvalue greater than 1 (Kaiser, 1960).

Algorithm selection

The RF algorithm was selected for supervised machine learning tasks. It can perform regression and classification tasks. The RF is an ensemble of many decision trees. It uses the bootstrap method to randomly select *M* samples from the raw dataset and then reinsert them to make them eligible for subsequent sampling. Repeat this process to obtain *K*-independent subsets of data and then train *K* decision tree learners accordingly. Finally, all decision tree learners are combined, and each learner is given an equal weight (Paul et al., 2018). In addition, RF is an ensemble learning method, its prediction performance is superior to the traditional supervised machine learning algorithm, and has a strong generalization ability (Breiman, 2001).

Grid search and cross validation

We used grid search method to adjust hyperparameters in regression and classification task. The grid search was used to generate a list of all possible values of each parameter in the estimation function, after which the values in each list were combined to generate a grid, and each grid was used as training a model. After the fitting function had tried all combination results, it returned the most suitable learner and automatically adjusted to the best parameter combination (Kim, et al., 2023).

Regarding the randomness of calculations, we have adopted the *k*-fold cross validation method in regression and classification tasks (Lin et al., 2022). The cross validation method uses randomly selected data subsets to fit multiple supervised machine learning models, thereby determining the reliability of the predicted outcomes (Marouf et al., 2022). In this study, we fitted an RF model for each cross validation in the regression and classification tasks, and repeated the process three times. Each model was trained on two-thirds of the data and used to predict the retained one-third of the samples. The regression task includes three RF regression models, and the classification task includes three RF classification models.

Evaluation indicators

Each RF regression model's predicted and actual results will be represented graphically in the regression task. In addition to reporting

Table 2

Confusion matrix.

Confusion matrix		Actual value	
		Fair	Better
Predictive value	Fair	TF	FF
	Better	FB	TB

Note: TF is the number of samples whose actual value is *fair*, and the model correctly predicts that it is *fair*. FF is the number of samples whose actual value is *better*, and the model incorrectly predicts that it is *fair*. TB is the number of samples whose actual value is *better*, and the model correctly predicts that it is *better*. FB is the number of samples whose actual value is *fair*, and the model incorrectly predicts that it is *better*.

the explained variance of each model to evaluate their predictive performance (Muñoz-Organero & Queipo-Álvarez, 2022), we will also report the average of the explained variance to evaluate the overall performance of the regression task. The confusion matrix for each RF classification model will be drawn in the classification task (Walther et al., 2022). The description of the confusion matrix is shown in Table 2. Similarly, the accuracy of each model will also be reported (Wu et al., 2023), and the final average accuracy will be reported to evaluate the overall performance of the classification task. In addition, the permutation test based on p -values will be used to verify the final classification result. In permutation test, class labels are permuted, and then randomly assign 'incorrectly' to the samples. This approach allows us to measure the probability that the predictions are the result of chance (Cruz-Tirado

et al., 2023).

For the regression task, the closer the explained variance is to 1, the better the model's predictive performance. For the classification task, the closer the accuracy is to 1, the better the model's predictive performance. The modeling flowchart of supervised machine learning is shown in Fig. 2.

Results

Regression task

With respect to demographic information, we did not observe significant sex-related difference concerning SCI outcome ($t_{(39)} = -0.383$, $p = 0.704$), and there was no significant correlation between age and SCI outcome ($r = 0.11$, $p = 0.495$). The regression task used Pearson correlation analysis to investigate the relationship between other candidate factors and the SCI outcome in ASD children. As shown in Fig. 3, in clinical information and behavioral factors, the SCI outcome of children with ASD were correlated with baseline RRB ($r = 0.343$, $p = 0.028$), muscle strength score ($r = -0.328$, $p = 0.036$), flexibility score ($r = -0.301$, $p = 0.056$), and speed-agility score ($r = 0.494$, $p = 0.001$). In the brain structure indicators, the GMV of 15 brain regions were correlated to the SCI outcome of ASD children, including the left orbital part of superior frontal gyrus ($r = 0.284$, $p = 0.072$), the right orbital part of superior frontal gyrus ($r = 0.295$, $p = 0.062$), the right opercular part of superior frontal gyrus ($r = 0.424$, $p = 0.006$), the left triangular part of

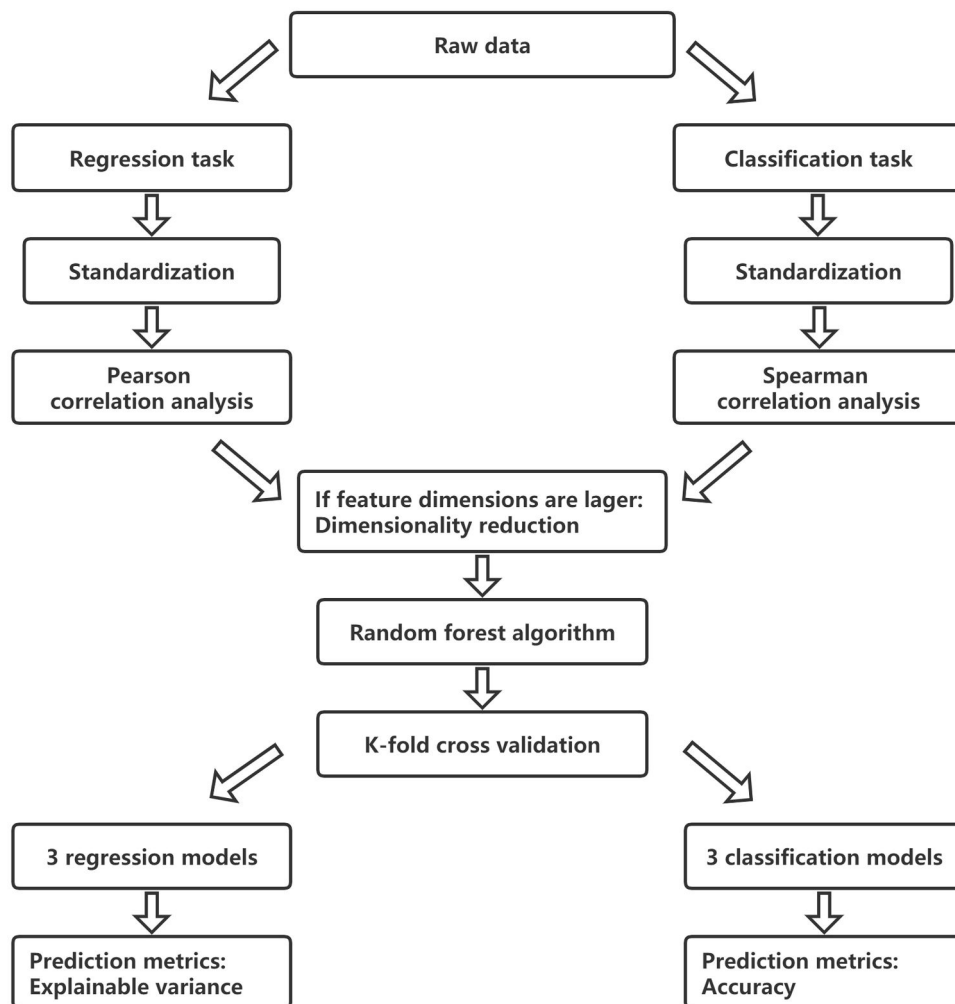


Fig. 2. The modeling flowchart of supervised machine learning.

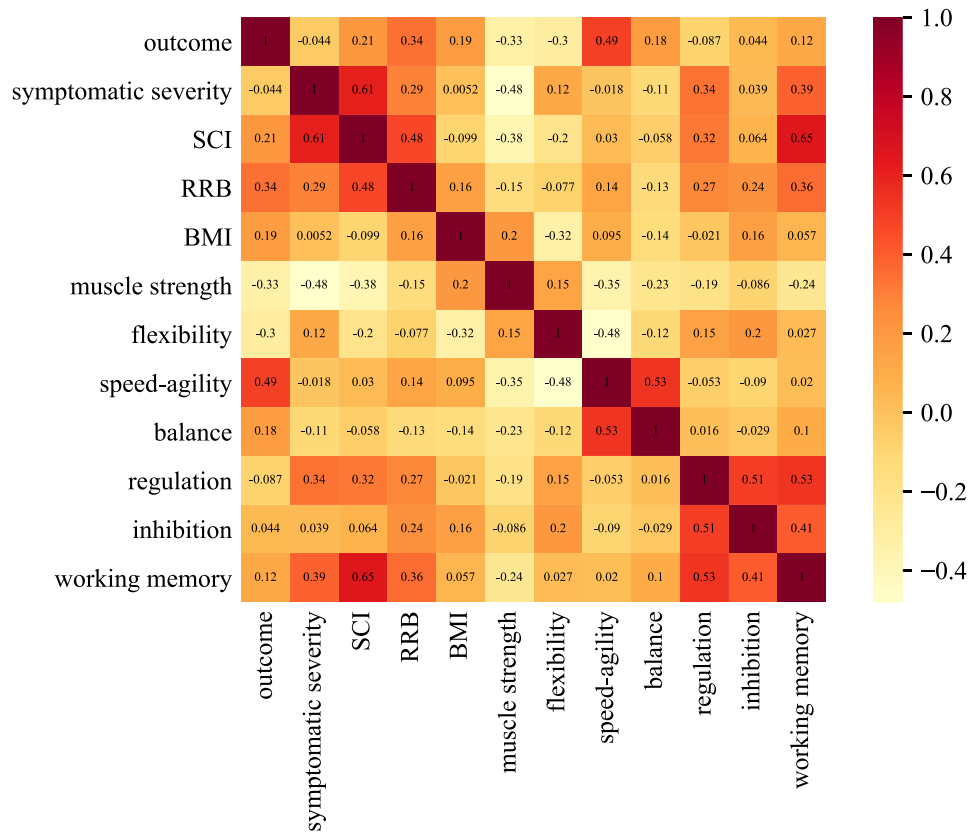


Fig. 3. The partial results of Pearson correlation analysis.

Note: SCI is social communication impairments, RRB is restricted and repetitive behaviors, and BMI is body mass index.

Table 3

The results of the principal component analysis.

Factors	PC 1	PC 2	PC 3	PC 4
RRB	0.042	0.354	-0.266	0.743
Muscle strength	0.031	-0.468	0.372	0.341
Flexibility	0.112	-0.44	-0.488	-0.032
Speed-agility	-0.009	0.614	0.21	-0.219
ORBsup.L	-0.281	-0.083	0.168	0.011
ORBsup.R	-0.277	-0.032	0.02	-0.073
IFGoperc.R	-0.211	0.043	0.094	0.025
IFGtriang.L	-0.262	-0.013	0.186	-0.031
IFGtriang.R	-0.27	-0.086	0.085	0.154
ORBinf.L	-0.267	-0.175	0.005	0.054
REC.L	0.269	-0.077	0.154	-0.093
REC.R	-0.275	-0.064	0.116	-0.075
LING.L	-0.274	-0.007	-0.225	-0.125
SOG.L	-0.251	0.027	-0.255	0.013
MOG.L	-0.251	0.084	-0.337	0.092
FFG.L	-0.266	0.017	-0.098	-0.044
SPG.R	-0.244	0.079	-0.308	-0.071
SMG.R	-0.21	0.112	0.235	0.231
THA.R	-0.221	0.03	-0.058	0.387
Eigenvalue	11.02	1.85	1.13	1.06
Variance contribution	56.59%	9.5%	5.82%	5.44%

Note: RRB is restricted and repetitive behaviors, ORBsup.L and ORBsup.R are the left and right orbital part of superior frontal gyrus, IFGoperc.R is right opercular part of inferior frontal gyrus, IFGtriang.L and IFGtriang.R are the left and right triangular part of inferior frontal gyrus, ORBinf.L is left orbital part of inferior frontal gyrus, REC.L and REC.R are the left and right gyrus rectus, LING.L is the left lingual gyrus, SOG.L is the left superior occipital gyrus, MOG.L is the left middle occipital gyrus, FFG.L is the left fusiform gyrus, SPG.R is the right superior parietal gyrus, SMG.R is the right supramarginal gyrus, and THA.R is the right thalamus.

inferior frontal gyrus ($r = 0.376, p = 0.015$), the right triangular part of inferior frontal gyrus ($r = 0.281, p = 0.075$), the left orbital part of inferior frontal gyrus ($r = 0.282, p = 0.074$), the left gyrus rectus ($r = 0.294, p = 0.063$), the right gyrus rectus ($r = 0.279, p = 0.077$), the left lingual gyrus ($r = 0.303, p = 0.054$), the left superior occipital gyrus ($r = 0.278, p = 0.079$), the left middle occipital gyrus ($r = 0.293, p = 0.063$), the left fusiform gyrus ($r = 0.357, p = 0.022$), the right superior parietal gyrus ($r = 0.342, p = 0.029$), the right supramarginal gyrus ($r = 0.293, p = 0.063$), and the right thalamus ($r = 0.309, p = 0.05$).

Initially, the Pearson correlation included 19 predictors, and we conducted PCA to retain the main data components. As shown in Table 3, we extracted four principal components with eigenvalues greater than 1, defined as PC 1, PC 2, PC 3, and PC 4. The cumulative variance contribution rate was 77.35%. Among them, GMV held a significant weight in PC 1, while clinical information and behavioral factors held a significant weight in PC 2. Finally, we selected four principal components for the regression task to predict the SCI outcome following the MBTP in children with ASD.

After determining the predictors, we used the cross validation method to fit three RF regression models to predict the SCI outcome following the MBTP in ASD children. In Fig. 4, each row shows an RF regression model established in each cross validation. Taking the first RF regression model as an example, the left figure shows the dataset division of the first cross validation. The first column contains samples with a data index of 1–10 (from top to bottom), and so on. The last column contains 11 samples with a data index of 31–41. The test set contains samples marked in red, while the rest is the training set. The graph in the middle is a regression tree in the first RF regression model. This regression tree randomly selected 18 samples, starting with node $PC\ 1 \leq 0.689$ is the first judgment condition, with six samples not meeting the condition and predicted to be -0.685 , and 12 samples meeting the condition. Then, with node $PC\ 4 \leq 0.189$ in the second judgment

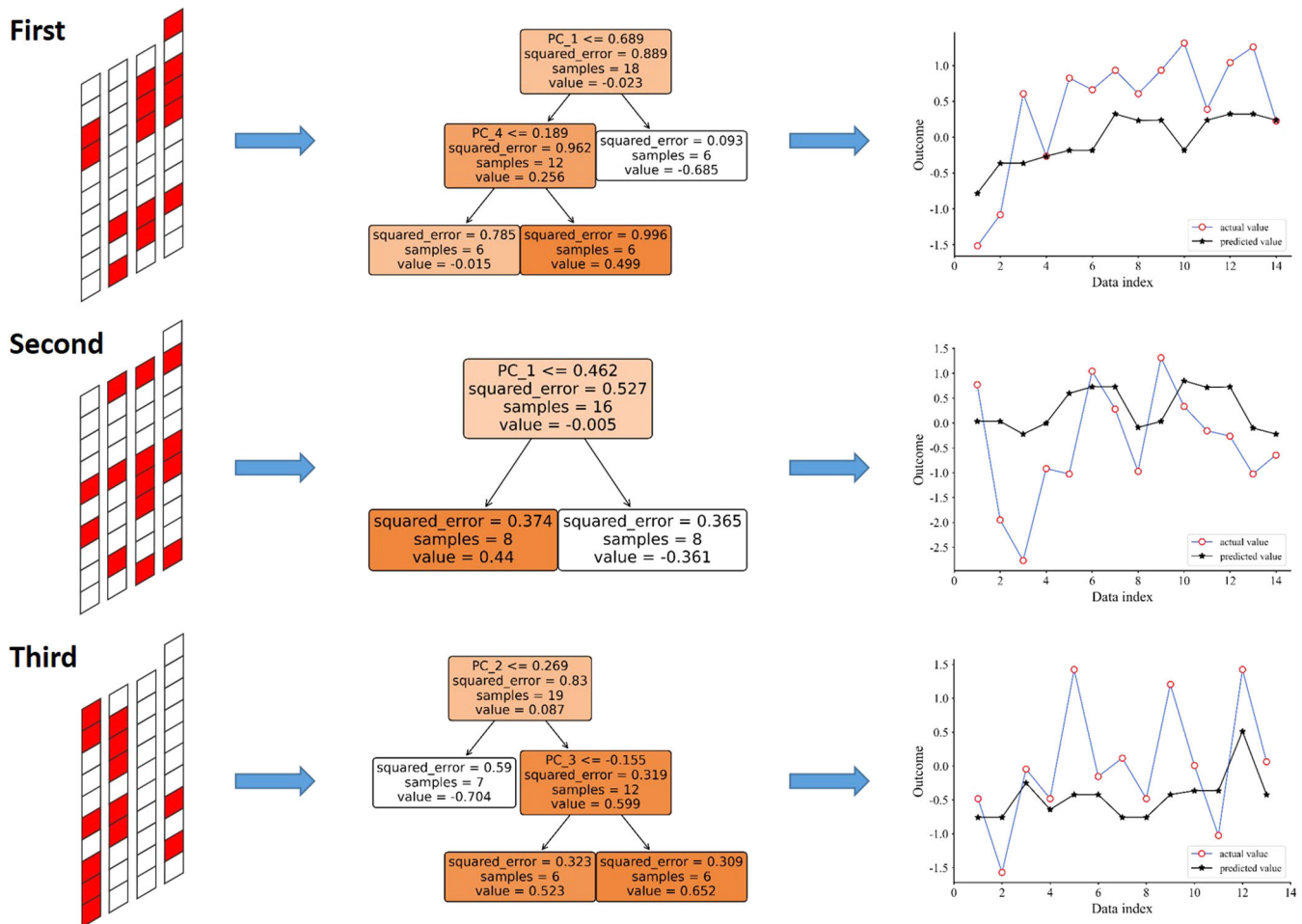


Fig. 4. The results of regression task.

Note: The optimal hyperparameters determined by grid search are mainly: (i) $n_estimators = 6$; (ii) $max_depth = 3$; (iii) $min_samples_leaf = 6$; (iv) $min_samples_split = 2$; (v) $max_features = 2$.

condition, with six samples met this condition and were predicted to be -0.015 , while six samples not met this condition and predicted to be 0.499 . The final prediction result of the first RF regression model is the average of the prediction results of all regression trees. The right figure shows the comparison between the prediction and the actual values. The abscissa represents the data index of the test set of the first RF regression model, while the ordinate represents the SCI outcome for these ASD children. In the regression task, the explained variance of three RF regression models was 40.35%, 21.1%, and 30.35%. The final average explained variance of the regression task was 30.58%. In addition, we used the `feature_importances_` function in RF to extract the feature importance (range: 0-1). The importance of PC 1, PC 2, PC 3, and PC 4 was 0.38, 0.31, 0.02, and 0.29, respectively.

Classification task

With respect to demographic information, the Chi-squared test result of sex and core symptom outcome had no statistical significance ($\chi^2 = 0.808$, $p = 0.369$), and there was no significant correlation between age and core symptom outcome ($r = 0.055$, $p = 0.731$). In the classification task, Spearman correlation analysis was used to quantify the relationship between other candidate factors and core symptom outcomes in children with ASD. As shown in Fig. 5, in clinical information and behavioral factors, the core symptom outcome of children with ASD was correlated to RRB ($r = 0.429$, $p = 0.005$), muscle strength score ($r = -0.287$, $p = 0.069$), and speed-agility score ($r = 0.282$, $p = 0.074$) at

baseline. In the brain structure indicators, no correlation was observed between the GMV of any brain region and the core symptom outcome of ASD children. Therefore, we selected RRB, muscle strength score, and speed-agility score as predictors for the classification task to predict core symptom outcomes following the MBTP in children with ASD.

Similarly, we used cross validation to fit three RF classification models to predict the core symptom outcome following the MBTP in children with ASD. In Fig. 6, each row shows an RF classification model established in each cross validation. Using the first RF classification model as an illustration, the figure on the left depicts the dataset partitioning of the initial cross-validation. The test set contains red-marked samples, while the remaining samples were used for training. The graph in the middle was a classification tree in the first RF classification model. This classification tree randomly selected 18 samples. The first criterion was the node $speed_agility \leq -0.695$, with seven samples met the condition and were predicted as a Fair outcome group, and 11 samples did not. Then, taking the node $muscle_strength \leq -0.487$ as the second criterion, five samples met this condition and were predicted as a Fair outcome group. Six samples did not meet this condition and were predicted as a Better outcome group. The final prediction result of the first RF classification model was the majority vote of all classification trees, as shown on the right side of the confusion matrix. In the classification task, the accuracy of the three RF classification models was 57.14%, 64.29%, and 76.92%. The final average accuracy of the classification task was 30.58% (permutation test, iterations = 1000, $p = 0.024$). We also used the `feature_importances_` function to extract the feature

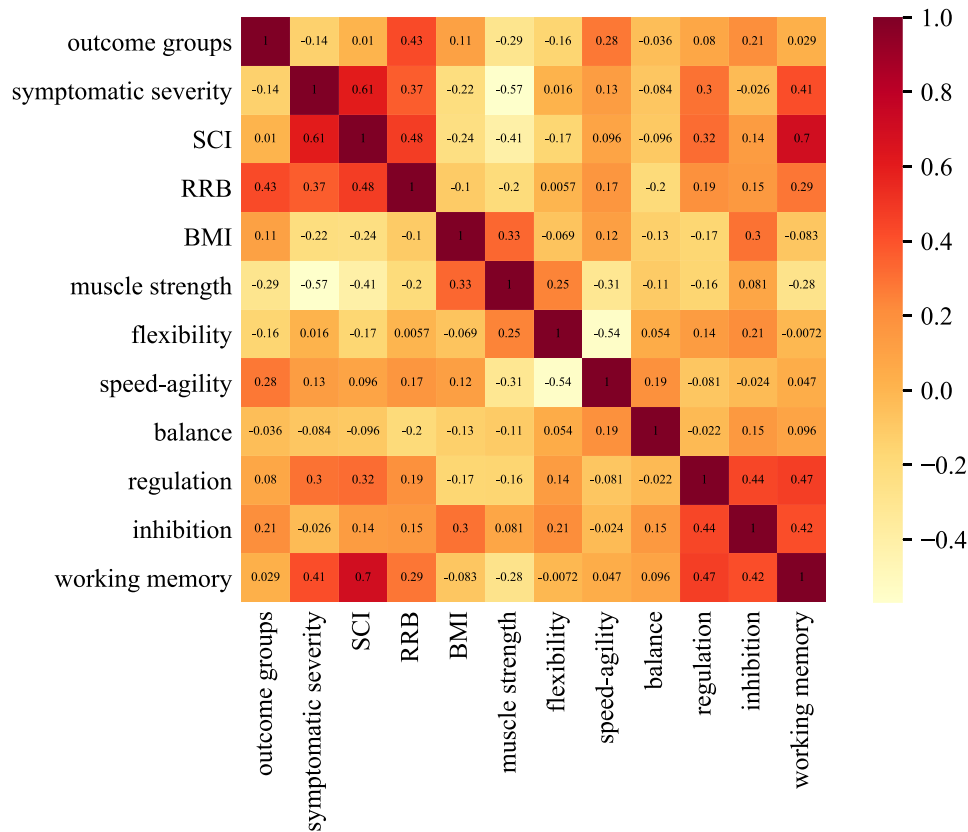


Fig. 5. The partial results of Spearman correlation analysis

Note: SCI is social communication impairments, RRB is restricted and repetitive behaviors, BMI is body mass index.

importance. The importance of RRB, muscle strength, and speed-agility was 0.57, 0.27, and 0.16, respectively.

Discussion

To maximize the benefits of exercise intervention, we used the study of the MBTP improving the symptom performance of ASD children as an example, applied the supervised machine learning method to predict the potential benefits of ASD children based on their uniqueness, and investigated and validated the efficacy of this method. Our findings indicated that the supervised machine learning method could predict potential benefits before an exercise intervention for ASD children.

We found that baseline RRB, muscle strength, flexibility, speed-agility, and GMV of 15 brain regions were predictors for the SCI outcome in ASD children. The baseline RRB, muscle strength, and speed-agility were predictors for core symptom outcomes in ASD children. To begin with, the SCI outcome and core symptom outcome of ASD children were positively correlated with baseline RRB, indicating that the more severe the RRB before the intervention, the better the outcomes of the MBTP. The RRB may limit the development of social skills in ASD children (Leekam et al., 2011). In contrast, ASD children with more severe RRB before intervention will achieve better intervention outcomes in the RRB symptom ($r = 0.489, p = 0.001$), providing them with more opportunities for communication. Besides, the SCI outcome and core symptom outcome of ASD children were negatively correlated with muscle strength and speed-agility at baseline, indicating that the worse the muscle strength and speed-agility before the intervention, the better the possible intervention outcomes. A previous study has shown that the MBTP improves muscle strength, speed agility, and SCI in ASD children (Cai et al., 2021). This may be because the MBTP has a larger capacity for improvement in ASD children with poorer physical fitness compared to those with better physical fitness, thereby better improving their core

symptoms. To confirm this potential relationship, further exploration is needed. Finally, the SCI outcome of ASD children was positively correlated with the GMV of 15 brain regions, indicating that ASD children with larger GMV in these brain regions may achieve better SCI outcomes. In addition, the three sub-dimensions of executive function were not correlated with the SCI outcome and core symptom outcome of ASD children, indicating that differences in executive function before receiving MBTP may not affect the benefits of ASD children.

We applied the RF algorithm in supervised machine learning to complete the regression and classification tasks. One advantage of choosing the RF algorithm is that it can visualize the decision tree learners, which opens the "black box" in the decision-making process (Dreiseitl & Ohno-Machado, 2002). In the regression task, we use the SCI outcome of children with ASD as the target variable. The explainable variance of the three RF regression models was 40.35, 21.1, 30.35%, and the average explainable variance was 30.58%, which predicted the SCI outcome following the MBTP in ASD children. In addition, we used dimensionality reduction methods to extract four principal components in feature engineering. Although their cumulative variance contribution rate reached 77.35% (Cao et al., 2022), some data information loss is still unavoidable. In the classification task, we used the core symptom outcome groups of ASD children as the target variable. The accuracy of the three RF classification models was 57.14, 64.29, 76.92%, and the average accuracy was 66.12%, which distinguished the core symptom outcome groups following the MBTP in ASD children. Few studies have applied the supervised machine learning method to predict the outcomes of exercise intervention for ASD children. There is no similar study for us to compare the model's prediction performance.

Several limitations of this study should be noted. Firstly, the data for this study comes from a longitudinal research project for a special population, and the sample size is insufficient. Future work on predicting the outcomes of exercise intervention in ASD children should use

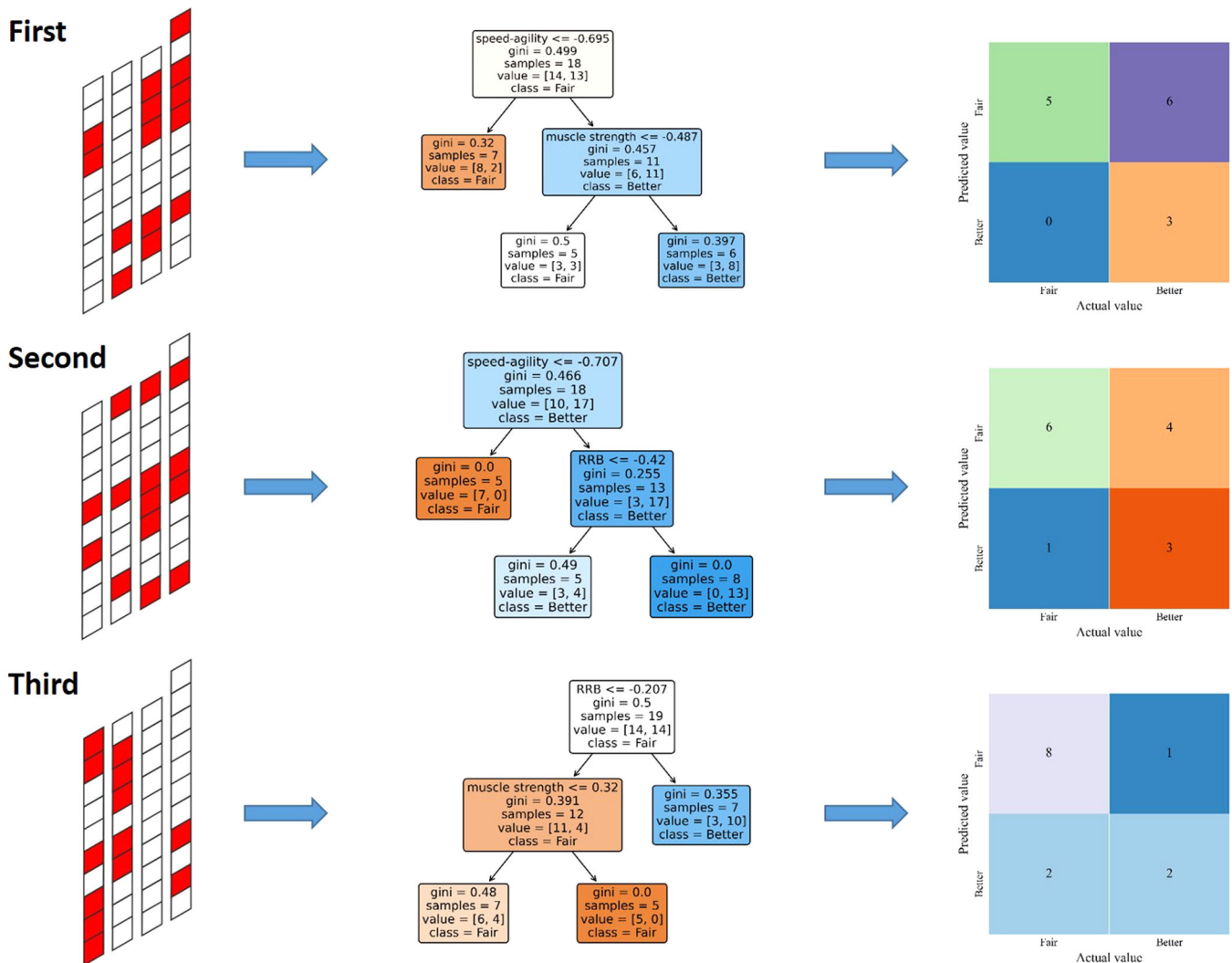


Fig. 6. The results of classification task.

Note: RRB is restricted and repetitive behaviors. The optimal hyperparameters determined by grid search are mainly: (i) $n_estimators = 6$; (ii) $max_depth = 3$; (iii) $min_samples_leaf = 5$; (iv) $min_samples_split = 2$; (v) $max_features = 2$.

larger samples, which will help us improve prediction performance using complex models such as convolutional neural networks (Marro-n-Esquivel et al., 2023). Secondly, it is unclear which variables may serve as reliable predictors of exercise intervention outcomes. We have selected some clinical information, behavioral factors, and brain structural indicators. Still, future research should not be limited to the selected factors. The genetic characteristics (Hodges et al., 2020) and neuroanatomical differences (Aglinskas et al., 2022) in ASD children should also be evaluated as potential predictors. Finally, since only one exercise intervention program was included in this study, we do not know how these ASD children would react to other exercise interventions. Ideally, different models can be established to predict the outcomes of different exercise intervention programs. Through further integration, recommend the most suitable exercise intervention program for each ASD child.

Conclusion

Our results confirmed that the supervised machine learning method can be applied to predict the outcomes of exercise intervention for children with ASD. Using the study on the efficacy of the MBTP for improving the core symptoms of ASD children as an example, the RF models can predict the SCI outcome and core symptom outcome

following the MBTP from the individual uniqueness in ASD children. Our findings provide a new and reliable method for identifying ASD children who are most likely to benefit from a specific exercise intervention program in advance and lay a solid foundation for establishing a personalized exercise intervention program recommendation system for ASD children.

Declaration of Competing Interest

All authors declare that they have no conflicts of interest to report regarding the present study.

References

- Aglinskas, A., Hartshorne, J. K., & Anzellotti, S. (2022). Contrastive machine learning reveals the structure of neuroanatomical variation within autism. *Science (New York, N.Y.)*, 376(6597), 1070–1074. <https://doi.org/10.1126/SCIENCE.ABM2461>
- Anzeneder, S., Zehnder, C., Martin-Niedeken, A. L., Schmidt, M., & Benzing, V. (2023). Acute exercise and children's cognitive functioning: What is the optimal dose of cognitive challenge? *Psychology of Sport and Exercise*, 66, Article 102404. <https://doi.org/10.1016/j.psychsport.2023.102404>
- Bakkelund, D. (2021). Order preserving hierarchical agglomerative clustering. *Machine Learning*, 111(5), 1851–1901. <https://doi.org/10.1007/s10994-021-06125-0>
- Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and*

- Developmental Disorders*, 30(3), 237–243. <https://doi.org/10.1023/A:1005596502855>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bremer, E., Crozier, M., & Lloyd, M. (2016). A systematic review of the behavioural outcomes following exercise interventions for children and youth with autism spectrum disorder. *Autism: The International Journal of Research and Practice*, 20(8), 899–915. <https://doi.org/10.1177/1362361315616002>
- Bruni, T. P. (2014). Test review: Social responsiveness scale–second edition (SRS-2). *Journal of Psychoeducational Assessment*, 32(4), 365–369. <https://doi.org/10.1177/0734282913517525>
- Cai, K., Wang, J., Liu, Z., Zhu, L., Xiong, X., Klich, S., et al. (2021). Mini-basketball training program improves physical fitness and social communication in preschool children with autism spectrum disorders. *Journal of Human Kinetics*, 73, 267–278. <https://doi.org/10.2478/hukin-2020-0007>
- Cai, K., Yu, Q., Herold, F., Liu, Z., Wang, J., Zhu, L., et al. (2020). Mini-basketball training program improves social communication and white matter integrity in children with autism. *Brain Sciences*, 10(11), 803. <https://doi.org/10.3390/brainsci10110803>
- Cao, Z., Yuan, H., & Li, H. (2022). Multi-source indexes risk identification of goaf based on PCA-PNN. *Journal of Safety Science and Technology*, 18(12), 104–109. <https://doi.org/10.11731/j.issn.1673-193x.2020.12.014>
- Castiglioni, I., Rundo, L., Codari, M., Di Leo, G., Salvatore, C., Interlenghi, M., et al. (2021). AI applications to medical images: From machine learning to deep learning. *Physica Medica*, 83, 9–24. <https://doi.org/10.1016/j.ejmp.2021.02.006>
- Chen, A., Xiong, X., Zhu, L., Dong, X., & Wang, W. (2021). Physical exercise and development of brain and mind in children and adolescents: Evidence and theory. *China Sport Science*, 41(11), 43–51. <https://doi.org/10.16469/j.css.202111006>
- Cheng, W., Sun, Z., Cai, K., Wu, J., Dong, X., Liu, Z., et al. (2023). Relationship between overweight/obesity and social communication in autism spectrum disorder children: Mediating effect of gray matter volume. *Brain Sciences*, 13(2). <https://doi.org/10.3390/brainsci13020180>
- Clemente, F. M., Moran, J., Ramirez-Campillo, R., Oliveira, R., Brito, J., Silva, A. F., et al. (2022). Recreational soccer training effects on pediatric populations physical fitness and health: A systematic review. *Children*, 9(11), 1776. <https://doi.org/10.3390/children9111776> (Basel, Switzerland).
- Cruz-Tirado, J. P., Lima Brasil, Y., Freitas Lima, A., Alva Pretel, H., Teixeira Godoy, H., Barbin, D., et al. (2023). Rapid and non-destructive cinnamon authentication by NIR-hyperspectral imaging and classification chemometrics tools. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 289, 122226. <https://doi.org/10.1016/j.saa.2022.122226>
- Dong, X., Chen, A., Liu, Z., Wang, J., Cai, K., & Xiong, X. (2020). Effects of mini-basketball training on repetitive behaviors and gray matter volume in preschool children with autism spectrum disorder. *China Sport Science And Technology*, 56(11), 25–31. <https://doi.org/10.16470/j.csst.2020126>
- Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, 35(5–6), 352–359. [https://doi.org/10.1016/s1532-0464\(03\)00034-0](https://doi.org/10.1016/s1532-0464(03)00034-0)
- Eisenberg, I. W., Wallace, G. L., Kenworthy, L., Gotts, S. J., & Martin, A. (2015). Insistence on sameness relates to increased covariance of gray matter structure in autism spectrum disorder. *Molecular Autism*, 6(1), 54. <https://doi.org/10.1186/s13229-015-0047-7>
- El-Sayes, J., Harasym, D., Turco, C. V., Locke, M. B., & Nelson, A. J. (2019). Exercise-induced neuroplasticity: A mechanistic model and prospects for promoting plasticity. *The Neuroscientist*, 25(1), 65–85. <https://doi.org/10.1177/1073858418771538>
- Greenacre, M., Groenen, P. J. F., Hastie, T., D'Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, 2(1). <https://doi.org/10.1038/s43586-022-00184-w>
- Haghighi, A. H., Broughani, S., Askari, R., Shahrabadi, H., Souza, D., & Gentil, P. (2022). Combined physical training strategies improve physical fitness, behavior, and social skills of autistic children. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-022-05731-8>
- Hodges, H., Fealko, C., & Soares, N. (2020). Autism spectrum disorder: Definition, epidemiology, causes, and clinical evaluation. *Translational Pediatrics*, 9(Suppl 1). <https://doi.org/10.21037/tp.2019.09.09>. S55–S65.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Kim, J. W., Sharma, V., & Ryan, N. D. (2015). Predicting methylphenidate response in ADHD using machine learning approaches. *International Journal of Neuropsychopharmacology*, 18(11). <https://doi.org/10.1093/ijnp/pyv052>. pyv052.
- Kim, Y., Kim, K. H., Park, J., Yoon, H. I., & Sung, W. (2023). Prognosis prediction for glioblastoma multiforme patients using machine learning approaches: Development of the clinically applicable model. *Radiotherapy and Oncology*, 183, 109617. <https://doi.org/10.1016/j.radonc.2023.109617>
- Kruse, C., Eiken, P., & Vestergaard, P. (2017). Clinical fracture risk evaluated by hierarchical agglomerative clustering. *Osteoporosis International*, 28(3), 819–832. <https://doi.org/10.1007/s00198-016-3828-8>
- Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders: A review of research in the last decade. *Psychological Bulletin*, 137(4), 562–593. <https://doi.org/10.1037/a0023341>
- Lenhard, F., Sauer, S., Andersson, E., Månsson, K. N., Mataix-Cols, D., Rück, C., et al. (2018). Prediction of outcome in internet-delivered cognitive behaviour therapy for paediatric obsessive-compulsive disorder: A machine learning approach. *International Journal of Methods in Psychiatric Research*, 27(1), e1576. <https://doi.org/10.1002/mp.1576>
- Li, X., Liu, J., Yang, W., Cao, B., He, X., Li, Z., et al. (2012). Executive function, theory of mind, and symptom in children with high functioning autism. *Chinese Mental Health Journal*, 26(8), 584–589. <https://doi.org/10.3969/j.issn.1000-6729.2012.08.005>
- Lin, Z., Lai, J., Chen, X., Cao, L., & Wang, J. (2022). Curriculum reinforcement learning based on K-fold cross validation. *Entropy (Basel, Switzerland)*, 24(12). <https://doi.org/10.3390/e24121787>
- Liu, X., Wang, Y., Fu, C., Xiao, H., & Xing, S. (2020). Auxiliary diagnosis of Alzheimer's disease based on feature extraction. *Chinese Journal of Medical Physics*, 37(5), 656–660. <https://doi.org/10.3969/j.issn.1005-202X.2020.05.024>
- Lord, C., Brugha, T. S., Charman, T., Frazier, T., Jones, E. J. H., State, M. W., et al. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers*, 6(1), 694–701. <https://doi.org/10.1038/s41572-019-0138-4>
- Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *Lancet (London, England)*, 392(10146), 508–520. [https://doi.org/10.1016/S0140-6736\(18\)31129-2](https://doi.org/10.1016/S0140-6736(18)31129-2)
- Maenner, M. J., Warren, Z., Williams, A. R., Amoakohene, E., Bakian, A. V., Bilder, D. A., et al. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *Morbidity and Mortality Weekly Report*, 72(2), 11–14. <https://doi.org/10.15585/MMWR.SS7202A1>
- Marouf, A. A., Mottalib, M. M., Alhajj, R., Rokne, J., & Jafarullah, O. (2022). An efficient approach to predict eye diseases from symptoms using machine learning and ranker-based feature selection methods. *Bioengineering (Basel)*, 10(1), 25. <https://doi.org/10.3390/bioengineering10010025>
- Marrón-Esquivel, J. M., Durán-López, L., Linares-Barranco, A., & Domínguez-Morales, J. P. (2023). A comparative study of the inter-observer variability on gleason grading against deep learning-based approaches for prostate cancer. *Computers in Biology and Medicine*, 159, Article 106856. <https://doi.org/10.1016/j.combiomed.2023.106856>
- Martínez-González, A. E., & Piqueras, J. A. (2018). Validation of the repetitive behavior scale-revised in Spanish-speakers participants with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 48(1), 198–208. <https://doi.org/10.1007/s10803-017-3276-0>
- Muñoz-Organero, M., & Queipo-Álvarez, P. (2022). Deep spatiotemporal model for COVID-19 forecasting. *Sensors (Basel, Switzerland)*, 22(9), 3519. <https://doi.org/10.3390/s22093519>
- Nematollahi, H., Moslehi, M., Aminolroayaei, F., Maleki, M., & Shahbazi-Gahrouei, D. (2023). Diagnostic performance evaluation of multiparametric magnetic resonance imaging in the detection of prostate cancer with supervised machine learning methods. *Diagnostics (Basel, Switzerland)*, 13(4), 806. <https://doi.org/10.3390/diagnostics13040806>
- Paul, A., Mukherjee, D. P., Das, P., Gangopadhyay, A., Chintia, A. R., & Kundu, S. (2018). Improved random forest for classification. *IEEE Transactions on Image Processing*, 27(8), 4012–4024. <https://doi.org/10.1109/TIP.2018.2834830>
- Pearson, R., Pisman, D., Meyer, B., Shumake, J., & Bevers, C. G. (2019). A machine learning ensemble to predict treatment outcomes following an internet intervention for depression. *Psychological Medicine*, 49(14), 2330–2341. <https://doi.org/10.1017/S003329171800315X>
- Popović Krneta, M., Šobić Šaranović, D., Mijatović Teodorović, L., Krajčinić, N., Avramović, N., & Bojović, Z., et al. (2023). Prediction of cervical lymph node metastasis in clinically node-negative T1 and T2 papillary thyroid carcinoma using supervised machine learning approach. *Journal of Clinical Medicine*, 12(11), 3641. <https://doi.org/10.3390/jcm12113641>
- Rosenblatt, L. E., Gorantla, S., Torres, J. A., Yarmush, R. S., Rao, S., Park, E. R., et al. (2011). Relaxation response-based yoga improves functioning in young children with autism: A pilot study. *Journal of Alternative and Complementary Medicine (New York, N.Y.)*, 17(11), 1029–1035. <https://doi.org/10.1089/acm.2010.0834>
- Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, 10(1), 91–103. <https://doi.org/10.1007/BF02408436>
- Sun, Z., Herold, F., Cai, K., Yu, Q., Dong, X., Liu, Z., et al. (2022). Prediction of outcomes in mini-basketball training program for preschool children with autism using machine learning models. *International Journal of Mental Health Promotion*, 24(2), 143–158. <https://doi.org/10.32604/ijmh.2022.020075>
- Thorell, L. B., & Nyberg, L. (2008). The childhood executive functioning inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology*, 33(4), 536–552. <https://doi.org/10.1080/87565640802101516>
- Toscano, C. V. A., Ferreira, J. P., Gaspar, J. M., & Carvalho, H. M. (2019). Growth and weight status of Brazilian children with autism spectrum disorders: A mixed longitudinal study. *Journal de Pediatria*, 95(6), 705–712. <https://doi.org/10.1016/j.jped.2018.06.008>
- Toscano, C. V. A., Ferreira, J. P., Quinaud, R. T., Silva, K. M. N., Carvalho, H. M., & Gaspar, J. M. (2022). Exercise improves the social and behavioral skills of children and adolescent with autism spectrum disorders. *Frontiers Psychiatry*, 13, Article 1027799. <https://doi.org/10.3389/fpsy.2022.1027799>
- Tse, C. Y. A., Pang, C. L., & Lee, P. H. (2018). Choosing an appropriate physical exercise to reduce stereotypic behavior in children with autism spectrum disorders: A non-randomized crossover study. *Journal of Autism and Developmental Disorders*, 48(5), 1666–1672. <https://doi.org/10.1007/s10803-017-3419-3>
- Tyler, K., MacDonald, M., & Meneer, K. (2014). Physical activity and physical fitness of school-aged children and youth with autism spectrum disorders. *Autism Research and Treatment*, 2014, Article 312163. <https://doi.org/10.1155/2014/312163>
- Walther, F., Heinrich, L., Schmitt, J., Eberlein-Gonska, M., & Roessler, M. (2022). Prediction of inpatient pressure ulcers based on routine healthcare data using

- machine learning methodology. *Scientific Reports*, 12(1), 5044. <https://doi.org/10.1038/s41598-022-09050-x>
- Wang, J., Cai, K., Liu, Z., Herold, F., Zou, L., Zhu, L., et al. (2020). Effects of mini-basketball training program on executive functions and core symptoms among preschool children with autism spectrum disorders. *Brain Sciences*, 10(5), 263. <https://doi.org/10.3390/brainsci10050263>
- Wei, W., Xie, Q., Zhu, J., He, W., & Li, Y. (2018). The psychometric characteristics of childhood executive functioning inventory among Chinese preschoolers. *Chinese Journal of Clinical Psychology*, 26(1), 26–29. <https://doi.org/10.16128/j.cnki.1005-3611.2018.01.006>
- Wu, M., Jiang, X., Du, K., Xu, Y., & Zhang, W. (2023). Ensemble machine learning algorithm for predicting acute kidney injury in patients admitted to the neurointensive care unit following brain surgery. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-33930-5>
- Xu, K., & Sun, Z. (2023). Predicting academic performance associated with physical fitness of primary school students using machine learning methods. *Complementary Therapies in Clinical Practice*, 51, Article 101736. <https://doi.org/10.1016/j.ctcp.2023.101736>
- Zhao, M., Chen, S., You, Y., Wang, Y., & Zhang, Y. (2021). Effects of a therapeutic horseback riding program on social interaction and communication in children with autism. *International Journal of Environmental Research and Public Health*, 18(5), 2656. <https://doi.org/10.3390/ijerph18052656>
- Zhou, H., Xu, X., Yan, W., Zou, X., Wu, L., Luo, X., et al. (2020). Prevalence of autism spectrum disorder in China: A nationwide multi-center population-based study among children aged 6 to 12 years. *Neuroscience Bulletin*, 36(9), 961–971. <https://doi.org/10.1007/s12264-020-00530-6>