International Journal on Semantic Web and Information Systems
Volume 20 • Issue 1

A Model for Predicting Physical Health of College Students Based on Semantic Web and Deep Learning Under Cloud Edge Collaborative Architecture

Yu Wang, Hohai University, China

Zhiyi Zhang, Hohai University, China

Peng Tang, Hohai University, China*

Shiyao Bian, Hohai University, China

ABSTRACT

A model for predicting physical health of college students based on semantic web and deep learning under cloud edge collaborative architecture is proposed to address the issue of most physical health prediction models being unable to fully describe the characteristics of sports performance changes and having large prediction errors. Firstly, the authors design a measurement data analysis system based on cloud edge collaboration architecture to improve data analysis efficiency. Then, they preprocess the data on the edge side, such as missing samples, and extract data features using an equal dimensional dynamic GOM model. Finally, they deploy the RBFNN-SSA model in the cloud center, input the characteristics of each indicator into the model for predictive analysis, and obtain the physical health status of college students. Based on the physical health test data of Hohai University from 2018 to 2021, an experimental analysis was conducted. The results showed that all three intervention measures had significant effects on maintaining and improving the physical health level of college students.

KEYWORDS

Cloud Edge Collaboration, College Students' Physical Health, Data Preprocessing, Equal Dimensional Dynamic GOM Model, Physical Fitness Prediction, RBFNN-SSA Model

College students are the guarantee of social development in China and represent the country's potential for national athletic strength. The physical health level of college students is not only related to their personal health and future development, but also has a strong correlation with the future development of the entire country, which to some extent determines the success or failure of the rejuvenation of the Chinese nation (Xinwen, 2023; Ming et al., 2020). The results of the 8th Chinese Student Physical Health Monitoring Survey jointly organized by six departments including the National Health Commission in 2019 showed that the overall level of physical health of college

DOI: 10.4018/IJSWIS.340379 *Corresponding Author

students (PHCS) is on the rise, but physical fitness indicators such as endurance, speed, and strength are still hovering at low levels (Liangyi, 2021; Computational Intelligence and Neuroscience, 2023). Therefore, it is very important to scientifically, promptly, and effectively predict the health status of college students and provide a methodological basis for university administrators to promptly grasp the current situation of students' physical health.

The traditional method of predicting physical examination scores for college students is to establish a discrete recursive model, which often has obvious limitations (Tan, 2023). Currently most quantitative methods for system analysis are mathematical statistics, such as regression analysis, etc. Their main requirements for samples are the following: first, a large sample size; second, samples that have good patterns (Wang, 2018). A mixed-effects, multi-logistic, normal model was designed in reference (Gerber & Craig, 2021) to predict player performance and quantify uncertainty. One study (Liang et al., 2022) proposes a jogging data analysis method based on intelligent computing and uses the analysis results to guide the improvement of jogging performance. Another study (Dubbs, 2018) uses a simple machine learning model—logical weighted regularized linear least-squares regression—to predict sports test results. Although the above methods have good predictive ability in specific scenarios, the quantitative methods of system analysis are greatly constrained, and the physical fitness test data of college students fluctuate, making it difficult to meet the distribution requirements of samples that must have good regularity in mathematical statistics.

Deep learning algorithms have been applied to predict the physical test scores of college students (Wang et al., 2023). Su and Chen (2022) used a deep learning method—Extreme Gradient Enhancement (XGB)—to predict player performance. Men (2022) proposed a sports action recognition method based on an improved Gaussian fuzzy algorithm. Zhu et al. (2023) proposed a rating prediction model based on lifting and decision tree regression (BDTR-SP) that used decision tree regression (DTR) as the ensemble learning structure of the learning machine to improve prediction accuracy. Korchi et al. (2023) utilized regression algorithms such as random forests and deep neural networks to predict student grades. With the powerful data analysis ability of deep learning, the above methods have achieved good prediction results, but their level of prediction accuracy struggles to meet the requirements of current physical education curriculum settings, and the prediction efficiency needs to be improved.

Therefore, a model for predicting PHCS health based on semantic web and deep learning under cloud edge collaborative architecture is proposed. The innovation points are summarized as follows:

1) Due to the low efficiency of traditional centralized data analysis models, the proposed method introduces a cloud edge collaborative architecture and deploys data preprocessing and feature extraction modules to the edge side, shortens transmission distance, and reduces computational pressure on cloud center servers, thereby shortening prediction time and ensuring data security.
2) In response to the problem of irregular distribution of student physical testing data, the proposed method uses semantic web to represent the correlation between physical health and various factors, in order to compensate for the shortcomings of mathematical statistics.
3) The proposed method uses Sparrow Search Algorithm (SSA) to optimize the parameters of Radial Basis Function Neural Network (RBFNN) and applies the optimized model to predict PHCS. The model is not affected by the characteristics of the data, and the prediction results are more accurate. This provides a scientific basis and practical reference for fully mobilizing the motivation of college students to participate in sports and improving their physical health level.

A DATA ANALYSIS SYSTEM BASED ON CLOUD EDGE COLLABORATION (CEC)

The physical form, function, and quality test results of college students are influenced by various factors and levels, and the fluctuation of physical test data is significant, even with great fluctuations. It is very difficult to predict the various results tested in college students' physical tests (Zhang, 2021;

Dedeliuk et al., 2018). To gain a more real-time, efficient prediction of PHCS, a prediction system based on CEC architecture was designed from a data-driven perspective. The architecture is shown in Figure 1.

First, preprocess the physical examination data of college students is preprocessed; for instance, missing values are processed. Then, the edge layer deploys an equal dimensional dynamic GOM model to extract corresponding features from various physical measurement indicators such as height and weight, and uploads them to the cloud center processing layer. Finally, the cloud center predicts various indicators and related influencing factors on the basis of the deployed prediction model and analyzes the physical and health status of college students, meeting the requirements of accurate analysis of large-scale college student physical test data and future physical and health status prediction.

Research Object

The survey included all current undergraduate students who began college from 2018 through 2021 at Hohai University, resulting in a total of 14,756 participants, comprising 10,186 male students and 4,570 female students. Five representative indicators were selected from the physical test data for this study, which included height, weight, vital capacity, sprint (50 m), and long-distance running (1,000 m for male students and 800 m for female students). The physical test data of undergraduate students at Hohai University who began college from 2018 through 2021 were used as the research subjects. Data collection and compilation took place between October and December 2022. The study was developed and implemented by integrating relevant literature and books on “physical education and health course teaching” and “physical fitness of college students.” The requirements of the National Physical Fitness Standards for Students were considered alongside the university’s specific circumstances.

The main objects of this study were the raw data from the 2018–2021 physical fitness tests, which included five indicators: height, weight, vital capacity, sprint, and long-distance running. Additionally, the derived indicator, body mass index (BMI), was also considered. All tests were conducted in strict accordance with the Work Manual of the National Student Physique and Health Survey, and the testing instruments met the standards set by the National Physical Health Standards for Students, referred to as the Standards. The grading and scoring of the tests were performed following the Standards, ensuring adherence to proper physical fitness testing procedures and the accuracy of the data obtained. The derived indicator, BMI, is defined as follows: BMI = weight/height^2 (kg/m^2). As indicated by the Standards, college students were classified into four categories: low weight, normal weight, overweight, and obesity. Table 1 displays the individual BMI evaluations for college students.

Figure 1. Data Analysis System Architecture Based on CEC

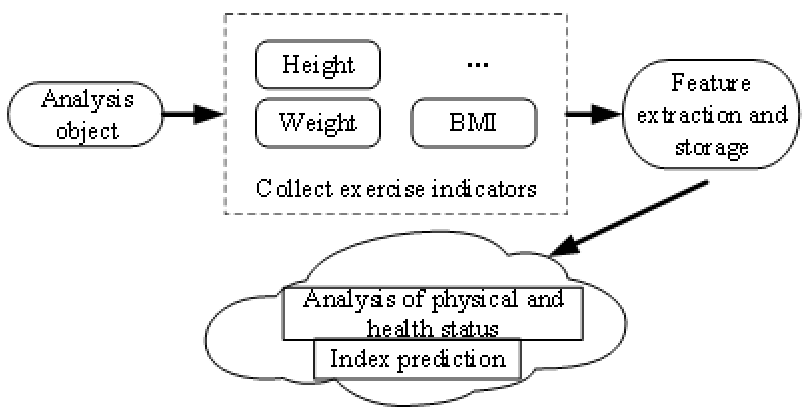


Table 1. Single BMI Evaluation for College Students (unit: kg/m²)

Grade	Male College Students	Female College Students
Low weight	≤17.8	≤17.1
Normal weight	17.9-23.9	17.2-23.9
Overweight	24.0-27.9	24.0-27.9
Obesity	≥28.0	≥28.0

Prior to the commencement of the study, data preprocessing was conducted on the physical test data of undergraduate students who began college in 2018. The students were initially categorized on the basis of the economic disparities of their respective places of origin, resulting in two groups: students from developed regions and students from less-developed regions. Subsequently, the students were further classified according to their gender. As a result, four distinct categories were formed: male students from developed regions, female students from developed regions, male students from less-developed regions, and female students from less-developed regions. The classification of regions was based on specific criteria, in which Beijing, Shanghai, Jiangsu, and Guangdong provinces were primarily classified as developed regions, while Anhui, Henan, Sichuan, and Jiangxi provinces were classified as less-developed regions.

Data Preprocessing

Missing Value Processing

Because of the large sample size of the research object, it was anticipated that unexpected situations might occur during physical testing, resulting in missing or empty values in the sample data. Therefore, it was necessary to preprocess the samples containing missing values (Christos et al., 2020). There are four main situations: 1) Manual filling: when the researchers have sufficient knowledge of the dataset, they can choose to fill in missing values. 2) Special value filling: researchers treat null values as a special attribute value. 3) K-means clustering: first, based on Euclidean distance, researchers determine the *K* samples closest to the sample with missing data and weighted average these *K* values to estimate the missing data of the sample (Wang Xiang et al., 2023). The Euclidean distance between points $x = \left(x_1, x_2, \cdots, x_n\right)$ and $y = \left(y_1, y_2, \cdots, y_n\right)$ is

$$d\left(x, y\right) = \sqrt{\left(x_1 - y_1\right)^2 + \left(x_2 - y_2\right)^2 + \cdots + \left(x_n - y_n\right)^2} = \sqrt{\sum_{i=1}^{n}\left(x_i - y_i\right)^2} \tag{1}$$

Handling of Outliers

Abnormal values refers to unreasonable values in the dataset that deviate from the normal range and are not erroneous values. Although outliers do not occur frequently, they can also have an impact on the actual project analysis, resulting in deviations in the results. The most common outliers in the research subjects are manual input errors in physical test data or individual specific indicator abnormalities. The proposed method adopts the Laida criterion (3 σ Principle) filter out outliers and eliminate them (Xiao Huiyue et al., 2021).

The classical 3σ rule is the following: let any dimension i have m data samples $x^i = \left(x_1^i, x_2^i, \cdots, x_m^i\right)$, abbreviated as $x = \left(x_1, x_2, \cdots, x_m\right)$, where the mean of x is μ and the standard deviation is σ. If

$$\left|x_j - \mu\right| \geq 3\sigma \tag{2}$$

where, x_j is considered an outlier.

Data Simplification

The amount of physical testing sample data was huge, and it was difficult to fully analyze all the data. It was necessary to simplify a large amount of data and determine a data indicator. The proposed method uses the arithmetic mean method to calculate the average of five physical measurement indicators. Due to the sensitivity of the average value to outliers and extreme values, the above 3σ It is necessary to handle outliers in principle.

The arithmetic mean is based on the least squares method and aims to describe the average trend. It is a representative statistic in the context of conventional sample probabilities. From the perspective of insight into causality, the mean is more able to provide direct answers to the problem of concern. Regardless of the distribution of the original data, the mean calculated by multiple data extracts will converge to a normal distribution. The general formula for arithmetic mean is

$$M = \frac{X_1 + X_2 + \cdots + X_n}{n} \tag{3}$$

Feature Extraction Method Based on Equal Dimensional Dynamic GOM Model in the Edge Side

The classic grayscale prediction feature models are the grayscale model GM (1,1) and GOM models. Generally, the longer the time series, the longer the prediction time, the larger the grayscale feature interval, and the corresponding prediction accuracy will also decrease (Yang & Cheng, 2021; Liu et al., 2023). To obtain the optimal gray level feature prediction interval for physical examination score prediction, an equal dimensional dynamic GOM gray level feature model was introduced to construct the prediction interval. The construction process was as follows:

1) Assuming the time series of physical examination scores was $X_i^0 = \left(x_{i1}^0, x_{i2}^0, \cdots, x_{it}^0\right)$, a GOM model was constructed based on this time series, and the corresponding predicted value $\hat{x}_{i.t+1}^0$ for the next time period could be obtained through this model.
2) Add the predicted results to the $X_i^0 = \left(x_{i1}^0, x_{i2}^0, \cdots, x_{it}^0\right)$. To ensure the same length of the $X_i^0 = \left(x_{i1}^0, x_{i2}^0, \cdots, x_{it}^0\right)$, it is necessary to remove the result x_{i1}^0 located at the beginning of the $X_i^0 = \left(x_{i1}^0, x_{i2}^0, \cdots, x_{it}^0\right)$ and construct a new time series $X_i^0 = \left(x_{i2}^0, x_{i3}^0, \cdots, x_{it}^0, \hat{x}_{i,t+1}^0\right)$. By using the $X_i^0 = \left(x_{i2}^0, x_{i3}^0, \cdots, x_{it}^0, \hat{x}_{i,t+1}^0\right)$, the GOM model can be reconstructed, and the next predicted value $\hat{x}_{i.t+2}^0$ can be calculated.
3) Add $\hat{x}_{i.t+2}^0$ to the $X_i^0 = \left(x_{i1}^0, x_{i2}^0, \cdots, x_{it}^0\right)$ again and repeat step 2) until the predicted features can reach the prediction target stopping.

After obtaining the grayscale feature prediction interval of physical measurement data, a semantic web was constructed using the correlation of various factors, providing a data foundation for subsequent prediction of PHCS.

PREDICTION OF COLLEGE STUDENTS' PHYSICAL HEALTH BASED ON DEEP LEARNING IN THE CLOUD CENTER

SSA-RBFNN Model

RBFNN is composed of an input layer (IL), a hidden layer (HL), and an output layer (OL) (Yadav Rajesh et al., 2023; Yu et al., 2023). The construction of HL space is to use radial basis functions as the basis of hidden layer units. Set 1 as the connection weight between the IL and the HL in RBFNN. The role of the HL is to optimize the parameters of the activation function, while the role of the OL is to optimize the connection weight. The center of the basis function, the width of the hidden layer, and the connection weight are the three parameters that RBFNN needs to solve for. The commonly used basis function in RBFNN is the Gaussian function, whose activation function is represented as

$$R_{(x_k - o_i)} = \exp\left(-\frac{1}{2\tau^2}\left\|x_k - o_i\right\|^2\right) \tag{4}$$

where the input sample is x_p, $p = 1, 2, \cdots, N$ and the total number is *N*. The centers of HL network node are c_i, $i = 1, 2, \cdots, H$, where *H* represents the number of HL nodes; τ is the variance of the Gaussian function.

The output y_k of the RBFNN network is represented as

$$y_k = \sum_{i=1}^{h} \omega_{i,k} \exp\left(-\frac{1}{2\tau^2}\left\|x_k - o_i\right\|^2\right) \tag{5}$$

where $\left\|x_k - o_i\right\|$ and o_i are the centers of the Euclidean norm and Gaussian function, respectively, while $\omega_{i,k}$ is the connection weight from the HL to the OL.

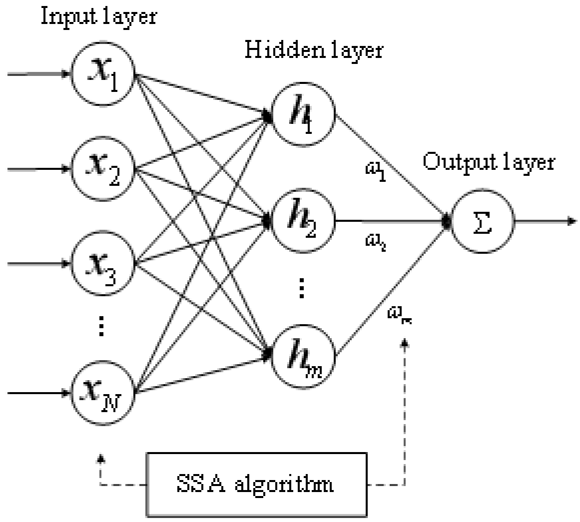
The learning ability of the RBFNN network is determined mainly by the center of τ, $\omega_{i,k}$, and basis functions. To ensure optimal network performance, SSA algorithm is used for parameter optimization; its optimization architecture is shown in Figure 2. The SSA algorithm simulates the foraging and anti-predation behavior of sparrows. This algorithm is relatively novel, with significant global search ability and high convergence speed (Wang., X., Ma, X., et al., 2023; Xue et al., 2023).

In the process of sparrows' foraging, they are divided into discoverers, followers, and vigilantes. As discoverers, sparrows actively seek food and provide foraging paths and directions. As followers, sparrows will obtain food through the directions provided by the enrollees, and the two can be exchanged. When the entire sparrow population is threatened by predators or becomes aware of danger, the vigilant sparrow will engage in anti-predatory behavior, and the sparrow on the periphery of the population will continuously adjust its location to obtain a better position.

The location and speed updates of the SSA algorithm are represented as follows:

1) Foraging behavior, updating discoverer location:

Figure 2. Optimization Architecture for Parameters of RBFNN Model Based on SSA



$$D_{ij}^{t+1} = \begin{cases} D_{ij}^{t} \cdot \exp\left(\dfrac{-i}{\chi \cdot T}\right), \Omega_1 L < \Omega_2 \\ D_{ij}^{t} + G \cdot L, \Omega_1 > \Omega_2 \end{cases} \tag{6}$$

where G is a random number; L is the unit row vector; χ is a random number between [0,1]; D_{ij}^{t} represents the position information; t is the current iteration number; Ω_1 and Ω_2 are warning values and safety values, respectively.

2) Update joiner location:

$$D_{ij}^{t+1} = \begin{cases} G \cdot \exp\left(\dfrac{D_{worst}^{t} - D_{ij}^{t}}{i^2}\right), i > N/2 \\ D_{P}^{t+1} + \left|D_{ij}^{t} - D_{P}^{t+1}\right| \cdot \kappa \cdot L, i \leq N/2 \end{cases} \tag{7}$$

where D_P represents the position of the current discoverer; D_{worse} is the global worst-case position; κ is a row vector that randomly contains only 1 and -1 elements.

3) Anti-predatory behavior, updating sparrow population position:

$$D_{ij}^{t+1} = \begin{cases} D_{best}^{t} + \delta \cdot \left| D_{ij}^{t} - D_{best}^{t} \right|, f_i > f_{best} \\ D_{ij}^{t} + S\left(\dfrac{\left| D_{ij}^{t} - D_{worst}^{t} \right|}{f_i - f_{worst} + \zeta} \right), f_i = f_{best} \end{cases} \tag{8}$$

where S is a random number between [-1,1], f_i is the individual fitness value, and ζ is a constant close to 0; D_{best} is the current global optimal position; δ is the step size parameter; f_{best} and f_{worse} are the global optimal and worst-case fitness values, respectively.

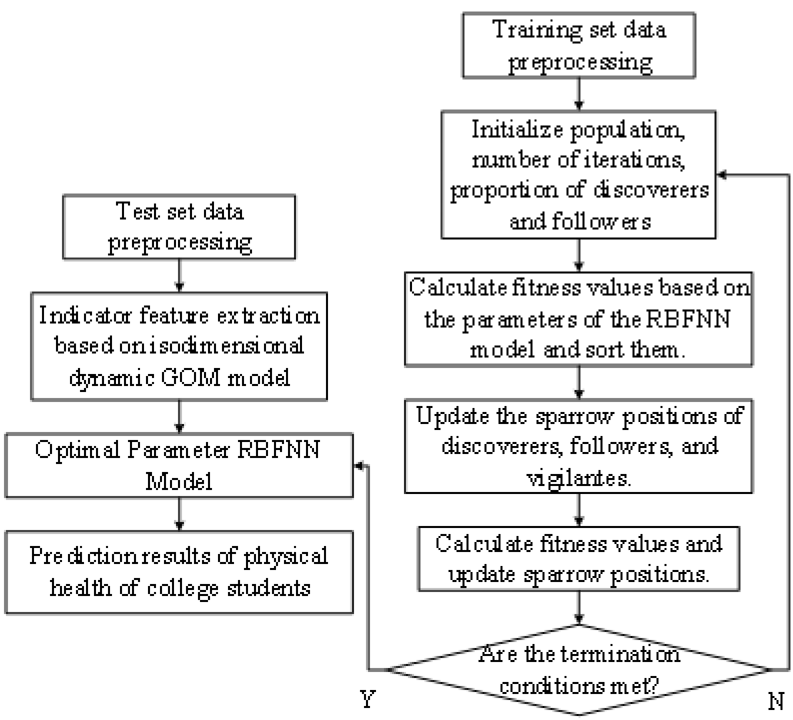
A Prediction Model for PHCS Based on SSA-RBFNN

In the system, the edge side uploads the various indicator features obtained from the equal dimensional dynamic GOM model to the cloud center, and the cloud center deploys the SSA-RBFNN model for predicting the PHCS. The overall process is shown in Figure 3.

The specific steps for predicting the PHCS are as follows:

1) Collect the physical test scores of college students from 2018 to 2021 at Hohai University and then construct a physical test score dataset (i.e. training and testing sets) and preprocess the data to obtain the time series of college student physical test data;
2) Using an equal dimensional dynamic GOM model, extract features from five indicators, as well as their derived BIM values;
3) Optimize the parameters of the RBFNN model using the SSA algorithm, using the following steps:

Figure 3. Prediction Process of PHCS Based on SSA-RBFNN



Step 1: Set the number of sparrow populations and iteration times, and set the ratio of sparrow discoverers and followers in advance;
Step 2: Calculate the fitness values of sparrows using the parameters ω, τ, and o of the RBFNN model, and sort them accordingly. The fitness value is the difference between the predicted output value and the actual value;
Step 3: Update the discoverer, follower, and predator positions of sparrows from equations (6) to (7);
Step 4: By calculating the fitness function value, continuously update it to the optimal individual position of all sparrows;
Step 5: Determine whether the requirements are met. If the set value is reached, the operation ends; Otherwise, go to Step 3 to continue iterating and updating;

4) Obtain the optimal RBFNN model based on the optimal parameters and input the characteristic values and influencing factors of each indicator into the model to predict the physical test scores of college students, including the physical test scores of college students from 2018 to 2021, as well as the physical test scores of 2023 and 2024. On the basis of the predicted results of physical examinations, the health status of students can be grasped, a finding which also provides technical support for the development of healthcare services (Nguyen et al., 2021; Xiao et al., 2022; Alamer et al., 2022; Vagelatos & Sarivougioukas, 2021; Sarrab & Alshohoumi, 2021).

RESEARCH RESULTS AND ANALYSIS

The experiment was conducted on AMDRyzen53350H, 16G memory, and 2 on a 10GHz computer; the program was written and implemented on MATLAB 2019b (Mo et al., 2023; Yen et al., 2023). In the SSA algorithm, the proportion of sparrow discoverers is 0.7, the proportion of followers is 0.3, the proportion of early warning sparrows is 0.2, the warning value is 0.6, the population size is 30, and the number of iterations is 100. The number of neurons in the input layer of the RBFNN network is 5, and the number of neurons in the HL is twice that of the input layer. Divide the data from the 2018–2021 physical health test (see "Research Object" section) into a training set and a testing set by 5:1. Simultaneously using commonly used encryption techniques in the system to ensure data security (Fatemidokht et al., 2021; Nguyen et al., 2021; Gupta et al., 2021; Xu et al., 2021).

Analysis of the Impact of Different Factors

Strong intervention measures were first used in sociology. Professor Shen Yuan, in his research on social transformation in China since the reform and opening up, referred to the intervention method of "strong intervention" as "exploring new ways, increasing efforts, and even trying to directly instill certain concepts into groups with slow development of social self-organization mechanisms, to promote the development of their autonomy."

Affected by COVID-19, students' health has declined significantly (Koweyes et al., 2021; Mouawad et al., 2021; Yu & Reiff-Marganiec, 2022; Narayan et al., 2022). This article draws inspiration from the theoretical framework of sociology and believes that schools adopt new methods and means to enhance students' emphasis on physical exercise, cultivate their lifelong sports awareness and habits, and continuously strengthen measures to promote physical health during the process. Specifically, there are three types of pre-intervention measures taken by Hohai University: 1) offering public physical education courses and strengthening physical fitness exercises during freshman and sophomore years; 2) Using mobile software, urging freshmen and sophomores to strengthen physical exercise in their spare time, and convert the effective exercise frequency or accumulated kilometers of long-distance running recorded in the background into public physical education course scores; 3) Incorporating the results of public physical education courses and physical fitness tests into the rules for evaluating scholarships for students.

BMI Status of College Students in Different Regions

The BMI index is an indicator that reflects an individual's obesity, thinness, and health level. The comparison of BMI among college students from 2018 to 2021 is shown in Figure 4.

From Figure 4 it is evident that the BMI of boys and girls in different regions declined significantly in 2019. Analysis of the domestic situation at that time shows that the outbreak of COVID-19 and the resultant decline in quality of life for residents led to a significant decline in the BMI index of students during this period. Because of the greater impact of the epidemic on underdeveloped areas compared to developed areas, the decrease in BMI of students in underdeveloped areas was significantly greater than that of students in developed areas. After 2020, students who began college in 2018 had basically completed their public physical education courses during their university years. Due to epidemic prevention and control, during their sophomore year, students could only participate in physical education courses online. In addition, limited activities and external environment have led to a significant decrease in the participation of most students in physical exercise. Moreover, the relatively fixed learning and life rhythm of students after returning to school has led to a significant increase in their BMI index in the following two years.

Impact of Public Physical Education Courses (PPEC) on Students' Physical Health

During their freshman and sophomore years, Hohai University offered a wide range of PPEC for undergraduate students, allowing them to independently choose sports to participate in their studies. During the epidemic, in order to ensure the quality of PPEC teaching, a mixed online and offline teaching mode was adopted to meet students' physical education learning and exercise needs (Hidalgo et al., 2022; Xiao et al., 2022; Salloum & Tekli, 2021). Taking lung capacity as an example, the comparison of the impact of PPEC on the PHCS is shown in Figure 5.

By analyzing the physical examination data of students who began college in 2018 for four years, it can be found that in addition to lung capacity indicators, the BMI, sprint, and long-distance test scores all improved during the second physical examination, and their second physical examination scores were also their best among the four college physical examinations. The second physical examination was conducted during the last academic year of the physical education course of students who began college in 2018. Subsequently, during the last two physical tests, the test scores of this group of students showed a downward trend and dropped to the lowest level during the fourth physical test.

After two academic years of compulsory PPEC and exercise, the physical examination results indicate that the public physical education courses at Hohai University have a certain effect on

Figure 4. Comparison of BMI Among College Students of Hohai University From 2018 to 2021

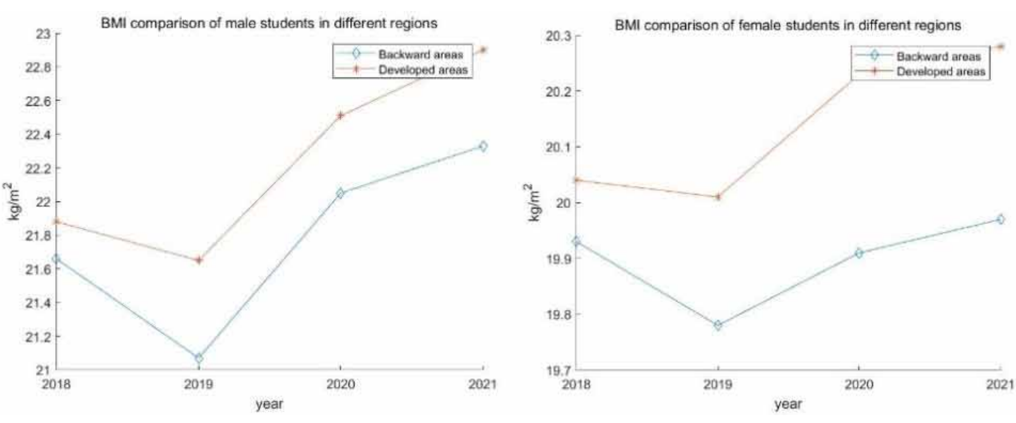
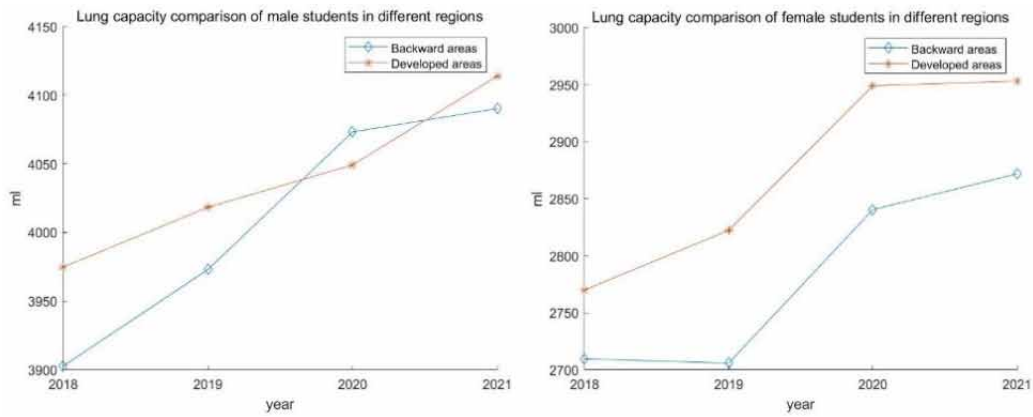


Figure 5. Impact of PPEC on Students' Physical Health



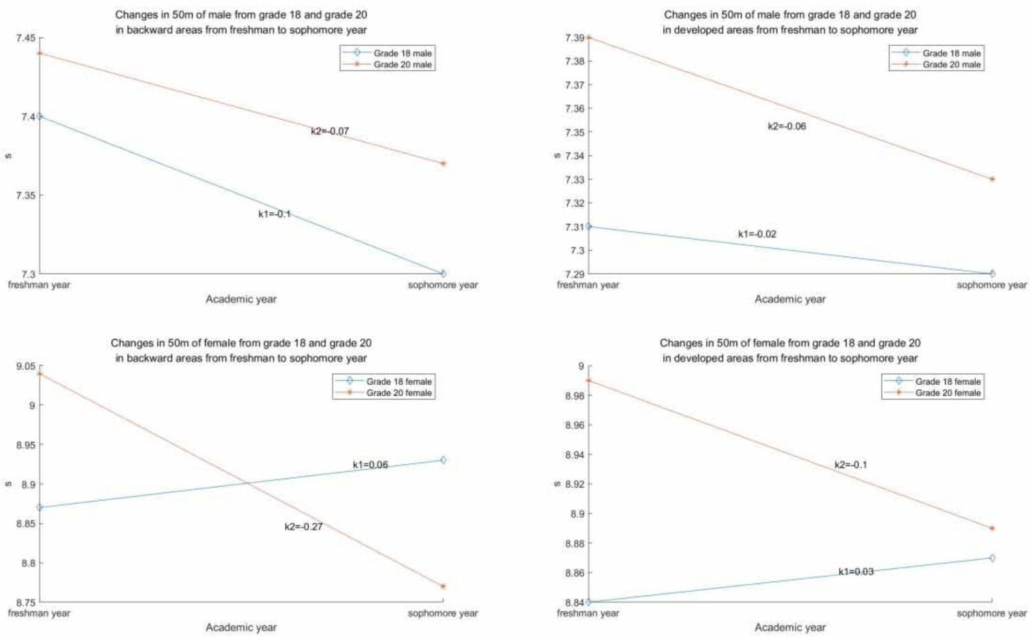
improving students' physical fitness. However, after students entered the public physical education courses that are not compulsory in their third and fourth years, their physical fitness level significantly decreased, reaching a level even worse than when they first enrolled. Therefore, this study suggests that PPEC in universities have a certain effect on maintaining and improving the physical fitness of students, but the cultivation of exercise habits is insufficient, failing to enable students to achieve the goal of sports participation.

Impact of the Implementation of the "Flashing" Plan on Students' Physical and Mental Health

Starting in 2020, Hohai University has planned and implemented the "Flashing" program, which stipulates that students must use the "Flashing Campus" app to complete the rated distance and number of long-distance runs within the semester. At the end of the semester, the cumulative effective number of long-distance runs will be converted into physical education grades. Students who fail to meet the requirements will have a maximum physical education grade of 70 points this semester, and this will be included in the teaching syllabus of Hohai University until 2020. Due to the implementation of "Flashing" only after the admission of undergraduate students who began college in 2020, the variable "whether or not to participate in Flashing "was used in data analysis. Students who began college in 2018 who did not participate in "Flashing" were used as the control group, and students who began college in 2020 who participated in" Flashing "were used as the experimental group. The intervention effect of "Flashing" on promoting the PHCS was compared and analyzed, as shown in Figure 6.

In Figure 6, k1 and k2 respectively represent the degree of changes in the physical testing data of students who began college in 2018 and 2020 during their first and second years. Overall, the improvement of the running level of students who began college in 2018 was slow or even decreased when "Flashing" was not conducted. However, students who began college in 2020 showed a significant improvement in their sprint test results after "Flashing," reflecting an improvement in their physical fitness in terms of explosive power. Comparing the running levels of students who began college in 2020 from the same region, it can be found that in the sprint event, boys in underdeveloped areas increased by 6.37s, boys in developed areas increased by 8.81s, while girls in underdeveloped areas increased by 8.09s and girls in developed areas increased by 9.37s. This can indicate that girls are more serious about "Flashing" than boys, and the improvement effect of their running skills is more significant.

Figure 6. Impact of the Implementation of the "Flash" Plan on Students' Physical Health



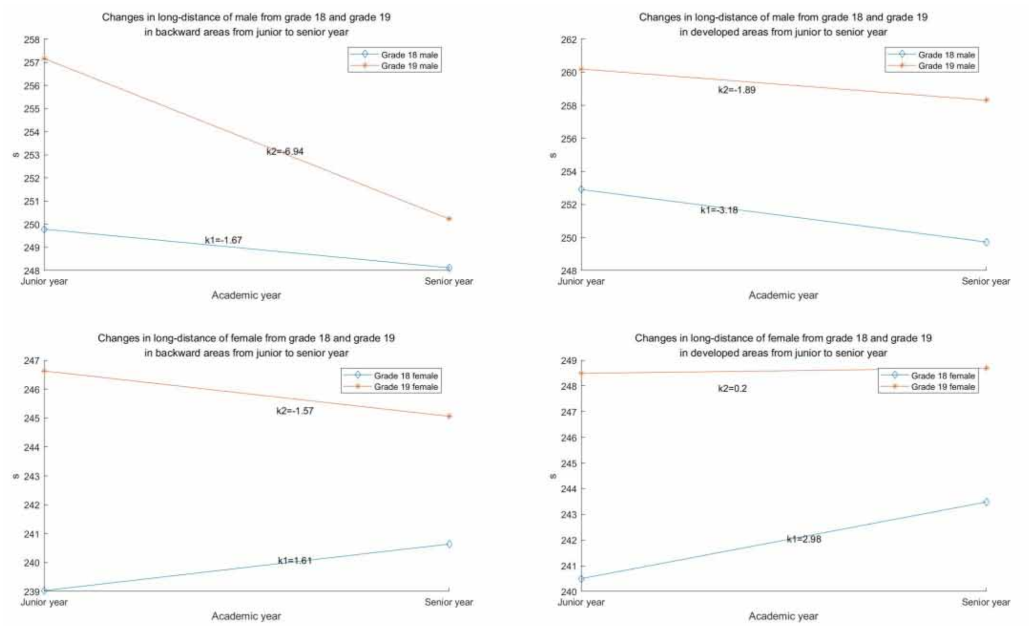
Impact of Scholarship Scoring System on Students' Physical Health

Scholarships are one of the important motivating factors for students to improve their academic performance in various aspects and also one of the strong measures taken by schools to promote students' physical health. During their junior and senior years at Hohai University in 2018, they did not include body side details in their scholarship policies. However, the scholarship policies for their junior and senior years in 2019 included body side details. By comparing the last two physical examination levels of 2018 and 2019, the intervention effect of the scholarship policy on maintaining and improving students' physical health without physical education courses can be judged, as shown in Figure 7.

As shown in Figure 7, overall, the scholarship policy can play a role in improving the physical health level of college students. In the long-distance running event, the performance improvement of students who began college in 2019 was significantly better than that of students who began college in 2018. In 2019, despite the absence of physical education classes in their third and fourth years, busy spare time, and scholarship policies, students could still actively engage in exercise to maintain or even improve their physical fitness level.

This study also found that there are gender differences in the intervention effects of scholarship policies. Comparing the data between males and females, it is evident that the physical condition of males did not significantly decrease without the incentive of scholarship policies, and the introduction of scholarship policies did not have a significant promoting effect on the physical intervention of males; However, in the absence of intervention from scholarship policies, the physical condition of girls significantly decreased. After the introduction of scholarship policies, the physical condition of girls improved significantly. Therefore, the introduction of scholarship policies has a stronger intervention effect on the physical health of female students.

Figure 7. The Impact of Scholarship Scoring System on Students' Physical Health



Analysis of the Impact of Different Factors

In order to demonstrate the predictive performance of the proposed model, it was compared with research by Dubbs (2018) and Zhu et al. (2023). The prediction accuracy and efficiency of the three methods for the data from 2018 to 2021 is shown in Table 2.

From Table 1 it can be seen that compared to other methods, the proposed method has a more accurate prediction, with an average accuracy of 96.95%. Because of the small fluctuation in lung capacity, the prediction accuracy is relatively high. Long-distance running was greatly influenced by various factors, resulting in a slight decrease in prediction accuracy. Dubbs used a logically weighted regularized linear least squares regression model to predict performance, but the method was more traditional, resulting in a low prediction accuracy of only 89.36%. However, the prediction time was relatively short, at 15.24ms. Zhu et al. proposed a performance prediction method based on the BDTR-SP model, which improved the prediction accuracy by 93.96% compared to Dubbs. However, the model complexity increased and the prediction time increased, reaching 25.08ms. Meanwhile, the DTR model lacked accurate feature extraction, resulting in a decrease in prediction accuracy compared to the proposed method. The proposed method combines the equal dimensional dynamic GOM model and RBFNN-SSA model for performance prediction, ensuring the reliability of the prediction. The cloud edge collaborative architecture is utilized to improve the prediction efficiency, with a prediction time of 19.71ms.

Analysis of the Trend of PHCS

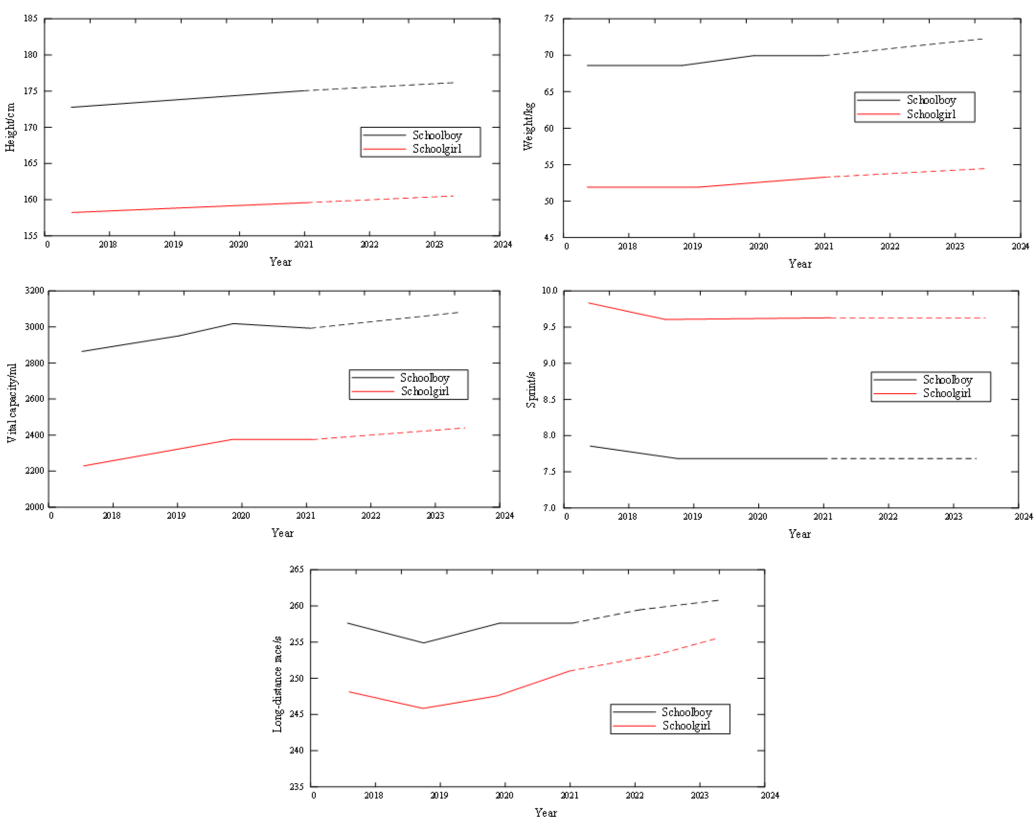
Based on the original average scores of physical fitness and health items for college students at Hohai University in the past four years from 2018 to 2021, and combined with the proposed model to predict the average scores for the next two years, the trend of each item from 2018 to 2023 is obtained. The five indicators are shown in Figure 8.

From Figure 8 it can be seen that the height and weight of male and female students are steadily increasing, while the upward trend of height and weight in female students is much

Table 2. Prediction Results of Different Models

Index	Prediction Accuracy /%		
	Alexander Dubbs, 2018	**Ling Zhu et al., 2023**	**Proposed Method**
Schoolboy height	90.54	95.62	98.13
Schoolgirl height	90.08	94.29	97.35
Schoolboy weight	89.81	94.51	97.46
Schoolgirl weight	89.46	93.89	96.97
Schoolboy vital capacity	91.37	95.73	98.53
Schoolgirl vital capacity	91.65	95.96	98.67
Schoolboy sprint	87.29	92.47	95.71
Schoolgirl sprint	86.83	91.52	94.88
Schoolboy long-distance race	87.95	92.68	95.93
Schoolgirl long-distance race	87.44	91.84	95.02
Schoolboy BMI	90.21	94.75	97.54
Schoolgirl BMI	89.72	94.31	97.16
Mean value	89.36	93.96	96.95
Time/ms	15.24	25.08	19.71

Figure 8. Prediction Results of PHCS: (a) Height, (b) Weight, (c) Vital Capacity, (d) Sprint, (e) Long-Distance Race



more moderate than that in male students; This indicates that girls pay more attention to physical symmetry than boys, but because of the impact of the epidemic at home, online physical education teachers being unable to supervise, and because of personal inertia, male and female students continue to gain weight.

The lung capacity of both male and female students showed a sharp increase from 2018 to 2020, but showed a slight decrease from 2020 to 2021, followed by a steady increase in the following two years; In 2020 and 2021, offline physical education classes and after-school physical activities were repeatedly interrupted by the epidemic, and it was difficult for students to practice more endurance projects. Teachers urged students to practice endurance running more to improve their physical fitness, and their lung capacity performance will continue to rise in the future.

The 50-meter race for male and female students showed an upward trend only from 2018 to 2019 and has remained relatively stable since then. Girls are generally more stable than boys in this indicator, with little fluctuation.

Long distance running events have shown a sharp upward trend for male and female students only from 2018 to 2019, followed by a continuous downward trend, and will continue to decline in the future. The decline for female students is greater than that for male students and should be taken seriously by schools, teachers, and students. This decline has a significant impact on online classes at home due to the epidemic, difficulties in ensuring continuous practice of endurance running after class, and students' subjective fear of endurance running, Teachers need to allocate more in-class and after-class tasks for students' endurance running in order to correct the trend of continuous decline in students' endurance quality in the future.

CONCLUSION

With the increasing emphasis on physical education in China and the application of artificial intelligence in the field of education, the importance of accurate prediction of PHCS is becoming increasingly prominent (Hashash et al., 2021; Sarivougioukas & Vagelatos, 2022). To this end, a prediction model for PHCS based on deep learning is proposed under the CEC architecture. In the CEC architecture, the edge side completes data preprocessing and feature extraction, and the cloud center uses the RBFNN-SSA model for prediction analysis to output the status of PHCS. On the basis of the physical health test data of Hohai University from 2018 to 2021, the experimental results show the following:

1) The proposed model utilizes the equal dimensional dynamic GOM model and RBFNN-SSA model for prediction, achieving good prediction results with an average accuracy of 96.95%.
2) All three intervention measures have a significant effect on maintaining and improving the physical health level of college students and can provide better scientific guidance for teachers to develop relevant physical education courses.

The prediction accuracy of the proposed model for long-distance running and other projects still needs to be improved. Therefore, in future research, more influencing factors will be taken into account to comprehensively improve the prediction accuracy of the proposed model. At the same time, in practical applications, it may be affected by geographical factors, leading to a decrease in prediction accuracy. It is necessary to further improve the universality of the proposed method.

AUTHOR CONTRIBUTIONS

Yu Wang: Conceptualization and Methodology; Zhiyi Zhang: Validation; Peng Tang: Writing of original draft; Shiyao Bian: Writing of review and editing.

FUNDING

1. Research results of Social Science Foundation of Jiangsu Province(22TYD009)
2. The Fundamental Research Funds for the Central Universities(B230207017)

CONFLICTS OF INTEREST

The authors declare that publication of this paper does not involve any conflicts of interest.

DATA AVAILABILITY STATEMENT

The data used to support the findings of this study are available from the corresponding author upon request.

REFERENCES

Alamer, S. A., Ilyas, Q. M., Ahmad, M., & Irfan, D. (2022). A metaphoric design of electronic medical record (EMR) for periodic health examination reports: An initiative to cloud's medical data analysis. *International Journal of Cloud Applications and Computing*, *12*(1), 1–18. doi:10.4018/IJCAC.2022010110

Dedeliuk, N., Kovalchuk, N., Vashchuk, L., Tomashchuk, O., Saniuk, V., & Savhuk, S. (2018). Model of students sports and recreation activity organisation in the conditions of higher educational institution. *Physical Education. Sports and Health Culture in Modern Society*, *1*, 29–35.

Dubbs, A. (2018). Statistics-free sports prediction. *Model Assisted Statistics and Applications : An International Journal*, *13*(2), 173–181. doi:10.3233/MAS-180428

Fatemidokht, H., Rafsanjani, M. K., Gupta, B. B., & Hsu, C.-H. (2021). Efficient and secure routing protocol based on artificial intelligence algorithms with UAV-assisted for vehicular ad hoc networks in intelligent transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), 4757–4769. doi:10.1109/TITS.2020.3041746

Gerber, E. A. E., & Craig, B. A. (2021). A mixed effects multinomial logistic-normal model for forecasting baseball performance. *Journal of Quantitative Analysis in Sports*, *17*(3), 221–239. doi:10.1515/jqas-2020-0007

Grentzelos, C., Caroni, C., & Barranco-Chamorro, I. (2020). A comparative study of methods to handle outliers in multivariate data analysis. *Computational and Mathematical Methods*, *3*(3), 1–1.

Gupta, B. B., Li, K. C., Leung, V. C., Psannis, K. E., & Yamaguchi, S, (2021). Blockchain-assisted secure fine-grained searchable encryption for a cloud-based healthcare cyber-physical system. *IEEE/CAA Journal of Automatica Sinica, 8*(12), 1877–1890.

Hashash, O., Sharafeddine, S., & Dawy, Z. (2021). MEC-based energy-aware distributed feature extraction for mHealth applications with strict latency requirements. *2021 IEEE International Journal on Communications Workshops (ICC Workshops), 1*, 1–6.

Hidalgo, A., Pérez, N., & Lemus-Aguilar, I. (2022). Factors determining the success of eHealth innovation projects. *International Journal of Software Science and Computational Intelligence*, *14*(1), 1–22. doi:10.4018/IJSSCI.309709

Korchi, A., Messaoudi, F., Abatal, A., & Manzali, Y. (2023). machine learning and deep learning-based students' grade prediction. *Operations Research Forum, 4*(4), 1–1.

Koweyes, J., Salloum, T., Haidar, S., Merhi, G., & Tokajian, S. (2021). COVID-19 pandemic in Lebanon: One year later, what have we learnt? *mSystems*, *6*(2), 10–1128. doi:10.1128/mSystems.00351-21 PMID:33879497

Liang, K., Wan, Y., & Jiang, Y. (2022). The effect of jogging on the physical health of college students based on intelligent computing. *Mobile Information Systems*, *2022*, 1–1. doi:10.1155/2022/7491839

Liangyi, M. (2021). Intervention effect of trail happy running app on college students' physique under the influence of COVID-19 epidemic situation. *Basic & Clinical Pharmacology & Toxicology*, *128*, 235–236.

Liu, Q., Zhu, H., & Qu, Z. (2023). Experiment and simulation analyses of GOM based salinity gradient osmotic energy conversion with nonsymmetric photothermal effect. *Energy Conversion and Management*, *298*, 117783. doi:10.1016/j.enconman.2023.117783

Men, Y. (2022). Intelligent sports prediction analysis system based on improved Gaussian fuzzy algorithm. *Alexandria Engineering Journal*, *61*(7), 5351–5359. doi:10.1016/j.aej.2021.08.084

Ming, S., Songmei, L., & Rilang, L. (2020). A Study on Educational Values of the Special Monitoring by College Students' Physiques. *Basic & Clinical Pharmacology & Toxicology*, *126*, 97–98.

Mo, M., Ji, S., Wang, X., Ghulam, M., & Ren, Y. (2023). Privacy data management mechanism based on blockchain and federated learning. *Cmc-Computers. Computers, Materials & Continua*, *74*(1), 37–53. doi:10.32604/cmc.2023.028843

Mouawad, P., Dubnov, T., & Dubnov, S. (2021). Robust detection of COVID-19 in cough sounds: Using recurrence dynamics and variable Markov model. *SN Computer Science*, *2*(1), 34. doi:10.1007/s42979-020-00422-6 PMID:33458700

Narayan, N., Jha, R. K., & Singh, A. (2022). A differential epidemic model for information, misinformation, and disinformation in online social Networks: CoVId-19 vaccination. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–20. doi:10.4018/IJSWIS.300827

Nguyen, G. N., Viet, N. H. L., Elhoseny, M., Shankar, K., Gupta, B. B., & El-Latif, A. A. A. (2021). Secure blockchain enabled Cyber–physical systems in healthcare using deep belief network with ResNet model. *Journal of Parallel and Distributed Computing*, *153*, 150–160. doi:10.1016/j.jpdc.2021.03.011

Retracted: Analysis of improving effect of running APP on college students' physique by using student data mining technology. (2023). *Computational Intelligence and Neuroscience*.

Salloum, G., & Tekli, J. (2021). Automated and personalized nutrition health assessment, recommendation, and progress evaluation using fuzzy reasoning. *International Journal of Human-Computer Studies*, *151*, 102610. doi:10.1016/j.ijhcs.2021.102610

Sarivougioukas, J., & Vagelatos, A. (2022). Fused contextual data with threading technology to accelerate processing in home UbiHealth. *International Journal of Software Science and Computational Intelligence*, *14*(1), 1–14. doi:10.4018/IJSSCI.285590

Sarrab, M., & Alshohoumi, F. (2021). Assisted-Fog-Based Framework for IoT-Based Healthcare Data Preservation. *International Journal of Cloud Applications and Computing*, *11*(2), 1–16. doi:10.4018/IJCAC.2021040101

Su, F., & Chen, M. (2022). Basketball players' score prediction using artificial intelligence technology via the Internet of Things. *The Journal of Supercomputing*, *78*(17), 19138–19166. doi:10.1007/s11227-022-04573-6

Tan, X. (2023). Enhanced sports predictions: A comprehensive analysis of the role and performance of predictive analytics in the sports sector. *Wireless Personal Communications*, *132*(3), 1613–1636. doi:10.1007/s11277-023-10585-z

Vagelatos, A., & Sarivougioukas, J. (2021). Using denotational mathematics for the formal description of home UbiHealth decision-support systems with knowledge flow. *International Journal of Software Science and Computational Intelligence*, *13*(4), 1–17. doi:10.4018/IJSSCI.2021100101

Wang, H. (2018). Design of college students' physique monitoring and service platform based on computer and network. *International Journal of Reasoning-Based Intelligent Systems*, *10*(2), 134–138. doi:10.1504/IJRIS.2018.092217

Wang, S., Wang, T., Pei, X., Wang, H., Zhu, Q., Tang, T., & Hou, T. (2023). Highway icing time prediction with deep learning approaches based on data from road sensors. *Science China. Technological Sciences*, *66*(7), 1987–1999. doi:10.1007/s11431-022-2230-8

Wang, X., Ma, X., Liu, X., Zhang, L., Tian, Y., & Ye, C. (2023). Research on optimal operation of cascade pumping stations based on an improved sparrow search algorithm. *Water Science and Technology: A Journal of the International Association on Water Pollution Research, 88*(8), 1982–2001.

Wang, X., Shao, Z., Shen, Y., & He, Y. (2023). Research on fast marking method for indicator diagram of pumping well based on K-means clustering. *Heliyon*, *9*(10), e20468–e20468. doi:10.1016/j.heliyon.2023.e20468 PMID:37842635

Xiao, H., Zhang, Y., Liu, X., Hao, Y., Liu, P., & Liu, D. C. (2021). A rapid ultrasound vascular disease screening method using PauTa criterion. *Journal of Physics: Conference Series*, *1769*(1), 012009. doi:10.1088/1742-6596/1769/1/012009

Xiao, J., Liu, X., Zeng, J., Cao, Y., & Feng, Z. (2022). Recommendation of healthcare services based on an embedded user profile model. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–21. doi:10.4018/IJSWIS.313198

Xinwen, M. (2023). Study on the effect of high intensity interval training on physical health of college students in plateau environment. *Frontiers in Sport Research*, *5*(7), 47–54.

Xu, Z., He, D., Vijayakumar, P., Gupta, B., & Shen, J. (2021). Certificateless public auditing scheme with data privacy and dynamics in group user model of cloud-assisted medical WSNs. *IEEE Journal of Biomedical and Health Informatics*, *27*(5), 2334–2344. doi:10.1109/JBHI.2021.3128775 PMID:34788225

Yadav, R., Yadav, D. K., & Kumar, A. (2023). RBF based some implicit–explicit finite difference schemes for pricing option under extended jump-diffusion model. *Engineering Analysis with Boundary Elements*, *156*, 392–406. doi:10.1016/j.enganabound.2023.08.021

Yang, Y., & Cheng, W. (2021). Research on the tourism industry impact of the national games based on grayscale models. *Tourism Management and Technology Economy, 4*(2).

Yen, C.-T., Chen, T.-Y., Chen, U.-H., Wang, G.-C., & Chen, Z.-X. (2023). Feature fusion-based deep learning network to recognize table tennis actions. *CMC-Computers*. *Computers, Materials & Continua*, *74*(1), 83–99. doi:10.32604/cmc.2023.032739

Yu, C., Lv, X., Huang, D., & Jiang, D. (2023). Reliability-based design optimization of offshore wind turbine support structures using RBF surrogate model. *Frontiers of Structural and Civil Engineering*, *17*(7), 1086–1099. doi:10.1007/s11709-023-0976-8

Yu, H. Q., & Reiff-Marganiec, S. (2022). Learning disease causality knowledge from the web of health data. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–19. doi:10.4018/IJSWIS.297145

Zhang, Y. (2021). The construction of college students' sports health management service platform based on artificial intelligence. *International Journal of System Assurance Engineering and Management, 13*(1), 1–9.

Zhu, L., Liu, G., Lv, S., Chen, D., Chen, Z., & Li, X. (2023). An intelligent boosting and decision-tree-regression-based score prediction (BDTR-SP) method in the reform of tertiary education teaching. *Information (Basel)*, *14*(6), 1–1. doi:10.3390/info14060317