Hindawi
Scientific Programming
Volume 2022, Article ID 8297739, 13 pages
https://doi.org/10.1155/2022/8297739

Research Article

Monitoring and Visualization of Physical Exercise Physiological Indicators Driven by Discrete Data

Ying Tang

School of Physical Education, Suzhou University, Suzhou 234000, China

Correspondence should be addressed to Ying Tang; tangying@ahszu.edu.cn

Received 19 January 2022; Accepted 16 February 2022; Published 11 March 2022

Academic Editor: Sheng Bin

Physical exercise physiological index monitoring has a wide range of applications in the fields of physiological index planning and design and organizational network evolution. Among the existing analysis methods for monitoring data points of physical exercise physiological indicators, the analysis error of point events under linear constraints is relatively large. Based on discrete data-driven datasets, this paper realizes the monitoring and visualization of sports physiological indicators. First, the principal component analysis of multivariate discrete data is used for dimensionality reduction. Second, the clustering of discrete physical exercise data uses the BIC criterion to preset the number of clusters, and the *R* software is used to visually realize the clustering results of physical exercise physiological indicators in each region in the text. The experiment solves the problem of mismatch of model parameter combinations when the physical exercise index monitoring quantity is used for the auxiliary analysis of the clustering results. Through the ARI index monitoring, the accuracy of the clustering physical exercise results of the method is increased to 89.7%, and the error rate is controlled within 4.3%. It promotes the superiority and effectiveness of multivariate discrete data-driven model clustering methods.

1. Introduction

In terms of discrete data processing, regression analysis uses mathematical statistics to reveal the interdependence (correlation) between two or more dimensions. If there is such a relationship, then when visualizing, all samples will be in a certain, and if this relationship does not exist, the visualization effect is a group of discrete points [1]. Cluster analysis is to divide the samples in the data set into several groups according to a certain relationship, the similarity in the same group is large, and the similarity between different groups is small [2–5], so this paper first adopts the principal component analysis (MFPCA) of multivariate functional data to reduce the dimension, and then uses the multivariate functional data clustering through. The BIC criterion presets the number of clusters, and the EM algorithm estimates the Gaussian mixture model density function parameter *R* and then uses *R* software to visually realize the clustering results of physiological indicators in various regions of the country in the text, and obtain different changing laws in different regions, and add the amount of data to assist in the analysis of the clustering results [6–8].

First, the B-spline basis function is selected to smoothly fit the original discrete data, and then, the functional dimension reduction is performed on the monthly average concentration curve based on the functional principal component analysis, and the functional principal component score matrix is obtained. The multivariate function Gaussian mixture model and the multivariate function K-means method are used to perform cluster analysis on the data quality of the country and region, and, respectively, describe the spatiotemporal characteristics of different physiological indicators affecting the data quality in different regions of the country in the past three years, and then use *R* software to realize the type of clustering results intuitively in the text [9–11]. Finally, the accuracy of the two clustering methods is compared through the simulation evaluation index ARI, and representative areas are selected for empirical analysis, the advantages and disadvantages of the two methods are summarized, and some measures and

suggestions for improving the status quo of physiological indicators are given. And different policies and measures are formulated according to different clusters. The amount of data is relatively large, and the continuous update of discrete data dimensions with time gradually presents functional characteristics, so it is suitable to be regarded as a curve changing with time in a long time, that is, an integral element in the function space for research and analysis. This not only reveals the dynamic change law of the research object with the development of independent variables but also relaxes the assumptions required for traditional data type modeling, has stronger universality, and shows its innovation in the application field, and these theoretical methods belong to Frontiers in Statistical Theory Research [12–15].

First, a method for identifying discrete physiological index line data is given. Second, the initial index system of discrete physiological index line data evaluation is established, and the initial index system is revised by interviewing; a discrete physiological index line data evaluation method based on rough set theory is proposed, and the discrete physiological index line data are established to evaluate the model to evaluate discrete physiological index line data. Third, according to the data evaluation results of discrete physiological index lines, combined with the ascertained data of discrete physiological index lines, the actual data of discrete physiological index lines are calculated, and the matching degree between the actual data and the manufacturing task capability requirements is analyzed. There may be excess or insufficient data. This paper builds a discrete physiological index line data adjustment model in a random demand environment, adjusts the actual data of the discrete physiological index line, formulates a reasonable data adjustment strategy, and realizes discrete physiological index lines. In addition, functional model clustering is used to conduct comprehensive clustering research and analysis on a national and regional scale. Compared with the K-means method, the correlation between multiple physiological indicators is more fully considered, and the variation of different physiological indicators is also considered. The law of time change is more in-depth in the application of statistical methods. Finally, through simulation and empirical analysis, the accuracy of the model-based clustering method is verified, and this paper summarizes and compares the multivariate functional model clustering and the existing multivariate functional K-means clustering methods. We provide relevant decision-making basis and theoretical support for future environmental governance, so as to monitor the data quality status and guide useful information in the future.

2. Related Work

Analyzing events or activities under the constraints of discrete linear space units has advantages in terms of accuracy, reliability, and application accuracy. Most of the existing analysis methods for discrete point elements in space use a single point as the main body of analysis, and there are few studies on the spatial analysis methods of point events under linear constraints, and the algorithms are more complex and difficult to implement. Taking the kernel density analysis method as an example, this paper uses multisource media data (mainly sports monitoring data) and road network data of different scales as data sources to study the kernel density analysis method under linear constraints. It aims to provide new effective and practical quantitative methods for planning management, personalized route recommendation, intelligent route recognition [16].

The visual representation of complex networks has brought great progress to bibliometric analysis. In the past, most of the literature studies used quantitative analysis methods. Such data analysis is generally carried out in the form of tabular statistics, which cannot quickly give people an intuitive visual perception and analysis. The keyword co-occurrence method based on a matrix network of the complex network makes the research of bibliometric analysis advance by leaps and bounds. The keywords of statistical scientific research literature or the correlation relationship of other cited literature, or the cooperative relationship between the authors of statistical papers, the co-occurrence of complex network form of network visualization can visually represent the intersection of disciplines, which greatly improves the efficiency of bibliometric analysis. Yu et al. [17] proposed a new node layout coordinate axis Hive plots for the visualization of the serving network caused by the surge in data volume. The node relationship is arranged in pairs on the coordinate axis to solve the spatial limitation of the absolute position of the node, and at the same time, the robustness of the network layout is enhanced, and even if the edges between some nodes are removed, the position of the relationship node N1 can remain unchanged. Bailey[18] proposed that the absolute position of the node is not important, but the relative position is important. It is the optimal solution to change the relative layout position of the node through the layout algorithm in the limited drawing space. In the ERP environment, through the comparative analysis of the application characteristics of rough capacity requirement planning (RCP) and capacity requirement planning (CRP), combined with the traditional capacity balancing method, a coarse capacity balancing method of load forward movement is proposed. This method can be used in the ERP system to make up for the defect that the traditional ERP system cannot carry out data balance. Sun et al. [19] combined with the actual project, focused on the data optimization of the data distribution of the shared equipment, and the reasonable planning of the workshop production, so that the physiological index system can meet the management goals of various physiological indicators, and studied and analyzed the multilevel physiological indicators of the enterprise. The planning model, combined with the physiological index ability balance and plan control method with the goal of delivering on time, developed the physiological index ability balance and plan control system.

The "average number" of connections of network nodes is gone, replaced by those hubs with important connection relationships, which is the verification of the law that the degree of nodes in complex networks has a power exponential function. Chen et al. [20] used the model construction method of the complex network to extract

structured, semistructured, and unstructured data in the database uniformly and provided a new method of data extraction and data management. Ahmad et al. [21] analyzed the difference between static network and time-sensitive network, explored the change of network structure during the period from information acquisition to access, and discussed that time-sensitive network is more controllable and requires less order of magnitude through the example of destroying the network path. The researchers are engaged in the performance analysis and control-related research of complex network dynamic systems, exploring the influence of the scale-free characteristics of the network on the network synchronization ability and the related research on the robustness and vulnerability of network synchronization. In the past, most of the literature studies used quantitative analysis methods. Such data analysis is generally carried out in the form of tabular statistics, which cannot quickly give people an intuitive visual perception and analysis. The keyword co-occurrence method based on a matrix network of the complex network makes the research of bibliometric analysis advance by leaps and bounds. The keywords of statistical scientific research literature or the correlation relationship of other cited literature, or the cooperative relationship between the authors of statistical papers, the co-occurrence of complex network form of network visualization can visually represent the intersection of disciplines, which greatly improves the efficiency of bibliometric analysis [22–25]. Scholars in the field of complex networks mainly focus on the dynamics of network propagation, the convergence and stability of neural networks, and other related research and make corresponding econometric models and algorithms.

3. Discrete Data-Driven Indicator Analysis

3.1. Discrete Data Level Statistics. Functional discrete data refer to data that is represented by a function and changes with a continuous set (time or space, etc.). From the perspective of the stochastic process, functional data ink (f) can be understood as a sample trajectory or a realization of stochastic process x (t). Its biggest feature is that the data have functional characteristics, and the external manifestations are various. The core idea is to regard the discrete observation data that are continuously updated with the time-discrete data dimension in a given interval as a function that changes with time. It can be considered that these data sample points come from an infinite dimensional space, which is reflected in the form of a smooth curve on the graph, revealing the dynamic change law of the research object with the development of independent variables.

$$\sum x_i \times w - \sum \exp(x) \times b - 1 = 0. \tag{1}$$

The requirements for data collection frequency are low; for example, the number of sample observation points can be the same or different, and the values at the same time point do not necessarily take values at equal intervals, so the data collection and processing methods are more flexible. From the peculiar properties of functional data, such as the ability to obtain higher-order function derivatives or differentials, and from the continuous point of view, a multiangle static and dynamic analysis of the operating law of the system can be performed to reveal the speed of the research object with the development and change of the independent variable. Deeper dynamic change laws such as acceleration and phase plane diagrams enhance the depth of potential data information mining: it can not only analyze traditional structured data but also process unstructured data such as audio and handwriting. The application gives a continuously changing analysis result over a period of time, has a more intuitive geometric interpretation, and reflects the practicability of functional data.

$$\begin{cases} \exp(x_i \cdot w) - \exp(x) * b + 1 & s_i = 1 \\ \exp(x_i \cdot w) + \exp(x) * b + 1 & s_i = -1 \end{cases}. \tag{2}$$

The feature selection algorithm mainly includes the Laplacian score, which is a second-order differential operator in the N-dimensional Euclidean space. The importance of the feature is measured according to the local retention ability of the special diagnosis. If the distance between samples is closer, the corresponding weight will be larger, and if the adjacent samples are far away, the weight will be smaller. For each feature, the Laplacian score represents its ability to maintain local structure, and the lower the score, the stronger the ability of the feature to maintain local structure. The features selected by the constraint score have better constraint retention ability, and in the dataset obtained in the real world, the pairwise relationship between samples is easier to calculate than the classification attributes of the samples. The Pearson correlation coefficient is a statistical calculation method used to measure the degree of linear correlation between two variables, and its value range is [-1, 1], which is usually used to measure the linear correlation between variables.

3.2. Linear Nesting of Indicators. The linear combination of uncorrelated indicators is the principal component, and the value is called the principal component score of the i-th observation, which is solved by the eigenvector $a - j$. In order to reduce dimensionality, a few principal components are generally selected, but most of the information of the original data can be retained, and information overlap can also be prevented. During the research, the variables themselves and the fluctuations between variables are represented by the variance of the principal components; that is, the variance can be used to define the principal components. The eigenvector 1 is normalized below, in order to eliminate any uncertainty increased by multiplying ai by a certain constant, and the principal component when the variance takes the maximum value is the first principal component.

$$\sum_m \sum_n w(m,n) \times x(i+m, j+n) = \begin{cases} x(i+m, j+n), m < n \\ w(m,n), m > n \end{cases}. \tag{3}$$

Before performing the principal component analysis on the functional data, the original discrete data obtained before

is preprocessed, fitted into functional data, and the discrete finite sample observation data points are converted into a functional curve form. Commonly used fitting methods are smoothing and interpolation. If the observed raw data are exact values with no observed error, interpolation is used; if the raw data have observed error, smoothing is used. Usually, the observed data have errors, so smoothing is used to fit the discrete raw data. Its use solves the storage problem of function information very well, and it also has the required flexibility and computing power when the amount of data is large or even hundreds of thousands of data points.

$$\sum f(i-1, j-1) - \sum g(i-1, j-1) = 0. \tag{4}$$

It transforms the original N-dimensional dataset into a new dataset called principal component by an orthogonal transformation of the discrete data dimension set of the dataset. The first principal component has the largest variance value for each subsequent component. Under the condition of orthogonality with the aforementioned principal components, it has the largest variance; the discrete data method uses the similarity between samples to construct a low-dimensional space and takes the similarity as the distance constraint and divides all sample points with a smaller distance.

Distortion is projected into a low-dimensional space, which is a perceptual map that observes the similarity of samples. It does not emphasize the attribute values of samples but only shows the relationship between samples. The optimal value can be the optimal value of each index value of the discrete physiological index line (if a certain index is the maximum value, the maximum value of the index in each discrete physiological index line is taken; on the contrary, take the the minimum value of the indicator in each discrete physiological indicator line), or it may be the optimal value recognized by the industry or recognized by the evaluator, or the weighted average of the above two in it.

The data packet format in Table 1 is defined as data packet header, convergence point ID, terminal node ID, data packet serial number, time stamp, total data packet length, data content, CRC checksum, and data packet tail. The request-response dialogue method is adopted between the server and the device terminal. The communication content between the sink node and the server consists of uplink data and downlink data; the uplink data are the data collected by the ZigBee terminal node from the physiological index workshop, including the operation of CNC machine tools. Information such as status, timestamp, number of materials discharged, number of materials returned, and downlink data is the confirmation of data reception sent by the application program on the server-side, and time calibration is loaded for downlink data.

3.3. Discrete Data Clustering. From a research perspective, the previous research and analysis of data quality based on traditional data types such as cross-sectional data, time series data, and relatively complex panel data have achieved some results, but they also have certain limitations. For example, if we rely too much on the independence of variables, and based on limited discrete sample data, the extension prediction of the model will be greatly limited and prone to the loss of important information. With the advent of the era of big data, massive data emerge, and functional data emerge as the times require, which also provides a data foundation for in-depth analysis of models. Its core idea is to regard the discrete observation data that are continuously updated with the time-discrete data dimension in a given interval as a function that changes with time. It reveals the dynamic change law of research objects with the development of white variables, and relaxes the assumptions required for modeling, and has stronger universality.

Table 1: Description of discrete physiological indicators.

Number	Indicator value	Mean value	Optimal value
1	0.13	3.43	36.66
2	8.65	3.54	31.17
3	4.29	1.57	34.04
4	4.90	5.14	30.37
5	9.67	8.77	38.25
6	6.26	3.46	33.85
7	6.57	7.19	32.99

$$\begin{vmatrix} w(m,n) & x(i+m) \\ x(i+m) & w(m,n) \end{vmatrix} = \begin{vmatrix} w(m,n) & 1 \\ 1 & w(m,n) \end{vmatrix} * \begin{vmatrix} 1 & x(i+m) \\ x(i+m) & 1 \end{vmatrix}. \tag{5}$$

Compare and analyze the actual data of the physiological index line with the capability of manufacturing task requirements and accurately grasp whether the actual data of the physiological index line can meet the needs of the manufacturing task, we can provide strong support for subsequent data adjustment decisions. The matching analysis of discrete physiological index lines between actual data and manufacturing task capability requirements is as follows. Each column of the decision matrix represents each attribute, namely, condition attribute C and decision attribute D, and each row represents each discrete physiological index line, namely, the universe of discourse U. First, we observe the decision matrix D, if there are two or more identical columns, keep any one of the columns in Figure 1, and delete the rest. That is, for the universe of discourse U, if the attributes a and c have the same resolving power, only one of them a or b or c is retained.

It is used to display data and receive data input by the user, providing an interactive operation interface for the user. The business logic layer is the part of the system architecture that reflects the core value. Its focus is mainly on the system design related to business requirements, such as the formulation of business rules and the realization of business processes, that is, the realization process of the functions of each module of the discrete data evaluation system. The data access layer is sometimes called the persistence layer, and its function is mainly responsible for database access and is the source of data import. The data evaluation module is the core part of the system. Using the established discrete physiological index line data evaluation

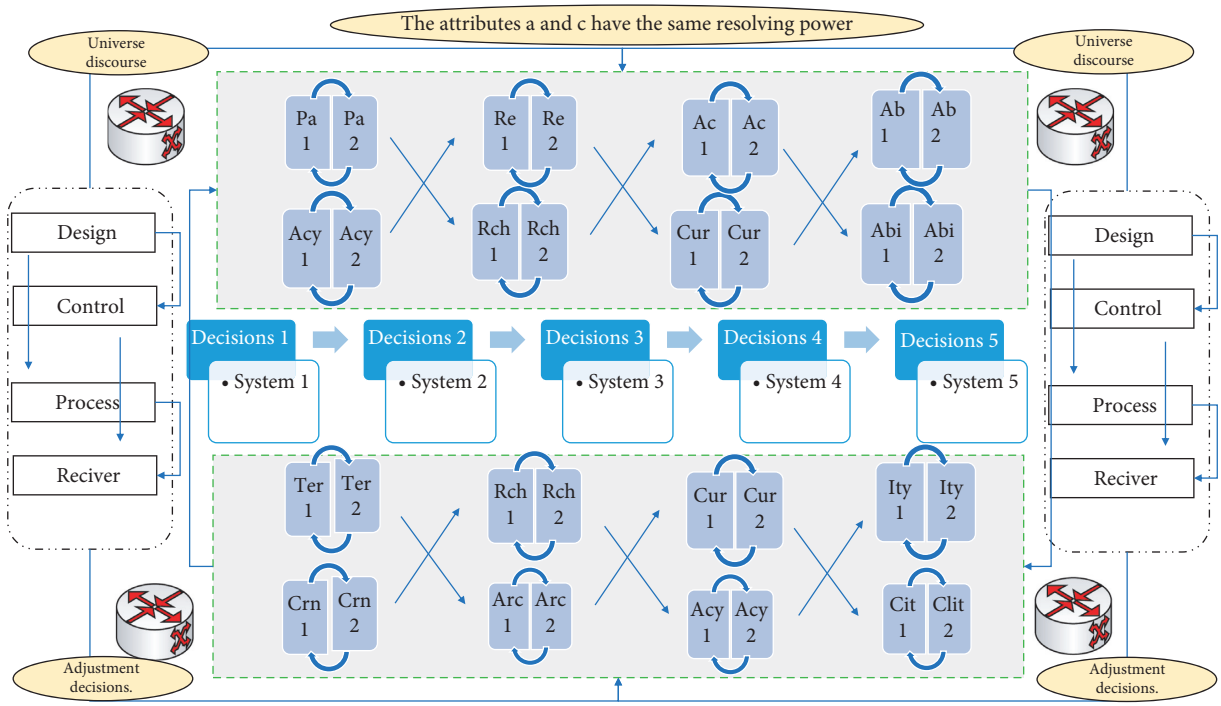

Figure 1: Hierarchical topology of physiological index lines.

index system, a discrete physiological index line data evaluation method based on rough set theory is designed, and the discrete physiological index line is evaluated by single element data. Combined with the factor of data evaluation, finally obtain the single factor data evaluation value and the comprehensive factor production capacity evaluation value.

3.4. Functional Data Analysis. Functional data are continuously updated according to the time-discrete data dimension, and it gradually presents functional characteristics, so it is suitable to be regarded as a curve in a relatively long period of time. In practical applications, these curves can be obtained by nonparametric fitting methods, and then, it is considered as an integral element in the function space for research and analysis, and the processing method is more flexible. Moreover, if tens of thousands of pieces of data are still aggregated and analyzed, it will not only be unfavorable to the extraction of the changing patterns of physiological indicators but also will lead to a waste of data resources.

$$\lim_{x\longrightarrow\infty}\sum_{m}\sum_{n} w(m,n)\times x(i+m,j+n)\Big/\lim_{x\longrightarrow\infty}\sum_{m}\sum_{n} w(m,n)=1. \tag{6}$$

It is a method that expresses the mapping relationship between samples through functions, so as to discover the interdependent quantitative relationship between two or more discrete data dimensions, and is widely used. Analysis can be divided into linear regression analysis and nonlinear regression analysis according to the type of relationship between independent variables and dependent variables. The relationship between one or more independent and dependent variables is modeled using the least-squares function of a linear regression equation. Such a function is a linear combination of one or more model parameters called regression coefficients. The case with only one independent variable is called simple regression, and the case with more than one independent variable is called multiple regression.

$$\sum u(P)-\sum u(i)\times[u(x_1),u(x_2),u(x_3),\ldots,u(x_{n-1}),u(x_n)]=0. \tag{7}$$

The reduction in the number of frequent itemsets will lead to a reduction in the number of association rules that satisfy the conditions. When the support degree is 0.3, the total number of association rules reaches 60; when it is greater than 0.4, the number of association rules obtained by association mining of knowledge points decreases sharply. Through the above analysis, it can be concluded that after discretizing the original data, constructing the sample data, and analyzing the association of the data set, the association between the knowledge points is obtained, and the association rule table is established. After the interest degree is introduced, by setting the threshold of the minimum interest degree, the association rules with higher reliability are screened out, which improves the reliability of the association rule table. Due to space limitations, here only select some association rules with higher confidence in the high scoring rate samples and the low scoring rate samples to analyze the results in Figure 2.

Hierarchical clustering is the hierarchical decomposition of a given data set according to certain characteristics (such as distance or density), which can be divided into two schemes: "bottom-up" and "item-down". Initially, all samples are grouped into separate groups, and in the iterative process, similar or adjacent groups are combined into one group until a certain condition is satisfied. The BIRCH

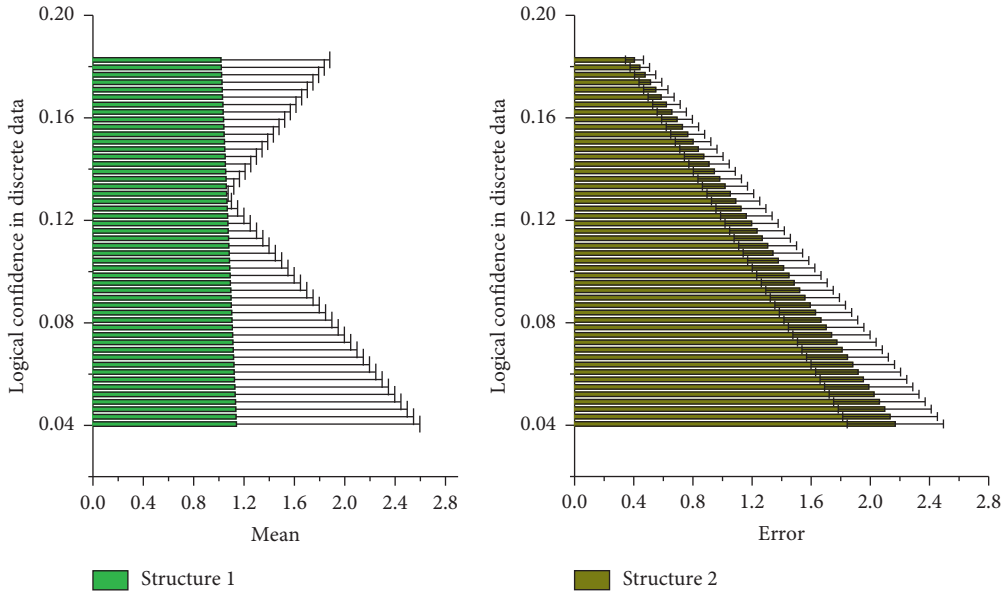

Figure 2: Hierarchical clustering confidence comparison.

clustering algorithm is applied to the recognition of color images. First, the BIRCH algorithm is used to extract the image with the highest similarity to the input image histogram in the database, and then, the corresponding image features are extracted, and then, the Euclidean distance is used to determine the attributes of the image and output the result, using the combination of text similarity algorithm and DBScan clustering algorithm to cluster the text, first perform feature extraction and word segmentation on the text.

4. Physical Exercise Physiological Index Monitoring and Visualization Model Construction Driven by Discrete Data

4.1. Discrete Data Visualization Factors. Aiming at the phenomenon that the latent information in a large amount of data is not mined, the experiment constructs a knowledge point association mining method based on association analysis. First, the algorithm suitable for knowledge point correlation mining is selected by comparison; then, the main steps of the constructed knowledge point correlation mining method are discussed in detail, the original data are discretized, and the knowledge point correlation analysis is carried out from two aspects. It includes constructing sample data for classification mining according to the score rate and hierarchical mining according to the proportion of knowledge point samples in the total data volume. Finally, the correlation between knowledge points in the database course is mined, and an association rule table between knowledge points with high reliability is obtained to measure the strength of the correlation of knowledge points.

$$H(P \mid (x,i) \in R(i)) = [(h(x_1), h(x_2), h(x_3), \ldots, h(x_{n-1}), h(x_n))]. \quad (8)$$

To mine the potential correlation between data through association analysis, in order to obtain the association rules that meet the needs, two thresholds need to be set: the minimum support degree and the maximum support degree. The parallel coordinate method overcomes the problem that the Cartesian coordinate system is difficult to express data above three dimensions and can express ultradiscrete data. In order to represent a discrete data set, parallel coordinates represent each discrete data dimension of the discrete data with a set of mutually parallel coordinate axes, and the discrete data dimension value of the sample is mapped to the relative position of the corresponding discrete data dimension axis, and each sample will be represented differently.

It is represented by a polyline formed by connecting points on the discrete data dimension; that is to say, after a sample point in the discrete space is mapped to a plane, it will be a polyline. In this system, the data transmission between the server-side application program and the aggregation node adopts the TCP protocol, and the server-side application program allocates a port for each aggregation node, so the aggregation node in Table 2 is connected to the industrial Ethernet through the RJ45 network port, and the data field bus or local area network realizes the data interaction between the aggregation node and the server-side application.

Parallel coordinates have a good mathematical foundation, so in the figure, each discrete data dimension value of the sample can be accurately mapped without any data loss, but as the discrete data dimension increases, the number of axes also increases accordingly, the higher the discrete data

Table 2: Server-side distribution of sink nodes.

Case node	Industrial value	Title node	Interaction value	Point node	Weight value
A1	0.36	B1	0.65	C1	0.65
A2	0.70	B2	0.12	C2	0.12
A3	0.13	B3	0.31	C3	0.31
A4	0.48	B4	0.55	C4	0.55
A5	0.77	B5	0.80	C5	0.80

dimension, the less conducive to analysis, and the parallel coordinates cannot read the correlation between samples, or the correlation between discrete data dimensions, or the correlation between samples and discrete data dimensions. The visualization method projects the discrete data dimensions of the discrete data set on the plane circumference at equal intervals in the form of points. The computational complexity of this visualization method is very low, the number of discrete data dimensions that can be expressed is arbitrary, and it is very intuitive and easy to understand, but it cannot express the correlation characteristics of discrete data dimensions, and the discrete data dimension values correspond to sample points. There is still a certain error in the distance comparison between discrete data dimension points.

4.2. Discrete Data-Driven Indicator Evaluation. Discrete data clustering is an important method of statistical research. It can mine traditional structured data information based on the function perspective, which is of great significance for enriching the information mining technology in the era of big data. At present, the research on national functional data clustering is still in its infancy, and there are mainly three categories: using traditional methods to cluster (such as K-means clustering, systematic clustering) after dimensionality reduction; nonparametric methods using special distances or curve differences; model-based clustering methods. The domestic functional data clustering methods for data quality are basically univariate functional data using the K-means clustering method after dimensionality reduction. The application of statistical methods is not deep enough, and there is no nationwide comprehensive clustering. However, there are many indicators that affect data quality, and the problems of changes in time and space are very complicated. Different types of changes in time in different regions and different interactions between different physiological indicators need to be considered.

$$\frac{\cos\left(T_i(i,j)/T_j(i,j)\right)}{\cos\left(T_i(i,j)/(i,j)\right)} \times \frac{\sin\left(T_i(i,j)/T_j(i,j)\right)}{\sin\left(T_i(i,j)/(i,j)\right)} - 1 = 0. \quad (9)$$

The choice of basis functions is important and affects subsequent function estimates. The actual selection of the appropriate basis function is determined by the user according to specific problems in practice. At present, common ones such as B-spline basis functions, Fourier basis functions, or wavelet basis functions are used, that is, using (f) to express the linearity of the basis function combination. The method is a dimensionality reduction algorithm that projects all points into a low-dimensional space at the expense of a small global physiological index error, constrained by the distance matrix. Table 3 mainly includes the definition and calculation of the distance matrix, and the generation of the initial position of the point, the calculation of the global physiological index error, and the iterative displacement strategy of points.

In the process of finding frequent itemsets, it is necessary to generate a large number of I/O operations, because it is necessary to judge whether the large number of candidate sets generated in the process of finding meet the conditions. Through the limitation of the support threshold, the number of eligible frequent itemsets will be much less than the number of candidate sets, which will greatly reduce the time cost in the next generation of association rules. On the other hand, the transaction database needs to be traversed many times when searching for frequent itemsets. In the process of the second step, it is not necessary to traverse the transaction database to generate association rules. Therefore, observing the whole process of association analysis, finding frequent itemsets takes a heavier proportion of the time and has a greater impact on the performance of the entire algorithm.

4.3. Physiological Index Monitoring of Physical Exercise. Physiological index monitoring adopts bottom-up implementation, and k-1 itemsets are used for mining to find out k itemsets. First, it is necessary to find the set of frequent 1-itemsets, and so on to repeat the loop operation until all frequent k-itemsets are found, and the iterative process ends here. Searching each algorithm is also a commonly used association analysis algorithm. No candidate set is generated during the mining process, and only two database scans are required. The frequent 1-itemsets are obtained in the first scan, and the frequent itemsets are arranged in descending order according to the obtained support. The second scan inserts the elements that meet the minimum support in each transaction into the frequent pattern tree, compresses it to generate an FP-Tree, and then mines the compressed FP-Tree to obtain the association rules that meet the conditions. The algorithm consists of two parts: the stored FP-Tree and the corresponding FP algorithm. The process is equivalent to the continuous iterative process of the construction and projection of a frequent pattern tree.

Table 3: Code implementation of global physiological indicators.

Number	Indicator content	Code text
1	Distance matrix $g_i(i, j)$	Var texts = '0123456789'.split('');
2	Many indicators sin θ cos θ	<scripttype = "text/javascript">
3	Basis functions $T(x_j)$	Var fontsize = 16;
4	Affect data quality $f_i(i, j)$	Var columns = nvas.width/fontsize;
5	Strategy of points $\lim_{i,j\longrightarrow\infty} \alpha(i, j)$	Body{margin: 0; padding:
6	Basis functions $f(x) - f(x_i)$	Var drops = [];
7	The global physiological	0; overflow: Hidden; }
8	The calculation of $\exp g_i$	For(var $x = 0$; $x <$ columns; x++){
9	The iterative displacement	Drops[x] = 1;
10	Code implementation	Function draw(){
11	Different regions $d(k, x_1)$	If(drops[i] $*$ fontsize > anvas.height
12	$x_1^3 - x_2^3$ index Error	Ctx.fillstyle = 'rgba(0, 0, 0, 0.05)';
13	Position of the point $w(a) \cdot x(t - a)$	Ctx.fillrect(0, 0, canvas.height);
14	In the process of $d(x_1)$	Setinterval(draw, 33)
15	The transaction database	Ctx.fillstyle = '#0F0';
16	$y_1^3 - y_2^3$ eligible Frequent	Ctx.font = fontsize + 'px arial';
17	Finding frequent itemsets	<html> <head> <title>kuaile
18	Through the limitation of $w(m, n)$	For(var $i = 0$; $i <$ drops.length;

$$\left\langle \begin{matrix} \exp\left(\dfrac{T_i(i,j)}{T_j(i,j)}\right) & 1 \\ -1 & -\exp\left(\dfrac{T_i(i,j)}{T_j(i,j)}\right) \end{matrix} \right\rangle^{T_i(i,j)T_j(i,j)} \rightleftharpoons \left\langle \begin{matrix} \exp\left(\dfrac{T_i(i,j)}{T_j(i,j)}\right) & \cos i \\ -\sin i & -\exp\left(\dfrac{T_i(i,j)}{T_j(i,j)}\right) \end{matrix} \right\rangle. \tag{10}$$

In order to test the influence of sports monitoring on the global physiological index error stress in the discrete data algorithm, in this experiment, we randomly generated 9 correlation coefficient matrices according to its method. P $\in$ {0.514, 0.571, 0.629, 0.686, 0.743, 0.8, 0.886, 0.943, 1). We perform one-dimensional discrete data dimensionality reduction and two-dimensional discrete data dimensionality reduction on the 9 distance matrices, respectively, to obtain the corresponding stress. In order to obtain the minimum value of the global physiological index error stressmjn, we run the discrete data algorithm 50 times repeatedly, and the number of iterations for each force-directed algorithm is set to 1000 times.

From R-CNN to Faster R-CNN, the reason for the continuous acceleration of computing speed is due to the fact that convolutional neural networks can learn features suitable for different subtasks according to the loss function. We first established the Euclidean distance matrix of the sample points according to the Pearson correlation coefficient and defined the sports monitoring. Then, through experiments, we discussed the relationship between the number of sample points, sports monitoring, and the global physiological index error in the discrete data algorithm. In the summary of the experimental results, we found that when using the discrete data algorithm to achieve dimensionality reduction, there is a positive correlation between the number of sample points in Figure 3 and the global physiological index error, and stressmin is an inverse correlation between sports monitoring and global stress error.

In discrete data dimensionality reduction, we use the method of randomly generating the initial position of the point and use the force-directed algorithm to adjust the position of the point. Due to the different initial positions of the points, the stress is different each time; that is, the physiological index value stress is a local optimum. We choose the minimum stress as the global minimum physiological index error stresslni in 50 repeated experiments, and compare them, as shown in it, and we find that the physical monitoring of the distance matrix is comparable to stressmi when the dataset is reduced to one-dimensional and two-dimensional by discrete data algorithms. There is an inverse correlation between them, and it is more obvious when it is reduced to 2 dimensions than reduced to 1 dimension.

5. Application and Analysis of Physical Index Monitoring and Visualization Model in Sports Driven by Discrete Data

5.1. Discrete Data Weight Update. Before mining discrete data weight correlations, data transformation needs to be performed to prepare for the next mining process. Combined with the specification of the data structure in the mining process, it is necessary to transform the data into a format with students as rows, knowledge points as columns, and data in the table as knowledge point-scoring rates. When discretizing the score rate of knowledge points, it is necessary to set a quantitative standard, that is, the line value, to segment the original data. The setting of the line

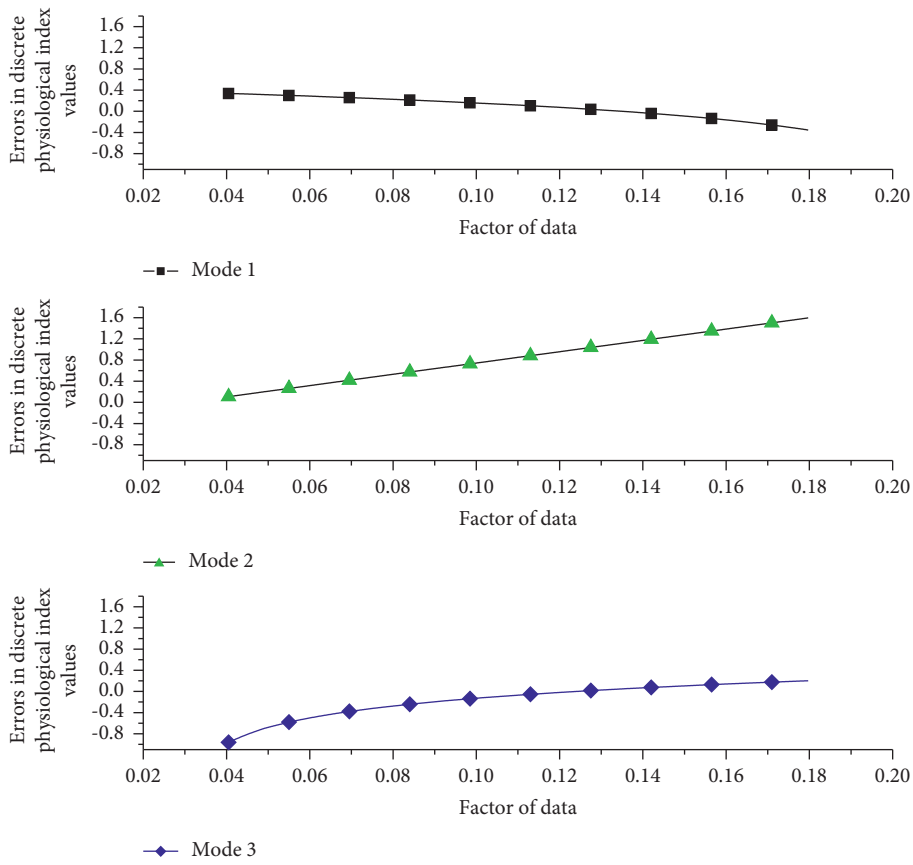

Figure 3: Error distribution of global physiological indicators in sports data.

value should be evaluated in combination with the passing line in the actual teaching and the average scoring rate of knowledge points. By taking the weighted average of these two factors, the appropriate line value can be obtained. On the one hand, in the actual teaching process, the measurement score is 60 points as the pass line. If the score rate of the knowledge point is higher than 0.6, it is considered to have reached the pass line.

$$f(i,j)\Big/\left\{\left[f_i(i,j)\right]^2+\left[f_i(i,j)\right]^2\right\}^3=\begin{cases}\exp\left(\dfrac{T_i(i,j)}{T_j(i,j)}\right), & i<j\\ \sin i+\cos i, & i>j\end{cases}. \tag{11}$$

On the other hand, the selection of the line value should also be based on the actual score rate. The average value of the knowledge point score rate can reflect the overall average level of the students to a certain extent, and students above the average level can consider this knowledge point. If the understanding is better and the mastery is sufficient, the quantification result is recorded as True (referred to as T); if it is lower than the average level, it can be considered that the knowledge point is insufficiently understood and mastered, and the quantitative result is recorded as False (referred to as F). Therefore, in the actual selection of line value, it is necessary to comprehensively consider the passing line and the actual knowledge point-scoring rate.

$$\max(d)-\frac{\lambda\times a}{\sin d}\sqrt{\frac{\max(Q(\text{on}))\times H(ide)\times\beta}{L(ide)}}=0. \tag{12}$$

Different from traditional cross-sectional data, time series data, and relatively complex panel data, functional data analysis based on functional data treats discrete data series as a whole with a unified internal structure, and in infinite dimensional functions, statistical analysis was performed in space. The algorithm is mainly realized by the random algorithm to realize the initial position generation of the point coordinates and the force-oriented algorithm to realize the coordinate optimization in two steps. In the experiment of projecting to a one-dimensional axis, we run the discrete data algorithm 50 times, and each force-directed algorithm iterates 1000 times to obtain experimental data; in the experiment of projecting to a two-dimensional plane, we also run 50 discrete data for the data algorithm, and each force-directed algorithm iterates 1000 times to obtain the experimental data as shown in Figure 4.

This system consists of two methods: PLC batch acquisition for centralized arrangement equipment and ZigBee wireless acquisition for point distribution equipment. Among them, the PLC batch acquisition system completes data transmission through Modbus protocol, and data acquisition nodes and servers are designed for Zigbee wireless mode. The terminal communication protocol realizes the real-time transmission of data. In-depth study of the key

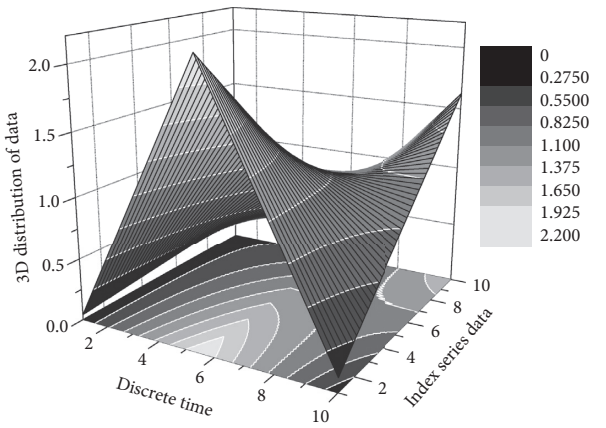

Figure 4: Three-dimensional distribution of discrete time series data.

technology of CNC machine tool state data acquisition, software and hardware are designed according to the system requirements, including the design of PLC acquisition module and Zigbee wireless acquisition module circuit, and the data acquisition module and data processing module are designed in software, and the automatic analysis and data verification of uploaded data, and establish a real-time connection with the database side.

5.2. Simulation of Sports Physiological Indicators Monitoring. Each layer of the sports physiological index monitoring protocol stack can provide services for the upper layer and adopts a layered structure. The framework of the ZigBee protocol stack is developed using the standard seven-layer reference model of system interconnection. One device contains a MAC layer and physical layer; MAC layer solves the access mode of the physical channel; physical layer includes RF transceiver and corresponding low-level control mechanism and provides configuration, message routing, and operation mode in network layer; ZigBee physical layer mainly deals with the following several tasks.

$$\iint \sin\theta \times T_j(i,j)didj = \iint \sin\theta \times A(i,j+1) - A(i,j-1)didj. \tag{13}$$

After selecting the basis system, it is necessary to determine the appropriate number of basis functions to approximate the function. The number of basis functions determines the fit and smoothness of the estimated function. When the number of basis functions is large, the function to be estimated can be well approximated, and the fitting degree is high, but the smoothness is low: when the number of basis functions is too small, the interesting feature information in the function may be missed. So in practical applications, we need to make a trade-off between fit and smoothness. When the properties of the basis functions are similar to the estimated functions, better fit and smoothness can be achieved at the same time with fewer basis functions. Of course, the number of basis functions can also be determined by methods such as model selection or crossvalidation.

$$\begin{cases} \dfrac{f(i,j)}{\left\{\left[f_i(i,j)\right]^2 + \left[f_i(i,j)\right]^2\right\}^3} = 0 \\ \dfrac{g(i,j)}{\left\{\left[g_i(i,j)\right]^2 - \left[g_i(i,j)\right]^2\right\}^3} = 0 \end{cases}. \tag{14}$$

Potential factors that affect operating efficiency include the size of the study area, the number of incident points, the complexity of the road network, the degree of connectivity between different programs, and the proficiency in ArcMap and the ArcPy module in *Python*. Among them, the complexity of the road network is the dominant factor. The more complex the road network, the greater the amount of data input for the spatial adjacency algorithm, the longer the spatial adjacency algorithm takes, and the longer the total time. The projection and kernel density algorithms consume less time and are basically negligible. Therefore, for the same road network, as long as the spatial adjacency relationship of the grid where the road network is located is determined first, and then, the kernel density of the point data in Figure 5 under the constraints of the road network is calculated, the operation efficiency will be greatly accelerated.

Since the original difficulty influencing factors are not obvious, it not only interferes with the teacher's work to mark the training set but also affects the accuracy of the neural network's difficulty prediction. In order to avoid the influence of the above problems, it is necessary to perform data preprocessing first, and give the standard of quantitative assignment, and then train the neural network model. Determining the criteria for quantitative assignment can perform feature normalization and subsequent series of processing on the influencing factors of the test items contained in the indicator data set.

On the other hand, the perspective of researchers in related fields is constantly changing: breaking the old sliding window model, and considering target detection as a category and position regression problem. (1) The spatial discrete point data are projected onto the line according to fixed rules to ensure that the point elements are on the line; (2) the linear cells are gridized to ensure that the spatial adjacency relationship of each grid in the linear cells after gridding is the same as that of the grid; (3) the gridded linear space unit is used as the calculation subject, the number of event points falling into each grid is calculated and used it as the generating element to calculate the kernel density value; (4) the kernel density value is used as a visualization attribute field, the visualization relies on the linear space unit, and the horizontal width of the line is used as an important indicator to identify the hot area. According to the above steps, the calculation and application of the kernel density value under linear constraints are basically realized.

5.3. Example Application and Analysis. The data set used in the experiment is 12 sports index data during the school year. Each set of sports index data set contains 45 test

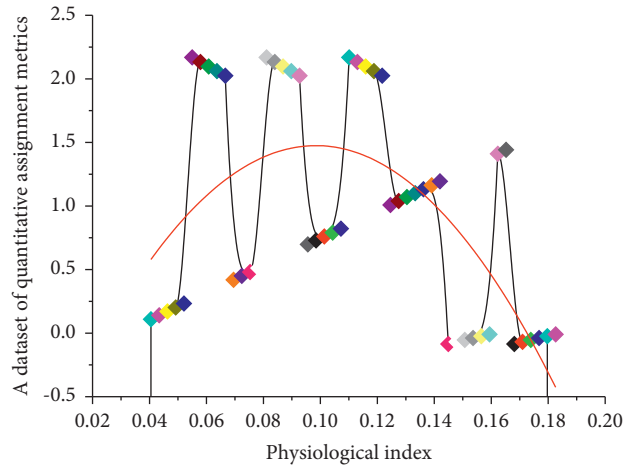

Figure 5: Quantitative assignment index data set distribution.

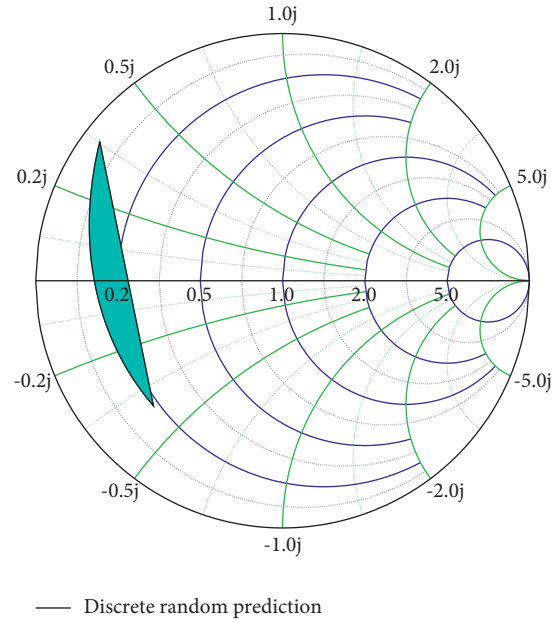

Figure 6: Polar coordinate distribution of discrete random prediction model.

questions, with a total of 540 questions, involving 48 knowledge points in the database course. The expert samples constructed from the dataset T are recorded as datasets. In order to measure the performance of the prediction model, the precision (P) and recall (R) of the experiment can be calculated according to the experimental results. The variable of the precision rate P is defined as the accuracy of the prediction model. In the data whose prediction result is a positive class, how many data are correctly predicted (the prediction result is a positive class), that is, the proportion of correct predictions; the variable of recall rate R is defined as the completeness of the model, how much of the data in the test set in the positive class is correctly predicted (the predicted result is the positive class). The precision rate and recall rate in Figure 6 can reflect the rationality of the prediction model to a certain extent and reflect the prediction ability of the prediction model.

Theoretically, the spatial point data abstracted from the spatial behavior of moving objects has random distribution on the two-dimensional plane. In this paper, idealized experiments are designed to preliminarily verify the feasibility of the gridded linear element kernel density algorithm. Using randomly generated 580 random points and random vectorized linear units, the performance effects of plane kernel density and linear kernel density algorithms under random distribution are compared and analyzed. The Average Nearest Neighbor tool calculates the exact degree of clustering of the data. In spatial statistics, the commonly used Moran index, P-value, and Z score can measure whether a piece of data belongs to discrete randomness or aggregation.

$$\begin{cases} x_1^3 - x_2^3 + y_1^3 - y_2^3 = \left[d\left(x_1\right) + d\left(k, x_1\right)\right]^3 \\ m_1^3 + m_2^3 - n_1^3 + n_2^3 = \left[t\left(x_1\right) - t\left(k, x_1\right)\right]^3 \end{cases}. \tag{15}$$

In order to collect the current state of the equipment, according to the analysis of the equipment fault parameters in the data, the meaning of the status signal word of the corresponding address is formulated. The MBT status signal word consists of 16 bits, and the collected status value is the value of each address corresponding to the PLC signal point, and the value range is 0–65535. In order to manage the status signal word corresponding to each device uniformly, each bit has the same meaning. The intuitive understanding from the name is that you only need to look at the object in the picture, but its original intention is to distinguish it from the two-step solution of target recognition and object positioning.

For example, the address of the device's running status is 16368, and the corresponding MBT status signal word 16-bit meaning is shown in the text. If multiple pieces of data are all clustered, which piece of data has a higher degree of aggregation, the nearest neighbor index can use a pure spatial clustering model to measure the degree of clustering of different data. According to the formula, the average nearest neighbor is very sensitive to the size of the study area, so it is most suitable for comparing different features in a fixed study area. It is generally believed that $\mathrm{NNR} \geq 1.5$ and $\mathrm{NNR} \leq 0.5$ are uniform distribution and aggregated distribution, respectively; $0.5 < \mathrm{NNR} \leq 0.8$ is aggregated random distribution; $08 < \mathrm{NNR} < 1.2$ is random distribution; $1.2 \leq \mathrm{NNR} < 1.5$ is random-discrete distribution.

Before analyzing the data in Figure 7, it is necessary to preprocess the data according to certain specifications in advance to ensure the validity of the results. Human spatial activities are constrained by linear units, and the abstracted spatial point data should be distributed on the line. However, due to the accuracy problem during data collection or the user's wrong choice of positioning, some of the obtained spatial point data are scattered on the line, thus affecting the analysis results.

In the algorithm, the computer can identify the target and the position of the target in the whole picture by only performing a forward calculation on the picture under the

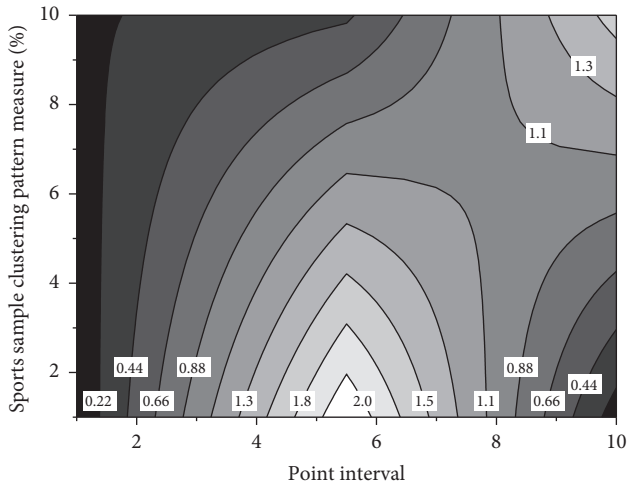

Figure 7: Measurement of spatial clustering pattern of sports samples.

designed network structure and the weights after training. The principle of the algorithm is to obtain the vertical point from the point to the line segment (if the vertical point is outside the line, the endpoint of the shorter end of the line segment shall prevail). Using simple *Python* code blocks to realize the preprocessing of discrete points outside the line to achieve the purpose of denoising and normalizing data.

$$L(y,g) = \left\langle \sum_{i=1}^{n} y_i \times \left[g_i + \ln \sum_{j=1}^{n} \exp(g_i) \right], \right. \\ \left. i > 1 \sum_{i=1}^{n} \sum_{j=1}^{n} \exp(g_i), i < 1. \right. \tag{16}$$

Aiming at the problem of the nonconvergence of the force-steering algorithm in the displacement strategy, we designed the SEFM (systematic error in the same direction first move) algorithm to replace the force-steering algorithm and further reduced the problem by weighting the force-steering algorithm and the SEFM algorithm. Finally, we compared the changes of the f global physiological index error when the initial position generation methods of different points and different displacement strategies were combined to achieve discrete data dimensionality reduction, so as to obtain a better discrete data algorithm.

6. Conclusion

In this paper, the discrete data dimension correlation matrix is first established by the physical exercise data correlation coefficient (each element in the matrix is the correlation coefficient between the physical exercise data discrete data dimensions), and then, the discrete data dimension correlation matrix is transformed into the physical exercise data dimension on the plane through the transformation function. For Euclidean distance matrix between physical exercise data dimension points (each element in the matrix is the Euclidean distance between discrete data dimension points on the plane), we use an algorithm to project the discrete data dimension onto a fixed-length line segment to become the physical exercise data dimension point and then map the line segment to the plane unit physical exercise index data set, so as to obtain the projection of the discrete data dimension of the data to the plane point and finally use an algorithm to adjust the position of the discrete data dimension point on the index data set to minimize the global physical exercise physiological index error. We define the physical exercise monitoring index of the distance matrix, and then randomly generate a number of distance matrices with different physical exercise monitoring values, and perform one-dimensional and two-dimensional dimensionality reduction calculations according to the algorithm to obtain the corresponding global physical exercise physiological index error, and obtain the sports monitoring. There is a negative correlation between the physical exercise index and the corresponding global physiological index error.Therefore, it is proposed to perform power function operation on the numerical value in the distance matrix under the condition of obeying the monotonicity constraint of the numerical value in the distance matrix to improve the physical exercise monitoring index and reduce the global physiological index error in the dimensionality reduction process of the discrete data algorithm.Firstly, the physical exercise data dimension is projected onto a one-dimensional line segment and mapped to the physical exercise index data set to realize the data. The discrete data dimension is projected to the plane point, and then, the algorithm is used to adjust the position of the discrete data dimension point on the physical exercise index data set, so as to minimize the error of the physical exercise physiological index and to realize the display of the discrete data dimension correlation in the method.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Anhui teaching demonstration course exercise physiology (2020sjjxsfk2553).

References

[1] X. Chu, X. Xie, S. Ye et al., "TIVEE: visual exploration and explanation of badminton tactics in immersive visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 118–128, 2021.

[2] C. Lourdais, E. Poirson, and L. Ma, "Emotional Responses to Health Data Visualization," in *Human-Computer Interaction. Human Values and Quality of Life*, pp. 61–74, Springer, Cham, Switcher land, 2020.

[3] X. Schelling, J. Fernández, P. Ward, J. Fernandez, and S. Robertson, "Decision support system Applications for scheduling in professional team sport. The team's

perspective," *Frontiers in Sports and Active Living*, vol. 3, Article ID 678489, 2021.

[4] A. Z. Chang, D. L. Swain, and M. G. Trotter, "Towards sensor-based calving detection in the rangelands: a systematic review of credible behavioral and physiological indicators," *Translational animal science*, vol. 4, no. 3, 2020.

[5] M. Paniccia, T. Taha, M. Keightley et al., "Autonomic function following concussion in youth athletes: an exploration of heart rate variability using 24-hour recording methodology," *Journal of Visualized Experiments: Journal of Visualized Experiments*, vol. 139, Article ID 58203, 2018.

[6] C. Towlson, C. MacMaster, B. Gonçalves, J. Sampaio, J. Toner, and N. Macfarlane, "The effect of bio-banding on physical and psychological indicators of talent identification in academy soccer players," *Science and Medicine in Football*, vol. 5, no. 4, pp. 11–13, 2021.

[7] P. Nosov, A. Ben, S. Zinchenko, I. Popovych, V. Matiechuk, and H. Nosova, "Formal approaches to identify cadet fatigue factors by means of marine navigation simulators," in *Proceedings of the ICTERI Workshops (16th International Conference on ICT in Research, Education, and Industrial applications)*, pp. 823–838, Kharkiv, Ukrine, Octobar, 2020.

[8] R. Dubois, N. Bru, T. Paillard et al., "Rugby game performances and weekly workload: using of data mining process to enter in the complexity," *PloS one*, vol. 15, no. 1, Article ID e0228107, 2020.

[9] N. Costadopoulos, M. Z. Islam, and D. Tien, "A knowledge discovery and visualisation method for unearthing emotional states from physiological data," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 3, pp. 843–858, 2021.

[10] R. Wang and J. Jia, "Design of intelligent martial arts sports system based on biosensor network technology," *Measurement*, vol. 165, Article ID 108045, 2020.

[11] E. A. Timme, A. A. Dayal, and Y. A. Kukushkin, "Training Plans Optimization Using Approximation and Visualization of Pareto Frontier," in *Proceedings of the 12th International Symposium on Computer Science in Sport (IACSS 2019)*, pp. 69–76, Springer, Moscow, Russia, January, 2019.

[12] Q. Xu, L. Chen, H. Chen, and B. Julien Dewancker, "Exercise thermal sensation: physiological response to dynamic-static steps at moderate exercise," *International Journal of Environmental Research and Public Health*, vol. 18, no. 8, p. 4239, 2021.

[13] J. E. Blanchfield, M. T. Hargroves, P. J. Keith et al., "Developing Predictive Athletic Performance Models for Informative Training Regimens," in *Proceedings of the 2019 systems and information engineering design symposium (SIEDS)*, pp. 5-6, IEEE, Charlottesville, VA, USA, April, 2019.

[14] R. Furukado and G. Hagiwara, "Examining the effects of digital gameplay of the racing genre on mood and heart rate," *Journal of Digital Life*, vol. 1, p. 5, 2021.

[15] P. Browne, A. J. Sweeting, C. T. Woods, and S. Robertson, "Methodological considerations for furthering the understanding of constraints in applied sports," *Sports Medicine-Open*, vol. 7, no. 1, pp. 11-12, 2021.

[16] J. Sheidin, J. Lanir, and T. Kuflik, "A comparative evaluation of techniques for time series visualizations of emotions," in *Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter*, pp. 4–9, Adova, Italy, September, 2019.

[17] Z. Yu, Y. Liu, and C. Zhu, "Comparative anesthesia effect of brachial plexus block based on smart electronic medical ultrasound-guided positioning and traditional anatomical positioning," *Journal of healthcare engineering*, vol. 202113 pages, Article ID 6676610, 2021.

[18] A. Bailey, A. Hughes, K. Bullock, and G. Hill, "A climber's mentality: EEG analysis of climbers in action," *Journal of Outdoor Recreation, Education, and Leadership*, vol. 11, no. 1, pp. 53–69, 2019.

[19] W. Sun, Y. Si, M. Guo, and S. Li, "Driver distraction recognition using wearable IMU sensor data," *Sustainability*, vol. 13, no. 3, p. 1342, 2021.

[20] S. Chen, K. Xu, X. Zheng et al., "Linear and nonlinear analyses of normal and fatigue heart rate variability signals for miners in high-altitude and cold areas," *Computer Methods and Programs in Biomedicine*, vol. 196, Article ID 105667, 2020.

[21] Z. Ahmad, M. N. Jamaludin, and A. H. Omar, "Development of wearable electromyogram for the physical fatigue detection during aerobic activity," *Movement, Heal. Exerc.*vol. 7, no. 1, pp. 15–25, 2018.

[22] R. M. Musa, A. P. P. A. Majeed, Z. Taha, S. W. Chang, A. F. A. Nasir, and M. R. Abdullah, "A machine learning approach of predicting high potential archers by means of physical fitness indicators," *PLoS One*, vol. 14, pp. 3–11, 2019.

[23] G. C. Millar, O. Mitas, W. Boode et al., "Space-time analytics of human physiology for urban planning," *Computers, Environment and Urban Systems*, vol. 85, Article ID 101554, 2021.

[24] K. Van Hoovels, X. Xuan, M. Cuartero, M. Gijssel, M. Swarén, and G. A. Crespo, "Can wearable sweat lactate sensors contribute to sports physiology?" *ACS Sensors*, vol. 6, no. 10, pp. 3496–3508, 2021.

[25] P. P. Y. Wu, N. Sterkenburg, K. Everett, D. W. Chapman, N. White, and K. Mengersen, "Predicting fatigue using countermovement jump force-time signatures: PCA can distinguish neuromuscular versus metabolic fatigue," *PLoS One*, vol. 14, no. 7, Article ID e0219295, 2019.