Research Article

Feature Extraction of National Physical Fitness Data Based on Data Mining

Ying Huang, Yang Li, and Wengang Ren

Department of Sports Arts, Hebei Sport University, Shijiazhuang 05000, Hebei, China

Correspondence should be addressed to Wengang Ren; rwg@hepec.edu.cn

Received 19 February 2022; Revised 22 March 2022; Accepted 25 March 2022; Published 2 May 2022

Academic Editor: Naeem Jan

In order to better understand and optimize the national physical fitness, this paper puts forward the national physical fitness data change feature extraction method based on data mining, uses the decision tree and association rule data mining algorithm to collect the national physical fitness data in recent years, constructs the database to realize the effective data management, and uses the data mining algorithm to construct the physical fitness change feature evaluation index. Finally, through experiments, it is confirmed that the national physique data change feature extraction method based on data mining has high effectiveness in the process of practical application. It can better understand the national physique change trend and put forward targeted suggestions for national physique health optimization.

1. Introduction

National physique is crucial to the future of the country, social progress, and personal development. If the promulgation and implementation of the outline of the national fitness plan are regarded as the new era of mass sports in China, then the "several opinions on accelerating the development of sports industry and promoting sports consumption" (hereinafter referred to as "opinions") issued by the State Council are more like a milestone event "opinions." It not only marks that mass sports have entered a new development period, but also raises national fitness as a fundamental goal of sports development [1]. As the country pays more and more attention to the national fitness work, various national fitness related plans and policies are constantly introduced and improved. At the same time, the mode of rapid development of mass sports in China has been slowly started [2]. Data preprocessing technology is an important means to improve data quality. How to combine data preprocessing technology with national physique monitoring and the data preprocessing technology (method) selected according to the characteristics of national physique monitoring and the objectives of each stage not only ensures the authenticity and reliability of national physique monitoring database data [3], but also objectively reflects the quality level of national physique monitoring database. In addition, comparing the national physique monitoring data of provinces (regions) and cities before and after data preprocessing can also be used as the standard to judge the quality of national physique monitoring of provinces (regions) and cities [4]. The application of this technology improves and ensures the data quality of national physique monitoring and puts forward new requirements and directions for the development of national physique monitoring in the next stage.

2. Feature Extraction of National Physical Fitness Data

2.1. National Physique Data Collection and Management. The national physique monitoring database, being China's sole countrywide physique database, allows researchers to...
analyze the country’s physique state and offers data to aid in the creation of government policies. Its data quality must be shown and approved by specialists at all levels, which means that, prior to the creation of the national physique monitoring database, a number of data preparation tasks must be completed. The data cleaning of the national physique monitoring database generally goes through several steps, according to an analysis of the construction process of the database over time: the determination of warehousing samples, the entry of relevant data, data integration, data cleaning, and data unification [5]. The determination of warehousing samples, the entry of relevant data, data integration, and data cleaning are among them. They clean and convert the coding structure and variable attributes of the part in accordance with the theme of building a data warehouse, in order to meet the overall design of the database and effectively and quickly meet visitor expectations [6]. As indicated in Figure 1, the major goal of data integration and data unification is to combine datasets from several sources into a single database.

Data integration is the process of logically or physically organically gathering database tables with different sources, formats, and characteristics and establishing data warehouse. There are many database tables after data cleaning [7]. In order to reduce the number of database tables and improve the convenience of data analysis and the consistency of analysis results, multiple database tables must be integrated.

After integrating the data of the questionnaire database and the measurement table into the same database, the existing data of the questionnaire database and the measurement table will be merged into the same database, resulting in the fact that the data of the questionnaire database and the measurement table will pass the test [8]. The national physique data collection and management process is shown in Figure 2.

Data entry includes completing questionnaires, digital entries, and other input activities within the allotted timeframes and in accordance with the established standards and amounts. Currently, national physical fitness monitoring data is mainly manually inputted by professional input employees. As illustrated in Figure 3, double-blind data entry is used from the commencement of data entry through the creation of the database to guarantee consistency between the data and the questionnaire data [9]. After entering the data, the national physical fitness monitoring center randomly picks 1% of the data to compare and recheck with the paper questionnaire [10]. Following the recheck, the firm will send a copy of the original input data sheet as the database’s original data.

After obtaining and entering the original data sheet of the company, we should start our data screening work. Data screening is an important means to ensure data integrity, accuracy, credibility, and interpretability, and it is also an important evaluation means to ensure data quality. Firstly, the data cleaning work is carried out on the original input data, mainly to judge the noise data, identify the outliers, and correct the inconsistency of some data. The noise data and outliers are identified through manual screening after being identified by professional software [11]. The original data table is imported into the “national physique monitoring data screening software” for monitoring index deviation detection and logical judgment. The data passing the deviation inspection and conforming to the logical judgment are directly included into the final logical screening database for data synchronization. In the application administration of national physical fitness monitoring, the data synchronization software is employed to resolve data consistency between the province data center and each ground station database [12]. The data synchronization application is installed on each ground station server and uses the Internet to link the provincial data center and the ground station database. It can monitor data changes on both sides on a regular basis and synchronize modified data to the other party’s database in a timely manner to maintain data integrity and consistency. The data synchronization service is built on the Web and does tasks on a regular basis using Windows services [13]. To determine if the data is synchronized or not, add three fields of modification time, synchronization time, and data source to the table to be synchronized and implement data management in the whole central database and each ground station database. Each record of each table is set with a unique ID to ensure the normal synchronization of data. After the synchronization service is started, first check the network condition. When the network is unobstructed, find the data whose modification time is greater than the synchronization time and synchronize the user data, physical examination data, and prescription data to the provincial data center in turn [14]. After the synchronization is successful, modify the synchronization time to the current time. Then, find the data whose modification time is greater than the synchronization time in the provincial data center. Once found, synchronize these data to the ground station database. After successful synchronization, modify the synchronization time to the current time [15]. When the network is blocked or synchronization fails, continue to check the network condition and execute the synchronization process after an interval of two minutes. The flow chart of synchronization program is shown in Figure 4.

In order to ensure the normal data synchronization, each table and record of the central database and each ground station database must be unique. The implementation process needs to identify the same user and the fraudulent use of user name, that is, to solve the conflict problem and the problem of user account combination and audit [16]. The data synchronization mechanism ensures the final consistency of data, makes the data between the provincial data center and each ground station database back up each other, and realizes the overall remote backup function.

2.2. Evaluation Algorithm of National Physique Data Characteristics. Physique data analysis is the core function of the model developed in this paper. This section is divided into two sections: one contains general data statistics from a large number of ordinary user physique data sets collected by the physique identification and health management program, and the other contains association rule mining of
Figure 1: Construction process of national physical fitness monitoring database.

Figure 2: National physique data collection and management process.

Figure 3: Double blind input process of national physique data collection and management data.
the data set to identify different ages, genders, and occupations. There is the law of health conditioning requirements of individuals with various characteristics in different locations [17]. There are the distribution of physical kinds and the rule of health conditioning demands of people with various characteristics in different regions. Python and the flask framework are used to construct the display layer and business logic layer of the whole data analysis function module. The flow chart of data analysis is shown in Figure 5.

The national physical fitness test database contains a large amount of data, which needs data cleaning, removing noise in the data and correcting data inconsistency. The main preprocessing contents include data cleaning and digitization [18]. After years of research, China has implemented and revised the national physical fitness measurement standards for many times. According to these standards, the data are processed to meet data mining needs. First, clean up the data. Eliminate missing values, identify and eliminate outliers, and check and correct errors in data \( x^{(0)}(n) \). This paper uses the data mining plug-in of data mining add-ins for Office 2017 to browse data and remove outlier data in Excel 2017 and preliminarily process the irregular original data through accumulation generation, subtraction generation, or weighted accumulation generation, so as to turn it into a more regular generation sequence. This paper uses cumulative generation to preprocess the original data. Cumulative generation: set the data of the series as the original data sequence at one time:

\[
\begin{align*}
x^{(0)}(t) &= \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n)\}. 
\end{align*}
\]

In order to analyze and compare the related information as a whole and weaken the dispersion of the related information \( \varepsilon_i \), it is necessary to process the related information centrally, for example, average; that is, the average value \( k \) of the correlation coefficient of the comparison series and the reference series at each time point \( n \) is calculated to reflect the corresponding correlation degree of each comparison series and the reference series.

\[
R_{ij} = \frac{1}{n} \sum_{k=1}^{n} \varepsilon_i(k) - x^{(0)}(t).
\]

Based on the above algorithm, the variables of physical fitness are mainly set by three-level indicators. The weight coefficients of physical fitness indicators of two and gender groups in the same year are shown in Table 1.

The weight coefficients of physical fitness secondary indicators of people of different ages and genders are shown in Table 2.

Now, the variable equation is based on the body shape “sub,” and the three-level index is set as the variable. Because the dimensions of the index are inconsistent, and the change direction is inconsistent, the original test value of the index should be converted first; that is, the score should be made according to the national physical fitness measurement standard, and the variable will be changed from the original variable to the score variable, which is expressed by “s + three-level index name”; for example, “s body fat rate” represents the score of body fat rate [19]. All three-level indicators represent the setting of conditional parameters of “s + three-level indicator name” of different ages, which can be represented by “if then” in STELLA software. Taking the physical condition of young people aged 20–24 as the evaluation standard, the scores of physical indicators are divided:

\[
\begin{align*}
\text{SBMI} &= \begin{cases} 
3, & \text{BMI} < 18.5 \\
4, & \text{BMI} \in (18.5, 23.9) \\
2, & \text{BMI} \in (24, 26.9) \\
1, & \text{BMI} \geq 27
\end{cases}
\end{align*}
\]

When SBMI is less than 18.5, 3 points will be obtained; when SBMI is between 18.5 and 23.9, 4 points will be obtained; when SBMI is between 24 and 26, 2 points will be obtained [20]; when SBMI is greater than 27, 1 point will be obtained. The flexibility is expressed by flexibility, and the expression function of SFlexibility is as follows:
If the flexibility is less than 41l, 1 point will be obtained; when the flexibility is between 411 and 693, 2 points will be obtained; when the flexibility is between 41.5 and 69.3, 3 points will be obtained; when the flexibility is between 13.14 and 15.96, 4 points will be obtained [21]; when flexibility is greater than or equal to 1596, 5 points will be obtained. The torso force is expressed in sit-up table, and the function is as follows:

$$ SP_{\text{flex}} = \begin{cases} 1, & \text{if } Flexibility < -5.6 \\ 2, & \text{if } Flexibility \in (-5.6, 2.1] \\ 3, & \text{if } Flexibility \in (2.2, 7.9] \\ 4, & \text{if } Flexibility \in (8.0, 14.8] \\ 5, & \text{if } Flexibility \geq 14.9 \end{cases} $$

$$ Sit - up = \begin{cases} 1, & \text{if } Sit - up < 17 \\ 2, & \text{if } Sit - up \in (18, 24] \\ 3, & \text{if } Sit - up \in (25, 38] \\ 4, & \text{if } Sit - up \in (39, 45] \\ 5, & \text{if } Sit - up \geq 45 \end{cases} $$

When sit-up is less than 17, 1 point will be obtained; when sit-up is between 18 and 24, 2 points will be obtained; when sit-up is between 25 and 38, 3 points will be obtained; when sit-up is between 39 and 45, 4 points will be obtained; when sit-up is greater than or equal to 45, 5 points will be obtained. The body fat rate is expressed as fat%, and the expression function of $SPAT\%$ is

$$ SPAT\% = \begin{cases} \text{Sit - up} / 100, & \text{if } Sit - up \leq 17 \\ \text{Sit - up} / 100 \times (85 - \text{Sit - up}), & \text{if } Sit - up > 17 \end{cases} $$

Table 1: Weight of three-level indicators of physical fitness.

| Model Parameter          | Weight coefficient (W) |
|--------------------------|------------------------|
|                          | Men aged 40–60         | Women aged 40–60 | Men aged 20–40 | Women aged 20–40 |
| Abdominal curl           | 0.142                  | 0.1391           | 0.1521         | 0.196            |
| Grip                     | 0.1605                 | 0.1621           | 0.1521         | 0.178            |
| Lower limb strength      | 0.1895                 | 0.2258           | 0.165          | 0.192            |
| Sitting body flexion     | 0.1725                 | 0.1925           | 0.2235         | 0.0183           |
| Reaction time            | 0.1601                 | 0.1421           | 0.1965         | 0.08             |
| Fitting degree (%)       | 95.7                   | 96.5             | 92.1           | 93.5             |
| Number of hidden layers  | 15                     | 13               | 15             | 15               |

Table 2: Weight coefficient of secondary indicators.

| Model Parameter          | Weight coefficient (W) |
|--------------------------|------------------------|
|                          | Men aged 40–60         | Women aged 40–60 | Men aged 20–40 | Women aged 20–40 |
| Abnormal body shape      | 0.255                  | 0.356            | 0.251          | 0.236            |
| Physiological function   | 0.191                  | 0.189            | 0.265          | 0.218            |
| Physical quality         | 0.568                  | 0.468            | 0.526          | 0.568            |
| Fitting degree (%)       | 95.32                  | 95.36            | 95.32          | 95.32            |
| Number of hidden layers  | 14                     | 14               | 19             | 18               |
When fat% is less than 12, 3 points will be obtained; when fat% is between 12.1 and 18, 4 points will be obtained; when fat% is between 18.1 and 24, 2 points will be obtained [22, 23]; when fat% is greater than or equal to 24, 1 point will be obtained. The systolic blood pressure is expressed by SBP, and the expression function of $SBP_H$ is

$$SBP_H = \begin{cases} 
\text{Fat}\% \times 12 & 3, \\
\text{Fat}\% \in (12.1, 18) & 4, \\
\text{Fat}\% \in (18.1, 24) & 2, \\
\text{Fat}\% \geq 24 & 1.
\end{cases}$$

(6)

When $SBP_H$ is less than 40; score 2 points when $SBP_H$ is between 41 and 60; when $SBP_H$ is between 60 and 90, get 3 points; when $SBP_H$ is greater than or equal to 90, get 1 point. The bone mineral density is expressed by SBMD, and the expression function of SBMD is

$$SBM\ D = \begin{cases} 
BM\ D \times 0.37 & 1, \\
BM\ D \in (0.38, 0.52) & 2, \\
BM\ D \in (0.53, 0.67) & 3, \\
BM\ D \geq 0.67 & 4.
\end{cases}$$

(8)

When SBMD is less than 0.37, 1 point will be obtained; when SBMD is between 0.38 and 0.52, get 2 points; when SBMD is between 0.53 and 0.67, get 3 points; when SBMD is greater than or equal to 0.67, get 4 points.

2.3. Implementation of Feature Extraction of National Physical Fitness Data. National fitness monitoring application management adopts a mature monitoring scheme based on zabbix software [24]. ZABBIX open source monitoring software has rich community support and provides rich plug-ins. ZABBIX offers more extensive development papers and better Chinese support than other monitors like Nagios and CACI. This article utilizes Zabbix software to build and construct a monitoring system to fulfil the criteria of server monitoring and application failure alert after a thorough comparison. Server host monitoring and application failure monitoring are two types of monitoring. Host monitoring, for example, may track the characteristics of the host’s hardware and services, whereas application fault monitoring can track the program’s potential problems. Server failure monitoring, application failure monitoring, stability guarantee monitoring, and monitoring alarm monitoring are the four modules that make up the monitoring section. These four modules address a wide range of national fitness monitoring application management monitoring needs. The high availability is achieved by the omnidirectional monitoring of four components. As illustrated in Figure 6, the physical fitness identification and data analysis model is built based on the monitoring data.

Physique identification and data analysis model mainly involves three main functions: physique identification, physique data analysis, and visual display of analysis results. In addition, it also needs some basic management functions. The overall functional structure of the model is shown in Figure 7.

There are many user roles involved in the whole model. In order to facilitate the function design and authority management of users with different roles, the users in the model are divided into four categories: (1) ordinary users, that is, users who use the model for physique identification and self-health management, and such users use the physique identification function; (2) data analysis user: that is, the user who uses the data analysis function in the model to analyze the data of physique identification results. This kind of user uses the physique data analysis and result visual display function; (3) ordinary administrators, that is, administrators who publish and maintain the contents of the scale and 9 kinds of physique related materials. Such users use the management function; (4) administrator, that is, the administrator who manages the basic information, roles, and permissions of users of other types of accounts. The management function is also used by these users. The table
structure in the database is developed based on the E-R model after defining the qualities and connections of each object in the E-R model. It is necessary to establish unified and standardized naming and coding rules for data tables and their fields when designing database table structures, and the table structure should meet the third paradigm as much as possible, retaining appropriate data redundancy in exchange for better time efficiency. Table 3 and Table 4 demonstrate the primary database table structure of the database. These data tables are implemented using MySQL as database management.

Comprehensive report service is a physical health description document provided to users by national physical fitness monitoring application management. The user's physical examination data is uploaded through the ground station and synchronized to the provincial data center through the data synchronization service. The user's detection data is then compared to official health standard data, and the diagnosis, risk status, and proposed management goals for the user's health status are provided. The user may access his own website and examine his own full analysis report as well as individual analysis reports and download and save the reports in PDF format. The user's comprehensive analysis and stratification proposed management objectives and the corresponding subreports. The function structure of comprehensive report generation is shown in Figure 8.

The flow chart of user's data characteristics comprehensive report generation is shown in the figure. When starting to generate the comprehensive report, first, you need to obtain the FTP information, database connection information, and management target service information in the configuration file to connect the database; then, check whether the comprehensive report has been generated. If it has been generated, download it directly from FTP and display it. Otherwise, continue to generate the report. In the process of generating the report, it is necessary to obtain the data of the first page and each subreport. The data of the first page includes the user's basic information, diagnosis results, cardiovascular exercise risk stratification, and suggested management objectives. The data for the subreport must be filled in selectively based on the parameter type. The information acquired is the key indicator data for each module over the last five years. Fill in the comprehensive report entity with the relevant data to produce a comprehensive report; obtain the corresponding itemized PDF report from FTP according to the configuration file's path for each itemized report, then combine the generated comprehensive report body and the obtained itemized PDF report into a PDF file of comprehensive analysis report, and upload the combined PDF file to the configuration file's FTP path. The created complete report file may be seen upon request.

3. Analysis of Experimental Results

The 11-item data acquisition program of national physique is used to collect 11 basic pieces of data of test users, including grip strength, vital capacity, height, weight, sitting forward flexion, reaction time, vertical jump, push-ups (male), eye closed one leg standing, step test, and sit-ups (female). Collecting users' national 11 pieces of data through the acquisition program is divided into three steps: card opening, acquisition, and card reading. The IC card model used is at24c16, which has an internal capacity of 16K (bits), which is 2048 bytes and can store the personal information
of the tested users and national 11 test pieces of data. Set the
write position of each test item as shown in Table 5.

If \( P \) is the national physical fitness test result, \( t \) is the test
time, \( X \) is the number of nationals, and \( N \) is the number of
experiments, then

\[
P_k = \frac{1}{R} \int (x)^k (x)^T = \frac{1}{T} \sum_{n=0}^{n} (n)^k - 1.
\]

According to the calculation formula, the national
physical quality can be judged. The statistical results of the
physical fitness model produced by the python package and
the logarithmic curve of the general physical fitness model
are displayed by the python package, and the statistical
results of the physical fitness model and the real curve
produced by the python package are displayed by the python
package. It provides rich interactive statistical charts such as
relationship charts and also provides the download function
of the generated charts, which can be saved to the computer
in the form of pictures. The general statistical results of
change data are shown in Figure 9.

Further, the traditional method and the feature ex-
traction method proposed in this paper are used for com-
parative detection. The above figure is a reference sample to
judge the accuracy of the extraction results of the two
methods. The results are shown in Table 6.

It can be seen from the table that the feature extraction
effect of this method is better than the traditional feature
extraction method in performance, and the extraction ac-
curacy is higher. Further, compare the traditional feature
mining method with the method in this paper in terms of
time consumption, as shown in Figure 10.

It can be seen from the figure that the traditional
feature mining method consumes significantly more time
than this grammar. It is concluded that this method has
high accuracy, less time consumption than traditional
methods, high mining speed, and accuracy and is
feasible.
### Table 5: Data storage location and format.

| Name                  | Page position | Byte address | Maximum number of bits of test data | Decimal point position | Test data unit     |
|-----------------------|---------------|--------------|-------------------------------------|------------------------|---------------------|
| Sitting body flexion   | 8             | 1975         | 4                                   | 3                      | centimeter          |
| Grip                  | 8             | 1931         | 4                                   | 3                      | kg.                 |
| Weight                | 8             | 1965         | 3                                   | 2                      | kg.                 |
| Step index            | 8             | 1956         | 4                                   |                         | Floating-point      |
| Height                | 8             | 1963         | 4                                   | 3                      | centimeter          |
| Abdominal curl        | 8             | 1961         | 3                                   | 2                      | second              |
| Vital capacity        | 8             | 1969         | 2                                   | 0                      | Nothing             |
| Longitudinal jump meter | 8            | 1987         | 4                                   | 2                      | Nothing             |
| Standing on one foot  | 8             | 2018         | 3                                   | 2                      | centimeter          |
| Reaction time         | 8             | 2041         | 2                                   | 2                      | second              |
| Push-up               | 8             | 2033         | 3                                   | 1                      | second              |

**Figure 9:** General statistical results of national body change data.

**Table 6: Feature extraction accuracy of two methods.**

| Number of experiments/time | Name       | Traditional method/% | Paper method/% |
|----------------------------|------------|-----------------------|----------------|
| 10                         | SBMI       | 83.1                  | 95.4           |
|                            | SFlexibility | 77.1                  | 91.3           |
|                            | Sit-up     | 72.3                  | 89.9           |
|                            | sfat%      | 83.1                  | 92.2           |
|                            | SBPH       | 66.1                  | 93.7           |
|                            | SBMD       | 65.5                  | 92.1           |
|                            | \( P_k \)  | 49.3                  | 88.2           |
| 20                         | SBMI       | 63.2                  | 89.3           |
|                            | SFlexibility | 71.2                  | 89.7           |
|                            | Sit-up     | 73.6                  | 89.8           |
|                            | sfat%      | 68.6                  | 95.2           |
|                            | SBPH       | 69.1                  | 93.3           |
|                            | \( P_k \)  | 64.3                  | 90.4           |
Data mining is applicable to the study of national physique. Through the data mining of physical fitness test data, we find some rules that are confirmatory and contain new knowledge, which proves that the data mining tool is suitable for physical fitness data analysis and serves the field of physical health. According to the mining results, different physical exercise methods are adopted for different populations.

### Data Availability

The data used to support the findings of this study are included within the article.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### References

[1] S. Liu, D. Liu, K. Muhammad, and W. Ding, “Effective template update mechanism in visual tracking with background clutter,” *Neurocomputing*, vol. 458, pp. 615–625, 2021.

[2] S. Liu, S. Wang, X. Liu et al., “Human memory update strategy: a multi-layer template update mechanism for remote visual monitoring,” *IEEE Transactions on Multimedia*, vol. 23, pp. 2188–2198, 2021.

[3] S. Liu, T. He, and J. Dai, “A survey of crf algorithm based knowledge extraction of elementary mathematics in chinese,” *Mobile Networks & Applications*, vol. 26, pp. 1–13, 2021.

[4] K. J. Lian, S.-y. Fang, and Y.-f. Zhou, “Model predictive control of the fuel cell cathode system based on state quantity estimation,” *Computer Simulation*, vol. 37, no. 7, pp. 119–122, 2020.

[5] A. Xe, A. Xd, and Z. B. Zhen, “Health level classification by fusing medical evaluation from multiple social networks,” *Future Generation Computer Systems*, vol. 114, no. 4, pp. 574–580, 2021.

[6] S. Gupta, B. L. Parne, and N. S. Chaudhari, “ISAG: IoT-enabled and Secrecy Aware Group-based handover scheme for e-health services in M2M communication network,” *Future Generation Computer Systems*, vol. 125, no. 3, pp. 168–187, 2021.

[7] K. G. Schmidt, A. Holtermann, M. B. Jørgensen, M. J. Svendsen, and C. D. N. Rasmussen, “Developing a practice and evidence-based guideline for occupational health and safety professionals to prevent and handle musculoskeletal pain in workplaces,” *Applied Ergonomics*, vol. 97, no. 4, Article ID 103520, 2021.

[8] S. Javid and A. Mirzaei, “Presenting a reliable routing approach in iot healthcare using the multiobjective-based multiagent approach,” *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5572084, 20 pages, 2021.

[9] M. Luhnen, S. S. Ormsland, A. Willemsen et al., “Developing a quality management system for the European network for health technology assessment (EU-netHTA): toward European HTA collaboration,” *International Journal of Technology Assessment in Health Care*, vol. 37, no. 1, p. e59, 2021.

[10] G. Bellini, M. Cipriano, S. Comai et al., “Understanding social behaviour in a health-care facility from localization data: a case study,” *Sensors*, vol. 21, no. 6, Article ID 2147, 2021.

[11] Z. Zhang and S. Zhang, “Application of Internet of things and naive bayes in public health environmental management of government institutions in China,” *Journal of Healthcare Engineering*, vol. 2021, Article ID 9171756, 7 pages, 2021.

[12] G. Dimauro, V. Bevilacqua, and L. Pecchia, “Bioelectronic technologies and artificial intelligence for medical diagnosis and healthcare,” *Electronics*, vol. 10, no. 11, p. 1242, 2021.

[13] F. Ullah, I. Ullah, A. Khan, M. I. Uddin, H. Alyami, and W. Alosaimi, “Enabling clustering for privacy-aware data dissemination based on medical healthcare-iot for wireless body area network,” *Journal of Healthcare Engineering*, vol. 2020, Article ID 8824907, 10 pages, 2020.

[14] R. Wang, “Exploration of data mining algorithms of an online learning behaviour log based on cloud computing,” *International Journal of Continuing Education and Life-Long Learning*, vol. 31, no. 3, p. 371, 2021.

[15] S. Baloch, S. S. Samsani, and M. S. Muhammad, “Fault protection in microgrid using wavelet multiresolution analysis and data mining,” *IEEE Access*, vol. 9, no. 99, p. 1, 2021.

[16] N. Bellin, E. Racchetti, C. Maurone, M. Bartoli, and V. Rossi, “Unsupervised machine learning and data mining procedures reveal short term, climate driven patterns linking physico-chemical features and zooplankton diversity in small ponds,” *Water*, vol. 13, no. 9, p. 1217, 2021.

[17] B. Yuan and S. Xu, “Applications of data mining in intelligent computer-aided athletic training,” *Computer-Aided Design and Applications*, vol. 18, no. S4, pp. 1–12, 2021.

[18] G. Fortino, R. Buyya, M. Chen, and F. Herrera, “Special issue on methods and infrastructures for data mining at the edge of internet of things,” *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10220–10221, 2021.

[19] Y. Shen, W. Zheng, W. Yin, A. Xu, and H. Zhu, “Feature extraction algorithm using a correlation coefficient combined with the VMD and its application to the GPS and GRACE,” *IEEE Access*, vol. 9, no. 99, p. 1, 2021.

[20] R. Dadsema, “Rotational moment shape feature extraction and decision tree based discrimination of mild cognitive impairment conditions using MR image processing,” *Biomedical Sciences Instrumentation*, vol. 57, no. 2, pp. 228–233, 2021.
[21] S. Nurmaini, B. Tutuko, M. N. Rachmatullah, A. Darmawahyuni, and F. Masdung, "Machine learning techniques with lowdimensional feature extraction for improving the generalizability of cardiac arrhythmia," *IAENG International Journal of Computer Science*, vol. 48, no. 2, pp. 369–378, 2021.

[22] Y. Wang, X. Song, G. Gong, and N. Li, "A multi-scale feature extraction-based normalized attention neural network for image denoising," *Electronics*, vol. 10, no. 3, p. 319, 2021.

[23] F.-P. An, J.-e. Liu, and L. Bai, "Pedestrian reidentification algorithm based on deconvolution network feature extraction-multilayer attention mechanism convolutional neural network," *Journal of Sensors*, vol. 2021, Article ID 9463092, 12 pages, 2021.

[24] R. Fonseca, O. Guarnizo, D. Suntaxi, A. Cadiz, and W. I. Creixel, "Convolutional neural network feature extraction using covariance tensor decomposition," *IEEE Access*, vol. 9, no. 99, p. 1, 2021.