Heliyon 10 (2024) e37294

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Advanced neurological activity status of athletes based on big data technology

Wenhui Ma [a,b], Bin Guo [c,*]

[a] College of Physical Education, China Three Gorges University, Yichang, 443002, Hubei, China
[b] Graduate School, Philippine Christian University, Malate, Manila, 1004, Philippines
[c] School of Physical Education, Dalian University, Dalian, 116622, Liaoning, China

ARTICLE INFO

Keywords:
Big data
Higher nervous activity
Function state
Machine learning

ABSTRACT

Currently, the application scope of big data (BD for short here) technology is relatively narrow, mostly used in the medical field, and the degree of application is relatively superficial, mostly for data statistical record analysis. Therefore, By combining the literature review, this article has decided to construct a system based on BD technology for analyzing the advanced neural activity status of athletes. The system is mainly divided into two parts, one is the biological information collection part. As an important source of system data, it is necessary to use professional equipment to collect ECG and EEG data and ensure the accuracy of the data through signal filtering, Gaussian noise elimination, salt and pepper noise, and exponential noise de-noising technology. The other is the algorithm problem of BD systems. Considering that the traditional algorithm can not deal with a large amount of data effectively, this paper chooses the BD spectral clustering algorithm based on core points as the main algorithm to cluster the data. By evaluating the efficiency of system learning, data collection and classification, system scheme construction, and error rate, this article ultimately determined the practical feasibility and effectiveness of the system. After completing the construction of the system, considering the gap between the system's performance and traditional data, this article analyzed the improvement data of various aspects of sports training. This paper compares the performance differences between the system based on BD technology and the traditional data analysis method under different indicators. In terms of data collection and classification, the accuracy of the system based on BD technology in the collection and classification of ECG and EEG data reached 100 % and 90 %, respectively, which was significantly higher than 60 % and 30 % of the traditional methods. By comparing the data from five training courses, it is found that the training efficiency of the conventional method has increased by 60 % in the first course, while the efficiency of the training method based on the BD system has increased by 85 % in the fifth course. For the activation efficiency of nerve function, the activation efficiency of brain nerve function reached 60 % and 90 % respectively in the two nerve function activation training based on the BD system, which was much higher than 30 % and 45 % of the traditional methods. Through a series of tests and comparative analysis of data, the effectiveness of the BD system is finally determined, which can achieve the goal of improving athletes' training efficiency.

* Corresponding author.
E-mail addresses: mawenhui@ctgut.edu.cn (W. Ma), 17640388598@163.com (B. Guo).

https://doi.org/10.1016/j.heliyon.2024.e37294
Received 1 April 2024; Received in revised form 27 August 2024; Accepted 30 August 2024
Available online 31 August 2024
2405-8440/

1. Introduction

Since the mid-20th century, the world has entered the third technological revolution centered on information technology. This historic change has not only driven the leapfrog development of technology, but also profoundly changed all aspects of human society. The rapid development and widespread application of electronic computer technology marks the beginning of a new technological era for humanity. During this period, the advancement of computer technology not only promoted the improvement of data processing capabilities, but also provided powerful technical support for information storage, transmission, and analysis. The accompanying big data technology, as an emerging product of the information age, has gradually emerged and become an important force driving social change. The rise of big data technology not only demonstrates enormous application potential in traditional fields such as business, healthcare, and finance, but also opens up new paths for scientific research and social development. The scale and complexity of data are rapidly increasing, and traditional data processing methods are gradually exposing their limitations in meeting the needs of modern information society. Therefore, the emergence of big data technology is seen as the key to solving this challenge. Its core value lies not only in its powerful data storage and processing capabilities, but also in how to extract valuable information from massive data and provide scientific support for decision-making through data analysis and modeling. The essence of big data technology is not only the improvement of computing power, but also a revolution in data cognition ability. It can reveal the deep relationships between data from multiple dimensions and scales through intelligent algorithms, providing new perspectives for understanding and optimizing complex systems. Although big data technology has made significant progress in fields such as business, finance, and healthcare, its application in sports science, especially in the analysis of advanced neural activity states in athletes, is still in the initial exploration stage. Traditional sports science research and training methods often rely on limited physiological data and empirical judgments, and the acquisition and processing of data are often single and lack systematicity, which poses many challenges to the evaluation and optimization of training effectiveness. Although traditional methods can provide guidance for athletes' training to a certain extent, their limitations make it difficult to comprehensively and accurately reflect the true state and training effects of athletes during the training process. With the continuous improvement of sports competition level, athletes' training is not limited to physical fitness, but also needs to pay attention to the functional status of the nervous system. The role of the nervous system in athlete performance is crucial, as it not only involves coordination, agility, and reaction time in movement, but also affects the athlete's endurance and adaptability to training.

How to effectively analyze and optimize the neural activity status of athletes has become a key research direction for improving athletic performance. In this context, based on big data technology, this article aims to construct a system for analyzing the advanced neural activity status of athletes. Through comprehensive collection and in-depth analysis of athletes' biological information, this article aims to reveal the complex relationship between neural function status and training effectiveness, and provide scientific basis for personalized training programs. This system significantly improves the accuracy of data and the depth of analysis, overcoming the limitations of traditional methods in processing large-scale and complex data. This study not only fills the gap in the application of big data technology in sports science, but also provides a new method and tool for future sports training, promoting the development of sports science towards a more accurate and scientific direction.

2. Related work

In the exploration of BD technology, many scholars have their views and applications on BD. Oussous, Ahmed believed that the core function of BD technology is primarily to conduct data surveys. Therefore, he conducted a detailed analysis of BD technology, listing the conventional types of data that BD can analyze, and concluded that BD has indeed improved people's work efficiency in data processing and analysis [1]. Dash, Sabyasachi has applied BD technology in the healthcare field. He used BD technology to manage and analyze healthcare data in healthcare and planned for the future development of healthcare technology. He believed that BD technology is the development trend of future healthcare information data analysis, and the application of BD technology should be promoted [2]. Choi and Tsan-Ming applied BD technology to operational management, believing that there is a large amount of data in operational management. Managing and analyzing such data through human resources can greatly affect work efficiency, so he attempted to improve data processing efficiency through BD technology [3]. Zhu, Li conducted data analysis on traffic data in intelligent transportation systems using BD technology. He believed that BD technology can quickly export and analyze traffic data from various places promptly, thereby determining whether there have been violations or traffic accidents in various places, and timely understanding the traffic conditions in various places [4]. Based on the above literature review, it is found that BD technology has been applied in various industries, and the presence of BD technology can be seen in many places. However, in these literature materials, there is no exploration of BD technology methods, and most of them directly use existing BD technologies to obtain results.

Considering that to truly and comprehensively understand BD technology, it is necessary to have an understanding of the theoretical methods of BD technology. After screening the literature, some relevant literature on BD technology theory and methods were selected and analyzed through consultation. In the literature, Shuangming, Wang believed that BD technology can be applied to sports. He proposed to build a sports resource information platform based on BD technology, which can enable athletes to quickly understand their weak points in sports and provide training and improvement [5]. Urbaczewski, Andrew has put forward a different perspective on BD technology, stating that although BD technology has improved the efficiency of sports training to a certain extent. However, this approach might reduce the enjoyment of sports and focus excessively on enhancing physical abilities, leading to a mechanistic view of athletes, and relevant research and investigation have been conducted for this purpose [6]. Beam, Andrew L. proposed the concept of applying BD to the healthcare field. He believed that BD technology is the future development direction and combined it with machine learning (ML for short here) technology to improve the efficiency of BD technology in processing data [7]. Wang C. Jason applied BD

Fig. 1. Partial system bio information collection equipment diagram.
Fig. 1a: Diagram of EEG signal transmission equipment Fig. 1b: Diagram of eye-ear concussion detection equipment.

technology to the COVID-19 pneumonia response in Taiwan. During the application process, he elaborated on the internal structure of BD technology, including but not limited to the collection, processing, and storage of BD. A series of COVID-19 pneumonia emergency plans and measures have been developed through this technology [8]. Ngiam, Kee Yuan's research direction in BD technology is healthcare services. In his research on BD technology, he compared BD technology with ML algorithm technology. He found that BD technology has some similarities with ML algorithms, and attempted to integrate the two technologies in an attempt to improve the working ability of BD technology [9]. Price, W. Nicholson conducted a detailed analysis of BD technology in the privacy of the medical BD era. He believed that BD technology is achieved by classifying and integrating all personal information, and drawing a three-dimensional image of individuals through data, to accurately control personal habits and develop corresponding push plans. This leads to individuals having no privacy in front of BD, and once the data is intercepted by criminals, it would cause losses to personal property and life safety [10]. In the literature review above, it was found that BD technology is often used in the medical field, and in the methodology section, many scholars have attempted to combine BD technology with ML algorithm technology. In terms of technical use, this type of literature technology is relatively fixed in use and only conducts further research on the existing foundation, which greatly affects the development of BD technology.

3. Theoretical methods of BD systems

BD technology, also known as massive data, refers to the massive amount of information involved that cannot be captured, managed, processed, and organized into information that helps enterprises make more positive business decisions within a reasonable time through mainstream software tools. BD technology has the following five characteristics: large volume, high speed, diversity, low-value density, and authenticity. The strategic significance of BD technology lies not in mastering massive data information, but in specialized processing of meaningful data [11,12]. From a technical perspective, the relationship between BD and cloud computing is like paper. BD is the front of the paper, while cloud computing is the back of the paper, and the two cannot be separated. Due to the large amount of internal data in BD, it is usually not possible for a single computer to perform BD operations. Therefore, a distributed architecture must be adopted, and the analysis of BD can only be completed through the joint assistance of multiple computers. The characteristic of BD technology is the distributed data mining of massive data. It must rely on the distributed processing, distributed database, cloud storage, and virtualization technology of cloud computing. This study adopts a combination of big data technology and machine learning algorithms, focusing on basketball players who collect their ECG and EEG biological information by wearing professional devices. Sampling techniques include signal filtering, noise cancellation, and signal denoising to ensure data accuracy. These data are then processed and analyzed using a spectral clustering algorithm based on core points.

3.1. System bio information collection

There are generally two ways to collect exercise information from sports trainers. One is to manually record the data, which is usually used on sports training data. Another method is to wear professional instruments and equipment to collect data through electronic detection. This data collection method is generally used in areas where data collection is more complex and difficult, such as ECG (electrocardiogram) data collection and EEG (electroencephalogram) data collection. This article focuses on basketball training as a sport and collects data from basketball players. Data collection items include electrocardiography, functional magnetic resonance imaging, eye-ear concussion response, and evoked potential. During the training process of athletes, the functional state of neural activity directly affects the performance of athletes' sports training level, so the collection of various data is very important. The following is a diagram of some system biological information collection devices, as displayed in Fig. 1: (a is an EEG signal transmission device, and b is an eye ear concussion detection device)

The corresponding data obtained through devices cannot be directly used by BD systems, and preprocessing operations need to be carried out on the obtained data. The data generated by preprocessing devices is mostly presented in the form of images, so corresponding image methods need to be applied in image data processing. The steps of image processing methods mainly include signal filtering, noise elimination, and signal denoising. The most important step is to denoise the signal image. At present, the common types of noise in image processing include Gaussian noise, salt and pepper noise, and exponential noise. Gaussian noise is the most common

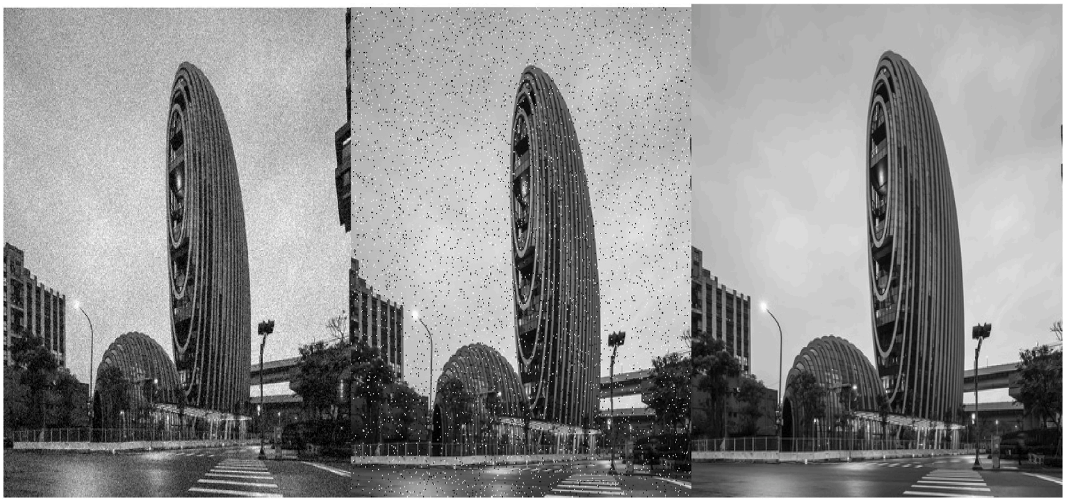

Fig. 2. Comparison of image data filtering results.
Fig. 2a: Gaussian noise image data Fig. 2b: Salt and pepper noise image data Fig. 2c: Raw image data.

and important noise. Many image-denoising algorithms are designed to reduce Gaussian noise. The probability density function is as follows Formula (1):

$$q(z) = \frac{1}{\sqrt{2\pi\varphi^2}} e^{-(\gamma-\beta)^2} \Big/ 2\varphi^2 \tag{1}$$

φ represents the standard deviation, and β represents the average of the grayscale values. Salt and pepper noise, also known as pulse noise or spike noise in related fields, mainly manifests as randomly distributed black and white dots in the image. Its probability density function is generally written as the following formula:

$$q(z) = \begin{cases} Q_a, & z = a \\ Q_b, & z = b \\ 1 - Q_a - Q_b & otherwise \end{cases} \tag{2}$$

In Formula (2), z represents the probability of grayscale values. Salt pepper noise can be eliminated by a median filter. The last common noise is exponential. The probability density function of exponential noise is as follows Formula (3) (4) (5):

$$q(z) = \begin{cases} ac^{-az}, & z \geq 0 \\ 0, & z \leq 0 \end{cases} \tag{3}$$

$$z^v = \frac{1}{a} \tag{4}$$

$$\partial^2 = \frac{1}{a^2} \tag{5}$$

In formula (3), c represents the function variable, In formula (4) z^v represents the mean value of the probability density function, and In formula (5) ∂^2 represents the variance of the probability density function. By using the above function formula, the image data can be filtered by radio waves, thereby making the image data clear and complete. To improve the filtering effect of image data, this article selected images with obvious visual differences to filter images containing Gaussian noise and salt and pepper noise. The following is a comparison of image data filtering results, as displayed in Fig. 2: (a is image data containing Gaussian noise, b is image data containing salt and pepper noise, and c is original image data)

3.2. BD system algorithm theory

The so-called BD system is an algorithm system that processes and analyzes various types of data. In the information technology era, the types of data processing algorithms have undergone rapid development. At present, a cluster analysis algorithm is adopted in data processing, and cluster analysis is also called group analysis. It is a statistical analysis method for studying classification problems and an important algorithm for data mining [13,14]. With time, the data structure has also become complex and ever-changing. Faced with various types of information data, a clustering algorithm can no longer meet the usage needs. Based on this, scholars have proposed different clustering algorithms based on the initial conditions and application criteria. In the traditional clustering algorithms, the

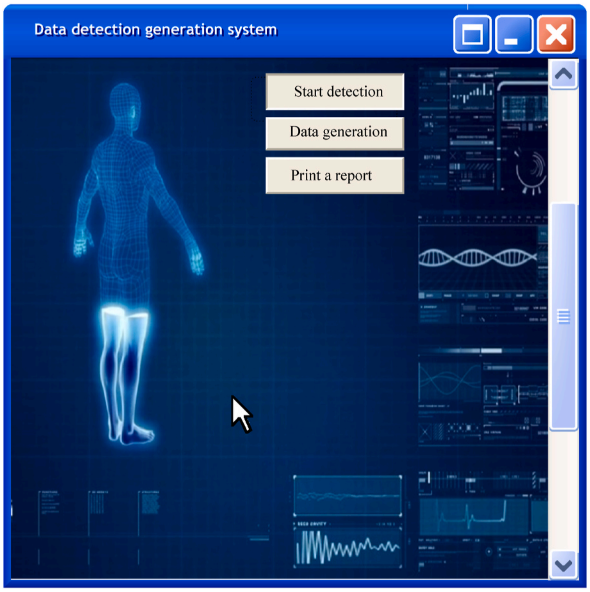

Fig. 3. Partial system operation interface of the BD system.

main algorithms can be divided into the following categories: partition clustering algorithm, hierarchical clustering algorithm, density-based clustering algorithm, and grid-based clustering algorithm. Although there are many types of algorithms proposed, the ability of classical clustering algorithms to handle large amounts of BD is still very limited. In the face of highly complex datasets, data processing systems that use traditional clustering algorithms exhibit very low data processing efficiency, and sometimes even experience failure and collapse [15]. To solve the problem of a large amount of data and difficult cluster analysis, this paper decided to use a spectral clustering algorithm based on core data as the main algorithm to cluster data. The spectral clustering algorithm is based on the spectral graph theory. Compared with the traditional clustering algorithm, it has the advantage of clustering in arbitrary shape sample space and converging to the global optimal solution. The spectral clustering algorithm can obtain the global optimal solution with arbitrary shape information data and is not affected by the input order of information data. By combining the two theories of athlete's body data collection and BD spectral clustering algorithm based on core points, and cooperating with ML, neural network, and other related information technologies, a BD technology-based analysis system for athletes' advanced neural activity functional state (hereinafter referred to as BD system) was constructed. The following is a partial system operation interface diagram of the BD system, as displayed in Fig. 3.

4. System performance test

Usually, BD technology is closely connected to the Internet, and BD needs to be collected and deeply learned through the network to improve data analysis capabilities. This can make the analyzed data more accurate, and the solutions constructed through the data are more three-dimensional and realistic, in line with user needs [16,17]. The BD system built in this paper combines information collection technology, cluster analysis algorithm technology, and network technology, and considers it from many aspects. Its purpose is to more comprehensively understand the functional status of high-level neural activities of sports personnel, and develop more scientific and effective sports training programs. Although in theory, various technologies are relatively mature and have a complete theoretical system, in practical combination and application, untested systems may not necessarily achieve perfect operation. Therefore, to detect the actual operation of the BD system, this article has decided to conduct a system functional test on the BD system and determine whether there are any problems or abnormalities in the system based on the test data.

4.1. System learning efficiency and data collection classification test

In BD systems, the identification and classification of data is not done manually, and a large amount of data can only be completed by system machines. However, the cognitive recognition of data by the system needs to be completed through ML. ML is a multi-disciplinary interdisciplinary field that specializes in studying how computers can imitate human behavior. Through ML, various data knowledge is continuously updated, and the knowledge structure of the machine system itself is improved to improve its performance. The BD system contains a large amount of detection data, with a wide variety of types. Therefore, whether the system can collect, recognize, and classify detection data has become the key first step in analyzing athlete status. The following are the data tested for system data collection, classification, and learning efficiency. The data collection classification test data table is displayed in Table 1, and the system data learning efficiency test data table is displayed in Table 2:

Table 1
Data collection classification test data table.

Unit(%)	Group 1	Group 2	Group 3	Group 4
Data 1	100 %	100 %	100 %	0 %
Data 2	99 %	100 %	100 %	0 %
Data 3	98 %	100 %	100 %	0 %

Table 2
System data learning efficiency test data table.

	Group 1	Group 2	Group 3	Group 4
Data 1	100 %	100 %	100 %	0 %
Data 2	100 %	100 %	100 %	0 %
Data 3	100 %	100 %	100 %	0 %

Table 3
System plan construction data test table.

	Scheme 1		Scheme 2	
	Data bulk	Construction rate	Data bulk	Construction rate
ECG test	60	100 %	100	100 %
EEG test	50	90 %	100	100 %
Eyesight test	10	30 %	100	100 %

In the data collection classification testing and system data learning efficiency testing, it is evident that there are abnormal situations. In the data collection and classification test, there were fluctuations in the first set of test data, with the highest classification rate of 100 % and the lowest being 98 %. This is within the normal range. Considering that there may be some combination of data in actual data records, it is necessary to manually debug and improve the system. Therefore, during the testing process, a 5 % collection and classification error is allowed. The first set of data clearly meets this standard, so there is no problem with the first set of data, but there is still a need to repair and improve the errors that occur. In contrast, there are obvious anomalies in the fourth group of data, with a collection and classification rate of 0 % for all data in the fourth group, indicating that there must be problems with the system or data. After discussion and analysis, it is believed that there are two possibilities for this situation to occur at present, one is a problem with the BD system, and the other is an abnormal data transmission. Based on the output of the first three sets of data, it is believed that there is a higher probability of system analysis problems caused by abnormal data transmission. Subsequently, a detailed inspection was conducted on the data collection device. After inspection, it was found that there was loose data wiring in the data collection device, and the fourth group of data came from the device. After debugging the equipment, the data collection and classification were conducted again, and the results showed that the fourth group of data collection and classification results were all 100 %. This also indicates that there are no technical anomalies in the BD system in terms of data collection and classification.

The efficiency of system data learning is an auxiliary function based on the data collection and classification function. Only after obtaining the corresponding data can the data be analyzed and learned, thereby achieving the goal of improving the system. The learning efficiency of the system's data learning function has reached 100 % for most data sets. Only the fourth set of data was affected by the collection device, and the corresponding data could not be collected. Therefore, the data could not be learned, resulting in a learning efficiency of 0 %. After restoring data collection and classification, the system conducted a second test on the fourth group of data, and the second test results showed that the learning efficiency of the fourth group also reached 100 %. The above test data can indicate that there are no abnormal situations in the BD system in terms of data collection, classification, and data learning, and the data can be analyzed, learned, and processed normally.

4.2. System solution construction and system error testing

The fundamental purpose of the system for collecting information on athletes is to understand their corresponding sports data cleaning, develop comprehensive sports improvement plans based on the data situation, and improve the efficiency of sports training. To improve athletes' sports performance, it is necessary to provide corresponding sports improvement plans. This plan is provided by the system after learning and analyzing sports data. Therefore, whether the system can construct and provide corresponding sports training plans is one of the important functions of BD systems. The following is the data test table for the construction of the system solution, as displayed in Table 3:

From the data test table of system scheme construction, it can be seen that this test compared the number of test data and scheme construction rate for three different schemes. It was found that in the two different schemes, the more data provided by the detection project, the higher the scheme construction rate. In Scheme 1, the number of test data for electrocardiogram testing is 60, and the construction rate of the scheme is 100 %. On the contrary, the number of test data for visual acuity testing is 10, and the construction

Table 4
System error test data table.

	Error test 1		
	Data bulk (number)	Processing time(second)	Error rate (%)
ECG test	100	4	0 %
EEG test	100	2	0 %
Eyesight test	100	3	0 %
Listening test	100	10	10 %
Muscle test	100	12	20 %
	Error test 2		
	Data bulk (number)	Processing time(second)	Error rate (%)
ECG test	1000	20	0 %
EEG test	1000	21	0 %
Eyesight test	1000	19	0 %
Listening test	1000	50	30 %
Muscle test	1000	55	50 %

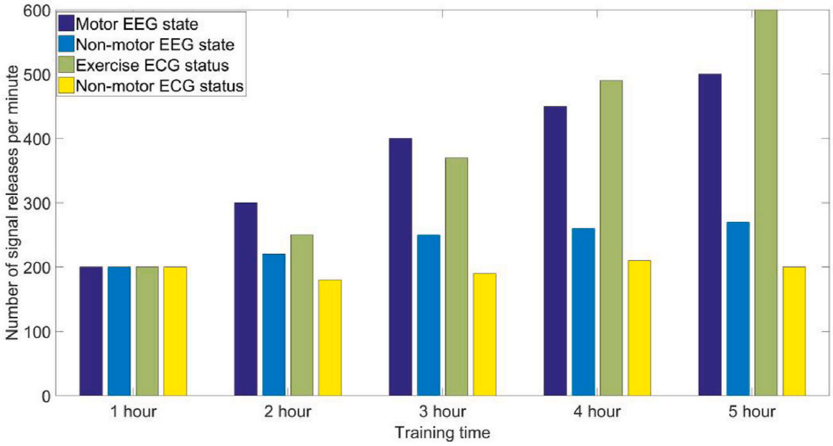

Fig. 4. Comparative analysis of changes in the influence of electrocardiogram and electroencephalogram neural data.

rate of the plan is only 30 %. Option 2 has increased the number of tests compared to Option 1. All test item data in Option 2 are 100, and the construction rate of Option 2 has also reached 100 %. This also indicates that the construction of the plan is influenced by the amount of data, and the larger the amount of data, the more complete the plan construction.

After a series of experimental tests, all functions have met the working requirements and standards, and there are no technical issues. However, considering that the system's work requires cooperation between modules. Therefore, to verify the actual usage of the system, it is necessary to conduct error testing on the system. By simulating various common data testing situations, it is necessary to check whether the system may experience errors due to certain reasons, which may lead to system testing failure. The following is the system error test data table, as displayed in Table 4:

From the system error test data table in Tables 4 and it can be seen that under certain detection data, the longer the detection takes, the higher the probability of system error. In System Error Test 1, 100 pieces of data were selected for each test, among which the duration of electrocardiogram (ECG), electroencephalogram (EEG), and visual acuity tests remained stable between 2 and 4 s, with an error rate of 0 %. The time required for listening and muscle testing is relatively long, ranging from 10 to 12 s, with an error rate of 10 %–20 %, especially when the test data volume is small. In System Error Test 2, the same test data increased the number of data by 10 times, with 1000 copies per test and error testing conducted on them. In the test results, there was a significant improvement in the testing time for all data, with the first three data testing times ranging from 19 to 21 s. The time required for hearing and muscle tests is longer, with a significant increase in error rate between 50 and 55 s, ranging from 30 % to 50 %. From the above data, it can be explained that the system error rate is related to the processing time of the system data. The longer the processing time, the higher the error probability. Therefore, it is necessary to stabilize the data processing time, reduce the time spent on data processing, and reduce system error reporting. After a detailed understanding of the reasons for the error, it was found that the reason for the long processing time of the data was due to conflicts in some combination data. The system was unable to learn from the combination data, resulting in the data being subjected to long-term operations in clustering algorithms, which increased the data processing time and led to errors due to the inability to analyze. In response to the issue of combined data, testers further improved and improved the system in the future, ultimately solving the problem of combined data and retesting the above data. According to the results, all data processing time

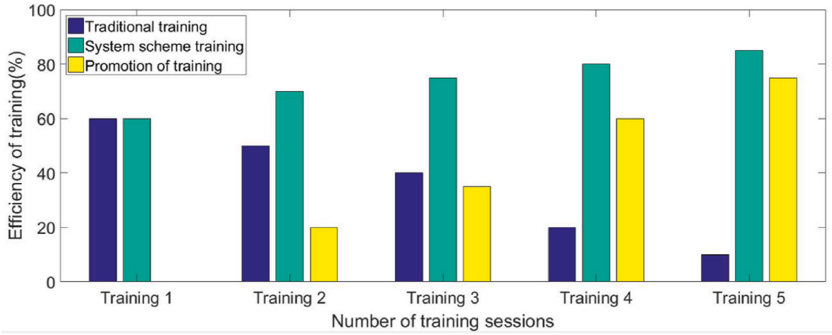

Fig. 5. Comparative analysis of efficiency data for five basketball training sessions.

is still affected by the amount of data, but the issue of system errors has been fundamentally resolved. Routine data testing does not generate errors.

5. Data of BD system for improving sports training

5.1. Comparative analysis of changes in the impact of electrocardiogram and electroencephalogram neural data and training efficiency data

During the process of basketball players' exercise, various physical neurological indicators would change with the time of exercise training, and the electrocardiogram and EEG nerves would also experience corresponding fluctuations. This fluctuation is not only influenced by sports training but also can be fed back to the athletes themselves, causing a chain reaction during sports training. To give a relatively simple example, during the training process, high-intensity and mismatched training can cause physical exhaustion and aversion among athletes. These types of emotions and physical changes can accelerate the decline of exercise status, leading to athletes developing resistance emotions in a very short period. Rapid diffusion and fermentation in subsequent training may ultimately lead to poor training effectiveness for athletes. The following is a comparative analysis of the changes in the impact of electrocardiogram (ECG) and electroencephalogram (EEG) neural data obtained from testing, as displayed in Fig. 4:

It can be seen that there are significant changes in the ECG and EEG data of basketball players in both exercise and nonexercise states. In the 5-h electrocardiogram data, the nonexercise data showed relatively small changes, with a minimum of 180 values per minute and a maximum of only 210 values per minute, with a gentle fluctuation. This can indicate that in this state, the electrocardiogram nerve is relatively stable and suitable for training. During the 5-h exercise state, the fluctuation amplitude of the electrocardiogram nerve values is significant and has been showing an upward trend. The lowest electrocardiogram nerve value is 200 values per minute, and the highest reaches 600 values per minute. According to the standard of electrocardiogram values, normal electrocardiogram values should be stable between 100 and 300, with a maximum limit of 400. In the state of exercise, the cardiac nerve function values significantly exceed this range, even reaching an astonishing 600 cardiac nerve values in the fifth hour. This indicates that the physical and mental health of athletes at that time was far beyond their capacity, which is highly likely to reduce training effectiveness and even endanger their physical and mental health. To improve unhealthy training methods, this article uses a BD system to analyze the data of ECG and EEG neural data, and forms corresponding training plans. This article compares the corresponding data recorded in previous training and obtains the following 5 comparative analyses of basketball training efficiency data, as displayed in Fig. 5:

In the training efficiency data graph, it can be seen that in traditional training methods, the training efficiency decreases with the increase in training times, indicating that conventional training methods have not scientifically considered and formulated training plans for basketball players. Among the 5 training efficiency data, the most significant improvement in training efficiency was the first training, with a 60 % improvement in training efficiency. The lowest improvement was the fifth training, with a training efficiency of only 10 %. By using a training plan that has been improved by the BD system for training, the training efficiency of athletes has significantly improved. The lowest training efficiency is the first training, with an efficiency increase of 60 %. As the functional state of the body's neural activity gradually stimulates, the training efficiency gradually improves. When conducting the fifth training, the training efficiency was improved to 85 %, which is 75 % lower than traditional training efficiency. From the data, it can be concluded that after analyzing and improving traditional training data using BD systems, the developed training plan can to some extent activate the athletes' neural activity status and improve their training efficiency.

5.2. Comparison and BD System's detection data on physical neurological activity function status

In the process of sports training, the purpose of sports training is to improve training efficiency and achieve the best training state in a very short time. The essence of achieving optimal physical activity in the human body is to fully activate the body's neural activity state, keeping the entire body's neural function in an active state. Therefore, how to effectively activate the overall neurological function is a key issue in improving training efficiency. Based on this, this article decides to compare the data of traditional training of body neurological function status with that of BD system training and determine the activation status of body neurological function

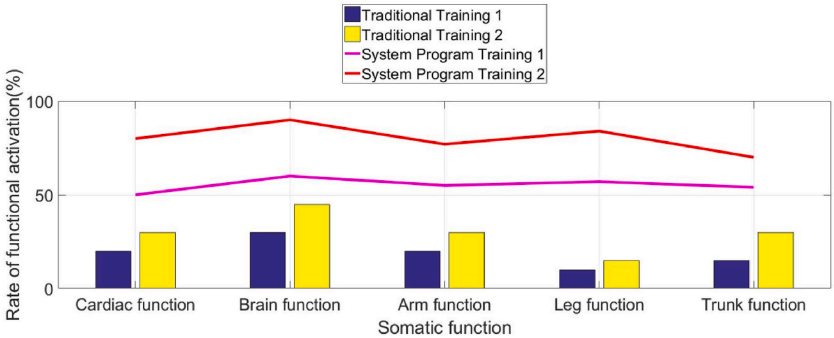

Fig. 6. Comparative analysis of data on the activation of bodily nervous functions.

between the two through the data. The following is a comparative analysis of the activation status data of the body's nervous system, as displayed in Fig. 6:

In the comparative analysis of data on the activation of body nervous function between two training sessions, it can be seen that the activation of body nervous function in traditional training is not ideal. The highest activation efficiency is in the brain's neural function. The activation efficiency of the first exercise training was 30 %, and the second one was 15 percentage points more active than the first one, increasing the activation efficiency to 45 %. Such activation efficiency is not ideal and has not met the expected activation efficiency of neural functions, while the activation of other neural functions is even more pessimistic. In contrast, the system training scheme is much more optimistic in activating the body's neural function. In the two neural function activation training sessions, the brain's neural function activation efficiency was also the highest, with the first activation efficiency reaching 60 %. The second time reactivated 30 % of neural function based on the first time, achieving a 90 % activation efficiency of neural function. At this point, the activity level of the nervous system has been fully aroused, and conducting exercise training in this state can quickly improve training efficiency. The above data indicate that BD technology does have a more scientific and efficient activation effect in terms of the activation status of the body's nervous system.

6. Conclusions

After system testing and comparative analysis of system data, it can be determined that the analysis system of advanced neural activity status of athletes based on BD technology is indeed feasible and effective in theoretical practice. Similarly, the training plan developed by the BD system based on the data has significantly improved compared to traditional training plans. It not only improves the training efficiency of athletes but also has a rapid activation effect on various neurological functions of the athlete's body. Based on the results of experimental testing and data comparison analysis, it can be concluded that the BD system has a good effect in analyzing the functional state of neural activity. It can greatly help athletes quickly activate their body functions, enter the optimal state of sports training as soon as possible, and improve training effectiveness. Future research should explore the application of this system in different sports and its long-term impact on exercise performance. And study the potential applications of the system in the fields of medical health and rehabilitation training. This will help further validate and improve the functionality of the system, and expand its application scope.

CRediT authorship contribution statement

Wenhui Ma: Writing – original draft. **Bin Guo:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Ahmed Oussous, Fatima-Zahra Benjelloun, Ayoub Ait Lahcen, Samir Belfkih, Big Data technologies: a survey, Journal of King Saud University-Computer and Information Sciences 30 (4) (2018) 431–448.
[2] Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma, Sandeep Kaushik, Big data in healthcare: management, analysis, and prospects, Journal of Big Data 6 (1) (2019) 1–25.
[3] Tsan-Ming Choi, Stein W. Wallace, Yulan Wang, Big data analytics in operations management, Prod. Oper. Manag. 27 (10) (2018) 1868–1883.
[4] Li Zhu, Fei Richard Yu, Yige Wang, Bin Ning, Tao Tang, Big data analytics in intelligent transportation systems: a survey, IEEE Trans. Intell. Transport. Syst. 20 (1) (2018) 383–398.
[5] Wang Shuangming, Jianwei Shi, Shiqiu Fu, Xu Wanlin, Research on the construction of sports resources information platform based on big data, International Journal of Advanced Network, Monitoring and Controls 3 (2) (2018) 28–32.
[6] Andrew Urbaczewski, Ryan Elmore, Big data, efficient markets, and the end of daily fantasy sports as we know it? Big Data 6 (4) (2018) 239–247.

[7] Andrew L. Beam, Isaac S. Kohane, Big data and machine learning in health care, JAMA 319 (13) (2018) 1317–1318.
[8] C. Jason Wang, Chun Y. Ng, Robert H. Brook, Response to COVID-19 in Taiwan: big data analytics, new technology, and proactive testing, JAMA 323 (14) (2020) 1341–1342.
[9] Kee Yuan Ngiam, Wei Khor, Big data and machine learning algorithms for health-care delivery, Lancet Oncol. 20 (5) (2019) e262–e273.
[10] W. Nicholson Price, I. Glenn Cohen, Privacy in the age of medical big data, Nat. Med. 25 (1) (2019) 37–43.
[11] Li Da Xu, Lian Duan, Big data for cyber-physical systems in industry 4.0: a survey, Enterprise Inf. Syst. 13 (2) (2019) 148–169.
[12] Yuji Roh, Geon Heo, Steven Euijong Whang, A survey on data collection for machine learning: a big data-ai integration perspective, IEEE Trans. Knowl. Data Eng. 33 (4) (2019) 1328–1347.
[13] Varun Grover, Roger H.L. Chiang, Ting-Peng Liang, Dongsong Zhang, Creating strategic business value from big data analytics: a research framework, J. Manag. Inf. Syst. 35 (2) (2018) 388–423.
[14] Ali Reza Alaei, Susanne Becken, Bela Stantic, Sentiment analysis in tourism: capitalizing on big data, J. Trav. Res. 58 (2) (2019) 175–191.
[15] Kathy A. Mills, What are the threats and potentials of big data for qualitative research? Qual. Res. 18 (6) (2018) 591–603.
[16] Alberto Ferraris, Alberto Mazzoleni, Alain Devalle, Jerome Couturier, Big data analytics capabilities and knowledge management: impact on firm performance, Manag. Decis. 57 (8) (2019) 1923–1936.
[17] Mehdi Mohammadi, Ala Al-Fuqaha, Sameh Sorour, Guizani Mohsen, Deep learning for IoT big data and streaming analytics: a survey, IEEE Communications Surveys & Tutorials 20 (4) (2018) 2923–2960.