

Retraction

Retracted: Sports Big Data Analysis Based on Cloud Platform and Its Impact on Sports Economy

Mathematical Problems in Engineering

Received 18 July 2023; Accepted 18 July 2023; Published 19 July 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] Y. Cheng and Y. Song, "Sports Big Data Analysis Based on Cloud Platform and Its Impact on Sports Economy," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6610000, 12 pages, 2021.

Research Article

Sports Big Data Analysis Based on Cloud Platform and Its Impact on Sports Economy

Ye Cheng¹ and Yan Song² 

¹Department of Physical Education, Hebei University of Economics and Business, Shijiazhuang 050061, Hebei, China

²College of Physical Education, Kunsan National University, Kunsan 54150, Jeonro-do, Republic of Korea

Correspondence should be addressed to Yan Song; 1920034@office365.kunsan.ac.kr

Received 26 December 2020; Revised 6 February 2021; Accepted 22 February 2021; Published 2 March 2021

Academic Editor: Sang-Bing Tsai

Copyright © 2021 Ye Cheng and Yan Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The service information system is constantly transforming to a networked information model, and domestic hardware equipment is constantly updated. Independent controllability has also become the basic requirement of the new information age. With the development of the information age and the new era of independent control, more and more services and applications will also be deployed on autonomous and controllable cloud platforms. With the rapid development of Internet technology in the information age and the resulting changes in productivity, people can record, store, and transmit more and more information. When information becomes recordable, storage, and easy to transmit, information becomes modern meaning nowadays, an era of information explosion characterized by massive, volatile, timely transmission, and diverse forms has truly come, forming what is now called the “big data era”. This article mainly introduces the analysis of sports big data based on the cloud platform and the research on the impact on the sports economy and intends to provide ideas and directions for the analysis of sports big data and the research on the impact on the sports economy. This paper proposes a cloud platform-based sports big data analysis and research methods for its impact on the sports economy, including the use of Hadoop cloud platform big data processing systems and support vector regression algorithms for cloud platform-based sports big data analysis and sports economy. The experimental results of this paper show that the average correlation between sports big data analysis and sports economic development is 0.5155, and appropriate cloud platform-based sports big data analysis plays a positive role in promoting sports economic development.

1. Introduction

Today's world is an era of data explosion. IDC report shows that the global digital volume was about 130 EB in 2005, and it will reach 40000 EB by 2020. There is no doubt that human society has entered the era of big data. Big data has penetrated into all walks of life and is an important part of the global economy. Without big data, it will be difficult for modern economy to have more innovation and vitality. The development and application of big data have become important sources to enhance the competitiveness of the country and enterprises.

In the 21st century, with the rapid development of science and technology and the gradual improvement of people's material and cultural level, people gradually realize

the importance of physical health. In recent years, people have begun to pay attention to the healthy development of the body. A series of fitness methods such as fitness, yoga, and marathon have become the icon of the new century. This shows that the overall healthy development of the body is inseparable from sports. The development of sports has received more and more attention from the government and the masses, and the economic development of the sports industry has made great progress. With the normalization of my country's economy, the development of the sports industry has gradually entered normalization. The “Thirteenth Five-Year Plan” for the Development of Sports Industry issued by the State Sports General Administration clearly pointed out the main tasks in expanding social supply, requiring the active promotion of “Internet + sports”. The State

Sports General Administration further emphasized the importance of the Internet for the development of sports, and the development of sports economy should keep up with the times and meet the needs of the masses. At the national level, enterprises and individuals are also encouraged to jointly provide the supply of sports public service information, and the development and utilization of mobile APP and other software are encouraged. Under the background of big data, the analysis of sports big data based on cloud platform and the research of its impact on sports economy is imminent.

Zhong et al. believe that, with the development of the Internet, recommendation systems are playing an increasingly important role in the field of big data processing such as e-commerce. Aiming at the big data processing problem in the recommendation system, Zhong et al. proposed a cloud platform collaborative filtering algorithm based on clustering and correlation. The algorithm uses k-medoids clustering and related multitree data structure to solve traditional the user-based collaborative filtering algorithm has been improved. First, the user-based cloud platform collaborative filtering technology is analyzed. On this basis, k-medoids clustering is used to propose a cloud platform collaborative filtering algorithm based on k-medoids, which can effectively solve the problem of data sparseness. Aiming at the problem of reduced recommendation accuracy caused by clustering technology, Zhong et al. propose a data structure of associative multitrees, which associates user information with neighbor information. It can be used to calculate the extended user project scores, making full use of the correlation between data on the cloud platform. This research is expensive and not suitable for popularization in practice [1]. Liu et al. conducted a lot of experiments on the Ali data set on the Hadoop cloud platform and found that the virtual opportunities on the cloud platform are affected by various factors, which can cause performance degradation and downtime, thereby affecting the reliability of the cloud platform. Traditional cloud platform anomaly detection algorithms and strategies have defects in detection accuracy, detection speed, and adaptability. Liu et al. proposed an SOM-based dynamic adaptive virtual machine anomaly detection algorithm and proposed an SOM-based unified modeling method for machine performance in the detection area, which avoided the cost of modeling a single virtual machine and improved the cloud platform the detection speed and reliability of large-scale virtual machines. In the process of SOM modeling, the important parameters that affect the modeling speed are optimized, which significantly improves the accuracy of SOM modeling, thereby improving the accuracy of virtual machine anomaly detection. The practicality of this research is weak [2]. Tang et al. believe that data-intensive analysis is the main challenge in smart cities due to the deployment of various sensors everywhere. The natural features of geographic distribution require a new computing paradigm to provide location awareness and delay-sensitive monitoring and intelligent control. Fog computing extends computing to the edge of the network to meet this demand. In the study, Tang et al. introduced a hierarchical distributed fog computing

architecture to support the integration of a large number of infrastructure components and services in future smart cities; a case study was analyzed using a smart pipeline monitoring system based on optical fiber sensors and sequential learning algorithms. To detect events that threaten pipeline security, a working prototype was constructed to evaluate the event detection performance of 12 different events through experimental evaluation. This research lacks experimental data support [3].

The innovations of this paper are as follows: (1) it proposes to use Hadoop-based sports big data analysis cloud platform to conduct experiments; (2) it proposes the establishment of sports economic data under the background of big data; (3) it conducts sports big data analysis cloud platform module design.

2. Sports Big Data Analysis Based on Cloud Platform and Research Method of Its Impact on Sports Economy

2.1. Hadoop-Based Sports Big Data Analysis Cloud Platform

2.1.1. Hadoop System. The Hadoop system is a top open-source software project under the name of the Apache open source organization. It evolved from Google's GFS distributed file system and MapReduce parallel computing framework. It is committed to creating an open source, scalable, and distributed large data processing platform. The infrastructure is given in [4]. Hadoop can be deployed on one to thousands of ordinary computer nodes, using distributed file systems to provide large amounts of data storage and using parallel programming models to process and analyze large amounts of data stored in distributed file systems [5]. Each node in the Hadoop cluster provides local storage and local computing, and the local storage and local computing of all nodes are unified to form a larger and more efficient storage and computing cluster. In many companies, such as Facebook and Yahoo, Hadoop clusters with a scale of thousands of nodes are deployed to integrate and streamline the large amounts of data generated by the company every day [6].

2.1.2. Overview of Hadoop. As the originator of big data processing, Hadoop has formed a complete ecological chain. Hadoop is mainly composed of two parts, namely, HDFS and MapReduce. HDFS is a distributed file system implemented by Hadoop, and MapReduce is a distributed parallel computing framework implemented by Hadoop [7, 8]. The Hadoop ecosystem also includes the following commonly used software: Hbase, Hive, Pig, Mahout, Zookeeper, Flume, and Sqoop.

Hbase is a distributed database based on HDFS that evolved from Google's BigTable. It provides real-time access to big data. It retrieves data through the primary keys Key and Range, which are more suitable for storing loose data [7].

Hive is a data warehousing tool for Hadoop. It provides complete Sql query operation commands. At the same time, MapReduce tasks can be executed with Sql statements,

which greatly reduce the cost of programmers running MapReduce tasks. MapReduce tasks can be executed through Sql-like statements to further improve the scalability of the Hadoop system [9].

Pig is the scripting language of Hadoop, which can query and process the data structure in the program. Pig has two operating modes: local mode and MapReduce mode. In MapReduce mode, Pig can automatically optimize the MapReduce program to improve the efficiency of program operation [10].

2.1.3. HDFS Reads Files

- (1) The client initiates an RPC request to read data to the remote NameNode by obtaining an instance of DistributeFileSystem [11].
- (2) The NameNode responds to the request and returns file block information related to the file. The file block information is mainly the address of the data node where the file block is located [12].
- (3) The client obtains an FSDataInputStream instance, starting from the first data block, and calls its read() method to read the file block data in the nearest data node [5].
- (4) After the client finishes reading the data of the current file block, it closes the connection with this data node, and, at the same time, searches for the data node address corresponding to the next file block and starts the data reading process of the file block [13].
- (5) After the client finishes reading all the file block data contained in the target file, it calls the close() method in the FSDataInputStream instance to complete the file reading process [14].

2.2. Support Vector Regression

2.2.1. Support Vector Machine. Suppose there is a training set $h = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, $x \in R^n$, $y \in$

$(-1, +1)$, y is a class label, and x is a vector with n attributes [15]. There is a general form of linear discriminant function in two-dimensional linear space:

$$f(x) = \langle \omega, x \rangle + b. \quad (1)$$

The normalized form of the equation of the optimal classification line L is

$$\langle \omega, x \rangle + b = 0. \quad (2)$$

If you want to make the points in the training set as far as possible from the classifier, find an optimal classifier to maximize the blank area on both sides of it [16]. Then, the optimal classifier should satisfy the following formula:

$$y_i(\langle \omega, x_i \rangle + b) \geq 1, \quad i = 1, 2, \dots, n. \quad (3)$$

$$\min J(\omega) = \frac{\|\omega\|^2}{2}.$$

Find by Lagrangian function:

$$L(\omega, b, a) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^n a_i(y_i(\omega x_i + b) - 1)a_i. \quad (4)$$

Among them, a_i is the Lagrange multiplier, and its partial derivative is taken to be zero. The following formula exists:

$$\frac{\partial L}{\partial \omega} = 0 \longrightarrow \omega = \sum_{i=1}^n a_i y_i x_i, \quad (5)$$

$$\frac{\partial L}{\partial b} = 0 \longrightarrow \sum_{i=1}^n a_i y_i = 0.$$

According to the KKT condition and duality principle, solving the maximum interval can be transformed into an optimization problem of finding the following function:

$$W(a)_{\min} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j a_i a_j (x_i x_j) - \sum_{j=1}^n a_j, \quad (6)$$

$$\text{subject to } \left\{ \begin{array}{l} a_i \geq 0, \quad i = 1, 2, \dots, n \\ \sum_{i=1}^n a_i y_i = 0 \end{array} \right\}.$$

Get the optimal solution $a' = (a'_1, a'_2, \dots, a'_i)^h$, then the modulus of the optimal classifier ω' has the following formula:

$$\|\omega'\|^2 = 2\omega(a') = \sum_{SV} a'_i a'_j (x_i x_j) y_i y_j. \quad (7)$$

The finally obtained classification function has the following expression:

$$f(x) = \text{sgn} \left(\sum_{SV} y_i a'_i (x_i x) + b \right). \quad (8)$$

2.2.2. Support Vector Regression Algorithm. The idea of support vector regression is very similar to classification; only the loss function is introduced. The goal is to find a function $f(x)$ whose error with the objective function y_i in all training sets is at most θ , while ensuring that θ is as small as possible [17, 18]. The following formula exists:

$$f(x) = \langle w, x \rangle + b, \quad w \in X, b \in \theta. \quad (9)$$

Minimize experience risk:

$$\text{minimize } \frac{1}{2} \|\omega\|^2, \text{ subject to } \begin{cases} y_i - \langle \omega, x \rangle - b \leq \theta \\ \langle \omega, x \rangle + b - y_i \leq \theta \end{cases}. \quad (10)$$

The kernel function is introduced to solve nonlinear problems. The simplest way is to map the data to a higher-dimensional space to make the data linearly separable and then use linear regression to solve the problem in the new high-dimensional space [19]. In fact, the kernel function is a kind of mapping. For different kernel functions, the data set will be mapped to different data spaces, which means different transformation functions, thereby improving the performance of various kernel function methods [20]. Under normal circumstances, there are mainly the following kernel functions:

The linear kernel function is expressed as

$$K(x_i, x_j) = (x_i x_j). \quad (11)$$

The polynomial kernel function is expressed as

$$K(x_i, x_j) = (x_i x_j + c)^d, \quad d > 0. \quad (12)$$

The Gaussian kernel function is expressed as

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}, \quad \gamma > 0. \quad (13)$$

Among them, because the feature space corresponding to the Gaussian kernel function is infinite dimensional, a limited sample set must be linearly segmented in the feature space, so the Gaussian kernel function is the most widely used kernel function in support vector regression [21].

The method part of this article uses the above method to analyze the sports big data based on the cloud platform and study the impact on the sports economy. The specific process is shown in Figure 1.

3. Sports Big Data Analysis Based on Cloud Platform and Research Experiment on Its Influence of Sports Economy

3.1. Establish a Sports Economic Database

3.1.1. Data Needs and Sources. Raw data needs: relevant information and data of local sports economic parks; China map data; statistics of various provinces; data of the level of sports development of various provinces [22].

Data analysis software requires Excel 2016, ArcGIS 10.2, and SPSS 24.0.

Sources of data acquisition: National Sports General Administration; National Bureau of Statistics; local sports

bureaus; local statistical bureaus; National Basic Geographic Center; National Geographic Surveying Information Bureau, etc. [23].

3.1.2. Data Collection and Processing. Taking the sports economic parks in 31 provinces, autonomous regions, and municipalities directly under the Central Government (excluding Hong Kong, Macao, and Taiwan) as the survey object, the relevant information and data of local sports economic parks are collected through the State Sports General Administration, and the addresses, names, approval year, leading sports industry category, and other information are sorted into Excel 2020 software [24, 25]. 1 : 4 million map spatial data of China are obtained through the National Basic Geographic Center. Use the Baidu picking coordinate system tool to collect the specific latitude and longitude of the local sports economic park and sort it into the Excel 2020 software. The statistical data of various provinces, municipalities, and autonomous regions are obtained from the official website of the National Bureau of Statistics [26].

3.2. Design of Cloud Platform Module for Sports Big Data Analysis

3.2.1. Design of Experimental Cluster Management Module

(1). **Cluster Virtual Network Design.** Use Hadoop's Neutron component to design the cluster network required for the experiment. The normal configuration of big data components when the cluster is created requires a specific virtualized network to support, and the cluster operation also requires a correct network environment. If the network is not designed in advance, users need to learn the steps of using the Hadoop virtualized network before deploying the cluster and create a corresponding network for their own experimental cluster, which will increase the cost of learning Hadoop, which is not beneficial to the experiment. Therefore, the predesign of the virtual network can pave the way for the rapid acquisition of experimental clusters in the experiment. When monitoring virtual machines and user experiment behaviors at the same time, it is also necessary to obtain data through the virtual network.

(2). **Cluster One-Click Deployment Design.** The user experiment is carried out on a virtualized big data cluster. Therefore, the deployment of the cluster is an essential part of the experiment. The experimental cluster management module starts with the quick response and convenient operation to complete the cluster one-click deployment design, thereby simplifying user operations and shortening the time for users to deploy experimental clusters.

The general steps for native deployment of big data clusters on Hadoop are upload image → register image → create node group template → create cluster template → deploy cluster. Uploading the image uses the Glance component of Hadoop, and the last four steps are completed by the Sahara component. From uploading a mirror image to

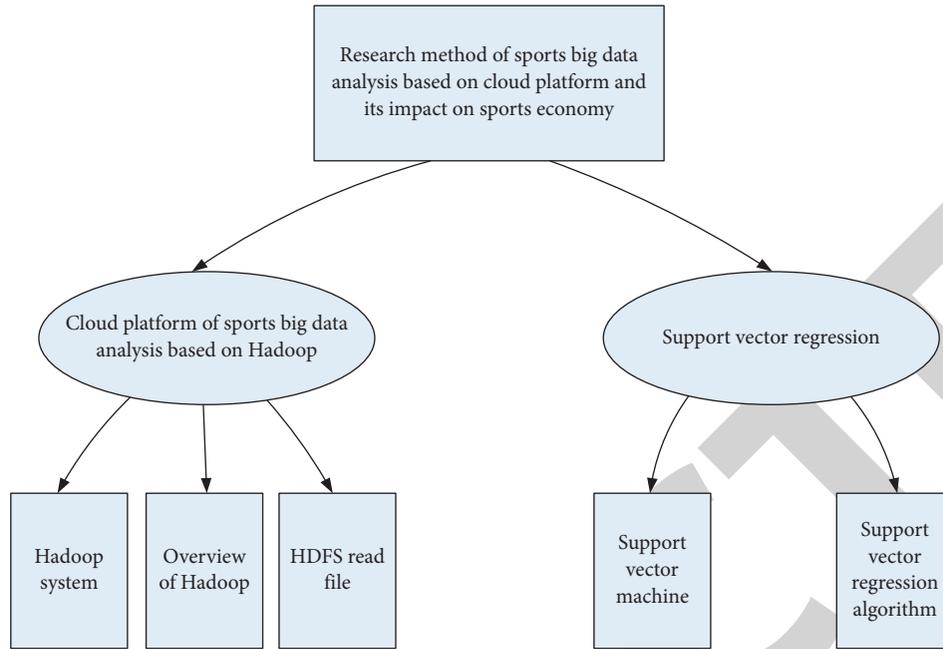


FIGURE 1: Part of the technical flow chart of this method.

creating a cluster template, it can be designed in advance based on the experimental content and provide a reusable template.

(3) *Import of Experimental Tools.* Prepare the tools and data sets that may be needed during the experiment in advance and save them to the server. Provide a one-click import button to remotely transfer from the server to the virtual machine.

(4) *Cluster Changes.* After the user successfully applies for the cluster, the cluster information and virtual machine information are displayed to the user, and operations for adding, deleting, modifying, and checking are provided. Including increasing or decreasing the number of nodes in the cluster can query cluster information through cluster naming or status, verify cluster status, and delete clusters owned by users.

3.2.2. *Platform Monitoring Module Design.* The granularity of Hadoop monitoring only reaches the resource usage allocated by the physical server, so it is necessary to design a fine-grained monitoring module for the user level including physical resource monitoring submodule, virtual machine resource monitoring submodule, and user experiment behavior monitoring submodule. The following three points need to be achieved: one is to refine the monitoring granularity and monitor from the three aspects of physical machines, virtual machines, and user experimental behaviors. The second is monitoring information processing. The resource monitoring information can be saved to the database, and historical warning records can be saved. Behavior monitoring, after categorizing users, is saved as a log file. The third is to isolate the monitoring information between users and show them to users.

For multinode physical server resource monitoring, the monitoring service is written as a script and automatically added to the Linux system service of each server. When the monitoring service transfers the information obtained by the query to the control node, it will bring its own host information, sort it according to the host name, and store it in the MySQL database. The control node will filter the information occupied by physical resources, filter out the time points when the resource occupancy rate exceeds the threshold, and store the resource shortage information in the historical warning file. Finally, the information occupied by physical resources is only displayed to the administrator on the interface after passing the authority judgment.

For virtual machine resource monitoring, a virtual machine information collector on the control node is used to query all virtual machines on all nodes. The information of all virtual machines is stored in the database of the control node. The computing node has no database to store the virtual machine information. If the virtual machine information collector is deployed in each computing node, the computing node also needs to obtain the virtual machine from the database of the control node. Machine information is used in order to collect resources. In order to reduce the number of remote database connections, a virtual machine information collector is deployed on the control node to complete the acquisition of all virtual machine information.

3.3. *Performance Testing.* After setting up the local cluster, perform the ParaView drawing test. First, perform a single-node test, and then perform a multinode test to compare the performance of cluster drawing. The tested data are VTK three-dimensional scalar data with sizes of 18.5 M, 291.8 M, 1425 M, and 8748 M, and the four sets of data are numbered data1-data4 in sequence.

3.3.1. Single-Node Test. This test uses a piece of test code written in Python to test. The measured parameters are the number of data points of the data to be tested, the number of data points drawn per second, and the total drawing time. To perform a single-node test, first start the ParaView server on node 1, then start the ParaView client on node 1, and connect to the server just started, and then you can run the test code to test the data.

3.3.2. Two-Node Data Test. To perform a two-node test, first start the ParaView server on node 1 and node 2 in MPI parallel mode, then start the ParaView client on node 1, connect to the ParaView server that has just been started, and run the test code to test the two-node data. Respectively draw data1-data4 just drawn on a single node on the ParaView client on node 1.

3.3.3. Four-Node Data Test. Finally, a four-node drawing test is performed. Start the four nodes of node 1 to 4 at the same time, then use MPI to run ParaView server in parallel, then start ParaView client on node 1 to connect to the server started, and run the test code to test the four nodes.

This part of the experiment proposes that the above steps are used for cloud platform-based sports big data analysis and research experiments on the impact on sports economy. The specific process is shown in Table 1.

4. Sports Big Data Based on Cloud Platform and Its Impact on Sports Economy

4.1. Development of Sports Economy Industrial Park

4.1.1. Time Development Characteristics. In 2006, the national sports industry base system was established; in 2011, under the guidance of the “Guiding Opinions of the General Office of the State Council on Accelerating the Development of the Sports Industry,” the State Sports General Administration issued the “National Sports Industry Base Management Measures (Trial)” to further clarify the national sports industry. The concept and type of the base standardize the management of the base; in 2014, the “Several Opinions of the State Council on Accelerating the Development of the Sports Industry and Promoting Sports Consumption” was formally promulgated, clearly proposing to “create a group of sports industries that conform to market laws and have market competitiveness.” In 2016, the State Sports General Administration issued a notice on further strengthening the construction of the national sports industry base and the “Thirteenth Five-Year Plan for Sports Development” put forward programmatic requirements and higher goals for the development of the sports industry base. In view of the national sports industry base’s leading role in the development of the regional sports industry and the policy needs of the development of the regional sports industry, various localities have begun to actively initiate the construction of local sports economic parks. The development of China’s local sports economic parks is drawn into a chart, as shown in Table 2 and Figure 2.

As can be seen from the chart, China was approved as the first batch of local sports economic parks in 2011. As of 2017, there were a total of 143 local sports economic parks. After that, it developed to 166 sports economic parks in 2018 and 176 in 2019 (a sports economic park). The growth rate generally showed a development trend of first increasing and then decreasing. The growth of local sports economic parks was mainly concentrated in 2016, 2017, and 2018.

4.1.2. Type Structure of Sports Economic Park. In the classification of the leading industries of local sports economic parks, we mainly refer to the National Sports Industry Statistical Classification issued by the National Bureau of Statistics in 2015. The National Sports Industry Statistical Classification defines the scope of the sports industry as sports management activities, sports competition performance activities, sports fitness, and leisure. There are eleven categories: activities, stadium services, sports intermediary services, sports training and education, sports media and information services, other sports-related services, sports goods and related product manufacturing, sports goods and related product sales, trade agency and rental, sports facilities construction. Statistics and sorting out the development types of national and local sports economic parks are shown in Table 3 and Figure 3.

It can be seen from the chart that the number of local sports economic parks belonging to the categories of sports management activities, sports competition performance activities, and sports fitness and leisure activities is the largest, with 95, 91, and 90 respectively, accounting for 11.85%, 11.35%, and 11.22 of the total (%); there are 86 sports venue services, accounting for 10.72% of the total; 68 sports intermediary services, accounting for 8.48% of the total; 63 sports training and education, accounting for 7.86% of the total; 64 sports media and information services, accounting for 7.98% of the total; other sports-related services, sports goods and related product manufacturing, sports goods and related product sales, trade agency and rental, and sports facilities construction accounted for 6.86%, 7.48%, 6.36%, and 5.74% of the total, respectively, 4.11%.

4.2. Sports Big Data Analysis Based on Cloud Platform and Its Impact on Sports Economic Development. (1) On the whole, the change trend of the total scale of national sports industry construction is basically consistent with that of the total scale under construction, both of which have basically maintained a rising trend. The specific situation is shown in Table 4 and Figure 4.

It can be seen from this that the development of the sports industry has always been on the rise during the period from 2012 to 2019, the sports economy is developing well, and the people are paying more and more attention to the importance of physical exercise and physical fitness.

(2) In 2008, the United States first broke out the sub-prime mortgage crisis and spread to the world. Affected by it, the vitality of my country’s real economy began to decline. The crisis also affected the development of the tertiary industry. At that time, the scale of the sports industry was

TABLE 1: Experimental steps in this article.

Sports big data analysis based on cloud platform and research experiment on its influence on sports economy	3.1	Establish a sports economic database	1	Data needs and sources
			2	Data collection and processing
	3.2	Design of cloud platform module for sports big data analysis	1	Design of experimental cluster management module
			2	Platform monitoring module design
	3.3	Performance testing	1	Single-node test
			2	Two-node data test
			3	Four-node data test

TABLE 2: Development of sports economic park.

Year of approval	Number of sports economic industrial parks	Growth	Growth rate
2011	9	9	
2012	9	0	
2013	15	6	66.67
2014	24	9	60.00
2015	40	16	66.67
2016	70	30	75.00
2017	143	73	104.29
2018	166	23	16.08
2019	176	10	6.02

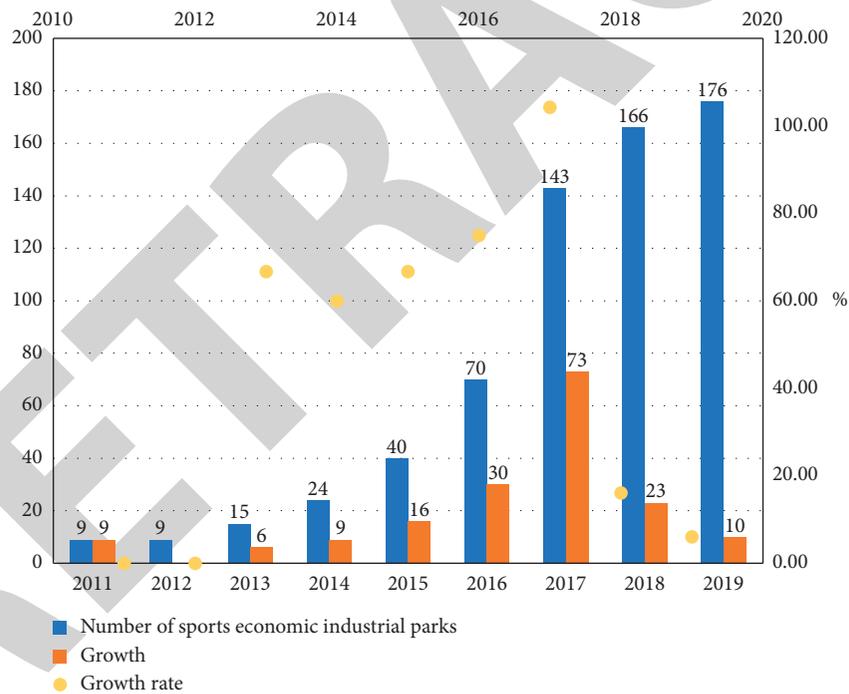


FIGURE 2: Development of sports economic park.

TABLE 3: Types of development of sports economic parks.

Type		Number of parks	Percentage (%)
1	Sports management activities	95	11.85
2	Sports competition performance activities	91	11.35
3	Sports fitness and leisure activities	90	11.22
4	Stadium services	86	10.72
5	Sports intermediary services	68	8.48
6	Sports training and education	63	7.86
7	Sports media and information services	64	7.98

TABLE 3: Continued.

Type		Number of parks	Percentage (%)
8	Other sports-related services	55	6.86
9	Sports goods and related product manufacturing	60	7.48
10	Sports goods and related products sales	51	6.36
11	Trade agency and rental	46	5.74
12	Sports venues and facilities construction	33	4.11

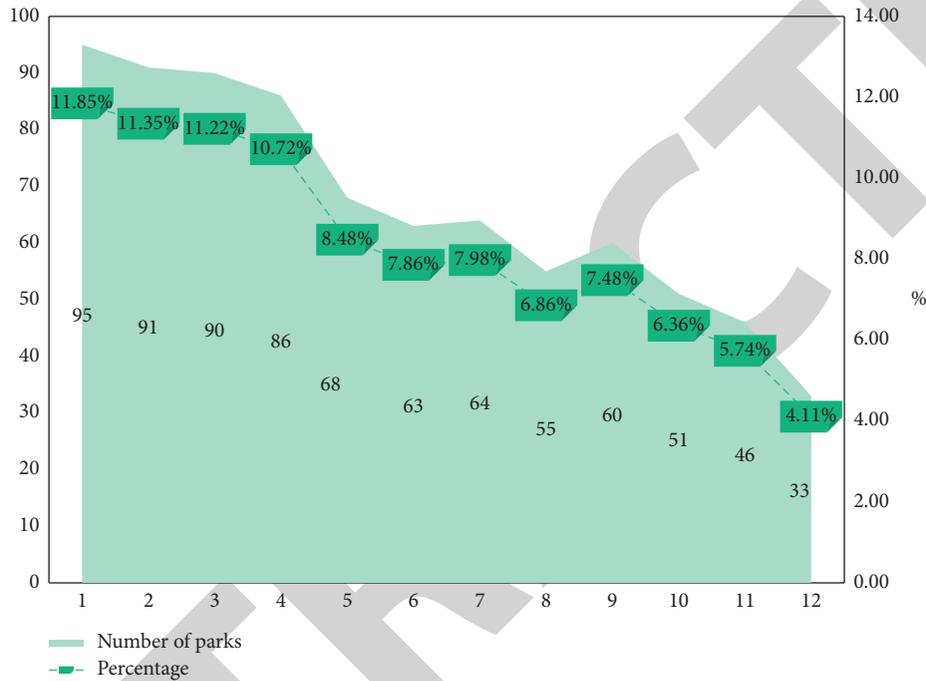


FIGURE 3: Development types of sports economic parks.

TABLE 4: Development scale of sports industry in recent years (unit: 100 million yuan).

Years	Sports industry scale under construction	Sports industry scale has been built
2012	2132.6	2791.7
2013	2167.4	2839.8
2014	2230.1	2978.4
2015	2287.5	3010.2
2016	2311.4	3233.4
2017	2379.6	3310.6
2018	2412.5	3419.1
2019	2420.7	3526.3

affected. In 2009, in order to stimulate the economic recovery and manage the economic crisis, the Chinese government launched a “4 trillion” large-scale stimulus policy. Driven by this, the vitality of sports-related companies has increased, and the total number of booths has rebounded sharply. Since then, it has basically maintained an upward trend. The specific situation is shown in Table 5 and Figure 5.

In 2019, the total number of booths in the sports industry-related commodity trading market reached 42,053, an increase of 1.05% compared with 41,611 in 2018. From the

perspective of the total number of booths in the sports economy the upward trend is basically maintained, but the increase is relatively small.

(3) The number of employees is also the most sensitive indicator reflecting the scale of the development of an industry or industry. This article uses the number of employees in the retail industry above the quota for sports goods and retail equipment to represent the number of employees in the sports industry, as shown in Table 6 and Figure 6.

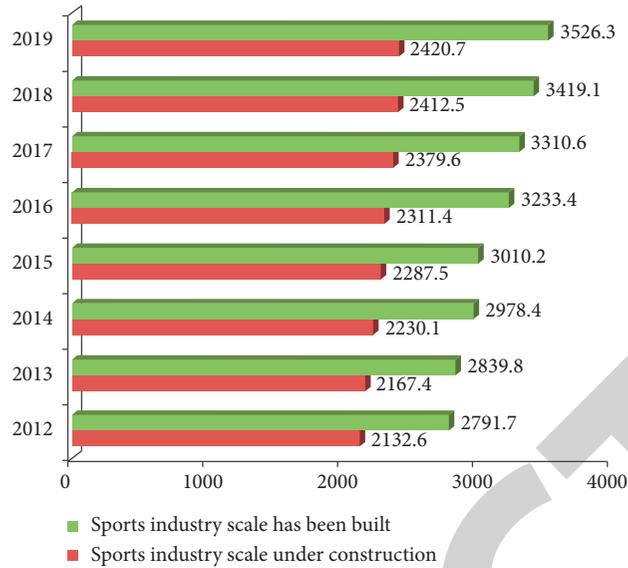


FIGURE 4: Development scale of sports industry in recent years (unit: 100 million yuan).

TABLE 5: Total number of stalls in the sports-related industry commodity trading market with a value of more than 100 million yuan in recent years.

Years	Total number of sporting goods booths	Total number of sports and entertainment booths	Total number of fitness exercise booths	Total number of booths for sports media and information services	Total number of booths for sports training and education	Total number of sports intermediary booths
2014	8411	8384	8124	6542	4589	5427
2015	8670	8706	8019	6431	4487	5306
2016	8331	8675	7941	6327	4931	5297
2017	8572	8531	7724	6540	4527	5441
2018	8431	8842	7609	6601	4761	5367
2019	8643	8967	7864	6527	4531	5521

It can be seen from Table 6 that from the retail perspective, the number of employees in the sports industry in my country has continued to increase, reaching 31,987 in 2019, an increase of 9.79% over the previous year; however, the annual growth rate of sports employees has fluctuated. Among them, the number of employees in 2014 has increased significantly, from 16,729 in 2013 to 19,947, an increase of 16.13%.

(4) According to the support vector machine regression algorithm described in the method section above, the original data are initialized through the calculation formula of the initial point zero image. The purpose of the initial value processing is to make each data correspond to the curve has a common point of intersection to facilitate comparison and analysis between various factors. Obtain the correlation between sports output value, sports consumption, number of sports industry employment, and sports big

data analysis, and the absolute correlation between sports big data analysis and sports economic development can be obtained, as shown in Table 7, Figure 7.

From the numerical calculation of the correlation degree in Table 7, the average correlation degree between sports big data analysis and sports economic development is 0.5155, which is relatively high. Appropriate sports big data analysis based on cloud platform plays a positive role in promoting the development of sports economy.

Based on the above statistical analysis on the scale of the sports industry, the scale of the national sports industry has been expanding, and the construction scale and the total scale under construction have shown a continuous expansion trend. The total number of stalls in the commodity trading market over 100 million yuan has also continued to rise. The number of employees in the retail

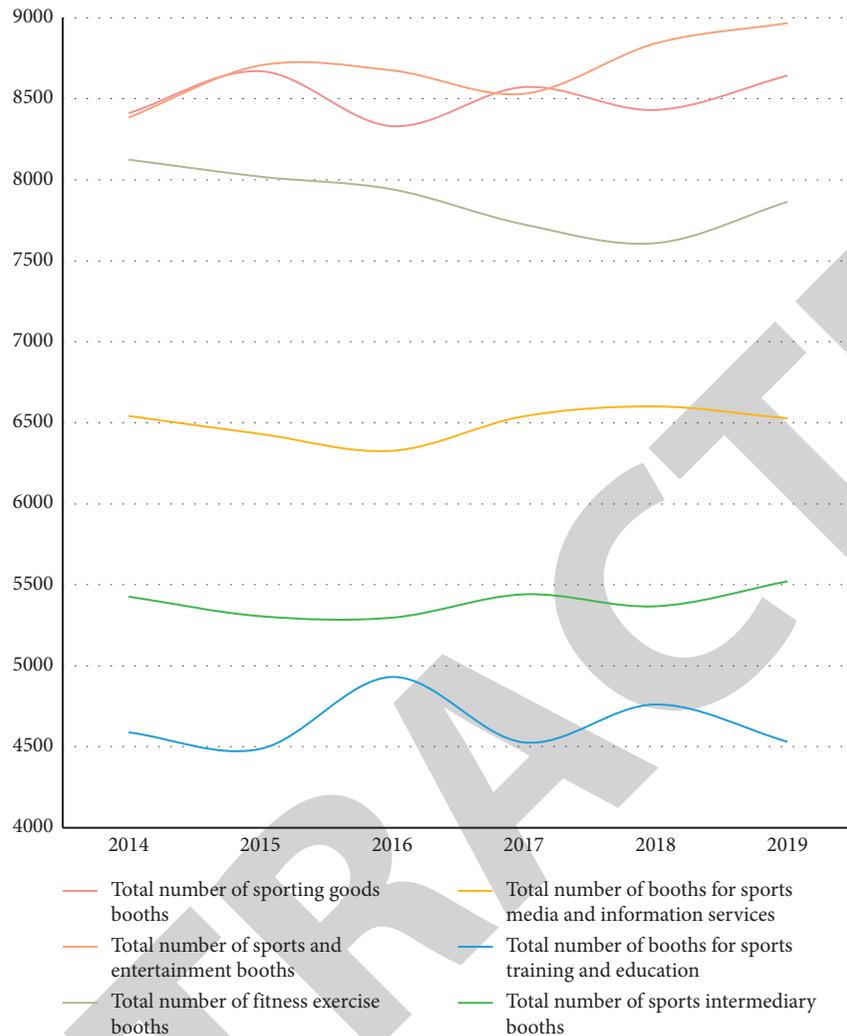


FIGURE 5: Total number of stalls in the sports-related industry commodity trading market with a value of more than 100 million yuan in recent years.

TABLE 6: The number of employees in retail industry above designated size for sporting goods and retail equipment in recent years.

Years	Number of employees	Growth rate (%)
2013	16729	11.17
2014	19947	16.13
2015	22021	9.42
2016	24437	10.97
2017	26014	6.06
2018	29135	12.00
2019	31987	9.79

industry above the designated size of retail equipment continues to increase. This shows that, with the development of the big data era, big data analysis has gradually been applied to the development of the sports economy. Through big data analysis, the current development situation, existing problems, and future development

directions of the industry can be better grasped, which is conducive to making more suitable. Decisions related to the development of sports economy promote the development of the sports economy industry and can also provide more employment opportunities and promote social development.

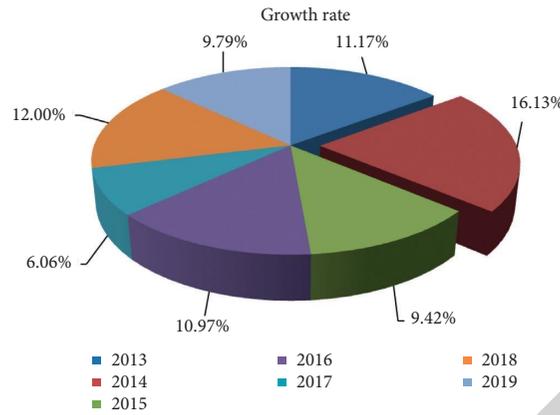


FIGURE 6: The number of employees in retail industry above designated size for sporting goods and retail equipment in recent years.

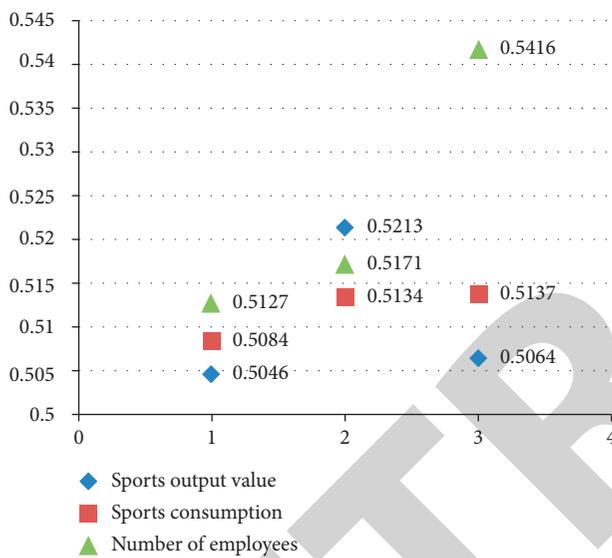


FIGURE 7: The absolute correlation between sports big data analysis and sports economic development.

TABLE 7: The absolute correlation between sports big data analysis and sports economic development.

	Total GDP of sports economy	Per capita GDP of sports economy	Per capita consumption expenditure of urban residents
Sports output value	0.5046	0.5213	0.5064
Sports consumption	0.5084	0.5134	0.5137
Number of employees	0.5127	0.5171	0.5416

5. Conclusions

Through the elaboration and comparison of data, information, big data and statistics research scope, characteristics, and analysis ideas, we believe that big data is information, but big data has given more meanings related

to change. It is the whole process of collecting, processing, converting, storing, transmitting, analyzing, algorithm, and application of data (all types of data) and even production and industrialization. This whole process not only changes the traditional data analysis, but also even will change people's work attributes and lifestyle. The basic purpose of data is to provide a basis for the collection and processing of information. The value of data is equivalent to the value of information. Big data has all the attributes of information. The key to data analysis is to discover new information from the complex data, thereby enhancing the understanding of things and making scientific and reasonable decisions.

With the rapid development of computer technology today, as computer hardware technology matures and the amount of information stored continues to increase, it becomes more difficult to screen, mine, and efficiently use large amounts of information. In the field of sports, with the rapid development of sports, the big data analysis of sports in my country is also facing the same problem. Big data analysis is conducive to discovering the potential connections and laws hidden in the massive data and is conducive to the deeper analysis of data, so as to play a certain role in guiding, predicting, and analyzing the development of sports economy.

Although of the thesis has conducted relevant research and analysis on the impact of sports big data analysis in the development of sports economy; these studies only consider the impact of relevant factors at the surface level and do not consider these methods in sports economic development statistics. Therefore, how to improve the coefficients in the models or formulas in these methods to make them more universal? This is also one of the contents worth studying in the future.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] X. Zhong, G. Yang, L. Li et al., "Clustering and correlation based collaborative filtering algorithm for cloud platform," *laeng International Journal of Computer Ence*, vol. 43, no. 1, pp. 108–114, 2016.
- [2] J. Liu, S. Chen, Z. Zhou, and T. Wu, "An anomaly detection algorithm of cloud platform based on self-organizing maps," *Mathematical Problems in Engineering*, vol. 2016, Article ID 3570305, 9 pages, 2016.
- [3] B. Tang, Z. Chen, G. Hefferman et al., "Incorporating intelligence in fog computing for big data analysis in smart cities," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2140–2150, 2017.
- [4] S. Liu, J. Tang, C. Wang, Q. Wang, and J.-L. Gaudiot, "A unified cloud platform for autonomous driving," *Computer*, vol. 50, no. 12, pp. 42–49, 2017.
- [5] J. Yang, H. Wang, Z. Lv et al., "Multimedia recommendation and transmission system based on cloud platform," *Future Generation Computer Systems*, vol. 70, pp. 94–103, 2016.
- [6] C. Cao, F. Cui, and L. Xu, "Research on intelligent traffic control model and simulation based on the Internet of things and cloud platform," *Journal of Computational and Theoretical Nanoscience*, vol. 13, no. 12, pp. 9886–9892, 2016.
- [7] W. J. Huang and H. L. Yi, "Advances in the application of cloud platform in diagnosis and treatment of OSA," *Lin Chuang Er Bi Yan Hou Ke Za Zhi=Journal of Clinical Otorhinolaryngology*, vol. 33, no. 4, pp. 310–312, 2019.
- [8] T. S. Priyeshkumar and D. Devaka, "A scalable and highly available distributed architecture for e-governance applications on private cloud platform," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 3, pp. 811–814, 2019.
- [9] X. Lu and S. Tan, "Cloud platform curriculum model design for nursing education under the integration of vocational education alliance background," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 8, pp. 239–248, 2017.
- [10] R. Xu, "Design of a bionic corn seed threshing machine based on rough sets under a cloud platform," *Agro Food Industry Hi Tech*, vol. 28, no. 1, pp. 2057–2061, 2017.
- [11] J. Wan, X. Yang, Z. Ren, and Z. Ye, "Performance evaluation and modeling method research based on IaaS cloud platform," *International Journal of Grid and Distributed Computing*, vol. 9, no. 10, pp. 141–152, 2016.
- [12] C.-m. Jiang, Y.-b. Li, and L. Zhi-Cong, "Review of the research on the optimization of the energy consumption of the cloud platform," *International Journal of Smart Home*, vol. 10, no. 8, pp. 241–250, 2016.
- [13] W. Yuan, P. Deng, T. Taleb, J. Wan, and C. Bi, "An unlicensed taxi identification model based on big data analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 6, pp. 1703–1713, 2016.
- [14] Z. Liu, K. K. R. Choo, and M. Zhao, "Practical-oriented protocols for privacy-preserving outsourced big data analysis: challenges and future research directions," *Computers & Security*, vol. 69, pp. 97–113, 2016.
- [15] H. A. Mahdiraji, E. Kazimieras Zavadskas, A. Kazeminia, and A. Abbasi Kamardi, "Marketing strategies evaluation based on big data analysis: a CLUSTERING-MCDM approach," *Economic Research-Ekonomska Istraživanja*, vol. 32, no. 1, pp. 2882–2898, 2019.
- [16] A. Ganesh, K. Vijiyakumar, K. Vijiyakumar, K. Premkumar, and P. Mathivanan, "To improve the accuracy in identifying breast cancer using various techniques of big data analysis," *International Journal of Advanced Research*, vol. 7, no. 10, pp. 491–498, 2019.
- [17] A. M. Abdelaziz, K. K. A. Ghany, T. H. A. Soliman, and A. A. E. M. Sewisy, "A parallel multi-objective swarm intelligence framework for big data analysis," *International Journal of Computer Applications in Technology*, vol. 63, no. 3, pp. 200–212, 2020.
- [18] C.-W. Yang, "A study on detecting the linked accounts by using big-data analysis," *Korean Journal of Financial Studies*, vol. 48, no. 1, pp. 73–103, 2019.
- [19] Y. Jiang, T. Wang, H. Zhao et al., "Big data analysis applied in agricultural planting layout optimization," *Applied Engineering in Agriculture*, vol. 35, no. 2, pp. 147–162, 2019.
- [20] Y. Liu, L. Pang, and X. Lu, "Click-through rate prediction based on mobile computing and big data analysis," *Ingénierie des systèmes d'information*, vol. 24, no. 3, pp. 313–319, 2019.
- [21] H. Li, "Estimation of stadium construction schedule based on big data analysis," *International Journal of Computers & Applications*, vol. 41, no. 3-4, pp. 268–275, 2019.
- [22] J. D. Yu, I. S. Lee, and I. S. Lee, "A prediction of stock price through the big-data analysis," *Journal of Society of Korea Industrial and Systems Engineering*, vol. 41, no. 3, pp. 154–161, 2018.
- [23] H. E. Naess, "Investment ethics and the global economy of sports: the Norwegian oil fund, Formula 1 and the 2014 Russian grand prix," *Journal of Business Ethics*, vol. 158, no. 2, pp. 535–546, 2019.
- [24] S.-H. Kim and E.-J. Jeong, "A study on the economic impacts of sports activities expenditure," *Korean Journal of Sport Management*, vol. 22, no. 5, pp. 1–13, 2017.
- [25] A. Gvercn, "The factors that affect the development of sports industry as an economy," *International Journal of Ence Culture and Sport*, vol. 4, no. 18, p. 515, 2016.
- [26] D. Cumming and S. Johan, "Venture's economic impact in Australia," *The Journal of Technology Transfer*, vol. 41, no. 1, pp. 25–59, 2016.