Wang X N and Dong G M. (2024) DESIGN OF SPORTS TRAINING REAL-TIME DATA ANALYSIS SYSTEM BASED ON RECONFIGURABLE DATA MINING TECHNOLOGY .Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 24 (96) pp. 68-85
DOI: https://doi.org/10.15366/rimcafd2024.96.005

ORIGINAL

DESIGN OF SPORTS TRAINING REAL-TIME DATA ANALYSIS SYSTEM BASED ON RECONFIGURABLE DATA MINING TECHNOLOGY

Xiao Ning Wang [1], Guo Ming Dong [2*]

[1] School of Physical Education, Shandong University, Jinan, Shandong 250061, China
[2] Postdoctoral Research Station of Suzhou University, Suzhou, Jiangsu 215021, China.
E-mail: 18366636288@126.com

Recibido 05 de octubre de 2023 **Received** October 05, 2023
Aceptado 05 de puede de 2024 **Accepted** May 05, 2024

ABSTRACT

The teaching materials generated by the promotion of education informatization are mostly idle and neglected, and the situation that education is decided entirely by experience is out of step with the new educational environment. In China, university sports are mainly conducted through physical education (PE) classes and amateur training, which are basically in a disorganized state. Due to the highly open nature of the modern sports system, changes in the way decisions are made in this information-based society have had a profound impact on the decision-making process in the field of sports. However, current school data processing is still at a rudimentary stage of data backup, querying and simple statistics, and does not provide an effective means to help teachers make decisions about school processes. Data mining (DM) is an in-depth approach to data analysis by mining useful information from large amounts of data, and this technique is now being used in an increasing number of fields. In this paper, we present the design of a real-time data analysis system for sports training based on DM technology, and use the corresponding mining tools of DM technology to discover relevant patterns or laws hidden in the data. Therefore, using the real-time data analysis system for sports training based on DM technology, useful information and patterns for improving examination performance can be obtained, which can improve targeted teaching methods and help students overcome learning difficulties, providing rational teaching. Synchronizing courses, establishing preparation, effectively guiding students in course selection, and improving course quality and educational effectiveness.

KEYWORDS: Physical training; Real-time data analysis system;

Reconfigurable DM.

1. INTRODUCTION

In school education, the quality of education affects the vitality of the school, whose fundamental task is to educate people, the main body of modern school development (Hui & Jin, 2021). Sport is a highly integrated discipline, including sports humanities, sports humanities and social sciences (Yin & Cui, 2021). In the past, the research field of sport application instant data analysis system was limited to sports training, sports evaluation, and sports management (Rajšp & Fister Jr, 2020). In recent years, under the guidance of the idea that science and technology is the main productive force, the awareness of science and technology development in sports departments at all levels in China has been increasing, and the level of scientific decision-making and management of sports has been improving (Bonidia et al., 2018). At the same time, as the scale of education expands, the training process generates a large amount of data, making it increasingly difficult for educational decision makers to understand and traditional data processing methods cannot cope with the accumulated data processing (Karachi et al., 2017). Considering this specificity, teachers must convert the non-percentage scores recorded in minutes, seconds, and meters to the percentage values of the students' national physical fitness standards at the end of the PE assessment (Afzali & Mohammadi, 2018). At the same time, each teacher must record the converted scores in the educational management software of the classroom (Zhang & Mao, 2021). Faced with the massive amount of data, the existing database management methods and data statistics methods are increasingly unable to adapt to the national proposal of "healthy exercise" and sports talent stratification (Hu, 2018). The data of various sports indicators cannot establish horizontal relationships (Aguilar et al., 2020; Wang, 2017). Therefore, if we do not adopt advanced management concepts, change the school philosophy, deeply understand the diverse needs of society, understand the characteristics of each student, and add a special specialty, it will be difficult to apply the original management methods and teaching methods, and sustainable development will become increasingly difficult (Gao et al., 2018). Existing processing methods only perform simple query and statistical processing of these data and do not perform in-depth analysis of the data to find out the factors affecting the students' exam performance (Wang & Liang, 2021). Data mining and data warehousing techniques achieve this better (Gamonales et al., 2021). DM has the dual property of problem identification and problem solving. It is the process of extracting hidden but potentially useful information and knowledge from large amounts of noisy real-world application data (Choi & Yoon, 2017). DM and knowledge discovery have taken data processing techniques to a more advanced stage (Kantilal & Sharma, 2020). It not only allows in-depth analysis of data, but also provides the necessary information in a timely and accurate manner, which helps to search deeply for

interrelationships between various elements in a huge amount of data (Park et al., 2020). As well as new rules to guide learning and classroom training in school sports. A large amount of sports data is being accumulated in the fields of sports competition and sports industry, and it has become an important task for sports researchers to use these data to discover important information that is useful but easily overlooked. The innovations of this paper are. (1) The purpose of the study of real-time data analysis system for PE in the basis of DM technology is to enhance the work efficiency and precision of teachers, thus freeing them from tedious work. (2) Utilization of DM allows for the identification of latent and relevant factors that impact teachers' teaching, thus providing suggestions for improving the quality of teaching. (3) DM technology is applied to a real-time data analysis system in the field of physical training, using a large amount of experimental data to establish a data warehouse that fits the physical quality of college students.

2. Related work

2.1 Sports real-time data analysis system

The survey of relevant data shows that the physical condition of Chinese youth has been declining continuously. How to apply data warehouse and DM technology in practical courses for information management and in-depth analysis of student sports data is an important topic for sports researchers to study and learn better. Some scholars have studied the instant data analysis system of sports data, the analysis of information of national fitness, and the intelligence of instant data analysis system, which is why more and more scholars are aware of the powerful support function of DM and want to apply it. Lili proposed a component structure for instant data analysis system, namely dialogue component, data component (database and database management system, model component, model-based and model-based management system (Lili, 2018). Pan proposed ID3 algorithm to mine and analyze data related to school administration and discover correlations between curricula to provide data for school decision making through an in-depth study of mining techniques reference (Pan, 2019). Rajput and Thakur proposed the structure of DSS systems, i.e., linguistic system, problem processing system, and knowledge system, which have the characteristics of "problem processing system" and "knowledge system" and influence to some extent (Rajput & Thakur, 2019). Li et al. used DM techniques to analyze a large amount of accumulated historical data. Factors affecting excavation, school sport research and education training were identified and the relationships between these factors were used to identify sport talents (Li et al., 2021). Patel and Shah proposed to combine instant data analysis system with expert system to play the role of DSS to reflect the decision making process comprehensively and solve the semi-structured and unstructured effectively. Initially, an intelligent instantaneous data analysis system was developed to solve the problem (Patel

& Shah, 2018; Yarahmadian et al., 2022). For the organic combination of all aspects of college students, scientific training theory and advanced training methods become possible to be applied to college sports training management, DM technology applied to college sports training support instant data analysis system is theoretically possible.

2.2 Data mining

DM is a new technology that unifies disciplines and a decision support process. In order to organically integrate all aspects of college students, scientific training theories and advanced training methods can be applied to college sports training management. DM mainly integrates database technology, artificial intelligence, machine learning, statistics, visualization techniques, pattern recognition and other information technologies to analyze data in a highly automated manner, perform inductive reasoning, discover potential patterns, and allow decision makers to adjust strategies and risk risk reduction and make the right decisions. Qin and Min use various DM methods such as classification, clustering, and sequential pattern analysis to mine information from online education databases and explore the relationship between student performance and learning behavior to help improve the quality of online education (Qin & Min, 2020). Using the DM technology provided by IBM, Xu et al. raised people's application of data, from low-level end-query operations, to providing beneficial decision support for business decision makers at all levels (Xu, 2017). Xu et al. used DM technology to implement personalized education in a distance education system so that the resources of the distance education system can be configured to support the personalized learning needs of students (Xu et al., 2014). Bandaru et al. investigated the application of principal component analysis techniques and Bayesian nearest neighbor algorithm in DM and used them in the evaluation of graduate students' comprehensive grades to reduce the number of indicators and information to be analyzed (Bandaru et al., 2017). Slater et al. studied faculty information databases and used association rule mining techniques to find relationships between various factors that influence academic development (Slater et al., 2017). Educational DM technology is a data processing technique that transforms raw data from various educational systems into useful information. The successful application of DM skills to sports training instant data analysis system will surely give you a satisfactory answer to this question.

3 Design Ideas of Real-time Data Analysis System for Physical Training Based on DM Technology

3.1 Design of real-time data analysis system based on DM

Nowadays, various schools in the country and abroad as well as various sectors of the society establish a series of online information services within

them which can also be called instant data analysis systems (Jassim & Abdulwahid, 2021). This system is designed to improve the efficiency and accuracy of PE teachers in higher education institutions, freeing them from tedious and boring work. In turn, it improves their management level teaching quality and meets the needs of three types of users: administrators, teachers, and students through a more accurate and complete information platform (Praveena et al., 2019). The SPSS analysis software integrates data collation, analysis, and result output, using a simple and convenient window interface, and the output can be tabulated graphically. The system allows the registration of information, inquiries, and online question and answer requirements, thus making it possible for internal personnel to learn various information related to the school or enterprise without having to leave home. Also when comparing and analyzing factor data of different dimensions, it is necessary to standardize them. the values of i and j factors in all samples are :

$$Z_{i,\ j} = \frac{X_{i,\ j} - \mu_j}{\sigma_j} \quad (1)$$

μ_j——Mean value of i th factor; σ_j——The standard deviation of j th factor.

First of all, the system adopts B/S (Browser/Server) model, which is a network structure model after the rise of WEB, and is a variation or improved structure of C/S structure. In the later stage of the algorithm search, the discovery probability is reduced in order to increase the convergence speed of the algorithm. Therefore, the discovery probability is improved as：

$$\text{pa}_t = exp(\ t/t\ cos(\ \text{pa}_{\text{begin}}{}_{max}{}_{max} \quad (2)$$

pa_t ——Discovery probability of t iteration; $exp(\ t/t\ cos(_{max}{}_{max}$ —— Function dynamic decreasing factor

Therefore, a mathematical model is required to abstractly describe the different network connections theoretically. Also, in order to facilitate analysis and processing in high level languages, it is necessary to find appropriate data structures from high level languages (Yang, 2021). It is very ineffective when used to find classes that differ greatly in size, or classes that present a non-convex shape, and it is sensitive to noise and outlier data, which can have a significant impact on the clustering effect even if only a small amount of that class exists. Judging the probability of classifying an instance into a class based on the value of the independent variable:

$$P(X) = P(Y = 1|X) = 1/(1 + e^{f(x)}) \quad (3)$$

X——Collection of data description attributes of; Y——Represents the category attribute of data. $P(X)$——Conditional probability

Second, the WEB browser is the main application software, and only one Browser, such as Internet Explorer, needs to be installed on the client machine as a unified client. The core part of system function implementation is centralized on the server, thus simplifying the development, maintenance and use of the system. It is necessary to determine whether test items exist under this test type, and if they do, they cannot be deleted, otherwise the deletion operation is executed. The program flow of function implementation is shown in Figure 1.

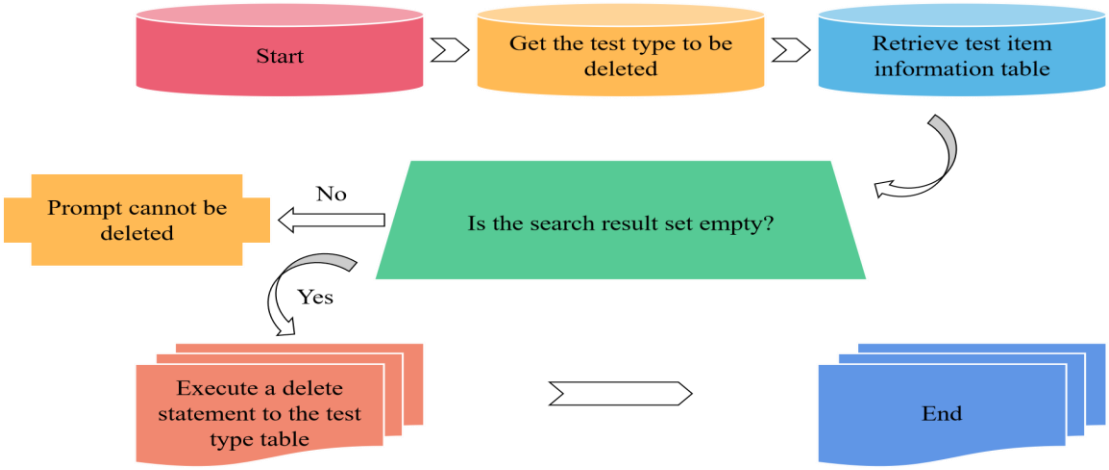

Figure 1: Process of test type deletion program

Under multiple test types, each test type occupies a certain percentage of the score, and the system should provide operations to add, modify, and delete these test types, and be able to set the weight of the score occupied by each type (Dai & Li, 2021). All tuning methods have limited effect on improving database performance if the database logic is poorly designed. The flow of DM is shown in Figure 2.

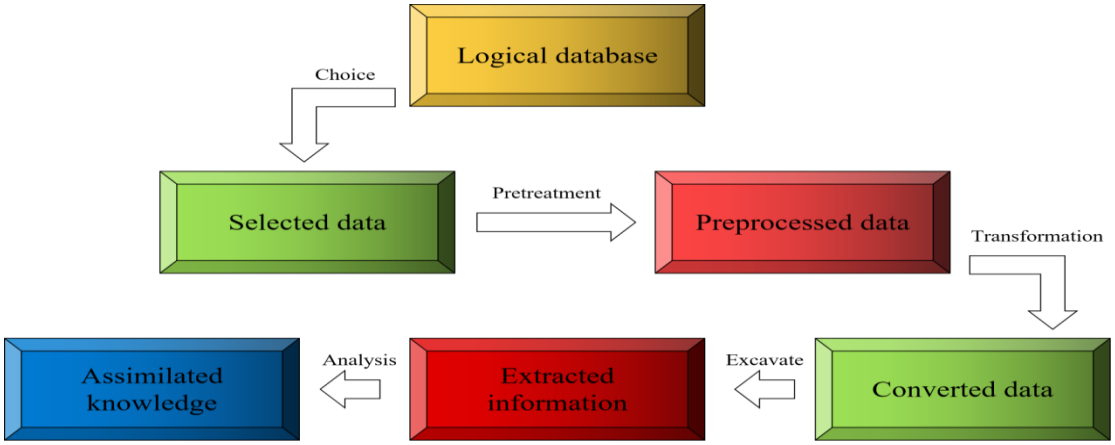

Figure 2: The process of DM

It provides the user with the operational status of the system so that the user can have a complete picture of its operational status, execution results and reasoning results. When describing the pattern of linking relationships between pages, a matrix representation can be used. Since its inception, DM technology has been oriented towards practical applications. The error is propagated in the reverse direction by constantly updating the weights and the bias of the prediction error of the representation network. The error is calculated as：

$$Err_j = O_j(1 - O_j)(T_j - O_j) \quad (4)$$

j——Output layer node; E_{rrj}——Error; O_j——Actual output; T_j——Based on the known target value of the given training sample

Finally, a three-tier B/S architecture is adopted, i.e., the client side uses a browser and the server side adopts a three-tier architecture design, which contains the representation layer, service application layer and data storage layer from top to bottom. When users log in to the system, they need to provide personal information such as user name and password, and only when they pass the verification can they log in to their respective system backend for related operations. In the case of merging two sets, it is only necessary to perform a bitwise summation operation on the binary sequences representing the two sets, and the binary sequences are used to perform the operation as follows:

$$N(h) = \frac{1}{0.77351}\sum_{i \in V} 2^{\frac{\sum_{i=1}^{k} b_i(i)}{k}} \quad (5)$$

i——Node; b_i——FMbitmask

Under each test type, new test items can be added, and these test items can be modified and deleted operations. The corresponding training programs are output according to the user's requirements, including individual training suggestions and collective grouping training programs. According to the central object of each cluster, i.e., the mean value of data records of that class, the size of each data record from those central objects is calculated; then the corresponding data records are reclassified according to the minimum distance.

3.2 Design of functional module of real-time data analysis system

DM query language, there are five basic DM primitives defined: task-related data primitives, kind of knowledge being mined primitives, background knowledge primitives, interest degree measurement primitives, representation of discovered patterns and visualization primitives (Liang, 2021). In this system, the three main categories of users targeted are administrators, teachers, and

students; therefore, for each category of users, their respective instant data analysis systems are created separately. Those functions that are mainly included in each subsystem are described below. The system functional module diagram is shown in Figure 3 below.

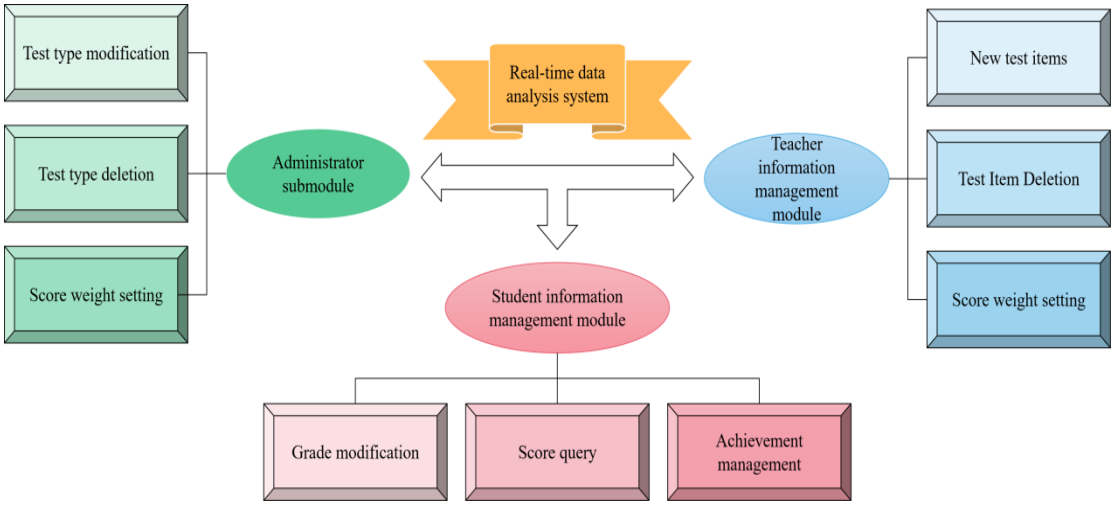

Figure 3: System function module diagram

The first is the administrator sub-module, which mainly includes permission management, teacher information management, student information management and grade analysis. The sub-functional modules can be processed as separate modules or their have special production cycles; the biggest advantage of this approach is the division of labor in software development, for example, in the target system, the student grade management module, can be divided into sub-functions such as student grade entry, grade modification, grade deletion and grade query, and different functions can be delivered to different developers for development. The mean substitution method is used to calculate the set-up to dig out the load data：

$$\overline{A}(d,t) = \frac{1}{n}\sum_{d=1}^{n} A(d,t) \quad (6)$$

$\overline{A}$(d,t)——Average value of load

The information needed for all decisions is provided and its systematic nature is emphasized. In contrast, the internal and external news and the experience and judgment adapted to individual decision making styles required by middle and senior decision makers are only partially provided, and it is impossible to reach the level of making the decision makers operate with ease. So data such as sports performance forms as well as physical fitness questionnaires of school students had to be collected and entered into them. This also shows that the solution of the linear programming problem is an integer solution. By transforming the diversity constraint into an upper and lower

bound constraint for the classification, the optimization problem is as follows：

$$max \sum_{i,j} w_{ij} x_{ij} \quad (7)$$

w_{ij}——ij th preference scoring matrix; x_{ij}——Final match.

The sub-functional modules can be circulated as separate products in the market; for example, in this system, the analysis of sports results can be completely independent of the sports results instant data analysis system itself. The collected source data is then subjected to a series of cleanups and transformations to remove invalid and erroneous data.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\sigma_i - \overleftrightarrow{\sigma_i})}{n}}, i = 1,2,\ldots,n \quad (8)$$

$\overleftrightarrow{\sigma_i}$——A predicted value in the set of predicted results; σ_i——The corresponding actual load value; n——The number of predicted values in the prediction result set

To reduce memory usage, we also negate the recursive approach and use a hierarchical reconstruction, using a Linked List to build FPTREE layer by layer, and when the related nodes are queued, only the information that establishes the association is retained, while other memory-occupying information is released, further reducing memory consumption. Finally, there is the student information management module, which mainly includes: admission results, class information, school study results. If possible, you can also add students' school rewards and punishments, social practice, education experience, graduation information and employment, etc. You can also add feedback from formerly employed students on their school learning situation. The sub-function module can be maintained as a separate module. When the amount of raw data is large, it is also possible to combine the division method so that an FP-Tree can be put into the main memory. The level-weighted support of the item set X={i_1, i_2,...,i_k} is defined as:

$$\text{Sup}_h = \max\{h_1, h_2, \ldots, h_k\} \times Sup(X) \quad (9)$$

Sup(X)——Traditional support count of item X; Max{ h_1, h_2..., h_k }——Weight of items.

Since the different functional subsystems can be logically independent, of course, independent maintenance can be achieved. Then a multidimensional data set based on this data source is established in SQL Server to facilitate the maintenance and analysis of the life cycle of the network service. mainly for

each period of the service, the system has to estimate the BER of the network service and retain the BER detection value. By calculating the correlation between different granularity N-gram features, the different granularity features are correlated with each other, and the corresponding weights are assigned to the different granularity features, and their representations are obtained through weighted summation.:

$$\alpha_i^t = \frac{exp(U_i^{t^T} u_w)}{\sum_t exp(U_i^{t^T} u_w)} \quad (10)$$

c_i^t——The i word representation of t feature

Therefore, the scattered data objects in the class are selected and later shrunk to the center of the class according to the shrinkage factor (which usually takes the value of a specific fraction). This allows DM systems to have a law to follow in terms of model definition and description, and various DM systems can share the model, and DM models can be nested in the middle of application systems. It is possible to make DM achieve the purpose of deep mining without separate development.

4. Application analysis of reconfigurable DM technology in real-time data analysis system of physical training

4.1 Data import DM algorithm processing and analysis

The data required for reconfigurable DM analysis is imported into this system from the designed and organized data warehouse for the corresponding DM operations, and the specific algorithms are processed as follows. First, the data described using the matrix method. A graph subject to the decision of its proximity live association can be represented in the form of a matrix. The network matrix can be formed by arranging each node in rows and columns, respectively. It is significant for the mining of attributes such as service name, rate, path and current BER, especially by analyzing the temporality of BER of optical channels, it can provide reference for the assurance of service quality of services and decision making of managers.

You can compare and analyze your academic performance with that of other students in your class, as well as with that of students in other classes and other majors. Understand your own learning deficiencies, so that you can improve your learning methods and create a free learning environment for yourself. Random sampling was performed by setting random seeds, and the valid information was divided into training and test samples by randomly selecting 80% and 20%, and the information gain and information gain rate of each attribute are shown in Figure 4 and Figure 5 below.

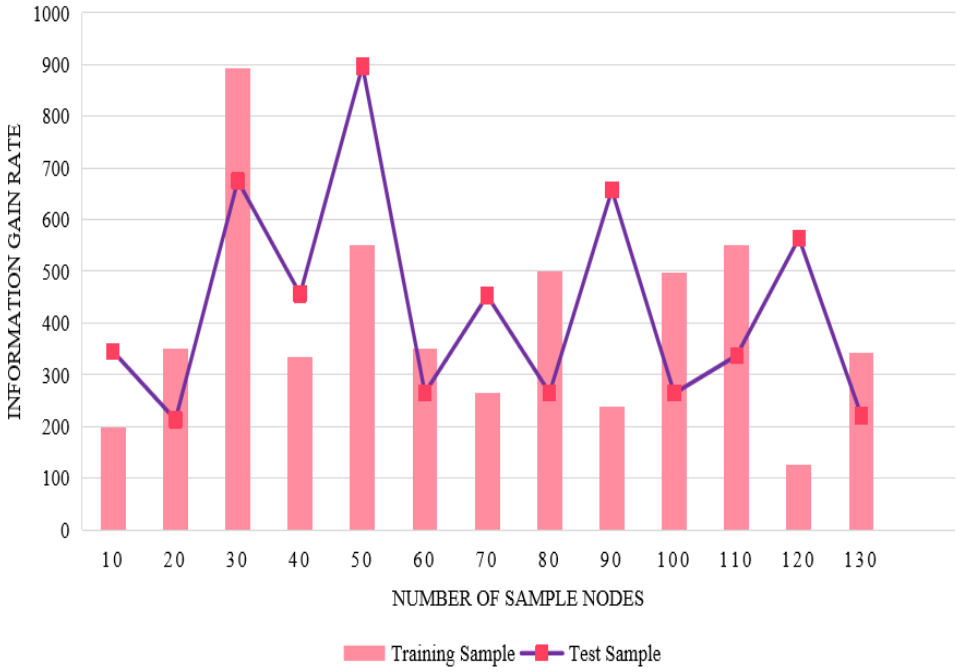

Figure 4: Information gain of each attribute

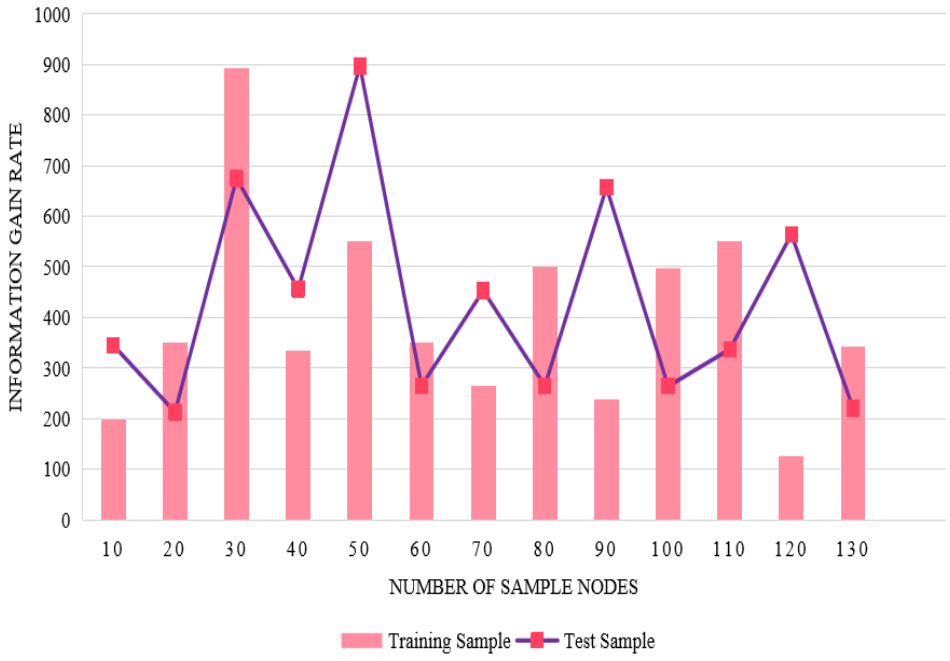

Figure 5: Information gain rate of each attribute

For a given new object, the COBWEB algorithm follows an appropriate path down and keeps modifying the count to find the best node that can classify the object. Second, all objects are placed in a cluster, and then gradually subdivided into smaller and smaller clusters until each object forms its own cluster, or until a user-defined stopping condition is reached. For example, a desired number of clusters is reached, or the distance between the two closest

clusters exceeds a certain threshold. The evaluation criterion for finding the best node is to temporarily place the object at each node position and then calculate the magnitude of the resulting classification utility, and the position with the maximum classification utility is a good choice for the object. A certain BER of the optical channel is selected for estimation, and the BER of the optical channel has the possibility of the same value, as shown in Table 1.

Table 1: Estimated data error rate and measured data error rate

BUSINESS NUMBER	**1520**	**1521**	**1523**	**1524**	**1525**	**1526**
ESTIMATED ERROR RATE	1.60e-5	1.60e-7	1.60e-3	1.60e-3	1.60e-9	1.60e-1
MEASURED ERROR RATE	1.60e-2	1.60e-6	1.60e-7	1.60e-4	1.60e-6	1.60e-2

First, the data is classified or divided into subsets of the data set, and then the decision tree algorithm is used to cycle through the subsets to generate a hierarchical decision tree from the downward direction. The decision tree has roots, nodes, and leaves of the tree, and the nodes and leaves in the tree are pruned according to the algorithm after the decision tree is constructed. Then the representative points are repeatedly replaced with non-representative points, and if the quality of clustering is improved, the replacement is saved. It is not only beneficial for one's own players to strengthen their weaknesses in sports training, and to give full play to their strengths when they face off with an individual player at their own limited level. Try to avoid the weaknesses that make you lose points, so that you can grasp the initiative of the situation and achieve better results.

Finally, again, it is recognized that in the process of clustering, if you want to perform efficient clustering, you must identify which are the important data that have a very important role in clustering. Which are the secondary data that can be treated as secondary to be clear separately. You can add school students for their own learning aspects at any time to evaluate the teacher's lectures feedback, to help teachers in the teaching methods to improve, but also for their own results analysis results to the teacher to give advice on teaching. It does not need to provide any additional a priori information beyond the set of data that the problem needs to deal with, and can discover potential and valuable knowledge from the information provided by the data itself, and obtain decision rules through knowledge reduction and dependency analysis. At the beginning, each object is considered to belong to a cluster, and then it is gradually merged to form larger clusters based on similarity until all objects are in the same cluster. Since the types of data objects stored in the database are transactions, text, etc., then it can be transaction database, text database, object-oriented database, etc.

4.2 Reconfigurable DM analysis in real-time train data analysis system

When applying the modularity principle, attention needs to be paid to information hiding and localization. For large-scale datasets, there may be huge correlation rules embedded. If mining is done blindly, not only is it inefficient, but also many irrelevant rules may be found. Therefore, it is necessary to train association rule mining analysis in instant data analysis system. First, using constraints, a specific association rule mining task can be controlled to a certain degree, thus making the mining work in the direction we expect. The capabilities of large data collections such as DISK-MINE, DRBFP-MINE and DM are measured by the composition of bar records generated from 30,000 different entries.

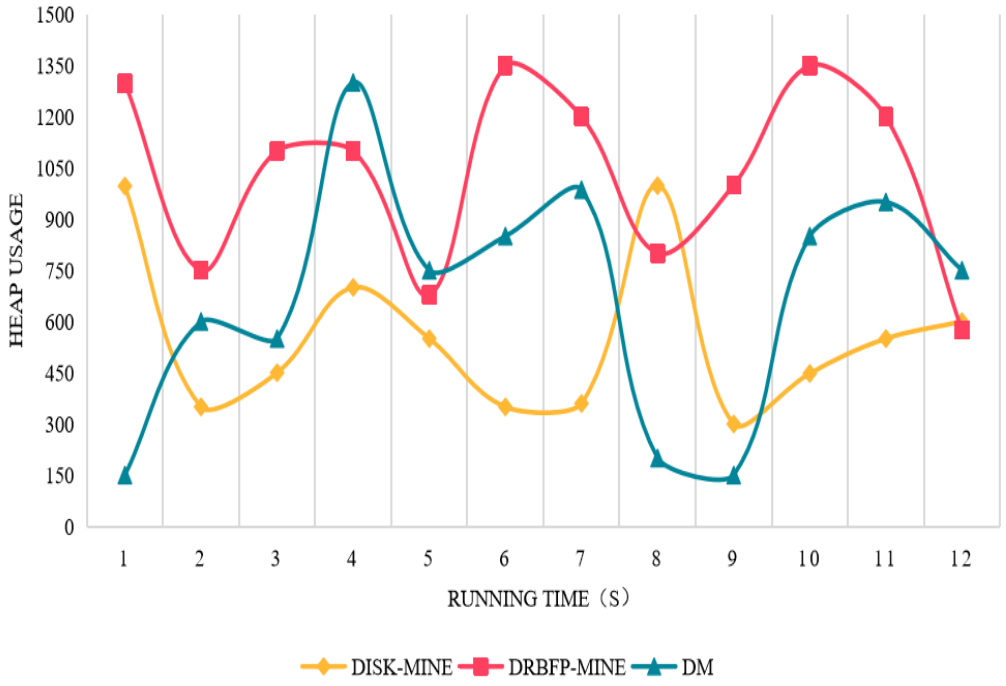

Figure 6: Comparative analysis of performance of large-scale data set algorithms

Similarly, different algorithms can be chosen for the same problem depending on the characteristics of different application domains. For example, the descriptive mining task is handled with conceptual abstraction such as data characterization and data differentiation and data summary outline processing. Instant score data belongs to structured big data, and random querying of web resource data may expose problems in data computation performance if SQL statements are used, which shows that new query techniques should be chosen. Secondly, constraints are used for querying frequent sets, and from the type of constraints used in the mining process, the constraints used for association rule mining can be classified as monotonicity constraints, anti-monotonicity constraints, transformable constraints and conciseness constraints. Thus this method visualizes the data rules, and its output is easier to understand, has

better accuracy and higher efficiency, and thus is more commonly used. There are often multiple levels of rectangular cells at different levels of resolution, and these different levels of rectangular cells form a hierarchy, i.e., each cell at a higher level is partitioned into multiple cells at a lower level.

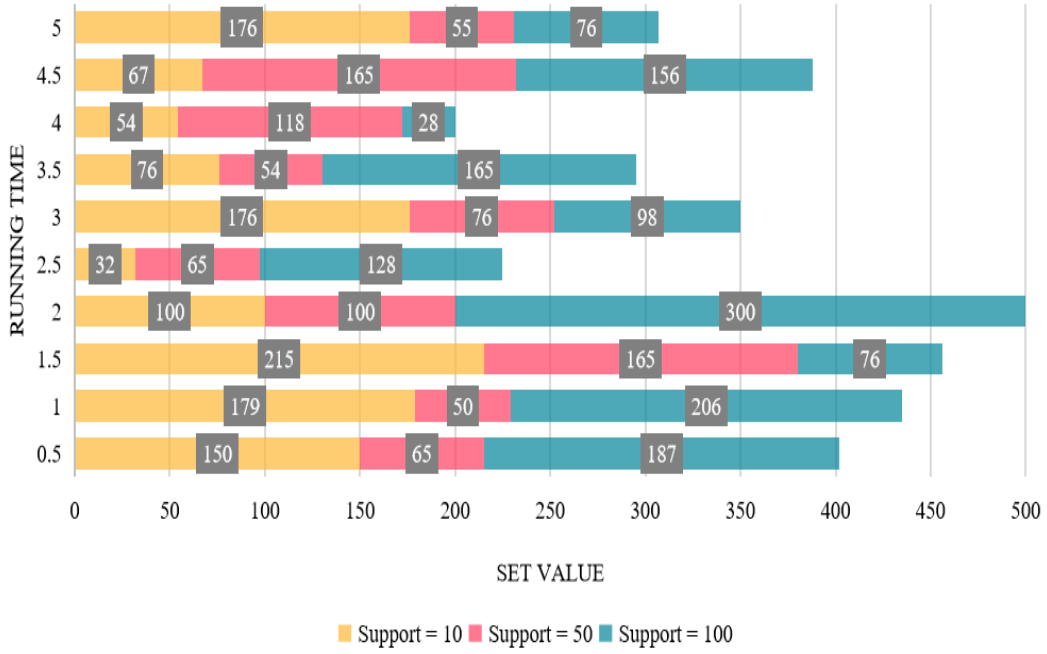

Figure 7: Running time comparison

In terms of data access, applications must stream access to datasets on top of HDFS. HDFS is suitable for handling large datasets with a focus on data throughput, so HDFS is designed with relaxed POSIX support to obtain greater data throughput. Finally, the merging of temporal intervals makes the fragments of temporal intervals that may be generated after filtering merge into a disjoint set of mining time zones. And memory-based mining is performed separately for each mining time zone to generate association rules. For the dataset 1, the mining results with DM and GF-DBSCAN are shown in Table 2 below.

Table 2: Running time of DM and GF-DBSCAN on Dataset 1

RUNNING TIME(S)	2000	4000	6000
DM	0.189	0.245	0.318
GF-DBSCAN	0.231	0.337	0.429

As can be seen from the above table, the average running time of DM in dataset 1 is reduced by 0.081 compared to GF-DBSCAN, so the advantage of DM in terms of time complexity becomes more and more obvious. For complex DM tasks, consisting of multiple data sources and DM modules, results need to be exchanged between the modules. The main components of the Predictive Model Markup Language PMML possess this flexible model exchange

capability and data format conversion capability, and enable the separation of the model from the data and tool parts. The remaining objects are then classified into the most similar clusters based on their similarity to the cluster centers. The mean value within the cluster is then used as the new centroid and the computation is re-iterated until the criterion function converges. The maximum available size of the program is set to demonstrate that the DM algorithm can mine association rules in large-scale datasets in a much smaller memory space. The CPU versus GC activity of the system is shown in Figure 8 below.

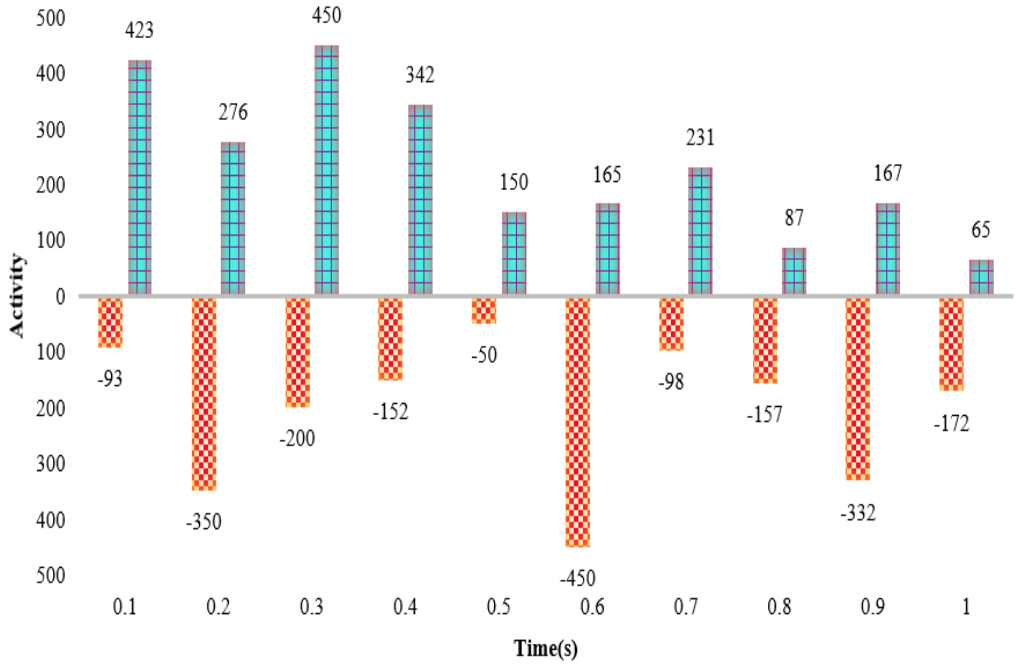

Figure 8: Comparison of CPU and GC activities

The statistical information stored in each cell provides the data in the cell without relying on the aggregated information of the query, thus making its computation process independent of the query. Its grid structure also facilitates parallel processing and incremental updates. The main tasks of the data warehouse include assisting users with the administration of the data warehouse server and handling their data query requests through the data storage architecture specific to the data warehouse. Since it comes from the field of statistics, the algorithm performs better and has higher robustness in the case of data sets less than a few hundred. And it can be dynamically allocated as the application load changes without downtime and restart, which makes configuration and management easier and more efficient.

5. Conclusions

PE is an important part of school education, and how to integrate computer and network technology into daily PE in schools is an inevitable requirement in today's information age. Reconfigurable DM Instant Data

Analysis Systems is a new discipline that takes research and development to the next level, helping sports coaches and athletes stay ahead in an increasingly competitive world. The need for data analysis and processing is considered from the perspective of undergraduate education, in the context of real-world schooling and how to determine the impact and contribution of each course to the ultimate level of professional competency throughout the overall professional education process at the university. Reconfigurable DM is not only considered by many researchers as one of the key research topics in pattern recognition and machine learning, but is an important area of research that can be lucrative for many in the field. The real-time data analysis system for PE based on reconfigurable DM technology proposed in this paper aims to improve the teaching methods and teaching quality of school teachers and to improve the learning style and learning efficiency of students.

REFERENCES

Afzali, G. A., & Mohammadi, S. (2018). Privacy preserving big data mining: association rule hiding using fuzzy logic approach. *IET Information Security*, *12*(1), 15-24.

Aguilar, B., Fang, P., Laubenbacher, R., & Murrugarra, D. (2020). A near-optimal control method for stochastic boolean networks. *Letters in Biomathematics*, *7*(1), 67. https://doi.org/10.30707/LiB7.1.1647875326.011975

Bandaru, S., Ng, A. H., & Deb, K. (2017). Data mining methods for knowledge discovery in multi-objective optimization: Part A-Survey. *Expert systems with applications*, *70*, 139-159.

Bonidia, R. P., Brancher, J. D., & Busto, R. M. (2018). Data mining in sports: a systematic review. *IEEE Latin America Transactions*, *16*(1), 232-239.

Choi, C., & Yoon, G. (2017). Cycling winner prediction model by using match information: application of decision tree analysis based on data mining. *The Korean Journal of Measurement and Evaluation in Physical Education and Sports Science*, *19*(4), 15-26.

Dai, X., & Li, S. (2021). Volleyball Data Analysis System and Method Based on Machine Learning. *Wireless Communications and Mobile Computing*(20), 1-11.

Gamonales, J. M., León, K., Rojas-Valverde, D., Sánchez-Ureña, B., & Muñoz-Jiménez, J. (2021). Data mining to select relevant variables influencing external and internal workload of elite blind 5-a-side soccer. *International journal of environmental research and public health*, *18*(6), 3155.

Gao, Q., Zhang, F.-L., & Wang, R.-J. (2018). Mining frequent sets using fuzzy multiple-level association rules. *Journal of Electronic Science and Technology*, *16*(2), 145-152.

Hu, Q. (2018). Research on the development of tennis sports based on data mining. *IPPTA*, *30*(8), 179-182.

Hui, D., & Jin, W. (2021). Design of real-time data analysis system for physical

training based on data mining technology. Journal of Physics: Conference Series,

Jassim, M. A., & Abdulwahid, S. N. (2021). Data mining preparation: process, techniques and major issues in data analysis. IOP conference series: materials science and engineering,

Kantilal, S. Y., & Sharma, Y. K. (2020). UNDERSTANDING ASSOCIATION RULE IN DATA MINING. *International Journal of Advanced Research*, *8*(6), 289-292.

Karachi, A., Dezfuli, M., & Haghjoo, M. (2017). Intelligent information and database systems. *Lecture Notes in Computer Science*, *5990*(6), 891-896.

Li, J., Lei, H., & Tsai, S. B. (2021). Online Data Migration Model and ID3 Algorithm in Sports Competition Action Data Mining Application. *Wireless Communications and Mobile Computing*(7), 1-11.

Liang, M. (2021). Optimization of Quantitative Financial Data Analysis System Based on Deep Learning. *Complexity*(1), 1-11.

Lili, X. (2018). Redundant processing algorithm of association rules based on hyper graph in data mining[J]. IPPTA. *Quarterly Journal of Indian Pulp and Paper Technical Association*, *30*(6), 683-689.

Pan, L. (2019). A big data-based data mining tool for physical education and technical and tactical analysis. *International Journal of Emerging Technologies in Learning (Online)*, *14*(22), 220.

Park, S.-U., Ahn, H., Kim, D.-K., & So, W.-Y. (2020). Big data analysis of sports and physical activities among Korean adolescents. *International journal of environmental research and public health*, *17*(15), 5577.

Patel, J., & Shah, P. (2018). A Review Approaches for Hiding Sensitive Association Rules in Data Mining. *INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING*, *6*(11), 920-924.

Praveena, K., Sirisha, G., Babu, S. S., & Rao, P. S. (2019). Efficient method in association rule hiding for privacy preserving with data mining approach. *Ingénierie des Systèmes d'Information*, *24*(1).

Qin, Y., & Min, G. (2020). Outlier data mining of multivariate time series based on association rule mapping. *International Journal of Internet Manufacturing and Services*, *7*(1-2), 83-96.

Rajput, V., & Thakur, P. (2019). Improving data analysis using data mining techniques for KSOMM and PAPLM. *International Journal of Advanced Research in Computer and Communication Engineering*, *8*(6), 28-32.

Rajšp, A., & Fister Jr, I. (2020). A systematic literature review of intelligent data analysis methods for smart sport training. *Applied Sciences*, *10*(9), 3013.

Slater, S., Joksimović, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics*, *42*(1), 85-106.

Wang, D., & Liang, F. (2021). Application of Artificial Intelligence and Big Data in Sports Event Service——Take Guilin as an Example. Journal of

Physics: Conference Series,

Wang, Z. (2017). 109. An Optimized Data Mining Algorithm Application in Volleyball Match Technique and Competition Tactics Analysis. *Boletín Técnico, ISSN: 0376-723X*, *55*(17).

Xu, L., Jiang, C., Wang, J., Yuan, J., & Ren, Y. (2014). Information security in big data: privacy and data mining. *IEEE access*, *2*, 1149-1176.

Xu, Y. (2017). 47. Research on the Framework of Data Processing and Sustainable Development of Sports Events Based on Data Mining Technology. *Boletín Técnico, ISSN: 0376-723X*, *55*(18).

Yang, N. (2021). Discussion on the application of enterprise business intelligence data analysis system. *Journal of Electronic Research and Application*, *5*(4), 23-26.

Yarahmadian, S., Oroji, A., & Williams, A. K. (2022). A hybrid differential equations model for the dynamics of single and double strand breaks of cancer cells treated by radiotherapy. *Letters in Biomathematics*, *9*(1), 141–158. https://doi.org/10.30707/LiB9.1.1681913305.295521

Yin, Z., & Cui, W. (2021). Outlier data mining model for sports data analysis. *Journal of Intelligent & Fuzzy Systems*, *40*(2), 2733-2742.

Zhang, S., & Mao, H. (2021). Optimization Analysis of Tennis Players' Physical Fitness Index Based on Data Mining and Mobile Computing. *Wireless Communications and Mobile Computing*(11), 1-11.